2015/47

Stationarity of Heterogeneity in Production Technology using Latent Class Modelling

Per J. Agrell and Humberto Brea-Solís

CORE

DISCUSSION PAPER

Center for Operations Research and Econometrics

Voie du Roman Pays, 34
B-1348 Louvain-la-Neuve
Belgium
http://www.uclouvain.be/core

CORE
Voie du Roman Pays 34, L1.03.01
B-1348 Louvain-la-Neuve, Belgium.
Tel (32 10) 47 43 04
Fax (32 10) 47 43 01
E-mail: immaq-library@uclouvain.be
http://www.uclouvain.be/en-44508.html

CORE DISCUSSION PAPER

2015/47

STATIONARITY OF HETEROGENEITY IN PRODUCTION TECHNOLOGY USING LATENT CLASS MODELLING

Per J. AGRELL[1] and Humberto BREA-SOLÍS[2]

November 2015

Abstract

Latent class modelling (LC) has been advanced as a promising alternative for addressing heterogeneity in frontier analysis models, in particular those where the individual scores are used in regulatory settings. If the production possibility set contains multiple distinct technologies, pooled approaches would result in biased results. We revisit the fundamentals of production theory and formulate a set of criteria for identification of heterogeneity: completeness (the inclusion of all data in the analysis), stationarity (the temporal stability of the identified production technologies), and endogeneity (no ad hoc determination of the cardinality of the classes). We also distinguish between the identification of a sporadic idiosyncratic shock, an outlier observation, and the identification of a time-persistent technology. Using a representative data set for regulation (a panel for Swedish electricity distributors 2000-2006), we test LC modelling for a Cobb-Douglas production function using the defined criteria. The LC results are compared to the pooled stochastic frontier analysis (SFA) model as a benchmark. Outliers are detected using an adjusted DEA super-efficiency procedure. Our results show that about 78 % of the distributors are assigned to a single class, the remaining 22 % split into two smaller classes that are non-stationary and largely composed of outliers. It is hardly conceivable that a production technology could change over this short horizon, implying that LC should be seen more as an enhanced outlier analysis than as a solid identification method for heterogeneity in the production set. More generally, we argue that the claim for heterogeneity in reference set deserves a more rigorous investigation to control for the multiple effects of sample size bias, specification error and the impact on functional form assumptions.

Keywords: Frontier analysis, latent class models, SFA, DEA, outliers, regulation
JEL Classification: D72, L51

1 Introduction

1.1 Background

Latent class (LC) models by Lazarsfeld and Henry (1968) have been introduced as promising solutions to the problem for regulatory models in for instance energy regulation, see (Cullmann, 2012; Agrell et al., 2014; Filippini and Orea, 2014; Llorca et al., 2014). The concept of an endogenous partition of the reference set ω in independent subsets, called *classes*, each represented by a separate cost function, is seducing and seems like a promising evolution of the state-of-the-art in regulatory benchmarking. Indeed, the numerical applications illustrating the cited works do indicate plausible and interesting classes that may correspond to differentiated production possibility sets. In this paper we compare LC models with non-parametric outlier detection methods. Both methodologies challenge the idea that all firms belong to a homogeneous production technology. LC models are useful to identify different groups of firms that implement a particular technology. Conversely, outlier detector methods deals with the inconsistencies or the prominence of a particular observation (Agrell and Niknazar, 2014) without presuming that these salient features are systematic in the sample.

We believe that the LC modeling is a welcome and potentially fruitful addition of to the regulatory toolbox. Nevertheless, the originality of this paper is that we define and test a set of conditions for the adoption of reference set partitioning for regulatory implementation on a panel data set for Swedish electricity distribution 2000-2006. Previous work has been concentrated at demonstrating examples of latent class models, or other clustering techniques, without framing the features of the method and the accompanying procedure in any regulatory setting. We show the empirical results for a real application, including highlights for the numerical and convergence problems that usually are implicit or omitted in similar work.

1.2 Outline

The outline of the paper is as follows: Section 2 describes the heterogeneity problem in production sets. Section 3 presents the underlying methods for latent class modeling and outlier detection. Section 4 presents the application and data from Swedish electricity distribution operations. The results are presented in Section 5 and the paper is closed with a discussion in Section 6.

[1]Louvain School of Management and and CORE, Université catholique de Louvain, B-1348 Louvain-la-Neuve, Belgium. E-mail:per.agrell@uclouvain.be.

[2]HEC Management School, University of Liege, B- 4000 Liege, Belgium. Email: ha.brea@ulg.ac.be.

The authors would like to thank Sebastien Van Bellegem et Maurice Queyranne for useful comments and suggestions. The work was supported by a grant from FNRS in the FRFC *Beyond Incentive Regulation* project.

The scientific responsibility is assumed by the authors.

2 Heterogeneity in production sets

Whereas the early work on heterogeneity (Førsund and Kittelsen, 1998) focused at the inclusion of environmental variables in models for electricity distribution, later work explored various statistical techniques to determine omitted variables (Farsi and Filippini, 2004). Applied post-calculation work for Data Envelopment Analsysis (DEA) initially utilized the second-stage approaches subject to extensive discussion in the literature (Simar and Wilson, 2007, 2011; Banker and Natarajan, 2011). In terms of model specification, the inclusion of a relatively complete[1] set of outputs and inputs has been presented as an (incomplete) remedy. Parametric solutions for this problem include the true fixed effects (TFE) and true random effects (TRE) formulations by Greene (2005). In TFE, we assume a set firm-specific time-invariant additive effects for each firm. This is less applicable to incentive regulation, where firms potentially change technology through investment or behavior during the period. In TRE, heterogeneity is captured by two elements, one time-invariant (as in TFE) and one time-variant component. The latter, assumed an i.i.d. random variable uncorrelated to the other variables, may absorb any persistent and repeated inefficiency as noise. Thus, the two parametric approaches are rarely directly applied in incentive regulation models. The conventional approach has instead been to split (partition) the overall set of comparators ω into smaller groups with higher comparability. Estimations of cost, efficiency and performance are then undertaken on the different subsets separately. We leave aside alternative ways of achieving such partitioning, focusing here at the latent class modelling.

2.1 Preliminaries

As mentioned above, a number of techniques could be used to partition a reference set through the identification of independent technologies. However, since we are interested in applications where this process is systematic and complete[2], we define three criteria for an adequate partitioning; completeness, stationarity and endogeneity.

