STATISTICS IN MEDICINE
Statist. Med. 2008; **27**:1539–1556
Published online 10 September 2007 in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/sim.3036

Use of instrumental variables in the analysis of generalized linear models in the presence of unmeasured confounding with applications to epidemiological research

K. M. Johnston[1,2,*,†], P. Gustafson[3], A. R. Levy[1,2] and P. Grootendorst[4,5]

[1]*Department of Health Care and Epidemiology, University of British Columbia, Vancouver, Canada V6T 1Z3*
[2]*Centre for Health Evaluation and Outcomes Science, St. Paul's Hospital, Vancouver, Canada*
[3]*Department of Statistics, University of British Columbia, Canada*
[4]*Faculty of Pharmacy, University of Toronto, Canada*
[5]*Department of Economics, McMaster University, Canada*

SUMMARY

A major, often unstated, concern of researchers carrying out epidemiological studies of medical therapy is the potential impact on validity if estimates of treatment are biased due to unmeasured confounders. One technique for obtaining consistent estimates of treatment effects in the presence of unmeasured confounders is instrumental variables analysis (IVA). This technique has been well developed in the econometrics literature and is being increasingly used in epidemiological studies. However, the approach to IVA that is most commonly used in such studies is based on linear models, while many epidemiological applications make use of non-linear models, specifically generalized linear models (GLMs) such as logistic or Poisson regression. Here we present a simple method for applying IVA within the class of GLMs using the generalized method of moments approach. We explore some of the theoretical properties of the method and illustrate its use within both a simulation example and an epidemiological study where unmeasured confounding is suspected to be present. We estimate the effects of beta-blocker therapy on one-year all-cause mortality after an incident hospitalization for heart failure, in the absence of data describing disease severity, which is believed to be a confounder. Copyright © 2007 John Wiley & Sons, Ltd.

KEY WORDS: instrumental variables analysis; generalized linear models; epidemiology generalized method of moments; unmeasured confounding

*Correspondence to: K. M. Johnston, Centre for Health Evaluation and Outcomes Science, St. Paul's Hospital, 620B-1081 Burrard St., Vancouver, BC, Canada.
†E-mail: karissaj@interchange.ubc.ca, johnston@stat.ubc.ca

Contract/grant sponsor: Canadian Institutes of Health Research
Contract/grant sponsor: Michael Smith Foundation for Health Research

Received 24 May 2006
Accepted 9 July 2007

Copyright © 2007 John Wiley & Sons, Ltd.

1. INTRODUCTION

Health scientists often use observational data to estimate treatment effects when controlled experiments are infeasible or unethical. A limitation of observational research is the non-random selection of subjects into different treatments, potentially leading to selection bias. The most commonly used solution to this problem is to identify, measure, and adjust for confounding variables that are correlated with both treatment and outcome. One limitation to covariate adjustment, however, is that the analyst might fail to adjust for pertinent confounding variables, because they are either unknown or not readily quantifiable. Instrumental variables analysis (IVA) can be a viable alternative. IVA exploits quasi-experimental variation in treatment assignment that is incidental to the health outcome being studied. This variation in treatment assignment is used to assess attendant effects on health outcomes. IVA is being increasingly used in epidemiology (see, for instance, [1–3]). However, it is not always used correctly. The standard IVA estimator requires that health outcome models be linear in parameters. Yet some analysts have applied the standard IVA estimator to non-linear models, such as the logistic regression model, despite the fact that, with a few exceptions, the IVA estimator is inconsistent in such contexts. Fortunately, generalized methods of moments (GMM) estimation can be used to consistently estimate the parameters of non-linear models in which selection bias is suspected.

In this paper, we present a simple method for applying IVA within the class of generalized linear models (GLMs) using GMM and identify a subset of GLMs in which GMM estimation is equivalent to maximum likelihood estimation (MLE). We illustrate the method within a simulation example as well as in the context of an epidemiological study of beta-blockers—a new long-term treatment for heart failure—among elderly persons in British Columbia, Canada. Specifically, we estimate the impact of beta-blocker therapy on one-year all-cause mortality after a first hospital episode for heart failure. The instrumental variable used was the time period, which we suggest may satisfy the criteria for a valid instrument.

2. INSTRUMENTAL VARIABLES

Given an outcome measure, a treatment variable, a set of unmeasured confounders, and, possibly, a set of measured confounders, a variable satisfies the conditions of an *instrumental variable* (also referred to as an *instrument*) if the following conditions hold: (1) it is correlated with the treatment variable; (2) given the treatment variable and all confounders (measured and unmeasured), it is conditionally independent of outcome; and (3) it is independent of the entire set of unmeasured confounders [4]. The strength of the instrument is determined by the correlation between treatment and instrument, with strong instruments being preferable. An intuitive example of a 'perfect' instrument would be to define the instrument as the randomization mechanism in a randomized controlled trial (RCT). Assuming 100 per cent compliance, randomization is perfectly correlated with treatment but is independent of all possible unmeasured confounders, and after accounting for differences in treatment, randomization should be independent of outcome. However, in most practical situations, identification of an instrument is less clear than this and is highly dependent on the specific problem at hand. Because the necessary assumptions cannot be verified empirically, the selection of an instrument is based on subject-matter knowledge, not statistical testing. Assuming that an appropriate instrument can be found, a number of analytical methods exist to use it to

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/sim

quantify the effects of treatment on outcome, with two-stage least-squares (2SLS) being a common example.

The 2SLS consists of applying least-squares regression twice. In the first model, the instrument and a set of explanatory variables are used to predict the treatment variable. The second model relates treatment and the same set of explanatory variables to the outcome of interest, with the treatment variable replaced by its predicted value from the first model. Although both models are linear, it has been shown that no assumptions are necessary regarding the respective distributions of the error terms [5], suggesting that 2SLS is appropriate for non-continuous outcomes, such as binary or count variables. While linear models can be useful in describing the direction and magnitude of additive effects, when working with such outcomes, it may be preferable to employ methods that specifically account for the form of the data (e.g. logistic and Poisson regression for binary and count data, respectively). Such outcome-specific methods make it more likely that model-based predictions will take on plausible values, and treatment effects are more likely to be directly comparable with effects reported by other investigators conducting similar studies. For example, in the case of a binary outcome, logistic regression produces parameter estimates that can be easily converted to odds ratios, and this is the standard form in which estimated treatment effects are reported. One approach to this problem is a method that has been called 'pseudo-2SLS' for binary data; it is analogous to 2SLS, but linear regression is replaced by logistic regression, allowing odds ratios to be estimated. This method may be satisfactory for a precursory assessment of effects or for estimating the direction of effects, but it is not based on statistical theory and the resulting estimated odds ratios have been shown to be inconsistent estimates of the true effects [6]. However, IVA is not restricted to 2SLS, and more general methods can be shown to explicitly accommodate non-linear models.

