University of Zurich

Socioeconomic Institute
Sozialökonomisches Institut

Working Paper No. 0404

Empirical Likelihood in Count Data Models: The Case of Endogenous Regressors

Stefan Boes

March 2004

Socioeconomic Institute
University of Zurich

Working Paper No. 0404

Empirical Likelihood in Count Data Models: The Case of Endogenous Regressors

March 2004

Author's addresses Stefan Boes
E-mail: boes@sts.unizh.ch.

Publisher Sozialökonomisches Institut
Bibliothek (Working Paper)
Rämistrasse 71
CH-8006 Zürich
Phone: +41-1-634 21 37
Fax: +41-1-634 49 82
URL: www.soi.unizh.ch
E-mail: soilib@soi.unizh.ch

Empirical Likelihood in Count Data Models: The Case of Endogenous Regressors

Stefan Boes*

University of Zurich

February 2004

Abstract

Recent advances in the econometric modelling of count data have often been based on the generalized method of moments (GMM). However, the two-step GMM procedure may perform poorly in small samples, and several empirical likelihood-based estimators have been suggested alternatively. In this paper I discuss empirical likelihood (EL) estimation for count data models with endogenous regressors. I carefully distinguish between parametric and semi-parametric methods and analyze the properties of the EL estimator by means of a Monte Carlo experiment. I apply the proposed method to estimate the effect of women's schooling on fertility.

Keywords: Nonparametric likelihood, Poisson model, endogeneity, fertility and education.

JEL-Classification: C14, C25, J13

* *Address for correspondence:* University of Zurich, Socioeconomic Institute, Zuerichbergstr. 14, 8032 Zurich, Phone: +41 1 634 2301, E-Mail: boes@sts.unizh.ch. I would like to thank Rainer Winkelmann, William Sander, and seminar participants at the University of Zurich for helpful comments.

1 Introduction

Count data models arise in many economic fields including health economics, demographic studies, labor economics, or industrial organization. Models for count data incorporate the special feature of the dependent variable being a nonnegative integer, or count. Examples are the number of doctor visits (Pohlmeier and Ulrich, 1995), the number of children born to women (Winkelmann and Zimmermann, 1995), the number of days a worker is absent from his job (Delgado and Kniesner, 1997), or the number of patents (Hausman, Hall and Griliches, 1984).

A serious problem occurring frequently in microeconomic applications is that of endogenous explanatory variables. This can be due to omitted variables, errors-in-variables, or more generally due to simultaneity, leading to inconsistency of parameter estimates obtained by standard methods. Within count data models endogeneity can be captured by additive or multiplicative errors in the mean function. Grogger (1990), Mullahy (1997), and Windmeijer and Santos Silva (1997) discussed nonlinear instrumental variables, or generalized method of moments (GMM) to estimate regression parameters consistently.

Recent work concerning the properties of GMM in small samples or with increasing degree of over-identification emphasizes the poor performance of the two step GMM procedure. Several alternative estimators were proposed, for example the continuous updating estimator (CUE) of Hansen, Heaton and Yaron (1996), the empirical likelihood (EL) estimator of Owen (1988), Qin and Lawless (1994) and Imbens (1997), and the exponential tilting (ET) estimator of Kitamura and Stutzer (1997) and Imbens, Spady and Johnson (1998). All of these estimators can be subsumed in the class of generalized empirical likelihood (GEL) estimators (Smith, 1997) and asymptotic equality of GEL and GMM was shown. Further studies of Newey and Smith (2000, 2004) and Imbens and Spady (2001) examined the higher order properties of GEL and GMM estimators and evidenced the relative advantage of EL compared to GMM and other GEL estimators in terms of smaller bias and higher order efficiency.

In this paper I discuss empirical likelihood estimation for count data models with endogenous regressors. I choose the empirical likelihood estimator due to its preferable properties and derive its first order conditions. I carefully distinguish between parametric and semi-parametric methods and analyze the properties of the estimator by means of a Monte Carlo experiment. As an empirical illustration of the proposed estimator I re-evaluate a study of Sander (1992) who analyzes the effect of women's schooling on fertility in the United States. Fertility is measured by the number of children ever born to a woman, thus the dependent variable is a count. Women's schooling might be endogenously determined because of non-zero correlation with unobservable traits. SANDER (1992) applies instrumental variables in a linear model, whereas WOOLDRIDGE (1997) tests in a Poisson model. In addition to that I estimate the model with GMM and EL, and compare the different methods.

The outline of the paper is as follows. Preliminaries are laid out in Section 2. Section 3 considers empirical likelihood estimation. Section 4 compares EL with GMM and maximum likelihood. Section 5 gives results of a Monte Carlo study and Section 6 applies EL to estimate the effect of women's schooling on fertility. Section 7 concludes.

2 Preliminaries

Econometric Modeling of Count Data

Let y_i, $i = 1, \ldots, n$ denote an independently distributed, nonnegative integer-valued variable with conditional mean specified as

$$E[y_i|x_i] = \mu_i(\beta) = \exp({x_i}'\beta), \tag{1}$$

where x_i is a k-dimensional vector of explanatory variables and β is a k-vector of unknown parameters.[1] A fully parametric assumption like the conditional distribution $y_i|x_i \sim \text{Poisson}(\mu_i(\beta))$ allows for straightforward application of maximum likelihood methods. In the particular example of a Poisson process the maximum likelihood (ML) estimator of β, namely $\hat{\beta}_{ML}$, solves the first order condition $\sum_i x_i(y_i - \mu_i(\beta)) = 0$. From standard maximum likelihood theory it follows that $\hat{\beta}_{ML}$ is consistent and $\sqrt{n}(\hat{\beta}_{ML} - \beta)$ converges in distribution to a normal with mean zero and estimated variance $n\{\sum_i \mu_i(\hat{\beta}_{ML})\, x_i {x_i}'\}^{-1}$, where $\mu_i(\hat{\beta}_{ML}) = \exp({x_i}'\hat{\beta}_{ML})$.

The standard Poisson model can be misspecified for several reasons. For example the assumption of equidispersion – the equality of mean and variance – is frequently violated in applied work and more general distributions are developed to cover features like over- or underdispersion.[2] However, GOURIEROUX, MONFORT and TROGNON (1984) showed that Poisson estimates are still consistent as long as the conditional mean is properly specified. Correct standard errors can be obtained by the estimated variance of the pseudo maximum likelihood (PML) estimator $\hat{\beta}_{PML}$ $(= \hat{\beta}_{ML})$, which

[1] For a discussion of count data models in respect of theory and practical applications see WINKELMANN (2003).