Condition 1. *Completeness*
A partition $(\tilde{\omega}_1, \tilde{\omega}_2, ..., \tilde{\omega}_J)$ *is called complete if*

$$\cup_j \tilde{\omega}_j = \omega$$

Condition 2. *Stationarity*
A partition $(\tilde{\omega}_1, \tilde{\omega}_2, ..., \tilde{\omega}_J)$ *on* ω *is defined over a balanced panel over the horizon* $t = 1, ..., T$. *For a horizon* T, *an observation claimed to belong to a subset* $\varpi_{j,t} \subseteq \omega_t$ *should be endogenously assigned to the same subset* ϖ_j *for all the period* $1, ..., T$. *The partition is stationary iff for all i*

$$i \in \tilde{\omega}_{j,t} \text{ for some } j \Rightarrow i \in \tilde{\omega}_{j,\tau} \text{ for all } \tau \neq t$$

[1]Completeness here is validated here by standard regression techniques based on the assumption that a frontier model should also be an average cost model with expected statistical performance.

[2]The intended applications here could be in economic regulation, performance assessment for arbitrage or determination of reference sets for future allocation of budgets or tasks.

Condition 3. *Endogeneity*
A partition $(\tilde{\omega}_1, \tilde{\omega}_2, ..., \tilde{\omega}_J)$ *on* ω *is called endogenous is there is a predefined criterion to determine the number of non-empty classes* J.

3 Methodology

In this section, we describe first the the latent class (LC) models and afterwards the non-parametric outlier detection method. Our application of LC models is based on the works of Greene (2002, 2005) and Orea and Kumbhakar (2004). These authors adapted the Aigner et al. (1977) framework to include the possibility that firms form groups or classes with different production technologies. The goal is to measure efficiency without neglecting the unobserved heterogeneity of the sample. Measuring efficiency requires the definition of a production possibilities set:

$$S = \{(x,q)|x \text{ can produce } q\}. \tag{1}$$

Where S is the production possibility set, $x \in \mathbb{R}^N_+$ is the inputs vector and $q \in \mathbb{R}^M_+$ is the outputs vector. The set S is made of $i = 1...I$ observations (i.e. (x_i, y_i) pairs). Without loss of generality, we focus at input contractions. An input distance function is defined as the maximum radial contraction of a given input vector while keeping the firm within the production possibilities set;

$$D(x,q) = \max\{\phi|(x/\phi, q) \in S\}. \tag{2}$$

where $D(x,q) = 1$ if the firm is efficient and $D(x,q) > 1$ if it is inefficient. Currently, there are two approaches for building an empirical production possibility set S using distance functions. Data envelopment analysis, a non-parametric method developed by Charnes et al. (1978), is deemed as the regulators' preferred methodology (Agrell and Niknazar, 2014). It has the advantage of not imposing a functional form and the disadvantage of not considering statistical noise. The frontier is made of the best performers of the industry. The alternative is the parametric approach of which stochastic frontier analysis (SFA) is the main example. In this framework the notion of statistical noise is introduced at the expense of the flexibility provided in the previous approach. As we mentioned, LC models are applied in the SFA context.

Following Bogetoft and Otto (2011), we rewrite $D(x,q)$ using its property of homogeneity of degree one in inputs and by defining a variable $u \geq 0$ such that $D(x,q) = e^u$:

$$\begin{aligned} D(x,q) &= e^u, \\ x_N D(\frac{x}{x_N}, q) &= e^u, \\ -\ln(x_N) &= \ln D(\frac{x}{x_N}, q) - u, \\ y &= f(\cdot) + v - u, \end{aligned} \tag{3}$$

where the normalization factor is x_N, $y = \ln(x_N)$, f is a function that represents $lnD(\frac{x}{x_N}, q)$, $v \sim N(0, \sigma_v^2)$ is the normally distributed statistical noise and $u \sim N_+(0, \sigma_u^2)$ is the truncated normal inefficiency term where efficiency corresponds to $u = 0$. The function f could take different functional specifications like translog or Cobb-Douglas (Coelli et al., 2005). In addition to the normalized inputs $\frac{x}{x_N}$, outputs q and parameters β; it is customary to include as an argument a vector of controls z to reduce the effects of other confounding variables. Therefore equation (3) can be rewritten as follows:

$$y = f(\frac{x}{x_N}, q, \beta, z) + v - u. \tag{4}$$

This framework is the standard SFA formulation for a homogeneous technology. A relaxation of this assumption was proposed by Greene (2002, 2005) and Orea and Kumbhakar (2004). Under this new scenario the co-existence of different technologies is assumed:

$$y_{i,t} = f(x_{i,t}/x_{i,t,N}, q_{i,t}, z_{i,t}, \beta_j) + v_{i,t|j} - u_{i,t|j}, \tag{5}$$

where the firm i observed in period t using technology j. It is possible to exploit the panel data configuration of the dataset as proposed by Orea and Kumbhakar (2004) by assuming that the inefficiency term varies as $u_{i,t} = \lambda_{i,t}(\cdot) \cdot u_{i|j}$. However, in this study we treat the sample as a cross-sectional dataset. The same treatment was followed by Llorca et al. (2014) in order to provide more flexibility in the computations. Although it usually yields similar results, our main interest is precisely these subtle differences. The next step is maximizing the log-likelihood function associated with equation (5). The conditional log-likelihood function can be written as $LF_{i,j}(\theta_j)$ where θ_j corresponds to all log-likelihood parameters for the class j (Orea and Kumbhakar, 2004). Let $P_{i,j} \in [0,1]$ be the probability that DMU i belongs to class j, then the overall unconditional log-likelihood function is equal to:

$$\ln LF(\theta) = \sum_{i=1}^{I} \ln \left\{ \sum_{j=1}^{J} LF_{i,j}(\theta_j) P_{i,j} \right\}, \quad 0 \le P_{i,j} \le 1, \quad \sum_{j=1}^{J} P_{i,j} = 1. \tag{6}$$

$P_{i,j}$ is modelled as a multinomial logit. Although it is possible to model these probabilities as depending on time-invariant firm-specific variables, we choose to use the simplest approach. The estimation process ends with the computation of the posterior probabilities for class membership, using Bayes' rule. These probabilities are later used to classify companies into different technological groups. The conditional probability $P(j|i)$ is given by the following expression:

$$P(j|i) = \frac{LF_{i,j} P_{i,j}}{\sum_{j=1}^{J} LF_{i,j} P_{i,j}}. \tag{7}$$

The procedure for selecting the right number of classes requires estimating several LC models for different values of j and then applying statistical tests for comparing the intermediate results. Rawlings et al. (1998) discuss model selection criteria with respect to size proposing criteria such as Mallow C_p, Akaike information criteria (AIC) and Bayesian information criteria (BIC) (Rawlings et al., 1998). Llorca et al. (2014) and Greene and Hensher (2013) suggest that BIC is preferable due to its overall fit features. We choose the BIC criterion which is computed using the following formula:

$$BIC = -2 \ln LF + p \ln I, \tag{8}$$

where LF is the result of the maximization of the log-likelihood function, p is the number of parameters and I is the number of observations. The estimation with the lowest BIC determines the optimum number of classes. The test weights the overall fitness of the model with respect to the total number of parameters estimated.