3. GENERALIZED METHOD OF MOMENTS ESTIMATORS

GMM estimators are an extremely broad class of estimators that include maximum likelihood, moment, and least-squares estimator as specific examples. GMM estimation involves identifying a set of *moment functions* of known variables and unknown parameters such that the functions have an expected value of zero. The moment functions can be simultaneously set to zero and solved, providing a consistent estimate of parameter values [7, 8]. In the case of linear and non-linear regression models, GMM estimation can be used as follows [9].

Consider a model defined by the orthogonality conditions:

$$E\{Z_j\boldsymbol{\varepsilon}\} = E\{Z_j[Y - h(\mathbf{X};\boldsymbol{\beta})]\} = 0, \quad j = 1 \ldots J \tag{1}$$

for an outcome variable Y, matrix of explanatory variables $\mathbf{X}$, a (possibly non-linear) function h, variables $Z^{(1)} \ldots Z^{(J)}$, and residuals $\boldsymbol{\varepsilon}$. Estimation of $\boldsymbol{\beta}$ using GMM proceeds by identifying variables $Z^{(1)} \ldots Z^{(J)}$ that are thought to satisfy orthogonality conditions (1), and $\hat{\boldsymbol{\beta}}$ can then be solved for using sample moment conditions:

$$\sum_{i=1}^{n} z_{ij} e_i = \sum_{i=1}^{n} z_{ij}[y_i - h(\mathbf{x_i};\hat{\boldsymbol{\beta}})] = 0, \quad j = 1 \ldots J \tag{2}$$

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/sim

Here, $\boldsymbol{\beta}$ is regarded as the target of inferential interest in the sense that the expectation of $\mathbf{Y}$ given $\mathbf{X}$ and unobserved confounders is $h(\mathbf{X}; \boldsymbol{\beta})$ plus an additive effect of the unobserved confounders. As a result of this unmeasured confounding, simply taking $Z_j = X_j$ will not satisfy (1) for all j. However, any Z_j that satisfies the assumptions for an instrumental variable will satisfy (1); further elaboration on this point is given in Appendix A. When the Z_j's are chosen to be a collection of instrumental variables, we will refer to this estimation method as the *generalized method of moments instrumental variables analysis* (GMM IVA). If h is a linear function, GMM IVA is equivalent to 2SLS; hence it can be thought of as a more general form of IVA that includes 2SLS as a special case [10]. Variables that are affected by unmeasured confounding can be replaced in (2) by an appropriate instrumental variable, whereas variables that are unaffected by unmeasured confounding satisfy the necessary orthogonality conditions and can serve as their own respective instruments. Thus, in the special case where there is assumed to be no unmeasured confounders, instrumental variables are not necessary, and (2) can be re-written as

$$\sum_{i=1}^{n} x_{ij}[y_i - h(\mathbf{x_i}; \boldsymbol{\beta})] = 0, \quad j = 1, \ldots, J \tag{3}$$

which is the standard specification for a regression model. If $\boldsymbol{\beta}$ is of dimension $p \times 1$, then J, the number of instruments, must be at least as large as p in order to identify estimates for $\boldsymbol{\beta}$. In the case where $J = p$, the system is said to be exactly identified, and system (2) can be solved exactly. Estimation of $\boldsymbol{\beta}$ is also possible when $J > p$, referred to as the over-identified case. In the over-identified case, it is possible to empirically test whether candidate instruments satisfy (1) [7], and to compare models using criteria analogous to Akaike, Bayesian, or Hannan–Quinn Information Criteria, respectively [11]. Whereas much work in econometrics does use multiple instruments, applications in epidemiology tend to involve a single instrument, likely reflecting the challenge in identifying valid instruments in typical epidemiological data sources. However, in studies where multiple instruments can be identified, some of the limitations to IVA can potentially be minimized. Below we present an extension of (2) that can be used to estimate $\boldsymbol{\beta}$ in both the exactly identified and over-identified cases.

Let

$$\mathbf{Z} = (\mathbf{z_1} \;\; \mathbf{z_2} \;\; \cdots \;\; \mathbf{z_n})'$$

$$\tilde{\mathbf{x}}_\mathbf{i} = -\frac{\partial \varepsilon_i}{\partial \boldsymbol{\beta}} = \left(-\frac{\partial \varepsilon_i}{\partial \beta_1} \;\; -\frac{\partial \varepsilon_i}{\partial \beta_2} \;\; \cdots \;\; -\frac{\partial \varepsilon_i}{\partial \beta_J} \right)'$$

$$\tilde{\mathbf{X}} = (\tilde{\mathbf{x}}_\mathbf{1} \;\; \tilde{\mathbf{x}}_\mathbf{2} \;\; \cdots \;\; \tilde{\mathbf{x}}_\mathbf{n})'$$

$$\mathbf{e} = (e_1 \ldots e_n)'$$

$$= (y_1 - h(\mathbf{x_1}; \hat{\boldsymbol{\beta}}) \ldots y_n - h(\mathbf{x_n}; \hat{\boldsymbol{\beta}}))'$$

$$\boldsymbol{\Omega}_\mathbf{ij} = (\mathrm{Cov}(\varepsilon_i, \varepsilon_j)), \quad i = 1, \ldots, n, \quad j = 1, \ldots, n$$

To summarize, $\mathbf{Z}$ is a matrix of instruments analogous to a design matrix; $\tilde{\mathbf{X}}$ is a matrix of negative residual derivatives with respect to each parameter element; $\mathbf{e}$ is a vector of sample residuals; and $\boldsymbol{\Omega}$ is the variance–covariance matrix of the residuals.

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/sim

In the over-identified case, the moment conditions given by (2) cannot be solved exactly, but estimates of $\boldsymbol{\beta}$ can instead be obtained by minimizing

$$q = \mathbf{e}'\mathbf{Z}[\mathbf{Z}'\boldsymbol{\Omega}\mathbf{Z}]^{-1}\mathbf{Z}'\mathbf{e} \tag{4}$$

This method can also be used in the exactly identified case, where q will minimize to exactly 0. In either case, the variance–covariance matrix of $\hat{\boldsymbol{\beta}}$ is [9]

$$([\tilde{\mathbf{X}'}\mathbf{Z}][\mathbf{Z}'\boldsymbol{\Omega}\mathbf{Z}]^{-1}[\mathbf{Z}'\tilde{\mathbf{X}}])^{-1} \tag{5}$$

In practice, $\boldsymbol{\Omega}$ is not known, and the estimator of White [12] is commonly used: $\mathbf{Z}'\boldsymbol{\Omega}\mathbf{Z} = \sum_{i=1}^{n} \mathbf{z_i}\mathbf{z_i}'\varepsilon_i^2$, which can be estimated using sample residuals as $\sum_{i=1}^{n} \mathbf{z_i}\mathbf{z_i}'e_i^2$. This presents a problem of circularity, as estimates of $\boldsymbol{\beta}$ are required to obtain sample residuals $\mathbf{e}$, and sample residuals are used to estimate $\boldsymbol{\beta}$. Greene [9] suggests a two-stage approach, first obtaining $\mathbf{e}$ by assuming $\mathbf{Z}'\boldsymbol{\Omega}\mathbf{Z}$ to be proportional to I_{nxn}, which is consistent but inefficient, then using these results to estimate $\boldsymbol{\beta}$.