[2] Examples are the negative binomial (negbin) and the logarithmic distribution, or mixture distributions, again see WINKELMANN (2003) for further details.

can be calculated by $\hat{V}(\hat{\beta}_{PML}) =$

$$\left(\sum_{i=1}^{n} \mu_i(\hat{\beta}_{PML})\, x_i x_i'\right)^{-1} \left(\sum_{i=1}^{n} \left(y_i - \mu_i(\hat{\beta}_{PML})\right)^2 x_i x_i'\right) \left(\sum_{i=1}^{n} \mu_i(\hat{\beta}_{PML})\, x_i x_i'\right)^{-1}, \tag{2}$$

where $\mu_i(\hat{\beta}_{PML}) = \exp(x_i'\hat{\beta}_{PML}) = \exp(x_i'\hat{\beta}_{ML})$.

Generalized Method of Moments

Since the seminal article of HANSEN (1982) generalized method of moments (GMM) has become a well-established estimation technique in applied and theoretical econometrics. GMM provides a general framework for dealing with moment conditions avoiding strong distributional assumptions. The specification of a conditional mean in (1) defines implicitly a conditional moment restriction $E[u_i|x_i] = 0$, where u_i is a regression error with $u_i = y_i - \mu_i(\beta)$. The law of iterated expectations proves that k unconditional moment restrictions can be constructed as

$$E[x_i(y_i - \mu_i(\beta))] = 0. \tag{3}$$

The GMM estimator $\hat{\beta}_{GMM}$ minimizes the weighted squared distance of sample and population moments, algebraically

$$\hat{\beta}_{GMM} = \arg\min_{\beta} \left(\frac{1}{n}\sum_{i=1}^{n} x_i(y_i - \mu_i(\beta))\right)' W \left(\frac{1}{n}\sum_{i=1}^{n} x_i(y_i - \mu_i(\beta))\right), \tag{4}$$

where W is weighting matrix. Since (3) is an exactly determined equation system, the GMM first order conditions are unaffected by the choice of W and identical to the ML first order conditions, hence $\hat{\beta}_{GMM} = \hat{\beta}_{ML}$. Under mild regularity conditions one can show the consistency and asymptotic normality of the GMM estimator.[3] The efficient GMM estimator can be obtained for the appropriate choice of weights which is $W = V[\sum_i x_i(y_i - \mu_i(\beta))]$. If the weighting matrix W is estimated by $n\{\sum_i (y_i - \mu_i(\hat{\beta}_{GMM}))^2 x_i x_i'\}^{-1}$, the variance of $\hat{\beta}_{GMM}$ is equal to that of $\hat{\beta}_{PML}$.

Endogenous Regressors

As mentioned above the consistency of ML and PML estimation crucially depends on the assumption of valid moment conditions, i.e. $E[y_i|x_i] = \mu_i(\beta)$ holds. In other words, a misspecified mean function leads to inconsistency of $\hat{\beta}_{ML}$ and $\hat{\beta}_{PML}$. The problem of endogeneity can be seen as one example in which the moment condition fails. Recent contributions on endogenous regressors

[3] For details see GOURIEROUX and MONFORT (1995: Ch. 9.5).

within count data models (Windmeijer and Santos Silva, 1997; Mullahy, 1997) consider two alternative approaches: (1) endogeneity with additive errors and (2) endogeneity with multiplicative errors.

The case of additive errors leads to a conditional mean function of the form

$$E[y_i|x_i, \xi_i] = \exp({x_i}'\beta) + \xi_i, \tag{5}$$

where endogeneity is present if $E[\xi_i|x_i] \neq 0$ and I assume the existence of a q-dimensional vector of instruments z_i ($q \geq k$) such that $E[\xi_i|z_i] = 0$. Again by the law of iterated expectations unconditional moment restrictions are

$$E[z_i(y_i - \mu_i(\beta))] = 0, \tag{6}$$

which can be estimated by GMM replacing the moment functions $x_i(y_i - \mu_i(\beta))$ in (4) by the functions $z_i(y_i - \mu_i(\beta))$.[4] A more intuitive approach to deal with endogenous regressors is to treat unobservable factors ε_i and regressors x_i symmetrically and to specify a mean function with multiplicative errors as

$$E[y_i|x_i, \varepsilon_i] = \exp({x_i}'\beta + \varepsilon_i) = \mu_i(\beta)\,\nu_i, \tag{7}$$

where $\nu_i = \exp(\varepsilon_i)$. The conditional expectation with respect to x_i is $E[y_i|x_i] = \mu_i(\beta)E[\nu_i|x_i]$ and endogeneity is present if $E[\nu_i|x_i] \neq 1$. In this case I assume that instruments z_i are available such that $E[\nu_i|z_i] = 1$ and conditional moment restrictions are given by $E[\nu_i - 1|z_i] = 0$. The law of iterated expectations shows that

$$E\left[z_i\left(\frac{y_i - \mu_i(\beta)}{\mu_i(\beta)}\right)\right] \;=\; E[z_i(\exp(-{x_i}'\beta)y_i - 1)] \;=\; E\left[z_i(\nu_i - 1)\right] \;=\; 0. \tag{8}$$

Windmeijer and Santos Silva (1997) emphasized that a set of instruments z_i cannot be orthogonal to ξ_i (the additive case) and $\nu_i - 1$ (the multiplicative case) at the same time since $y_i - \mu_i(\beta)$ and $\mu_i(\beta)$ are correlated. In this paper I concentrate on endogeneity with multiplicative errors, although all results are easily extended to the additive case.

The moment conditions in (8) can be estimated by GMM as presented in the preceding paragraph. An interesting case arises when the number of instruments q exceeds the number of regressors k (the over-determined case), and one can apply a two step GMM procedure to estimate the parameters β in (8). The efficient estimator $\hat{\beta}_{GMM2}$ is the argument β that minimizes the objective function

$$\left(\frac{1}{n}\sum_{i=1}^{n} z_i\left(\frac{y_i - \mu_i(\beta)}{\mu_i(\beta)}\right)\right)' \tilde{V}^{-1} \left(\frac{1}{n}\sum_{i=1}^{n} z_i\left(\frac{y_i - \mu_i(\beta)}{\mu_i(\beta)}\right)\right), \tag{9}$$

[4] Nonlinear instrumental variable (NLIV) estimation in count data models is discussed in Grogger (1990).

where $\tilde{V}^{-1} = n^{-1}\sum_i\{(y_i - \mu_i(\tilde{\beta}))/\mu_i(\tilde{\beta})\}^2 z_i z_i{}'$ are the optimal weights with $\mu_i(\tilde{\beta}) = \exp(x_i{}'\tilde{\beta})$, and $\tilde{\beta}$ is a first step GMM estimator using any weighting matrix W, e.g. the q-dimensional identity matrix. Under mild regularity conditions $\hat{\beta}_{GMM2}$ is consistent and the stabilizing transformation $\sqrt{n}(\hat{\beta}_{GMM2} - \beta)$ is asymptotically normal with mean zero and estimated variance

$$\left\{\left(\frac{1}{n}\sum_{i=1}^{n}\frac{z_i\, y_i\, x_i{}'}{\mu_i(\hat{\beta}_{GMM2})}\right)'\left(\frac{1}{n}\sum_{i=1}^{n}\left(\frac{y_i - \mu_i(\hat{\beta}_{GMM2})}{\mu_i(\hat{\beta}_{GMM2})}\right)^2 z_i z_i{}'\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^{n}\frac{z_i\, y_i\, x_i{}'}{\mu_i(\hat{\beta}_{GMM2})}\right)\right\}^{-1}, \quad (10)$$

where $\mu_i(\hat{\beta}_{GMM2}) = \exp(x_i{}'\hat{\beta}_{GMM2})$.