3.1 DEA and super-efficiency

Distance functions can also be computed using non-parametric methods. Charnes et al. (1978) were the pioneers in the introduction of the data envelopment analysis (DEA) in the estimation of distance functions. Banker (1984) and Banker et al. (1984)[3] extended the analysis to multi-output settings and Färe et al. (1985) provided a more detailed treatment within the context of production theory. The departure point is equation (1), the definition of the production possibilities set and equation (2), the formal definition of an input distance function. In this particular context, we assume a common technology for all the firms in the sample. The input distance function for firm k at period t under r returns to scale, $D^t(x_{k,t}, q_{k,t})$ is then obtained by solving the following linear programming problem:

$$\begin{aligned} [D^t(x_{k,t}, q_{k,t})]^{-1} = \quad & \min_{\Theta, \lambda \in \mathbb{R}^I_+} \Theta \\ & \text{s.t.} \quad \Theta x_{k,t} \geq \textstyle\sum_{i=1}^{I} \lambda_i x_{i,t} \\ & \qquad q_{k,t} \leq \textstyle\sum_{i=1}^{I} \lambda_i q_{i,t} \\ & \qquad \lambda \in \Gamma(r) \end{aligned} \tag{9}$$

where $\Gamma(crs) = \mathbb{R}^I_0$,$\Gamma(vrs) = \left\{\lambda \in \mathbb{R}^I_0 | \sum_i \lambda_i = 1\right\}$. Efficiency is defined as the inverse of the input distance function $E^t(x_{k,t}, q_{k,t}) = [D^t(x_{k,t}, q_{k,t})]^{-1}$. Therefore $E^t(x_{k,t}, q_{k,t}) \leq 1$ and $E^t(x_{k,t}, q_{k,t}) = 1$ when the firm is efficient. One of the consequences of the DEA formulation is that the discriminatory ability may be low for low I and/or large models. Andersen and Petersen (1993) formalize an approach to better rank decision making units that were classified as efficient using the standard DEA method [4]. The authors propose to exclude the analyzed unit from the reference set. In this way, efficiency scores could be larger than one for some units, meaning that they are "super-efficient". Banker and Chang (2006) perform a simulation experiment and found that this methodology is better suited for detecting outliers than for ranking units. Formally, super-efficiency is obtained by modifying the linear programming problem (9) as follows:

$$\begin{aligned} E^{Super(k)}(x_{k,t}, q_{k,t}) = \quad & \min_{\Theta, \lambda \in \mathbb{R}^{I-1}_+} \Theta \\ & \text{s.t.} \quad \Theta x_{k,t} \geq \textstyle\sum_{i \neq k} \lambda_i x_{i,t} \\ & \qquad q_{k,t} \leq \textstyle\sum_{i \neq k} \lambda_i q_{i,t} \\ & \qquad \lambda \in \Gamma(r) \end{aligned} \tag{10}$$

[3] First mention of superefficiency was made in Banker and Gifford (1988).

[4] Banker et al. (1989) also use the super-efficiency method to detect outliers using a sample of hospitals

4 Data

We apply the model on panel data for electricity distribution system operators (DSOs) from Sweden for the period 2000 to 2006 obtain from the national regulatory authority, the Energy Market Inspectorate. This data form a relevant test case not only because they are audited and used in regulation, but also due to the plausible presence of heterogeneity in a complete sample from the fourth largest country in Europe with large variations in operating conditions, climate and urban density. Moreover, the sample contains operators with different governance structures, private corporations, public corporations, cooperatives and public utilities, see Agrell and Bogetoft (2010) for a discussion about the sector.

The initial unbalanced panel contains 277 operators and 1,454 observations. We apply several cleaning procedures in order to work with a balanced panel. All merged companies are treated as single entities. We eliminate operators from the sample that are not present in all years of the period. Operators with incomplete or odd information (i.e. no labor expenditure) are also removed. After all data treatment, the balanced panel dataset contains 118 different operators, in all 826 observations.

All the estimated frontier models have two inputs, two outputs and a subset of control variables. Physical labor (LABOR), measured as deflated labour expenditure[5] and the installed capacity of substations (TRAFO) are used as inputs. TRAFO is used as a proxy for capital[6]. We define four output variables: energy delivered at low voltage (DELLV), energy delivered at high voltage (DELHV), total energy delivered (DELTOT), and number of connected clients (CLIENTS). The control variables are injections of energy from decentralized generation (DGINJ), customer density (DENSITY, measured as number of low voltage clients per total circuit length at low voltage), length of low voltage overhead lines (LINELV), cable length at low voltage (CABLELV) and circuit length of high-voltage cables and overhead lines (CIRCHV). Table 1 contains the descriptive statistics of the dataset. The standard deviation with respect to the mean is quite high, meaning that a wide diversity of firms is represented in the sample.

5 Analysis

5.1 Latent class models results

We calculate[7] three different latent class models that are described in Table 2. The models differ in their outputs and control variables; although they share the same Cobb-Douglas functional form. The first model (A) corresponds to a pure energy model; type B and C include the number of

[5]Labor cost index for sector C+E (SNI2002) 1996-2006 and average monthly salary cost for employees (sector E, SNI92) for 2001, www.scb.se.

[6]Convergence problems arise when total capital was used as input; hence we decided to use the installed capacity of substations as a proxy. The correlation between these two variables is 0.97.

[7]All computations were done in Limdep. We used the library flexmix in R to validate our results. Flexmix does not have the frontier specification.

Table 1: Descriptive statistics, electricity DSO, Sweden 2000-2006, $n = 826$.