Note that (1) specifies a model that is averaged across possible values for unmeasured confounders but does not specify the relationship between $\mathbf{x_i}$, $\mathbf{y_i}$ and unmeasured confounders $\mathbf{u_i}$ for specific values of $\mathbf{u_i}$. This has been suggested as a potential limitation to using GMM IVA in non-linear models, as the method may not yield a consistent estimator of $\boldsymbol{\beta}$ for a plausible form of the relationship, referred to as *symmetric* [13, 14]. In Appendix A, however, we demonstrate that, even in the case of a symmetric model, $\hat{\boldsymbol{\beta}}$ estimated with GMM IVA can be shown to be an approximately consistent estimator of $\boldsymbol{\beta}$.

4. GENERALIZED LINEAR MODELS

We will restrict our attention to GMM IVA within the class of GLMs, which encompass most commonly used linear and non-linear models in epidemiology. Briefly, a GLM is a model of the form of (1) that is defined by three components, the random component, defined by a distribution from the exponential family with mean μ_i; the linear predictor, η_i; and the link function, $\eta_i = g(\mu_i)$, with the corresponding mean function, $\mu_i = h(\eta_i)$ [15]. This function h is equivalent to the h function in (1), with $h(\mathbf{x_i}; \boldsymbol{\beta}) = h(\mathbf{x_i}'\boldsymbol{\beta})$.

The exponential family refers to a class of distributions whose density functions can be written as

$$f(y; \theta, \phi) = \exp\left\{\frac{y\theta - b(\theta)}{a(\phi)} + c(y, \phi)\right\} \tag{6}$$

for parameters θ and ϕ, and functions a, b, and c. If a link function is chosen so that $\eta_i = \theta_i$, it is said to be a *canonical* link. Examples of GLMs using a canonical link are ordinary linear regression, logistic regression, and Poisson regression using a log link.

Table I provides formulas for the elements of $\tilde{\mathbf{X}}$ for some of the most commonly encountered GLMs.

4.1. MLE of β

If we assume no unmeasured confounders, estimation of $\boldsymbol{\beta}$ in a GLM proceeds *via* MLE by setting the derivative of the log-likelihood of the data to 0 and solving for $\boldsymbol{\beta}$. It can be shown that this is

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/sim

Table I. Generalized method of moments instrumental variables analysis formulas for some common generalized linear models.

Model	$\varepsilon_i = y_i - h(\mathbf{x_i'\beta})$	$\tilde{\mathbf{x}}_{\mathbf{ij}} = -\frac{\partial \varepsilon_i}{\partial \beta_j}$
Linear	$y_i - \mathbf{x_i'\beta}$	x_{ij}
Logistic	$y_i - \frac{1}{1+e^{-\mathbf{x_i'\beta}}}$	$\frac{x_{ij}e^{-\mathbf{x_i'\beta}}}{(1+e^{-\mathbf{x_i'\beta}})^2}$
Poisson (log link)	$y_i - e^{\mathbf{x_i'\beta}}$	$x_{ij}e^{\mathbf{x_i'\beta}}$
Gamma (inverse link)	$y_i - \frac{1}{\mathbf{x_i'\beta}}$	$\frac{x_{ij}}{(\mathbf{x_i'\beta})^2}$

equivalent to solving the following system of equations:

$$\sum_{i=1}^{n} x_{ij}[y_i - h(\mathbf{x_i'\beta})](\mathrm{Var}(y_i))^{-1}\left(\frac{\partial \mu_i}{\partial \eta_i}\right) = 0, \quad j = 1, \ldots, J \tag{7}$$

Clearly, system (7) is not necessarily equivalent to system (3). Although both are valid methods for estimating $\boldsymbol{\beta}$ under the assumption of no unmeasured confounders, (7) is the standard method used to fit a GLM. If GMM IVA is to be used to adjust for unmeasured confounding in a GLM, it is desirable that (3) be equivalent to (7) in the special case of no unmeasured confounders. In Appendix B we show that (3) and (7) are equivalent for all standard GLMs using a canonical link. We also explore the difference between (3) and (7), both theoretically and empirically, for some commonly used non-canonical links (probit and complementary log–log links for binomial data and log link for negative binomial data).

5. ILLUSTRATION OF GMM IVA

5.1. Applied epidemiological example

Background: Here we illustrate GMM IVA, using an example describing the effect of beta-blocker use on mortality after an incident hospitalization for heart failure among elderly individuals in British Columbia, Canada, during 1993–2000. Prior to 1995, beta-blockers were thought to be contraindicated in individuals with heart failure. However, in 1995, based on the outcomes of several large RCTs [16, 17], consensus guidelines were altered to tentatively recommend beta-blockers as a therapy for individuals with heart failure [18]; in 2001, a full reversal in guidelines occurred, and beta-blockers became recommended as routine treatment for heart failure [19].

RCTs are the gold standard in estimating medication efficacy, as random allocation to treatment group can control for the effects of all confounders, both measured and unmeasured. However, results from an RCT may be of limited generalizability if the sample were not representative of the target population in clinical practice [20]. This has been found to be the case for RCTs of treatments for heart failure, with the elderly, women, and minorities systematically underrepresented [21]. Advantages in validity offered by RCTs are not offset by potential limitations in generalizability, and randomization remains the only way to ensure that all potential confounders are balanced between

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/sim

treatment groups. However, well-designed observational studies can complement the knowledge gained through RCTs by describing the actual use of medications in routine settings [22].

The objectives here were to quantify the effect of beta-blocker use on mortality in regular practice and to potentially account for unmeasured confounders including disease severity. Confounding by severity could plausibly occur in either direction. Physicians may have been concerned that individuals with more severe heart failure would be unable to tolerate beta-blocker therapy, particularly given that until recently it had been considered to be contraindicated. If this were the case, using standard logistic regression to quantify the effect of beta-blocker prescription on mortality without adjusting for disease severity would be expected to produce a biased estimate in which beta-blockers would appear to be more effective in preventing mortality than they actually are, given that individuals with the most severe disease are also most likely to experience mortality within one year. Conversely, it is also plausible that physicians would be reluctant to prescribe beta-blockers to their healthier patients for fear of unintended negative effects and would tend to prescribe them as a 'last resort' for their sickest patients. In this case, the expected bias would be in the opposite direction, with beta-blockers prescribed to the patients most likely to experience mortality and thus appearing to be less effective in preventing mortality than they actually are.

Methods: The data were population based for the province of British Columbia (BC) and consisted of prescription records linked with hospital separation abstracts and vital statistics. We included records for all individuals over the age of 65 years who were released from hospital after an incident episode of heart failure during 1993–2000. Multiple hospitalizations for a single individual were not included, and records from 1990 to 1992 were used to identify and exclude prevalent cases from the study period. The separation abstracts contained information describing the length of hospitalization, demographics, and co-morbidities, but no description of disease severity. Treatment was defined as an beta-blocker prescription within 30 days of hospital separation. The outcome of interest was defined as an all-cause mortality within one year of hospital separation. To remove bias due to censoring, observations were restricted to individuals discharged from hospital alive, surviving for at least 30 days, and discharged from hospital at least one year prior to the study cut-off date and therefore eligible to contribute one year of study time.