3 Empirical Likelihood Estimation

Based upon the work of Owen (1988, 1991, 2001) and Qin and Lawless (1994, 1995) I now develop the empirical likelihood (EL) estimator of β for the conditional mean specification (1) taking into account that x_i may be endogenous in a multiplicative sense, thus considering unconditional moment restrictions (8).

Let p_i denote the unknown probability assigned to the sample outcome (y_i, x_i, z_i) of one observation i with $0 \leq p_i \leq 1\ \forall i$ and impose a normalization $\sum_i p_i = 1$. Furthermore, let $p = (p_1, \ldots, p_n)'$ denote the n-dimensional vector of probabilities. Then a nonparametric likelihood estimator of p is obtained from maximizing a nonparametric log-likelihood function,

$$\max_p\ n^{-1}\sum_{i=1}^{n}\ln p_i \quad \text{s.t.} \quad \sum_{i=1}^{n}p_i = 1. \quad (11)$$

Without further restrictions optimal probability weights in (11) are given by $p_i = n^{-1}$, the empirical density function. To incorporate special features of the data generating process impose empirical moments as additional restrictions. From (8), a q-dimensional vector of empirical moment conditions can be specified as

$$\sum_{i=1}^{n}p_i z_i\left(\frac{y_i - \mu_i(\beta)}{\mu_i(\beta)}\right) = 0. \quad (12)$$

Note the difference between sample moments, where each observation is weighted by n^{-1}, and the empirical moments in (12), where each observation is weighted by p_i. The optimization problem in (11) using the additional restrictions in (12) can be rewritten as

$$\max_p\ n^{-1}\sum_{i=1}^{n}\ln p_i \quad \text{s.t.} \quad \sum_{i=1}^{n}p_i z_i\left(\frac{y_i - \mu_i(\beta)}{\mu_i(\beta)}\right) = 0, \quad \sum_{i=1}^{n}p_i = 1, \quad (13)$$

which implies the optimality conditions

$$p_i(\beta) \quad = \quad \frac{1}{n\left[1-\lambda(\beta)'z_i\left(\frac{y_i-\mu_i(\beta)}{\mu_i(\beta)}\right)\right]} \qquad \text{and} \tag{14}$$

$$\lambda(\beta) \quad = \quad \arg_\lambda\left[\sum_{i=1}^{n}\left(\frac{z_i\left(\frac{y_i-\mu_i(\beta)}{\mu_i(\beta)}\right)}{n\left[1-\lambda(\beta)'z_i\left(\frac{y_i-\mu_i(\beta)}{\mu_i(\beta)}\right)\right]}\right)=0\right], \tag{15}$$

where $\lambda(\beta)$ is a q-dimensional vector of Lagrangean multipliers with respect to the empirical moment restrictions. Plugging (14) and (15) into the objective function in (13) yields an empirical log-likelihood function,

$$\ln L_{EL}(\beta) = -\ln(n) - n^{-1}\sum_{i=1}^{n}\ln\left[1-\lambda(\beta)'z_i\left(\frac{y_i-\mu_i(\beta)}{\mu_i(\beta)}\right)\right]. \tag{16}$$

The maximum of (16) is the value of β, namely the empirical likelihood estimator $\hat{\beta}_{EL}$, that simultaneously solves

$$\sum_{i=1}^{n}\left(\frac{-x_i\,\frac{y_i\,{z_i}'\lambda(\hat{\beta}_{EL})}{\mu_i(\hat{\beta}_{EL})}}{n\left[1-\lambda(\hat{\beta}_{EL})'z_i\left(\frac{y_i-\mu_i(\hat{\beta}_{EL})}{\mu_i(\hat{\beta}_{EL})}\right)\right]}\right) \quad = \quad 0 \tag{17}$$

$$\sum_{i=1}^{n}\left(\frac{z_i\left(\frac{y_i-\mu_i(\hat{\beta}_{EL})}{\mu_i(\hat{\beta}_{EL})}\right)}{n\left[1-\lambda(\hat{\beta}_{EL})'z_i\left(\frac{y_i-\mu_i(\hat{\beta}_{EL})}{\mu_i(\hat{\beta}_{EL})}\right)\right]}\right) \quad = \quad 0, \tag{18}$$

where (18) follows directly from the optimality condition (15). Since (17) and (18) build up a highly non-linear equation system, numerical methods have to be applied to obtain the value of $\hat{\beta}_{EL}$.

Under similar regularity conditions as in the GMM framework Qin and Lawless (1994) showed the consistency of the empirical likelihood estimator and proved the asymptotic normality of the stabilizing transformation $\sqrt{n}(\hat{\beta}_{EL}-\beta)$ with mean zero and estimated variance

$$\begin{aligned}&\left\{\left(\sum_{i=1}^{n}p_i(\hat{\beta}_{EL})\,\frac{z_i\,y_i\,{x_i}'}{\mu_i(\hat{\beta}_{EL})}\right)'\right.\\ &\times\left.\left(\sum_{i=1}^{n}p_i(\hat{\beta}_{EL})\left(\frac{y_i-\mu_i(\hat{\beta}_{EL})}{\mu_i(\hat{\beta}_{EL})}\right)^2 z_i{z_i}'\right)^{-1}\left(\sum_{i=1}^{n}p_i(\hat{\beta}_{EL})\,\frac{z_i\,y_i\,{x_i}'}{\mu_i(\hat{\beta}_{EL})}\right)\right\}^{-1},\end{aligned} \tag{19}$$

where $p_i(\hat{\beta}_{EL})$ is given by (14) evaluated at $\hat{\beta}_{EL}$. Note that the terms in (19) are estimated by probability weights obtained from an empirical likelihood optimization whereas the terms in (10) are estimated by probability weights n^{-1}. One important feature of EL and efficient GMM is, relating to the work of Chamberlain (1987), that both estimators reach the semiparametric efficiency bound.