Variable	Code	Units	Average	Min	Max	StdDev
Substations	TRAFO	MVA	104.29	2.00	3,676.00	299.62
Labor	LABOR	man months	381.19	3.18	10,104.96	758.80
Energy Delivered LV	DELLV	MWh	303,847	1,933	10,052,301	834,278
Energy Delivered HV	DELHV	MWh	109,332	-	3,384,869	316,013
Energy Delivered Total	DELTOT	MWh	413,179	1,933	13,154,231	1,139,605
Total Connections	CLIENTS	#	21,627	290	739,374	63,408
Injections	DGINJ	MWh	10,978.62	-	476,188.00	38,618.96
Density	DENSITY	km^{-1}	17.99	5.49	56.28	8.34
Cables LV	CABLELV	km	926.77	11.00	32,030.00	2,655.22
Overhead lines LV	LINELV	km	461.04	-	20,703.00	1,808.43
Circuits length HV	CIRCHV	km	848.83	9.00	29,794.00	2,924.61

connections as output. All models have the same inputs: labor and the number of substations, which is the normalization factor. All variables were normalized with respect to the geometric mean and expressed in natural logarithms[8].

Table 2: Specification of SFA models A, B & C.

Model	Output variables	Control variables
A	DELLV; DELHV	DGINJ; DENSITY
B	DELTOT; CLIENTS	DGINJ; CABLELV; LINELV
C	DELTOT; CLIENTS	DGINJ; CIRCHV

Table 3 contains the standard SFA coefficients for the analyzed models obtained without the application of the LC method. There are some common features among these models. Labor elasticity for the average firm is below 15%; the technological regress is around 1.3% and decentralized injections seem to reduce input requirements. The three models present increasing returns to scale. When connections are included, they become the most explanatory output. The length of LV cables and lines increase input requirements, whereas HV circuits do not show up as significant. Finally, as expected, density reduces the input requirements.

The main results of the LC method are presented in Table 4. The BIC criterion was used for determining the number of classes. Three classes were identified in the three models. On average the first class represents 6% of the sample; the second class 81.7% and the third class 12.23%. There are several shared patterns across models that are worth emphasizing. The smallest class

[8] We tested a large number of specifications but all of them had convergence problems. The following alternative specifications were tested: 1. Exploitation of the panel data structure; 2. The use of unbalanced panel in order to gain more observations 3. Total capital as input 4. Operating expenditures as input instead of labor costs; 5 Inclusion of the cost of metering; 6. Specification of the control variables as ratios; 7. Use of the control variables as explanatory variables for the probabilities. 8 Elimination of the frontier component as in Llorca et al. (2014). 9. Inclusion of three outputs: energy delivered LV, energy delivered HV and number of connections

Table 3: Results for standard SFA models A, B & C.

Variable	A	B	C
LABOR	0.1179***	0.1142***	0.1319***
DELLV	-0.8189***		
DELHV	-0.1487***		
DELTOT		-0.1171***	-0.1508***
CLIENTS		-0.7711***	-0.7467
LABOR^2	0.2124***	0.1906***	0.2018
DELLV^2	0.0071		
DELHV^2	-0.0269***		
DELTOT^2		0.1534***	0.1592***
CLIENTS^2		-0.1682***	-0.1681***
TREND	-0.0129***	-0.0135***	-0.0154***
TREND^2	-0.0054*	-0.0033	-0.0027
DJINJ	0.0031**	0.0047***	0.0059***
DENSITY	0.0056***		
CABLELV		-0.0900***	
LINELV		-0.0255***	
CIRCHV			-0.1038
CONSTANT	0.2824***	0.1170***	0.1026***
LAMBDA	2.0910***	1.3418***	1.2322***
SIGMA	0.2122***	0.1729***	0.1655***

Notes: ***$p < 0.001$; **$p < 0.05$; *$p < 0.01$. $n = 826$

is always labor-intensive while the largest is capital intensive. The trend coefficient that captures technical change is negative in all the models across classes. Our assessment of the effects of decentralized generation is inconclusive; the signs differ between classes.

Expanding the analysis, we find that the pure energy model (A) (Table 4) has several interesting features. First, the second class exhibits constant returns to scale in the technology, where the elasticity of energy delivered for the average firm is close to 90% for low-voltage and about 10% for high-voltage, respectively. Furthermore, the firms in the first class show decreasing returns to scale and focus more on LV distribution while the firms in the third category are more concentrated in HV and present increasing returns to scale.

Type B and C models (Table 4) have some common characteristics. The number of connections is more relevant for the second and third classes, while energy delivered is more important factor for the first class. All estimated categories exhibit increasing returns to scale. LV Cables, LV Lines and HV Circuits increase input requirements.

5.2 Class characteristics and stationarity

After partitioning the sample into different classes, the next step is to describe their features. We compare the classes within each model by running hypothesis tests for the mean differences with respect to their characteristics. Table 5 reports the levels of significance (10%, 5% and 1%) for

Table 4: Results LC models A, B & C

	A			B			C		
Variable/Class	1 (6.4 %)	2 (77.4%)	3(16.2 %)	1 (7.0%)	2 (84.7%)	3 (8.2%)	1 (4.6 %)	2 (83.1 %)	3 (12.3 %)
LABOR	0.3869***	0.0016	0.1316***	0.3865***	-0.0004	0.1252***	0.3611***	0.0084	0.1209***
DELLV	-1.1055***	-0.8982***	-0.6106***						
DELHV	-0.0474**	-0.1060***	-0.2458***						
DELTOT				-0.4245***	-0.0613***	-0.0090	-0.5417*	-0.1206***	-0.1955***
CLIENTS				-0.4424***	-0.8825***	-0.6059***	-0.3689	-0.8266***	-0.6450***
LABOR^2	-0.0464	-0.0027	0.2444***	-0.0092	-0.0171	0.0958**	-0.1697	-0.0062	0.2440***
DELLV^2	0.0347	0.0085**	-0.0082*						
DELHV^2	-0.0127***	-0.0183***	-0.0422***						
DELTOT^2				-0.0761	0.0505***	0.0420*	-0.2882	0.0641***	0.1304***
CLIENTS^2				0.0955	-0.0769***	-0.083***	0.3243*	-0.0757***	-0.1714***
TREND	-0.0344***	-0.0065***	-0.0095**	-0.0245***	-0.0077***	-0.0091***	-0.0100	-0.0069***	-0.0214***
TREND^2	-0.0055	-0.0070***	0.0010	-0.0045	-0.0044**	-0.0017	-0.0120	-0.0063***	0.0001
DGINJ	-0.0085	-0.0023**	0.0121***	0.0076	-0.0011	0.0253***	0.0170	0.0005	0.0054**
DENSITY	0.0087**	0.0043***	0.0038**						
CABLELV				-0.1944***	-0.0313***	-0.2584***			
LINELV				-0.0758***	-0.0187***	-0.0663***			
CIRCHV							-0.1795***	-0.0486***	-0.1155***
CONSTANT	0.0030	0.2147***	0.2755***	-0.1166	0.0817***	0.1243***	-0.1182	0.1080***	0.1528***
SIGMA	0.1100***	0.0990***	0.0526***	0.0808***	0.0785***	0.0498***	0.1112***	0.1111***	0.0874***
LAMBDA	0.3001	2.5608***	0.0000	0.0000	0.7127	2.9507	0.0000	3.0585***	2.5782