Time period was suggested as a valid instrument. Time period was treated as a binary variable, indicating hospitalizations occurring on or before 1998 versus after 1998. This year was chosen because, given the strong reversal in guidelines, uptake of beta-blocker therapy would be expected to be delayed. While we observed an increase in beta-blocker prescription rates every year after 1995, the largest increases were observed after 1998 (Figure 1). We also performed the analysis treating time as a continuous variable, as well as a binary variable using 1997 as a cutpoint.

Because of the changes in guidelines, beta-blocker use was expected to have increased over time, implying that time period should satisfy the condition that an instrumental variable be associated with treatment. A method has been proposed to formally test this assumption [23], but we felt the observed relationship (Figure 1) was sufficient to establish instrument strength.

With respect to the second assumption required of an instrument, i.e. that it be independent of outcome, conditional on treatment and unmeasured confounders, there may have been other innovations during the time period of interest which affected the observed mortality rate over time. However, we would suggest that the change in guidelines regarding beta-blockers was the most substantial. Given treatment status and disease severity, we are not aware of other major temporal patterns in heart failure diagnosis or treatment in BC during the study period. Empirically, this hypothesis is also supported by these data in that, for standard logistic regression models, there was a statistically significant association observed between time and mortality in models that did

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/sim

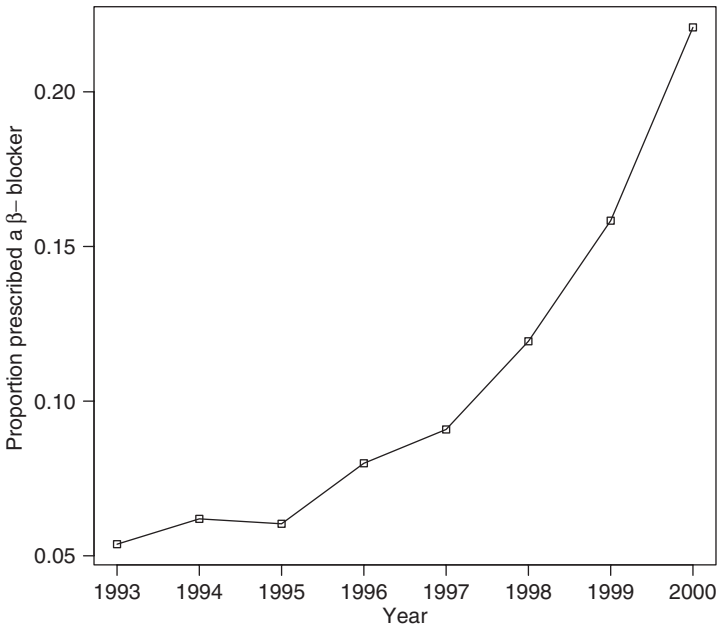

Figure 1. Proportion of elderly individuals prescribed a beta-blocker within 30 days of separation after incident hospitalization for heart failure, 1993–2000.

not adjust for beta-blocker use, but this association was dampened and became statistically non-significant in models that did adjust for beta-blocker use. This occurred consistently for every form of the time variable that was considered here. While we agree that these logistic regression models may have been biased due to unmeasured confounding, this finding does provide preliminary evidence that beta-blocker uptake was the most important contributor to temporal patterns in heart failure mortality and that the second assumption may be approximated if not met.

The third assumption requires that time be independent of all unmeasured confounders. When considering disease severity as a likely important unmeasured confounder, this is equivalent to assuming that incident hospitalized cases of heart failure in BC were of similar severity throughout the study period. This assumption may have been violated due to increased pressure on hospital beds over this time period [24], possibly resulting in less severe cases not being admitted to hospital during the later time period. We empirically tested this by examining temporal patterns in in-hospital case-fatality, co-morbidity scores, and length of stay, three potential correlates of severity. All three showed a small positive relationship with time, suggesting that severity of hospitalized heart failure may have increased in time. If this is the case, while the third assumption would be violated, the resulting bias in the GMM IVA estimate should be conservative, since increasing severity over time would imply increased mortality, the opposite effect to that expected from increased beta-blocker uptake. Conversely, subtle changes over time in diagnosis, classification, and patterns in referral may have occurred, resulting in changes to the definition of heart failure over time. If such changes were also related to the probability of being prescribed a beta-blocker and, further, resulted in improved prognoses for heart failure patients, the GMM IVA estimate would be biased away from the null hypothesis of no effect.

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/sim

Table II. Distribution of explanatory variables ($n = 21\,217$).

Variable		
Beta-blocker use	Per cent prescribed a beta-blocker	10.5
Sex	Per cent male	50.2
Age group	65–74 y	32 per cent
	75–84 y	46 per cent
	$\geqslant$85 y	22 per cent
Contraindication	Per cent contraindicated	24.3
Polanczyk co-morbidity score	Mean	5.2
	95 per cent CI	(0.9–15.5)
	Range	(3.0–27.0)

Table III. Estimated odds ratios (95 per cent confidence intervals) with one-year mortality after incident hospitalization for heart failure in using logistic regression and generalized method of moments instrumental variables analysis (GMM IVA).

	Estimated odds ratio (95 per cent CI)			
Variable	Logistic regression		GMM IVA	
Beta-blocker use within 30 days	0.68	(0.60–0.77)	0.23	(0.04–1.09)
Sex*	0.69	(0.64–0.74)	0.70	(0.65–0.75)
Age group†				
75–84 y	1.33	(1.23–1.45)	1.32	(1.21–1.44)
$\geqslant$85 y	1.85	(1.68–2.04)	1.81	(1.64–2.01)
Contraindication	1.27	(1.17–1.37)	1.26	(1.17–1.36)
Polanczyk co-morbidity score	1.10	(1.09–1.12)	1.10	(1.09–1.12)

*Reference group: males.

†Reference group: 65–74 y.

The explanatory variables considered were age, sex, contraindication to beta-blocker therapy (at least one of asthma, diabetes, or chronic obstructive pulmonary disease (COPD)), and Polanczyk co-morbidity score [25]. Briefly, the Polanczyk co-morbidity score is a validated heart failure-specific case-mix adjustment, with higher scores indicating greater co-morbidity burden. Table II describes the observed distributions of these variables.

Two models were fit: a logistic regression model that did not account for unmeasured confounders and a GMM IVA extension of logistic regression, using time period as an instrument. Both models adjusted for all covariates discussed above. In the GMM IVA model, all variables besides beta-blocker therapy served as their own respective instruments.