4 Interpretation of Empirical Likelihood Estimates

Comparing the First Order Conditions of GMM and EL

In the spirit of Newey and Smith (2004) and to get a deeper understanding of generalized method of moments and empirical likelihood estimation in the count data framework, I compare the first order conditions of both estimators. The optimization in (9) gives the first order conditions of the two step GMM estimator, algebraically

$$\left(\sum_{i=1}^{n}\frac{1}{n}\,\frac{z_i\,y_i\,x_i{}'}{\mu_i(\hat\beta_{GMM2})}\right)'\left(\sum_{i=1}^{n}\frac{1}{n}\left(\frac{y_i-\mu_i(\tilde\beta)}{\mu_i(\tilde\beta)}\right)^2 z_i z_i{}'\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} z_i\left(\frac{y_i-\mu_i(\hat\beta_{GMM2})}{\mu_i(\hat\beta_{GMM2})}\right)\right)=0, \quad (20)$$

where $\tilde\beta$ is the first round estimator. In the context of empirical likelihood estimation one can show that conditions (17) and (18) imply first order conditions[5]

$$\begin{aligned}&\left(\sum_{i=1}^{n}p_i(\hat\beta_{EL})\,\frac{z_i\,y_i\,x_i{}'}{\mu_i(\hat\beta_{EL})}\right)'\\ &\quad\times\left(\sum_{i=1}^{n}p_i(\hat\beta_{EL})\left(\frac{y_i-\mu_i(\hat\beta_{EL})}{\mu_i(\hat\beta_{EL})}\right)^2 z_i z_i{}'\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} z_i\left(\frac{y_i-\mu_i(\hat\beta_{EL})}{\mu_i(\hat\beta_{EL})}\right)\right)=0. \quad (21)\end{aligned}$$

Equations (20) and (21) show the main difference between GMM and EL. Each estimator sets a linear combination of sample moments equal to zero, where the sample moments are given by the right brackets in both equations. GMM and EL differ in the way of calculating these linear combinations. GMM uses sample moments for the Jacobian matrix (left brackets) and the matrix of second moments (middle brackets). Furthermore, the weighting matrix depends on a first step (inefficient) estimator. In contrast to that EL uses empirical moments for the Jacobian term and the matrix of second moments, whereby the probability weights p_i are chosen efficiently.

Relationship to Maximum Likelihood Estimation

In a standard Poisson model the conditional probability function $f(y_i|x_i;\beta)$ is given by

$$f(y_i|x_i;\beta)=\frac{\exp(-\exp(x_i'\beta))\exp(y_i x_i'\beta)}{y_i!}, \quad (22)$$

which follows directly from the conditional mean specification (1) and the distributional assumption $y_i|x_i \sim \text{Poisson}(\mu_i(\beta))$. The (parametric) sample likelihood function can be written as $L(\beta;y,x)=\prod_{i=1}^{n} f(y_i|x_i;\beta)$ and the maximum likelihood estimator chooses the value of β such that the observed sample is most likely. Once estimates of β are obtained, the parametric specification allows for a

[5]For a general derivation and interpretation see Newey and Smith (2004).

discussion of marginal probability effects, i.e. the effect of a small ceteris paribus increase in one regressor on the probability of observing a certain outcome of y. Furthermore, one can calculate probabilities of outcomes that are not observed in the sample.

Within empirical likelihood estimation parameters p_i of a joint probability mass function $\prod_{i=1}^{n} p_i$ are defined, and this function is maximized with respect to constraints defined in terms of empirical moment conditions. The probability mass function can be interpreted as a multinomial distribution with n parameters p_i, one parameter for each data outcome (y_i, x_i, z_i). Moreover, one can think of constrained maximum likelihood estimation of p with constraints represented by empirical moments and a natural normalization for probability functions, $\sum_i p_i = 1$. As noted in the previous section probability weights n^{-1} are optimal if empirical moments are absent or moment restrictions (8) display an exactly determined equation system. The latter follows since optimal Lagrangean multipliers are zero in this case. If the number of instruments is larger than the number of parameters (the over-determined case), $\lambda(\hat{\beta}_{EL})$ differs from zero and $p_i(\hat{\beta}_{EL})$ differs from n^{-1}. Information theoretic approaches show that $p_i(\hat{\beta}_{EL})$ is chosen as close as possible to n^{-1} taking into account that the empirical moments have to be fulfilled (see Kitamura and Stutzer (1997) and Imbens, Spady and Johnson (1998)).

Two important conclusions can be drawn from the preceding discussion. First, we cannot compare probabilities p_i with a parametrically specified conditional probability function like the Poisson, or any other count data distribution. A conditional probability function $f(y_i|x_i;\beta)$ gives the probability of observing one of the values $y_i = 0, 1, 2, \ldots$ given a vector of explanatory variables x_i, whereas p_i gives the sample probability of one observation. Second, empirical likelihood estimates of p_i do not allow for calculation of marginal probability effects or prediction of probabilities of outcomes that are not observed in the sample.

5 Monte Carlo Evidence

In this section I illustrate the theoretical advantage of empirical likelihood estimation by means of a Monte Carlo experiment. I choose different sample sizes (n = 100, 500, and 1000), and for each sample size 1000 vectors of y are drawn from a Poisson distribution with parameter $\mu_i = \exp(0.5 - x_{1i} - 0.5\,\tilde{x}_{2i} + \varepsilon_i)$, hence the true parameter vector is $\beta^0 = (0.5 \;\; -1 \;\; -0.5)'$. The regressors x_{1i} and $\tilde{x}_{2i}$ are independently and uniformly distributed on the interval $[0, 1]$, the unobservable factors ε_i are independent drawings from a normal distribution with mean zero and variance 0.7. I assume that

$\tilde{x}_{2i}$ cannot be observed but $x_{2i} = \tilde{x}_{2i} + v_i$, where v_i is a classical measurement error independently and normally distributed with mean zero and variance equal to one. To complete the story I generate different instruments z_i with properties $\text{corr}(z_i, x_{2i}) = \rho$ and $\text{corr}(z_i, \varepsilon_i) = 0$ which allows us to vary the number and quality of instruments by choosing different values of ρ.

There is a large number of possible combinations of instruments and sample sizes and therefore I picked out just a few. Precisely, for each sample size I define weakly identified setups (3 instruments with $\rho = 0.1,\ 0.1,\ 0.5$), partial identified setups (3 instruments with $\rho = 0.5,\ 0.5,\ 0.5$ or $\rho = 0.1,\ 0.5,\ 0.9$), and strong identified setups (3 instruments with $\rho = 0.9,\ 0.9,\ 0.9$). For $n = 500$ observations and partial identification I increase the number of instruments to five and ten (all instruments with $\rho = 0.5$ or evenly distributed between 0.1 and 0.9). All setups are estimated by GMM and EL, and for the sake of completeness the PML estimator is calculated for all sample sizes. As mentioned above numerical methods have to be applied to obtain the EL estimator. I use the BFGS (Broyden, Fletcher, Goldfarb, Shanno) algorithm as it is implemented in the *constrained optmum* procedure of GAUSS[6] and maximize the empirical likelihood function in (16) with respect to β and λ, and subject to the constraint in (15). PML and GMM estimators are calculated by the same algorithm, maximizing the parametric likelihood function of a Poisson model and minimizing the GMM objective function in (4) respectively. The results are reported in Tables 1 to 4.