Notes: $^{***}p < 0.001$; $^{**}p < 0.05$; $^{*}p < 0.01$. $n = 826$

these tests as well as the direction (larger or smaller). The results show that excluding the number of connections as an output variable reduces the discriminatory power of the LC model. Model A produces classes that in general have similar features. Conversely, classes in model B are very dissimilar between each other. In particular, class 1 is made of the largest entities; class 3 medium size distributors and finally class 2 contains the smallest units. Once the length of LV lines and cables are no longer control variables, the differences across classes are significantly reduced (model C).

There are several issues regarding the content of Table 5. First, the dissimilarities across models are noteworthy, highlighting the sensibility of the method with respect to the functional specification. This underlines the importance of a prior analysis of the functional form for the sector, potentially using other research or complementary data. Another striking finding in Table 5 is the importance of the variables capturing the influence of the firm size. Substations and labor are always relevant in the specification of the classes and the inclusion of the length of LV lines and cables increases the discriminatory power of the LC method. Hence, the results offer some support for classifying companies according to size.[9] Finally, we note the relevance of decentralized generation in the classification of the operators. Even though we cannot identify a clear pattern regarding decentralized injections, our results show that there is a large heterogeneity across firms that was not detected in the standard models.

[9] Ad hoc size classification is a common set partitioning in regulation.

Table 5: Results of the Test of Difference in means across classes, LC models A,B & C

	A			B			C		
	1-2	1-3	2-3	1-2	1-3	2-3	1-2	1-3	2-3
TRAFO	$+^{*}$			$+^{**}$		$-^{*}$	$+^{**}$	$+^{*}$	
LABOR				$+^{***}$			$+^{**}$	$+^{**}$	
DELLV						$-^{*}$			
DELHV			$-^{*}$	$+^{*}$	$-^{*}$	$-^{***}$			$-^{*}$
CLIENTS						$-^{*}$			
CLIENTSLV						$-^{*}$			
DGINJ				$+^{**}$	$-^{**}$	$-^{***}$	$+^{***}$		* _
CABLELV					$-^{*}$	$-^{**}$			
CABLEHV					$-^{*}$	$-^{**}$			
LINELV					$-^{**}$	$-^{**}$			
LINEHV					$-^{*}$	$-^{**}$			
PEAKLOAD						$-^{*}$			
CAPITAL					$-^{*}$	$-^{*}$			
REVENUE					$-^{*}$	$-^{**}$			
OPEX								$+^{***}$	$+^{***}$

Notes: + means that the class with the lowest number is the largest; - is the opposite

Returning to the preliminaries in Section 2.1, we review the critical assumption regarding the stationarity of the classes. Given the type of data set and the application, it heterogeneity in production technology would mean that there would be unobserved (excluded) environmental characteristics influencing the efficiency of the production process. One might imagine that for instance that operators in the arctic region have higher costs of perforing line maintenance due to temperature and access conditions. However, since the concession areas are stable, an operator exhibiting such conditions should be classified as accordingly throughout the period if the heterogeneity is stationary.

An alternative could be an idiosyncratic shock for an operator a specific period, e.g. a flooding or a restructuring of the load. Such influence could be identified as an anomaly in the sense of an *outlier* rather than as a heterogeneity in the production possibility set. The difference between the identification of the idiosyncratic outlier and the partitioning in technology sets is to be found in the time and unit persistence of the classification.

Table 6 shows that on average 83 firms out of 118 remain in the same category through time and 15 of them change classes only once. Therefore, 17% of the firms change classes twice or more. Given that the sum of the average size of the smaller classes is approximately 18%, the previous number is quite high. It is important to emphasize that less than 50% of the members of the first and third classes are stationary, meaning that they are mostly made by members that at some point belonged to other classes.

Table 6: Stationarity of classes across models, h = number of class changes.

	A	B	C	Average
Stationary class 1 share	40%	48%	37%	42%
Stationary class 2 share	71%	83%	78%	77%
Stationary class 3 share	37%	31%	41%	36%
# Stationary firms ($h = 0$)	75	90	84	83
# Non-stationary firms ($h \geq 1$)	43	28	34	35
# Single-change firms ($h = 1$)	19	12	14	15
# Volatile firms ($h \geq 2$)	24	16	20	20

5.3 Non-parametric outlier detection methods

In the previous section, several indications were given to support the idea of the smaller classes being essentially composed by outliers. The next step is compare these classes with our super-efficiency estimates to verify whether there indeed is such a correspondence. Given that the super-efficiency methodology neglects the inefficient outliers as well as producing many observations above one, we follow Agrell and Niknazar (2014) and implement the methodology used by the German regulator to trim down the number of ”upper outliers”. For our case, we do not only identify the ”extremely” super-efficient DMUs but also determine those who are extremely inefficient. A DMU is extremely super-efficient if:

$$E^{Super(k)}(x_{k,t}, q_{k,t}) > q(0.75) + \eta(q(0.75) - q(0.25)), \tag{11}$$

for a given η where q is the quintile of the distribution of the super-efficiencies (i.e. $q(0.75)$ is the third quintile). Analogously, for the case of extremely inefficient firms, we implement the following equation:

$$E^{Super(k)}(x_{k,t}, q_{k,t}) < q(0.25) + \delta(q(0.75) - q(0.25)) \tag{12}$$

for a given δ. Both η and δ can be obtained using a search process such that the number of observations that are considered extreme stabilizes. Figures 1 and 2 show this process. For the case of η the chosen number is 1, for δ is 0.75. Once the upper and lower outliers are identified the next step is to compare them with the LC model results.