Results: Results of the logistic regression and GMM IVA models are reported in Table III. Odds ratios for all explanatory variables were similar between the two models, although confidence intervals for beta-blocker use were wider for the GMM IVA model; it has been noted previously that IVA tends to have reduced efficiency [6, 9]. However, the point estimate for the odds ratio between beta-blocker therapy and mortality decreased more than two-fold in the GMM IVA

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/sim

model, supporting the hypothesis that the treatment effect resulting from the first model may have been underestimated due to the confounding effects of disease severity. When time period was categorized in alternative ways, results remained similar, with estimated odds ratios between beta-blocker use and mortality ranging from 0.23 to 0.41, suggesting that treatment effects were robust to the specific categorization chosen for the instrument. The wide confidence interval surrounding treatment effect prevents results from being conclusive, and discrepancy from smaller effect sizes estimated in RCTs [16, 17] suggests that it may be an overestimate. It is possible that violations of the assumptions required for a valid instrument were partially responsible for the magnitude of the point estimate. However, results from the GMM IVA model support the conclusion that beta-blocker therapy is associated with decreased mortality and suggest that, if the odds ratios estimated through ordinary logistic regression were in fact affected by unmeasured confounding, the resulting estimate was likely biased towards the null hypothesis of no association.

5.2. Simulation example

While the applied example above illustrates a potential real-world application of GMM IVA in epidemiology, it is not ideal for exploring the fully empirical properties of the method, as it addresses an open question for which the true association remains unknown. We therefore examined the properties further using a simulated example for which the true association was known *a priori*.

Methods: Data were simulated for a Poisson-distributed response variable, which could potentially represent a count of clinical events. Different levels of correlation between the treatment and the unmeasured confounder and the instrument, respectively, were considered. For each combination of pre-specified correlations, the following steps were carried out:

1. Generate instrument (Z) and confounder (C) variables from independent standard normal distributions.
2. Generate variable $T^* = aZ + bC + \varepsilon$, where a and b are chosen to satisfy the desired correlations between T^* and Z and C, respectively, and $\varepsilon \sim \mathrm{N}(0, 1)$. Define the treatment variable T so that
$$T = \begin{cases} 1 & \text{if } T^* > 0 \\ 0 & \text{otherwise} \end{cases}$$
3. Generate the outcome variable y from a Poisson distribution, with the expected value depending on T, C, and covariates $\mathbf{X}$. The relative risk associated with T was 3.0; the relative risk associated with C was 0.5. Normally distributed covariates $X_1 \sim \mathrm{N}(-2, 2)$ and $X_2 \sim \mathrm{N}(1, 3)$ contributed relative risks of 0.75 and 1.5, respectively.
4. Estimate the relative risk associated with T both by fitting a GLM with a Poisson link, not accounting for C, and by using GMM IVA, using Z as an instrument for T, and X_1 and X_2 as their own respective instruments.

The correlations considered were $\mathrm{cor}(T^*, Z) \in \{0.5, 0.7\}$ and $\mathrm{cor}(T^*, C) \in \{0.5, 0.7\}$. For each combination, the steps outlined above were carried out for 1000 iterations; for each iteration, estimates of $\boldsymbol{\beta}$ using each method were recorded, as well as the empirical correlations $\hat{\mathrm{cor}}(T, Z)$ and $\hat{\mathrm{cor}}(T, C)$.

Results: For both Z and C, correlations with T^* of 0.5 corresponded to an average Pearson correlation coefficient of 0.40 with T, while correlations of 0.7 corresponded to an average Pearson

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/sim

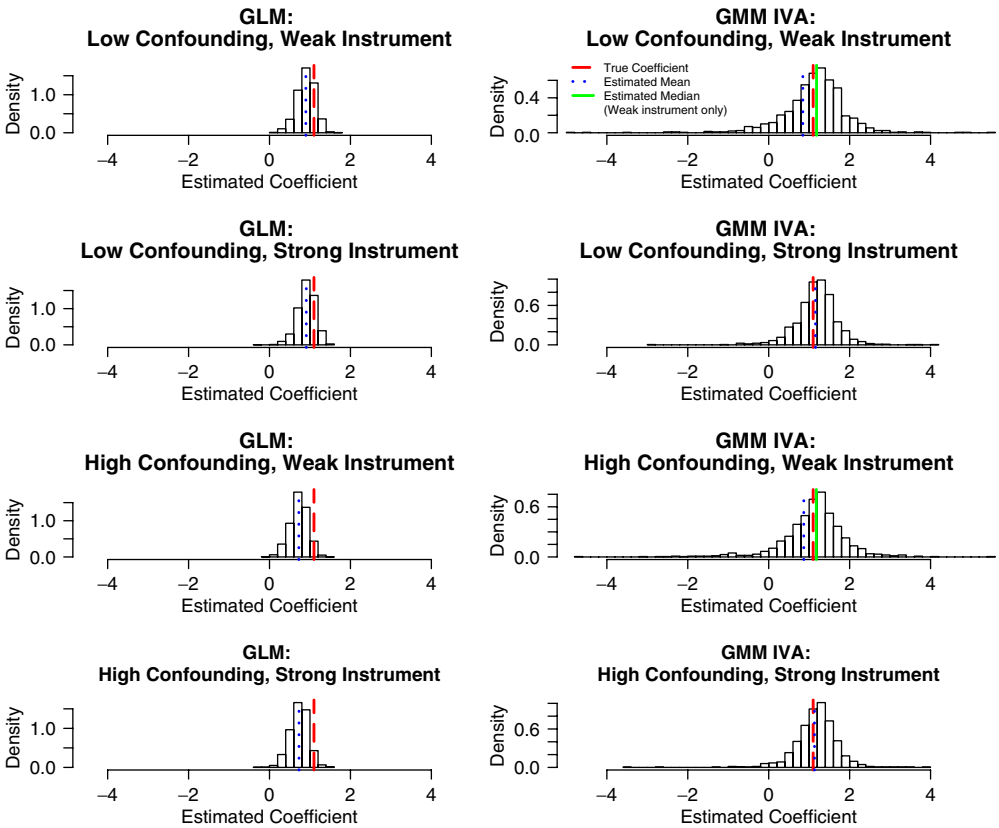

Figure 2. Empirical distributions of estimated treatment coefficients using GMM and GLM IVA estimation for different levels of confounding and instrument strength. Note that 'extreme' outliers from GLM IVA estimation have been omitted from GMM IVA plots, but not from estimation of mean or median.

correlation coefficient of 0.56. Both methods consistently estimated the coefficients associated with the unconfounded covariates X_1 and X_2.