— Table 1 about here —

Table 1 gives the results for a sample size of $n = 100$ observations. For each estimator I calculate the mean and standard deviation of the beta's (standard deviation in parentheses). As we would expect, the PML estimate of β_1 is unaffected by the measurement error, but the estimate of β_2 is biased and inconsistent. As in the linear model with classical errors-in-variables we have a regression to the mean, i.e. the parameter estimate is biased towards zero. GMM and EL estimates of β_1 do not differ considerably from PML estimates, but estimates of β_2 are closer to the true value of -0.5. Moreover, GMM and EL estimates are substantially different. The bias of EL is less than the bias of GMM, particularly in the case of weakly correlated instruments, and standard deviations are smaller.

— Tables 2 and 3 about here —

[6] Aptech Systems, Inc., http://www.aptech.com .

Tables 2 and 3 give the results for the same set of instruments but two different sample sizes, $n = 500$ and $n = 1000$. EL as well as GMM estimates of β_2 get closer to the true value and standard deviations of all estimates decrease. Only in the case of 500 observations and weak identification EL is notably closer to -0.5 than GMM. With 1000 observations differences between both estimators disappear independently of varying the quality of instruments. I emphasize that both estimators perform much better in the case of strong identification, or in other words the presence of weak instruments causes both estimators to be more biased. Another important result is that throughout different setups EL has smaller standard deviations than GMM.

— Table 4 about here —

Table 4 shows the results for partial identification with five or ten instruments and a sample size of 500 observations (only 500 replications are calculated). Five instruments with $\rho = 0.5$ cause the GMM estimator to be more biased compared to the case of three instruments, whereas the EL estimator performs approximately the same. With correlations between 0.1 and 0.9 discrepancies between EL and GMM disappear. Ten instruments with $\rho = 0.5$ again cause the GMM estimate of β_2 to be more biased than EL. Parameter estimates are similar to the case of five instruments, but standard deviations are larger. Evenly distributed correlations between 0.1 and 0.9 show approximately the same estimates as with five instruments, but standard deviations of the GMM estimator increase substantially. These results support theoretical findings of Newey and Smith (2004) and Imbens and Spady (2001) that with increasing number of instruments and decreasing number of observations GMM may be more biased than EL.

So far, I have only considered the standard deviation of the beta's and not the mean of the estimated standard errors based on the asymptotic distribution. Hence, I distinguish between robust point estimates and inference. In fact I calculated both values, and with a sample size of 1000 observations they are approximately the same. But with sample sizes of 100 or 500 observations the two values differ substantially, in particular in the case of weak identification. Therefore, the classical normal asymptotic approximations to the finite-sample distributions are very poor. This result is not surprising since recent work concerning the properties of GMM and EL estimators under weak identification shows nonstandard distributions (see e.g. Stock, Wright and Yogo, 2002, and Guggenberger and Smith, 2003). Inference can be improved by using bootstrapped standard errors or by applying methods proposed in the above-quoted literature, such as tests based on the objective function (16).

6 An Empirical Example

As an illustration of empirical likelihood estimation in count data models with endogeneity, I consider a data set similar to that used by Sander (1992) and Wooldridge (1997) taken from the National Opinion Research Center's *General Social Survey*. I used the same independent and pooled cross-sections across even years from 1972 to 1984 as in Wooldridge (1997) with the number of siblings (*sibs*) as additional variable.[7] The variables of interest are the number of children ever born to women (*childs*) as dependent count variable, and years of schooling (*schooling*) as explanatory variable to determine the effect of women's schooling on fertility. Additional controls are quadratic age (*age*, *agesq*), a dummy for race (*black*), region at age sixteen (relative to south), type of residence at age sixteen (relative to big cities), and time dummies for the even years from 1974 to 1984. As discussed in Sander (1992) schooling might be endogenous in the fertility regression, e.g. due to unobservable traits correlated with *schooling*.

In the sample of 992 women between the ages of 35 and 54 the average number of children is 2.7 (standard deviation 1.6) and on average a woman attends 12.9 (2.6) years of schooling. First of all assuming schooling is exogenous I estimate a linear model and a Poisson model with robust standard errors. The results are reported in Table 5, columns OLS and PML. The coefficient on schooling in the linear model is -0.11 with a t-statistic of -5.70, thus the estimate is highly significant. Economically, given one more year of schooling the expected reduction in the number of children is 0.11. In other words, attending a university for 5 years reduces the expected number of children about a half compared to a woman who does not attend a university. Note that interpretation in the linear model is somewhat misleading because negative predicted outcomes for the dependent count variable are possible. In the Poisson model the estimated coefficient on schooling is -0.0428 with a standard error of 0.0086 which implies that the coefficient is statistically significant and each additional year of schooling reduces the number of children about 4.3 percent. Multiplying the coefficient on schooling in the Poisson model by the average number of children shows that the implied marginal effect at the sample means is about the same as in the linear model.

— Table 5 about here —

[7] Sander (1992) uses data from 1985 to 1991. The data set is available from the data archive of Wooldridge's (2003) textbook for even years 1972 to 1984 (without *sibs*). The whole data set collected for almost all years between 1972 and 1994 is freely available online at http://www.soc.qc.edu/QC_Software/GSS.html. Comprehensive information on the *General Social Survey* including the data set for almost all years between 1972 and 2002 can be found online at http://www.norc.uchicago.edu/projects/gensoc.asp and http://www.icpsr.umich.edu:8090/GSS/homepage.htm.

For possibly endogenous schooling I use father's and mother's schooling, and the number of siblings as instruments. A simple OLS regression of *schooling* on the instruments controlling for the other variables shows highly significant instruments. WOOLDRIDGE (1997) tests for endogeneity by adding the residuals from this regression to his original Poisson model and computing the corresponding t-statistic. Within a linear model SANDER (1992) tests for endogeneity with a simple Hausman test. I replicate these tests and conclude that *schooling* is not highly endogenous in the fertility equation, a result similar to those of SANDER (1992) and WOOLDRIDGE (1997).