Table 7 contains the results of comparing LC models with the outliers obtained through the super-efficiency method. In this table, all years and models are represented. For each year there are two columns; the first one shows the number of firms that fall into each class while the second informs about the percentage of those firms that have been identified as outliers. The last combined column ”sample” provides additional information. The first column counts the numbers of firms that have been categorized as class one or three at least once. Meanwhile, the last column corresponds to the percentage of these operators that were classified as outliers. For example, in the first row of Table 7, we find that 14 firms were classified into class 1 using model A; 86% of those are identified as outliers by the super-efficiency method at some time.

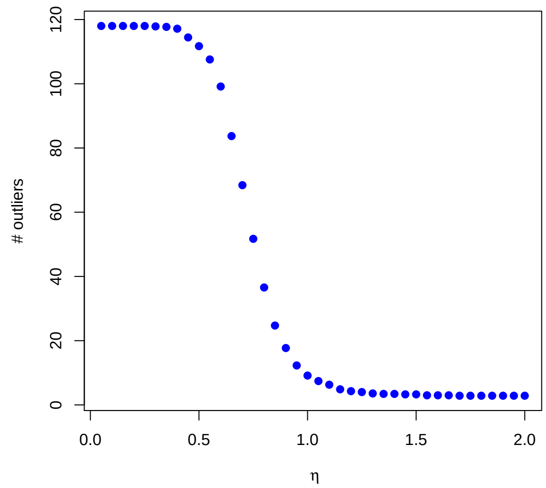

Figure 1: Number of super-efficient DMUs as a function of η.

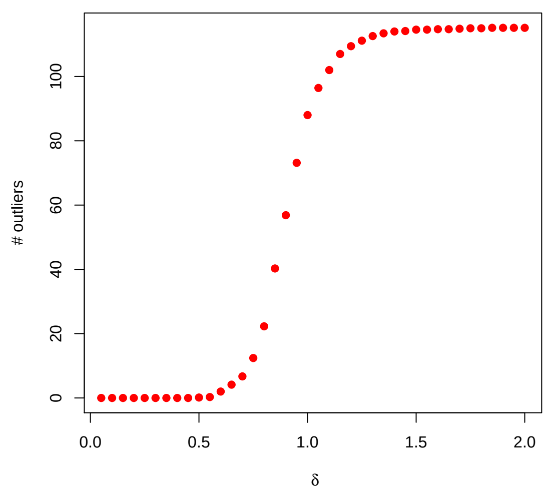

Figure 2: Number of super-inefficient DMUs as a function of δ.

An indisputable result in Table 7 is that the class 1 seems to be made of outliers in all models. On average, class 1 has less than seven members, thereof 83% classified as outliers. Model C has the highest discriminatory power in terms of establishing class 1 as a set of outliers. These

Table 7: Comparing LC models with outlier detection (superefficiency).

		2000		2001		2002		2003		2004		2005		2006		Sample*	
		#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%
	Class 1	7	57%	9	44%	6	83%	9	67%	8	88%	7	100%	7	86%	14	86%
A	Class 2	94	9%	86	8%	94	20%	87	15%	91	8%	92	18%	95	18%		
	Class 3	17	24%	23	13%	18	17%	22	27%	19	16%	19	11%	16	19%	41	37%
	Class 1	6	67%	7	57%	10	70%	9	89%	9	78%	9	89%	8	88%	16	75%
B	Class 2	104	8%	99	6%	99	16%	100	14%	102	8%	99	15%	97	15%		
	Class 3	8	50%	12	33%	9	44%	9	33%	7	29%	10	30%	13	31%	22	59%
	Class 1	2	100%	4	100%	5	100%	6	100%	7	86%	6	100%	8	88%	9	100%
C	Class 2	101	9%	93	9%	102	8%	101	9%	103	9%	99	8%	87	6%		
	Class 3	15	33%	21	19%	11	27%	11	9%	8	13%	13	15%	23	17%	34	35%

results provide evidence that the LC method is a partial substitute for outlier detection within the parametric framework. On the other hand, the conclusion is less straightforward with respect to class 3, the other small category. On average the percentage of class 3 members that are classified as outliers is higher than the those on class 2, but commonly below 50%. Class 3 may be capturing a persistent technological difference. Nevertheless, these results should be analyzed in conjunction with those provided in Table 6. On average the third class is the most volatile across models. Less than 40% of the firms belonging to this third category are stationary meaning that they remain in the same class throughout the seven-year period. Therefore, even though the third class seems not to be a ”pure outlier” category the lack of stationarity of the firms raises the question of whether in fact represents a ”true category of firms.”

In summary, given the difficulties of implementing a full LC frontier model with panel data structure (Llorca et al., 2014), due to convergence problems, it is important to report the stationarity for each class and their characteristics, specially if the methodology is going to be used to partition the reference set for regulatory purposes. The LC models used as a mechanism to identify technological groups could have the downside of classifying a firm into a wrong category in a semi-permanent way. It is advisable to run a non-parametric outlier detection method to contrast the results. Furthermore, clustering firms into groups can be used as complement rather than a substitute for a general benchmarking procedure.

6 Discussion

Several methods have been developed to address the problem of heterogeneity in frontier analysis. One approach is to partition the reference set into groups or classes, whose elements share similar features, using identical technology. We propose criteria to assess whether these partitions can lead to a meaningful differentiation between heterogeneous technologies and idiosyncratic shocks (outliers). We postulate that partitions must be complete, stationary, and endogenous.

Latent class models create endogenous partitions that could be the solution for the heterogeneity problem. In this study we use panel data from Swedish electricity distributor operators in order to assess the features of this approach. The first difficulty that we faced in our endeavour was the convergence problem that entails implementing a panel data stochastic frontier analysis using the latent class models. This computational problem has also been reported in other studies and it has occasionally been addressed by disregarding the panel structure of the dataset and/or the estimation of the firm-specific efficiency component.