As would be expected, instrument strength showed no effect on results from the GLM model; in addition, GMM IVA performed similarly for both levels of confounding considered here. However, the bias of the GLM estimate increased with increased confounding, and, while the GMM IVA model was similarly unbiased for both instrument strengths, it displayed greater variability than the GLM model, particularly when using a weaker instrument. There were also a small number of cases for which GMM IVA produced extremely skewed estimates of treatment effect, particularly when using a weak instrument. With an actual coefficient of $\log(3) = 1.10$, estimated coefficients below -5 were considered to be extreme outliers—these cases occurred 1.8 per cent of the time with a weak instrument and low confounding, and 2.4 per cent of the time with a weak instrument and high confounding. When using a strong instrument, there was a single extreme outlier when considering low confounding (0.1 per cent), and no extreme outliers when considering high confounding. Figure 2 displays the distribution of estimated treatment coefficients using GLM and GMM IVA estimation, for all combinations of confounding and instrument strength considered

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/sim

here. The dashed lines represent the true coefficient value of log(3). The extreme outliers were removed from the GMM IVA plots, so that the bulk of the distributions could be viewed on a more appropriate scale. However, the means of the respective sampling distributions, represented in Figure 2 by dotted lines, were calculated with the extreme outliers included. As a result of these outliers, mean estimates resulting from GMM IVA with a weaker instrument were negatively skewed. However, the corresponding medians, represented in Figure 2 by solid lines, remained unbiased estimators of the true coefficient. The medians are shown only for the results of GMM IVA with a weak instrument; for all other cases the mean and median were similar.

6. DISCUSSION

Here we described an application of GMM estimation of IVA to the class of GLMs, examined the theoretical properties of GMM IVA used in this setting, and illustrated the method through epidemiological and simulation examples. IVA has long been used in econometrics, and, more recently, to account for measurement error in generalized estimating equations [26]. IVA can be used to obtain consistent parameter estimates in the presence of unmeasured confounding, a problem that is commonly encountered in epidemiological research with no widely accepted solution; however, to date, IVA has not been widely used in this setting. A possible reason for this is the dearth of published literature describing the application of IVA to non-linear models commonly applied within epidemiology. GMM IVA presents a theoretically sound method for performing IVA on a wide class of non-linear models. We chose to focus on the class of GLMs here because the majority of statistical models used within epidemiology are members of this class.

There are a number of limitations to all forms of IVA, including GMM IVA. Foremost is the difficulty in choosing an appropriate instrument. For the results of IVA to be valid, the assumptions listed in Section 2 must hold. If IVA is being used to control for suspected unmeasured confounding, it suggests that the available data may be missing important variables, making it likely that a limited number of variables are available as candidate instruments. Given that the assumptions cannot be tested empirically, a logical argument must be presented justifying the choice of instrument. If such an instrument cannot be found, IVA is not possible. Further, recent results have suggested that IVA assumptions are more likely to be violated in the presence of strong unmeasured confounders and that IVA is most effective when unmeasured confounding is expected to be at most moderate [27].

In our example of the effect of beta-blockers on heart failure mortality, the IVA assumptions may not have been fully justified. While we presented heuristic arguments to support their plausibility, the large effect size reported here is valid only to the degree that the assumptions were met. However, in observational settings where unmeasured confounding is suspected, given careful consideration of the potential flaws of an instrument and biases that may result, analysis using an imperfect instrument can still help in providing a more complete picture than regression alone. Future approaches to this example might consider the use of additional instruments so that econometric methods developed for the overidentified case could be used to assess instrument validity. A potential instrument, used in other examples of health interventions, e.g. [28], is a place of residence relative to the admitting hospital. Unfortunately, this information was not included in the data source utilized here.

Assuming that a valid instrument can be found, a second limitation to IVA is that of decreased efficiency. In our example, the association between beta-blocker use and mortality was estimated to be relatively weak but statistically significant when using standard logistic regression. A stronger

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/sim

association was identified using the GMM IVA extension of logistic regression, but the associated standard error was also larger and the association was no longer statistically significant. Although this may have been resolved by using a larger sample size, given the large number of individuals included here ($n = 21\,217$), problems of efficiency are likely to remain in a majority of epidemiological studies undertaken in practice. While it is true that the power associated with such a large sample was reduced by the fact that only 10.5 per cent of individuals were prescribed a beta-blocker, this is not a factor that can be controlled by investigators when performing an observational study. The efficiency is also dependent on the association between treatment and instrument; hence, there is less cause for concern when working with a strong instrument. In our example, all variables other than beta-blocker use served as their own respective instruments, implying maximum possible instrument strength, and the standard errors associated with these parameters were nearly identical when using GMM IVA compared with standard logistic regression (Table III).

A third limitation of GMM IVA is the potential for 'extreme' solutions, particularly when using a weak instrument—in the simulation example here, such solutions occurred about 2 per cent of the time when using a weak instrument, but only once out of 2000 iterations when using a strong instrument. In our example, however, these outliers always fell in the opposite direction of the true effect, i.e. estimated relative risks close to 0 when the actual relative risk was 3. If such highly implausible results were obtained in practice, it is likely that alternative estimation techniques would be employed for comparison purposes and that the instability of the GMM IVA solution would be identified.

A limitation of the GMM specific to the analysis of non-linear models is the potential for inconsistent estimators due to the nature of unmeasured confounding, particularly in the plausible scenario of a symmetric model in which unmeasured confounders are related to outcome in a manner equivalent to that of measured confounders [14]. However, we derive in Appendix A that GMM IVA is approximately consistent for symmetric models. Further, in contrast to the results of Terza [14], our simulation example specified a symmetric model, and we did not detect a major bias in GMM IVA estimators, particularly for a strong instrument. A potential area for future research is the identification of specific conditions under which GMM IVA is likely to provide acceptable estimates of parameters from a non-linear symmetric model.

7. CONCLUSION

Unmeasured confounding is a ubiquitous concern in observational studies of medical therapy. Instrumental variables have long been recognized as a method of accounting for unmeasured confounding. GMM IVA, which extends 2SLS to the analysis of GLMs, provides unbiased estimation of treatment effects in the presence of unmeasured confounding and can therefore be considered when evaluating epidemiological data. By using GMM IVA, the analyst trades off greater accuracy with reduced precision. The usefulness of GMM IVA increases with increasing confidence in the strength and validity of the instrument.

APPENDIX A

In the GLM scenario, population-level versions of the estimating equations being solved take the form

$$E[\mathbf{X}(Y - h(\mathbf{X}'\boldsymbol{\beta}))] = 0 \tag{A1}$$

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/sim

when using direct regression without regard for unobserved confounding, and

$$E[\mathbf{Z}(Y - h(\mathbf{X}'\boldsymbol{\beta}))] = 0 \tag{A2}$$

when using IVA in the just-identified case. To shed light on how (A2) deals with unobserved confounding, assume that the true distribution of outcome Y given $\mathbf{X}$ and a scalar unobserved confounder U follows

$$E[Y|\mathbf{X}, U] = h(\mathbf{X}'\boldsymbol{\beta}^*) + U \tag{A3}$$

with $E[U] = 0$ (such centering is reasonable if an intercept term is included in the model). Here we might interpret $\boldsymbol{\beta}^*$ as the true parameter vector of interest as it describes the treatment–outcome relationship for a 'typical' subject having the average value of $U = 0$ for the unobserved confounder.