Nevertheless, I apply instrumental variable estimation of the fertility equation using two stage least squares (2SLS), GMM, and EL. The results are reported also in Table 5. Estimating the linear model with 2SLS yields a highly significant coefficient on *schooling* of -0.1661 (0.0400). Thus the effect of schooling is more negative compared to OLS meaning that schooling and unobservable traits are positively correlated. The GMM estimate of the effect of schooling is -0.0669 (0.0168), the EL estimate is -0.0740 (0.0166). Both are significant on the one percent level, EL implying a 7.40 percent decrease in the number of children given one more year of schooling (GMM: 6.69 percent). Of particular interest in the context of instrumental variable estimation and over-determined restrictions is the validity of moments. In the GMM framework this can be tested based on the value of the objective function (9) evaluated at the GMM estimator (the J-Test). Within empirical likelihood estimation a test can be based on a likelihood ratio statistic, namely the scaled by $2n$ difference of the log-likelihood function (16) evaluated at $p_i = n^{-1}$ and $p_i(\hat{\beta}_{EL})$ respectively. Both test statistics are asymptotically chi-squared distributed under the null hypothesis of valid moment equations, with a critical value of $\chi^2_{(2),0.95} = 5.99$ at 5 percent level of significance. The values of the over-identifying test statistics are reported in Table 5 (GMM: 0.39, EL: 0.67), thus both tests cannot reject the null of valid moment restrictions.

7 Conclusions

In this paper I developed the empirical likelihood estimator for a count data model with standard mean specification taking into account that regressors might be endogenously determined. I considered the case of multiplicative errors in the mean and derived the first order conditions for the EL estimator. Based on Monte Carlo simulations I showed that empirical likelihood can improve upon GMM, particularly in situations when samples are small and instruments are weak. In such cases the use of EL is therefore strongly recommended.

Empirical likelihood as applied here estimates a count data model without assuming a conditional distribution function, and only specifying the mean function. These weak assumptions allow for robust estimation of the parameters of interest and results of the Monte Carlo experiment support this argument. On the other hand, we forego the possibility of predicting (out of sample) probabilities and calculating marginal probability effects.

In an empirical application, the EL method was used to estimate the effect of women's schooling on fertility based on 992 pooled cross sectional observations of the *General Social Survey* for even years from 1972 to 1984. To account for potential endogeneity of schooling, parent's schooling and the number of siblings were used as instruments. The EL point estimate of the schooling effect was substantially below the standard Poisson estimate. However, the null hypothesis of no endogeneity could not be rejected.

References

Chamberlain, G. (1987): "Asymptotic Efficiency in Estimation with Conditional Moment Restrictions," *Journal of Econometrics*, 34, 305–334.

Delgado, M., and T. Kniesner (1997): "Count Data Models with Variance of Unknown Form: An Application to a Hedonic Model of Worker Absenteeism," *Review of Economics and Statistics*, 79, 41–49.

Gourieroux, C., and A. Monfort (1995): *Statistics and Econometric Models*, vol. 1. Cambridge University Press, Cambridge.

Gourieroux, C., A. Monfort, and A. Trognon (1984): "Pseudo Maximum Likelihood Methods: Applications to Poisson Models," *Econometrica*, 52, 701–720.

Grogger, J. (1990): "A Simple Test for Exogeneity in Probit, Logit, and Poisson Regression Models," *Economics Letters*, 33, 329–332.

Guggenberger, P., and R. Smith (2003): "Generalized Empirical Likelihood Estimators and Tests under Partial, Weak and Strong Identification," cemmap working paper 08/03.

Hansen, L., J. Heaton, and A. Yaron (1996): "Finite-Sample Properties of Some Alternative GMM Estimators," *Journal of Business & Economic Statistics*, 14, 262–280.

Hansen, L. P. (1982): "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50, 1029–1054.

Hausman, J., B. Hall, and Z. Griliches (1984): "Econometric Models for Count Data with an Application to the Patents - R&D Relationship," *Econometrica*, 52, 909–938.

Imbens, G. (1997): "One-Step Estimators for Over-Identified Generalized Method of Moment Models," *Review of Economic Studies*, 64, 359–383.

Imbens, G., and R. Spady (2001): "The Performance of Empirical Likelihood and its Generalizations," paper presented at 2001 NSF-Berkeley Econometrics Symposium on "Identification and Inference for Econometric Models.

Imbens, G., R. Spady, and P. Johnson (1998): "Information Theoretic Approaches to Inference in Moment Condition Models," *Econometrica*, 66, 333–357.

Kitamura, Y., and M. Stutzer (1997): "An Information-Theoretic Alternative to Generalized Method of Moments Estimation," *Econometrica*, 65, 861–874.

Mullahy, J. (1997): "Instrumental-Variable Estimation of Count Data Models: Applications to Models of Cigarette Smoking Behavior," *Review of Economics and Statistics*, 79, 586–593.

Newey, W., and R. Smith (2000): "Asymptotic Bias and Equivalence of GMM and GEL Estimators," *unpublished working paper no. 01/517*, University of Bristol.

——— (2004): "Higher Order Properties of GMM and Generalized Empirical Likelihood Estimators," *Econometrica*, 72, 219–255.

Owen, A. (1988): "Empirical Likelihood Ratio Confidence Regions for a Single Functional," *Biometrika*, 75, 237–249.

——— (1991): "Empirical Likelihood for Linear Models," *Annals of Statistics*, 19, 1725–1747.

——— (2001): *Empirical Likelihood*. Chapman & Hall/CRC, Boca Raton.

Pohlmeier, W., and V. Ulrich (1995): "An Econometric Model of the Two-Part Decisionmaking Process in the Demand for Health Care," *Journal of Human Resources*, 30, 339–361.

Qin, J., and J. Lawless (1994): "Empirical Likelihood and General Estimating Equations," *Annals of Statistics*, 22, 300–325.

——— (1995): "Estimating Equations, Empirical Likelihood, and Constraints on Parameters," *Canadian Journal of Statistics*, 23, 145–159.

Sander, W. (1992): "The effect of women's schooling on fertility," *Economics Letters*, 40, 229–233.

Smith, R. (1997): "Alternative Semi-Parametric Likelihood Approaches to Generalised Method of Moments Estimation," *Economic Journal*, 107, 503–519.

Stock, J., J. Wright, and M. Yogo (2002): "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments," *Journal of Business & Economic Statistics*, 20, 518–529.

Windmeijer, F. A. G., and J. M. C. Santos Silva (1997): "Endogeneity in Count Data Models: An Application to Demand for Health Care," *Journal of Applied Econometrics*, 12, 281–294.

Winkelmann, R. (2003): *Econometric Analysis of Count Data*. Springer Verlag, Berlin.

Winkelmann, R., and K. Zimmermann (1994): "Count Data Models for Demographic Data," *Mathematical Population Studies*, 4, 205–221.

Wooldridge, J. (1997): "Quasi-Likelihood Methods for Count Data," in *Handbook of Applied Econometrics Vol. 2 - Microeconomics*, ed. by M. Pesaran, and P. Schmidt. Blackwell.

