Parametric and Semiparametric Estimation in Evaluating Agro-Environmental Schemes

Stima parametrica e semiparametrica per la valutazione di politiche agro-ambientali

Barbara Pacini¹, Meri Raggi¹, Davide Viaggi²

¹ Dipartimento di Scienze Statistiche “P. Fortunati”, Università di Bologna
² Dipartimento di Economia e Ingegneria Agrarie, Università di Bologna

e-mail: meri.raggi@unibo.it

Keywords: semiparametric estimation, single index models, policy evaluation

1. Introduction

In this paper we focus on parametric and semiparametric estimation of binary choice models as applied to the participation of farms in Agro-Environmental Schemes (AES) in Europe. EU policies are increasingly aimed at the integration of environmental considerations into the Common Agricultural Policy and at the development of agricultural practices preserving the environment and safeguarding the countryside. AES were introduced into EU agricultural policy during the late 1980s as an instrument to support specific farming practices that help to protect the environment.

A large survey on European farms, both involved and not involved in AES, has been conducted within the EU project “Integrated Tools to design and implement AES (ITAES)”, in order to characterize the individual farmer and the farming system, analyze perceptions and attitudes towards AES, preferences for contract attributes and general behavior with respect to environmental issues. For our analysis we use a subsample of 1093 farms from 5 different regions; the proportion of AES participants is quite the same in all regions. AES involve specific prescriptions, which leads to additional costs of production. We compare a parametric specification with a semiparametric estimator of binary response models to identify the key factors underlying farmers’ decision to apply AES and to correctly predict participants’ behavior. Our main goal is to check the performances of parametric and semiparametric estimators, to be then used both as a selection correction term and/or as a propensity score in a matching procedure, in order to correctly evaluate the costs of AES comparing participants and not participants with the same characteristics.

2. Methodology

Consider the following model: \(Y^*_i = X_i'\beta + u_i\), where \(X\) a \(k \times 1\) vector of explanatory variables, \(\beta\) is \(k \times 1\) vector of parameters and \(u_i\) is a random error term, for \(i = 1, \cdots, n\).

We observe \(Y_i = 1\) if \(Y^*_i > 0\) and \(Y_i = 0\) otherwise. The likelihood function

\[
lnL(\beta) = \sum_{i=1}^{n} Y_i ln[F(X'_i \beta | X_i)] + (1 - Y_i) ln [1 - F(X'_i \beta | X_i)]
\]

\((⋆)\) Work supported by ITAES-6FP-SSPE-CT-2003-502070.
depends on the specification of \( P(Y = 1|X, \beta) = F(X'\beta|X) \) with \( F \) the cumulative distribution function of \( u \) conditional on \( X = x \); assuming a probit specification, for example, we have \( F(X'\beta|X) = \Phi(X'\hat{\beta}/\sigma_u) \), where \( \hat{\beta} \) is the usual MLE of \( \beta \) and \( \sigma_u \) is assumed equal to 1. We do not have enough prior knowledge to justify a specific distributional assumption, so we need to relax some parametrization. In the same time, we avoid a direct estimation of \( P(Y = 1|X) \) by nonparametric regression (so relaxing the linear index constraint), mainly for two reasons: 1) with a high number of explanatory variables the nonparametric regression estimator suffers of curse of dimensionality; 2) we are interested in interpreting the parameters \( \beta \), which carry information on the participation mechanism and are not identified by nonparametric regression.

Following Klein and Spady (1993) we assume that the choice probability function depends on a parametrically specified index function \( P(Y = 1|X) = E(Y|X) = G(X'\beta) \). The function \( G \) can be estimated by a kernel regression estimator \( G_n \):

\[
G_n(X'_i\beta) = \frac{\sum_j Y_j[X'_i\beta - X'_j\beta]/h_n}{\sum_j k[X'_i\beta - X'_j\beta]/h_n}
\]

By inserting \( G \) in (1) we have a quasi-loglikelihood function, which can be efficiently estimated posing some suitable conditions on \( K \) (kernel function) and \( h_n \) (bandwidth parameter). For identification purposes, at least one continuous regressor is needed and one parameter in vector \( \beta \) is fixed equal to one, note also that the semiparametric specification does not allow to disentangle the intercept from the estimated index function.

3. Empirical results

Using ITAES survey data, we first estimate a model for participation in AES. A preliminary comparison between the performances of parametric (probit) and semiparametric estimators can be done in terms of confusion matrices. An example is shown in Table 1, where we can observe a better overall correct classification ratio for the semiparametric single index model (88.7% versus 67.8%).

Table 1: Comparison between predictions and observations (numbers)

<table>
<thead>
<tr>
<th></th>
<th>Parametric</th>
<th>Semiparametric</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Classified</td>
<td>Classified</td>
</tr>
<tr>
<td>True</td>
<td>NP P</td>
<td>True</td>
</tr>
<tr>
<td>NP</td>
<td>285 199</td>
<td>NP 417 67</td>
</tr>
<tr>
<td>P</td>
<td>153 456</td>
<td>P 57 552</td>
</tr>
</tbody>
</table>

References