It is easy to check that if (A3) holds then the solution to (A1) will have $\boldsymbol{\beta} \neq \boldsymbol{\beta}^*$, if U and $\mathbf{X}$ are correlated. In this sense, a bias arises from unobserved confounding. On the other hand, under the IV assumptions (A2) can be written as

$$\begin{aligned}
0 &= E\{E[\mathbf{Z}(Y - h(\mathbf{X}'\boldsymbol{\beta}))|\mathbf{X}, U]\} \\
&= E\{E[\mathbf{Z}|\mathbf{X}, U]E[(Y - h(\mathbf{X}'\boldsymbol{\beta}))|\mathbf{X}, U]\} \\
&= E\{E[\mathbf{Z}|\mathbf{X}, U](h(\mathbf{X}'\boldsymbol{\beta}^*) - h(\mathbf{X}'\boldsymbol{\beta}) + U)\} \\
&= E\{E[\mathbf{Z}(h(\mathbf{X}'\boldsymbol{\beta}^*) - h(\mathbf{X}'\boldsymbol{\beta}) + U)|\mathbf{X}, U]\} \\
&= E[\mathbf{Z}(h(\mathbf{X}'\boldsymbol{\beta}^*) - h(\mathbf{X}'\boldsymbol{\beta}) + U)]
\end{aligned} \tag{A4}$$

which is indeed solved by $\boldsymbol{\beta} = \boldsymbol{\beta}^*$, since $\mathbf{Z}$ and U are assumed independent. Thus, GMM IVA deals with the unobserved confounding of the form (A3).

However, as stressed in [14], models of the form (A3) may not seem the most natural way in which unobserved confounding might be present. If (A3) is replaced by the perhaps more natural form

$$E[Y|\mathbf{X}, U] = h(\mathbf{X}'\boldsymbol{\beta}^* + U) \tag{A5}$$

then (A4) does not yield $\boldsymbol{\beta} = \boldsymbol{\beta}^*$ exactly. This is still the case if (A5) is replaced by the Taylor approximation:

$$E[Y|\mathbf{X}, U] = h(\mathbf{X}'\boldsymbol{\beta}^*) + h'(\mathbf{X}'\boldsymbol{\beta}^*)U \tag{A6}$$

However, (A6) can be further approximated by

$$E[Y|\mathbf{X}, U] = h(\mathbf{X}'\boldsymbol{\beta}^*) + h'(E[\mathbf{X}'\boldsymbol{\beta}^*])U \tag{A7}$$

which is of the same form as (A4), suggesting that if the actual underlying model is of the form of (A5), the estimates for $\boldsymbol{\beta}$ produced using GMM IVA are at least approximately consistent for parameters governing the treatment–outcome relationship given the unobserved confounder.

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/sim

APPENDIX B

Recall (systems (3) and (7)), that, in the case of no unmeasured confounding, GMM estimation of model parameters is based on

$$\sum_{i=1} x_{ij}[y_i - h(\mathbf{x_i}; \boldsymbol{\beta})] = 0$$

while MLE estimation is based on

$$\sum_{i=1}^{n} x_{ij}[y_i - h(\mathbf{x_i'}\boldsymbol{\beta})](\mathrm{Var}(y_i))^{-1}\left(\frac{\partial \mu_i}{\partial \eta_i}\right) = 0$$

Therefore, the two sets of estimates are equivalent if, for some constant c,

$$(\mathrm{Var}(y_i))^{-1}\left(\frac{\partial \mu_i}{\partial \eta_i}\right) = c \tag{B1}$$

Recall that for all GLMs, the outcome variable y belongs to the exponential family, implying that its density function can be written in the form of (6). It can be shown that [15]

$$E(y_i) = \mu_i = b'(\theta_i)$$

and

$$\mathrm{Var}(y_i) = b''(\theta_i)a(\phi_i)$$

In the class of models using a canonical link, $\eta_i = \theta_i$, this can be rewritten as

$$E(y_i) = \mu_i = b'(\eta_i) \tag{B2}$$

and

$$\mathrm{Var}(y_i) = b''(\eta_i)a(\phi_i) \tag{B3}$$

As a result of (B2) and (B3), $\partial \mu_i / \partial \eta_i = b''(\eta_i)$, and $(\mathrm{Var}(y_i))^{-1}(\partial \mu_i / \partial \eta_i) = a(\phi_i)$. Therefore, (B1) holds only if $a(\phi_i) \equiv c$ for some constant c. This can be shown to hold for all common GLM distributions using a canonical link, including normal, Poisson, binomial, gamma, exponential, and inverse Gaussian distributions. In general, it holds for any distribution that is either characterized by a single parameter or for which GLM estimation proceeds by estimating one parameter and treating all others as nuisance parameters.

Thus, the only commonly encountered GLMs for which GMM and MLE estimation differ are those which use non-canonical links. In these cases, let $(\mathrm{Var}(y_i))^{-1}(\partial \mu_i / \partial \eta_i)$ be referred to as the *inflation factor*, since it is a measure of the inflation of system (7) relative to system (3). Table BI reports inflation factors for three of the most widely used GLMs with non-canonical links, and Figure B1 displays them graphically. Where the inflation factor is near 1.0, the two systems are numerically similar. For binomial data with a complementary log–log link and negative binomial data with a log link, GMM and MLE systems appear to be similar for negative values of η_i, but less so for positive values of η_i. For binomial data with a probit link, the two systems are dissimilar when η_i departs from 0 in either direction.

However, note that inflation factors refer to the systems of equations that are to be solved to obtain parameter estimates, not the estimates themselves. To compare differences between

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/sim

Table BI. Inflation factors for some common GLMs using non-canonical links.

Model	Inverse link function		Inflation factor	
	$\mathrm{Var}(y_i)$	$(\mu_i = h(\eta_i))$	$\frac{\partial \mu_i}{\partial \eta_i}$	$(\frac{1}{\mathrm{Var}(y_i)} \frac{\partial \mu_i}{\partial \eta_i})$
Binomial (probit link)	$\mu_i(1-\mu_i)$	$\mu_i = \Phi^{-1}(\eta_i)$	$\phi^{-1}(\eta_i)$	$\frac{\phi(\eta_i)}{\Phi(\eta_i)(1-\Phi(\eta_i)}$
Binomial (complementary log–log link)	$\mu_i(1-\mu_i)$	$\mu_i = 1-\exp(-e^{\eta_i})$	$\exp(\eta_i - e^{\eta_i})$	$\frac{e^{\eta_i}}{1-\exp(-e^{\eta_i})}$
Negative binomial (log link)	$\frac{\mu_i(\mu_i+\theta)}{\theta}$	$\mu_i = e^{\eta_i}$	e^{η_i}	$\frac{\theta}{e^{\eta_i}+\theta}$