Table 1: Monte Carlo results part 1 ($n = 100$, 3 instruments for x_2)

	PML		GMM		EL	
				$\rho = 0.5, 0.5, 0.5$		
β_0	0.5964	(0.2830)	0.8614	(0.7278)	0.8526	(0.4237)
β_1	-0.9820	(0.5083)	-0.9845	(0.7684)	-0.9690	(0.5512)
β_2	-0.0406	(0.1353)	-0.1960	(0.7319)	-0.3235	(0.4947)
				$\rho = 0.1, 0.5, 0.9$		
β_0			0.9868	(1.2391)	0.7908	(0.2675)
β_1			-1.0510	(1.0880)	-0.9507	(0.3710)
β_2			-0.2790	(0.7600)	-0.3409	(0.2937)
				$\rho = 0.9, 0.9, 0.9$		
β_0			0.8972	(0.7687)	0.8502	(0.3793)
β_1			-1.0616	(0.7844)	-0.9965	(0.4797)
β_2			-0.2584	(0.6170)	-0.3312	(0.3986)
				$\rho = 0.1, 0.1, 0.5$		
β_0			0.8850	(0.8265)	0.7700	(0.2171)
β_1			-1.0337	(1.0563)	-0.9797	(0.3080)
β_2			-0.1050	(0.6959)	-0.2860	(0.2976)

first value: mean of β's, second value (in parentheses): standard deviation of β's; 1000 replications

true model: $E[y_i|x_i] = \exp(0.5 - x_{1i} - 0.5\,\tilde{x}_{2i} + \varepsilon_i)$, $\varepsilon_i \sim N(0, 0.7)$

Table 2: Monte Carlo results part 2 ($n = 500$, 3 instruments for x_2)

	PML		GMM		EL	
			$\rho = 0.5, 0.5, 0.5$			
β_0	0.6235	(0.1260)	0.9405	(0.3473)	0.9198	(0.2095)
β_1	-1.0021	(0.2216)	-1.0086	(0.2713)	-0.9878	(0.2317)
β_2	-0.0403	(0.0628)	-0.4389	(0.2826)	-0.4438	(0.2106)
			$\rho = 0.1, 0.5, 0.9$			
β_0			0.9186	(0.4016)	0.8893	(0.2046)
β_1			-1.0285	(0.2713)	-1.0064	(0.2260)
β_2			-0.3817	(0.3454)	-0.3928	(0.2344)
			$\rho = 0.9, 0.9, 0.9$			
β_0			0.9500	(0.3413)	0.9077	(0.1863)
β_1			-0.9967	(0.3215)	-0.9787	(0.2347)
β_2			-0.4561	(0.2884)	-0.4409	(0.1837)
			$\rho = 0.1, 0.1, 0.5$			
β_0			0.8938	(0.4709)	0.8361	(0.2231)
β_1			-1.0004	(0.2915)	-0.9616	(0.2650)
β_2			-0.3106	(0.4290)	-0.3834	(0.2498)

first value: mean of β's, second value (in parentheses): standard deviation of β's; 1000 replications

true model: $E[y_i|x_i] = \exp(0.5 - x_{1i} - 0.5\,\tilde{x}_{2i} + \varepsilon_i)$, $\varepsilon_i \sim N(0, 0.7)$

Table 3: Monte Carlo results part 3 ($n = 1000$, 3 instruments for x_2)

	PML		GMM		EL	
			$\rho = 0.5,\ 0.5,\ 0.5$			
β_0	0.6239	(0.0899)	0.9473	(0.3278)	0.9021	(0.1546)
β_1	-0.9963	(0.1620)	-0.9959	(0.2045)	-0.9791	(0.1749)
β_2	-0.0376	(0.0432)	-0.4450	(0.2703)	-0.4359	(0.1736)
			$\rho = 0.1,\ 0.5,\ 0.9$			
β_0			0.9524	(0.2543)	0.8753	(0.1398)
β_1			-1.0004	(0.1833)	-0.9739	(0.1727)
β_2			-0.4616	(0.2062)	-0.4285	(0.1475)
			$\rho = 0.9,\ 0.9,\ 0.9$			
β_0			0.9656	(0.2448)	0.9239	(0.1399)
β_1			-1.0038	(0.1867)	-0.9808	(0.1587)
β_2			-0.4799	(0.2002)	-0.4563	(0.1354)
			$\rho = 0.1,\ 0.1,\ 0.5$			
β_0			0.9174	(0.3330)	0.8517	(0.1557)
β_1			-0.9985	(0.2050)	-0.9649	(0.1885)
β_2			-0.3807	(0.3377)	-0.3975	(0.1806)

first value: mean of β's, second value (in parentheses): standard deviation of β's; 1000 replications

true model: $E[y_i|x_i] = \exp(0.5 - x_{1i} - 0.5\,\tilde{x}_{2i} + \varepsilon_i)$, $\varepsilon_i \sim N(0, 0.7)$

Table 4: Monte Carlo results part 4 ($n = 500$, 5 and 10 instruments for x_2)

	PML		GMM		EL	
$n = 500$			$\rho = 0.5, 0.5, 0.5, 0.5, 0.5$			
β_0	0.6235	(0.1260)	1.0647	(0.2933)	0.8685	(0.1781)
β_1	-1.0021	(0.2216)	-0.8783	(0.2055)	-0.9897	(0.2106)
β_2	-0.0403	(0.0628)	-0.2947	(0.2259)	-0.4627	(0.1864)
			$\rho = 0.1, 0.3, 0.5, 0.7, 0.9$			
β_0			0.7703	(0.1397)	0.8506	(0.1086)
β_1			-1.0261	(0.5246)	-0.9551	(0.1481)
β_2			-0.4067	(0.1648)	-0.4065	(0.1224)
			$\rho = 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5$			
β_0			0.7599	(0.1524)	0.8395	(0.1785)
β_1			-1.0219	(0.3242)	-0.9944	(0.2084)
β_2			-0.2794	(0.4579)	-0.4262	(0.2947)
			$\rho = 0.1, 0.2, 0.3, 0.4, 0.5, 0.5, 0.6, 0.7, 0.8, 0.9$			
β_0			0.9349	(0.3076)	0.8305	(0.1209)
β_1			-0.9975	(0.4601)	-0.9662	(0.1527)
β_2			-0.4219	(0.3042)	-0.3902	(0.1420)

first value: mean of β's, second value (in parentheses): standard deviation of β's; 500 replications

true model: $E[y_i|x_i] = \exp(0.5 - x_{1i} - 0.5\,\tilde{x}_{2i} + \varepsilon_i)$, $\varepsilon_i \sim N(0, 0.7)$