To overcome the convergence problem, we estimated three latent class model specifications without the balanced panel assumption. These models vary in terms of outputs and control variables. Even though all the models have three classes with some common features, we identified large discrepancies. In summary, class formation seems highly dependent of modelling specification. Furthermore, some of the classes were not stationary casting doubts about the compliance of the partitions with respect to the neutrality condition. Another important finding is the overlap of the smallest classes with the outliers detected with the super-efficiency procedure. This finding weakens the model in terms of uniqueness. The latent class modelling instead of identifying technological clusters might just be a parametric outlier detection method. Of course, this coincidence in conjunction with the lack of stationary reinforces our reservations about the nature of these classes. Thus, latent class modeling is no panacea that can be uncritically applied to the incumbent incentive regulation approaches in use. Subdivision of the dataset in smaller subset aggravates the dimensionality, convergence and discriminatory problems previously mentioned. The confusion of outliers and stationary technologies may lead to arbitrary windfall profits for 'lucky' firms, at the expense of lowered procedural robustness and trust.

Despite its apparent shortcomings, latent classes modelling provided some interesting insights about the Swedish electricity distribution industry. One of the most important is the inconclusiveness about the influence of decentralized generation on the input requirements. In contrast with the standard SFA specification, the LC model detected some heterogeneity with respect to this variable. Given the current integration of renewables and decentralized generation, our results suggest that the cost impact may be ambiguous, contrary to the assumption in e.g. the German network regulation (Agrell and Bogetoft, 2015). Moreover, we do agree that latent class modelling has a potentially useful complementary role to play in the arsenal of regulatory analysis tools.

Future research should focus on trying to understand and reduce the convergence problems that arise in latent class estimation. From a normative viewpoint, it would also be interesting to investigate whether any set partitioning method, even integrated in a larger regime, could satisfy the conditions we have formulated. Another statistical issue to be resolved is whether the a priori assumption of the existence of multiple independent technologies can be validated in alternative manner, in particular when the functional form is unspecified.

7 References

References

Agrell, P. J., Bogetoft, P., 2010. Harmonizing the nordic regulation of electricity distribution. In: Energy, Natural Resources and Environmental Economics. Springer, pp. 293–316.

Agrell, P. J., Bogetoft, P., 2015. Theory, Techniques and Applications of Regulatory Benchmarking. Oxford University Press, Ch. 13, forthcoming.

Agrell, P. J., Farsi, M., Filippini, M., Koller, M., 2014. Unobserved heterogeneous effects in the cost efficiency analysis of electricity distribution systems, lecture notes in energy Edition. Vol. 54. Springer, Ch. 12, pp. 281–302.

Agrell, P. J., Niknazar, P., 2014. Structural and behavioral robustness in applied best-practice regulation. Socio-Economic Planning Sciences 48 (1), 89–103.

Aigner, D., Lovell, C., Schmidt, P., 1977. Formulation and estimation of stochastic frontier production function models. Journal of econometrics 6 (1), 21–37.

Andersen, P., Petersen, N. C., 1993. A procedure for ranking efficient units in data envelopment analysis. Management Science 39 (10), 1261 – 1264.

Banker, R., Gifford, J., 1988. A relative efficiency model for the evaluation of public health nurse productivity. Mimeo, Carnegie Mellon University.

Banker, R. D., 1984. Estimating most productive scale size using data envelopment analysis. European Journal of Operational Research 17 (1), 35 – 44.

Banker, R. D., Chang, H., 2006. The super-efficiency procedure for outlier identification, not for ranking efficient units. European Journal of Operational Research 175 (2), 1311 – 1320.

Banker, R. D., Charnes, A., Cooper, W. W., 1984. Some models for estimating technical and scale inefficiencies in data envelopment analysis. Management Science 30 (9), 1078–1092.

Banker, R. D., Das, S., Datar, S. M., 1989. Analysis of Cost Variances for Management Control in Hospitals. A Research Annual, Carnegie Mellon U, pp. 269 – 291.

Banker, R. D., Natarajan, R., 2011. Statistical tests based on dea efficiency scores. In: Handbook on data envelopment analysis. Springer, pp. 273–295.

Bogetoft, P., Otto, L., 2011. Benchmarking with DEA, SFA, and R. Springer, New York.

Charnes, A., Cooper, W., Rhodes, E., 1978. Measuring the efficiency of decision making units. European journal of operational research 2 (6), 429–444.

Coelli, T., Rao, D., O'Donnell, C., Battese, G., 2005. An Introduction to Efficiency and Productivity Analysis. Springer US.

Cullmann, A., 2012. Benchmarking and firm heterogeneity: a latent class analysis for german electricity distribution companies. Empirical Economics 42 (1), 147–169.

Färe, R., Grosskopf, S., Lovell, C., 1985. The Measurement of Efficiency of Production. Studies in Productivity Analysis. Springer.

Farsi, M., Filippini, M., 2004. Regulation and measuring cost-efficiency with panel data models: Application to electricity distribution utilities. Review of Industrial Organization 25 (1), 1–19.

Filippini, M., Orea, L., 2014. Applications of the stochastic frontier approach in the analysis of energy issues. Economics and Business Letters 3 (1).

Førsund, F., Kittelsen, S., 1998. Productivity development of norwegian electricity distribution utilities. Resource and Energy Economics 20, 207–224.

Greene, W., 01 2005. Fixed and random effects in stochastic frontier models. Journal of Productivity Analysis 23 (1), 7–32.

Greene, W. H., 2002. Alternative panel data estimators for stochastic frontier models. Working paper, Department of Economics, Stern School of Business, NYU.

Greene, W. H., Hensher, D. A., 2013. Revealing additional dimensions of preference heterogeneity in a latent class mixed multinomial logit model. Applied Economics 45 (13-15), 1897 – 1902.

Lazarsfeld, P. F., Henry, N. W., 1968. Latent Structure Analysis. Houghton Mifflin, Boston.

Llorca, M., Orea, L., Pollitt, M. G., 2014. Using the latent class approach to cluster firms in benchmarking: an application to the us electricity transmission industry. Operations Research Perspectives 1 (1), 6–17.

Orea, L., Kumbhakar, S. C., 2004. Efficiency measurement using a latent class stochastic frontier model. Empirical Economics 29 (1), 169–183.

Rawlings, J. O., Pantula, S. G., Dickey, D. A., 1998. Applied regression analysis: a research tool. Springer Science & Business Media.

Simar, L., Wilson, P., 2007. Estimation and inference in two-stage, semi-parametric models of production processes. Journal of Econometrics 136, 31–64.

Simar, L., Wilson, P. W., 2011. Two-stage dea: caveat emptor. Journal of Productivity Analysis 36 (2), 205–218.