Table 5: Estimates of children ever born to women age 35 to 54

3 instruments within 2SLS, GMM, and EL (standard errors in parentheses)					
	OLS	2SLS	PML	GMM	EL
schooling	-0.1133 ***	-0.1661 ***	-0.0428 ***	-0.0669 ***	-0.0740 ***
	(0.0199)	(0.0400)	(0.0086)	(0.0168)	(0.0166)
age	0.5592 ***	0.5391 ***	0.2180 ***	0.2465 ***	0.2300 ***
	(0.1469)	(0.1480)	(0.0567)	(0.0592)	(0.0579)
age^2	-0.0061 ***	-0.0059 ***	-0.0024 ***	-0.0027 ***	-0.0024 ***
	(0.0017)	(0.0017)	(0.0006)	(0.0007)	(0.0007)
black	0.9450 ***	0.9334 ***	0.3147 ***	0.3029 ***	0.3012 ***
	(0.1985)	(0.1994)	(0.0678)	(0.0726)	(0.0712)
west	-0.1095	-0.1047	-0.0373	-0.0325	-0.0320
	(0.2033)	(0.2041)	(0.0719)	(0.0770)	(0.0755)
north central	0.1132	0.1171	0.0461	0.0599	0.0581
	(0.1584)	(0.1590)	(0.0548)	(0.0580)	(0.0568)
east	-0.2846 **	-0.2863 **	-0.1048 **	-0.0841 *	-0.0766 *
	(0.1543)	(0.1549)	(0.0539)	(0.0576)	(0.0568)
farm	-0.2156 *	-0.2667 **	-0.078 *	-0.1167 **	-0.1135 **
	(0.1550)	(0.1592)	(0.0581)	(0.0636)	(0.0622)
other rural	-0.0157	-0.0897	-0.0021	0.0010	-0.0004
	(0.1885)	(0.1953)	(0.0694)	(0.0738)	(0.0738)
town	0.0792	0.0650	0.0288	0.0270	0.0239
	(0.1277)	(0.1285)	(0.0496)	(0.0516)	(0.0507)
smallcity	0.2411 *	0.2386 *	0.0898 *	0.0909 *	0.0781
	(0.1653)	(0.1659)	(0.0583)	(0.0633)	(0.0626)
intercept	-8.2404 ***	-7.0952 **	-3.3272 ***	-3.7096 ***	-3.3308 ***
	(3.2394)	(3.3369)	(1.2573)	(1.3234)	(1.2960)
time dummies:	yes				
number of observations:	992				
instruments for *schooling*:	father's schooling, mother's schooling, siblings				
over-identifying test statistic:				0.3860	0.6717

significance levels : * 10% ** 5% *** 1%

Working Papers of the Socioeconomic Institute at the University of Zurich

The Working Papers of the Socioeconomic Institute can be downloaded from http://www.soi.unizh.ch/research/wp/index2.html

0404	Empirical Likelihood in Count Data Models: The Case of Endogenous Regressors Stefan Boes, March 2004, 22 p.
0403	Globalization and General Worker Training Hans Gersbach and Armin Schmutzler, February 2004, 37 p.
0402	Restructuring Network Industries: Dealing with Price-Quality Tradeoffs Stefan Bühler, Dennis Gärtner and Daniel Halbheer, January 2004, 18 p.
0401	Deductible or Co-Insurance: Which is the Better Insurance Contract under Adverse Selection? Michael Breuer, January 2004, 18 p.
0314	How Did the German Health Care Reform of 1997 Change the Distribution of the Demand for Health Services? Rainer Winkelmann, December 2003, 20 p.
0313	Validity of Discrete-Choice Experiments – Evidence for Health Risk Reduction Harry Telser and Peter Zweifel, October 2003, 18 p.
0312	Parental Separation and Well-Being of Youths Rainer Winkelmann, October 2003, 20 p.
0311	Re-evaluating an Evaluation Study: The Case of the German Health Care Reform of 1997 Rainer Winkelmann, October 2003, 23 p.
0310	Downstream Investment in Oligopoly Stefan Buehler and Armin Schmutzler, September 2003, 33 p.
0309	Earning Differentials between German and French Speakers in Switzerland Alejandra Cattaneo and Rainer Winkelmann, September 2003, 27 p.
0308	Training Intensity and First Labor Market Outcomes of Apprenticeship Graduates Rob Euwals and Rainer Winkelmann, September 2003, 25 p.
0307	Co-payments for prescription drugs and the demand for doctor visits – Evidence from a natural experiment Rainer Winkelmann, September 2003, 22 p.
0306	Who Integrates? Stefan Buehler and Armin Schmutzler, August 2003, 29 p.
0305	Strategic Outsourcing Revisited Stefan Buehler and Justus Haucap, July 2003, 22 p.
0304	What does it take to sell Environmental Policy? An Empirical Analysis for Switzerland Daniel Halbheer, Sarah Niggli and Armin Schmutzler, 2003, 30 p.
0303	Mobile Number Portability Stefan Buehler and Justus Haucap, 2003, 12 p.
0302	Multiple Losses, Ex-Ante Moral Hazard, and the Non-Optimality of the Standard Insurance Contract Michael Breuer, 2003, 18 p.
0301	Lobbying against Environmental Regulation vs. Lobbying for Loopholes Andreas Polk and Armin Schmutzler, 2003, 37 p.
0214	A Product Market Theory of Worker Training Hans Gersbach and Armin Schmutzler, 2002, 34 p.
0213	Weddings with Uncertain Prospects – Mergers under Asymmetric Information Thomas Borek, Stefan Buehler and Armin Schmutzler, 2002, 35 p.

0212 Estimating Vertical Foreclosure in U.S. Gasoline Supply
Zava Aydemir and Stefan Buehler, 2002, 42 p.

0211 How much Internalization of Nuclear RiskThrough Liability Insurance?
Yves Schneider and Peter Zweifel, 2002, 18 p.

0210 Health Care Reform and the Number of Doctor Visits ? An Econometric Analysis
Rainer Winkelmann, 2002, 32p.

0209 Infrastructure Quality in Deregulated Industries: Is there an Underinvestment Problem?
Stefan Buehler, Armin Schmutzler and Men-Andri Benz, 2002, 24 p.

0208 Acquisitions versus Entry: The Evolution of Concentration
Zava Aydemir and Armin Schmutzler, 2002, 35 p.

0207 Subjektive Daten in der empirischen Wirtschaftsforschung: Probleme und Perspektiven.
Rainer Winkelmann, 2002, 25 p.

0206 How Special Interests Shape Policy - A Survey
Andreas Polk, 2002, 63 p.

0205 Lobbying Activities of Multinational Firms
Andreas Polk, 2002, 32 p.

0204 Subjective Well-being and the Family
Rainer Winkelmann, 2002, 18 p.

0203 Work and health in Switzerland: Immigrants and Natives
Rainer Winkelmann, 2002, 27 p.

0202 Why do firms recruit internationally? Results from the IZA International Employer Survey 2000
Rainer Winkelmann, 2002, 25 p.

0201 Multilateral Agreement On Investments (MAI) - A Critical Assessment From An Industrial Economics Point Of View
Andreas Polk, 2002, 25 p.