Recent titles

CORE Discussion Papers

2015/12 Mathieu PARENTI, Philip USHCHEV, Jacques-François THISSE. Toward a theory of monopolistic competition.
2015/13 A.B. TAYLOR, J.M. HENDRICKX and F. GLINEUR. Smooth strongly convex interpolation and exact worst-case performance of first-order methods.
2015/14 Christian HAFNER, Hans MANNER and Léopold SIMAR. The "wrong skewness" problem in stochastic frontier models: A new approach.
2015/15 Paul BELLEFLAMME, Nesssrine OMRANI Martin PEITZ. The Economics of Crowdfunding Platforms.
2015/16 Samuel FEREY and Pierre DEHEZ. Multiple Causation, Apportionment and the Shapley Value.
2015/17 Mehdi MADANI and Mathieu VAN VYVE. A MIP framework for Non-convex uniform price day-ahead electricity auctions.
2015/18 Yukihiro NISHIMURA and Pierre PESTIEAU. Efficient taxation with differential risks of dependence and mortality.
2015/19 Louis N. CHRISTOFIDES, Michael HOY, Joniada MILLA and Thanasis STENGOS. Nature or nurture in higher education? Inter-generational implications of the Vietnam-Era lottery.
2015/20 Philippe J. DESCHAMPS, Alternative lormulations of the leverage Effect in a stochatic volatility model with asymmetric leavy-tailed errors.
2015/21 Jean HINDRIKS and Yukihiro NISHIMURA. Equilibrium leadership in tax competition models with capital ownership: a rejoinder.
2015/22 Frédéric VRINS and Monique JEANBLANC. The Φ-Martingale.
2015/23 Wing Man Wynne LAM. Attack-dettering and damage dontrol investments in cybersecurity.
2015/24 Wing Man Wynne LAM. Switching costs in two-sided markets.
2015/25 Philippe DE DONDER, Marie-Louise LEROUX. The political choice of social long term care transfers when family gives time and money.
2015/26 Pierre PESTIEAU and Gregory PONTHIERE. Long-term care and births timing.
2015/27 Pierre PESTIEAU and Gregory PONTHIERE. Longevitiy variations and the welfare State.
2015/28 Mattéo GODIN and Jean HINDRIKS. A review of critical issues on tax design and tax administration in a global economy and developing countries
2015/29 Michel MOUCHART, Guillaume WUNSCH and Federica RUSSO. The issue of control in multivariate systems, A contribution of structural modelling.
2015/30 Jean J. GABSZEWICZ, Marco A. MARINI and Ornella TAROLA. Alliance formation in a vertically differentiated market.
2015/31 Jens Leth HOUGAARD, Juan D. MORENO-TERNERO, Mich TVEDE and Lars Peter ØSTERDAL. Sharing the proceeds from a hierarchical venture.
2015/32 Arnaud DUFAYS and Jeroen V.K. ROMBOUTS. Spare change-point time series models.
2015/33 Wing Man Wynne LAM. Status in organizations.
2015/34 Wing Man Wynne LAM. Competiton in the market for flexible resources : an application to cloud computing. 2015/35 Yurii NESTEROV and Vladimir SHIKHMAN. Computation of Fisher-Gale equilibrium by auction.
2015/36 Maurice QUEYRANNE and Laurence A. WOLSEY. Thight MIP formulations for bounded up/down tiems and interval-dependent start-ups.
2015/37 Paul BELLEFLAMME and Dimitri PAOLINI. Strategic promotion and release decisions for cultural goods.

Recent titles
CORE Discussion Papers – continued

2015/38 Nguyen Thang DAO and Julio DAVILA. Gender inequality, technologial progress, and the demographic transition.
2015/39 Thomas DEMUYNCK, Bram DE ROCK and Victor GINSBURGH. The transfer paradox in welfare space.
2015/40 Pierre DEHEZ. On Harsanyi dividends and asymmetric values.
2015/41 Laurence A. WOLSEY. Uncapacitated lot-sizing with stock upper bounds, stock fixed costs, stock overloads and backlogging: A tight formulation.
2015/42 Paul BELLEFLAMME. Monopoly price discrimination and privacy: the hidden cost of hiding.
2015/43 Pierre PESTIEAU and Gregory PONTHIERE. Optimal fertility under age-dependent labor productivity.
2015/44 Jacques DREZE. Subjective expected utility with state-dependent but action/observation-independent preferences
2015/45 Joniada MILLA, Ernesto SAN MARTÍN and Sébastien VAN BELLEGEM. Higher education value added using multiple outcomes.
2015/46 Helmuth CREMER, Pierre PESTIEAU and Kerstin ROEDER. Social long-term care insurance with two-sided altruism.
2015/47 Per J. AGRELL and Humberto BREA-SOLÍS. Stationarity of Heterogeneity in Production Technology using Latent Class Modelling.

Books

W. GAERTNER and E. SCHOKKAERT (2012), *Empirical Social Choice*. Cambridge University Press.
L. BAUWENS, Ch. HAFNER and S. LAURENT (2012), *Handbook of Volatility Models and their Applications*. Wiley.
J-C. PRAGER and J. THISSE (2012), *Economic Geography and the Unequal Development of Regions*. Routledge.
M. FLEURBAEY and F. MANIQUET (2012), *Equality of Opportunity: The Economics of Responsibility*. World Scientific.
J. HINDRIKS (2012), *Gestion publique*. De Boeck.
M. FUJITA and J.F. THISSE (2013), *Economics of Agglomeration: Cities, Industrial Location, and Globalization.* (2nd edition). Cambridge University Press.
J. HINDRIKS and G.D. MYLES (2013). *Intermediate Public Economics.* (2nd edition). MIT Press.
J. HINDRIKS, G.D. MYLES and N. HASHIMZADE (2013). *Solutions Manual to Accompany Intermediate Public Economics.* (2nd edition). MIT Press.
J. HINDRIKS (2015). *Quel avenir pour nos pensions ? Les grands défis de la réforme des pensions.* De Boeck.
P. BELLEFLAMME and M. PEITZ (2015). *Industrial Organization: Markets and Strategies* (2nd edition). Cambridge University Press.

CORE Lecture Series

R. AMIR (2002), Supermodularity and Complementarity in Economics.
R. WEISMANTEL (2006), Lectures on Mixed Nonlinear Programming.
A. SHAPIRO (2010), Stochastic Programming: Modeling and Theory.