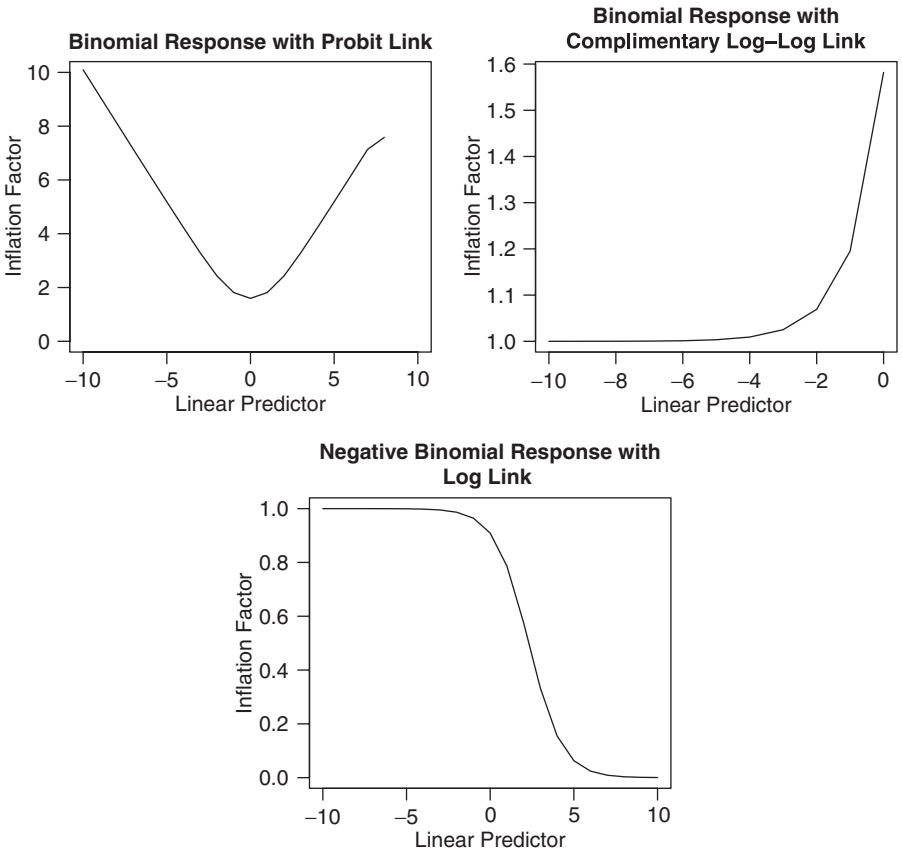

Figure B1. Inflation factors for three common generalized linear models using non-canonical links.

parameter estimates obtained from the two methods, we performed a simulation experiment. For each of the three non-canonical links considered here, we generated data from a 3-parameter model, with 1000 different sets of true parameter values. For each of the 1000 simulations, parameters were estimated using the GMM and MLE methods, respectively, and the two resulting sets of

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/sim

parameters were compared, characterized by the maximum absolute per cent difference, i.e.

$$\max_{i=1,2,3}\left[100\left|\frac{\hat{\beta}_i^{(\mathrm{MLE})}-\hat{\beta}_i^{(\mathrm{GMM})}}{\hat{\beta}_i^{(\mathrm{GMM})}}\right|\right] \tag{B4}$$

Parameters were generated from a uniform distribution on the interval $[-1,-0.5]\cup[0.5,1]$, chosen so that (B4) would not reflect instability due to parameter estimates being near zero. The percentage of the 1000 simulations in which the maximum absolute per cent difference between estimated parameters was no more than 10 per cent was 99.7, 86.0, and 62.6 per cent for the probit, complementary log–log, and negative binomial distributions, respectively. Thus, while the two methods of estimation are not identical for any of these distributions, there is empirical evidence that they produce numerically similar results in the case of the probit link, and possibly the complementary log–log link. In the case of negative binomial data, in which the log link is the most commonly used, this simulation suggests that GMM and MLE may not produce similar results; hence, it is not recommended that GMM IVA results in this case be treated as comparable with results from standard GLM analysis.

REFERENCES

1. Beck CA, Penrod J, Gyorkos TW, Shapiro S, Pilote L. Does aggressive care following acute myocardial infarction reduce mortality? Analysis with instrumental variables to compare effectiveness in Canadian and United States patient populations. *Health Services Research* 2003; **38**:1423–1440.
2. Brooks JM, Chrischilles EA, Scott SD, Chen-Hardee SS. Was breast conserving surgery underutilized for early stage breast cancer? Instrumental variables evidence for stage II patients from Iowa. *Health Services Research* 2003; **38**:1385–1402.
3. Linden A, Adams JL. Evaluating disease management programme effectiveness: an introduction to instrumental variables. *Journal of Evaluation in Clinical Practice* 2006; **12**(2):148–154.
4. Greenland S. An introduction to instrumental variables for epidemiologists. *International Journal of Epidemiology* 2000; **29**:722–729.
5. Angrist JD. Estimation of limited dependent variable models with dummy endogenous regressors: simple strategies for empirical practice. *Journal of Business and Economic Statistics* 2001; **19**:2–16.
6. Foster EM. Instrumental variables for logistic regression: an illustration. *Social Science Research* 1997; **26**: 487–504.
7. Hansen LP. Large sample properties of generalized method of moments estimators. *Econometrica* 1982; **50**: 1029–1054.
8. Imbens GW. Generalized method of moments and empirical likelihood. *Journal of Business and Economic Statistics* 2002; **20**:493–506.
9. Greene WH. *Econometric Analysis* (5th edn). Prentice-Hall: Englewood Cliffs, NJ, 2003.
10. Hogan JW, Lancaster T. Instrumental variables and inverse probability weighting for causal inference from longitudinal observational studies. *Statistical Methods in Medical Research* 2004; **13**:17–48.
11. Andrews DWK. Consistent moment selection procedures for generalized method of moments estimation. *Econometrica* 1999; **67**(3):543–664.
12. White H. A heteroskedasticity consistent covariance matrix and a direct test for heteroskedasticity. *Econometrica* 1980; **48**:817–838.
13. Mullahy J. Instrumental-variable estimation of count data models: applications to models of cigarette smoking behavior. *The Review of Economics and Statistics* 1997; **79**:586–593.
14. Terza JV. Estimation of policy effects using parametric nonlinear models: a contextual critique of the generalized method of moments. *Health Services and Outcomes Research Methodology* 2006; **6**:177–198.
15. McCullagh P, Nelder JA. *Generalized Linear Models* (2nd edn). Chapman & Hall: London, 1989.
16. Fisher M, Gottlieb S, Plotnick G. Beneficial effects of metoprolol in heart failure associated with coronary artery disease: a randomized trial. *Journal of the American College of Cardiology* 1994; **23**(4):943–950.

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/sim

17. Bristow M, O'Connell J, Gilbert E, French W. Dose–response of chronic beta-blocker treatment in heart failure from either idiopathic dilated or ischemic cardiomyopathy. Bucindolol investigators. *Circulation* 1994; **89**: 1632–1642.
18. Williams JF, Bristow MR, Fowler MB, Francis GS, Garson A, Gersh BJ, Hammer DF, Hlatky MA, Leier CV, Packer M, Pitt B, Ullyot DJ, Wexler LF, Winters WL, Ritchie JL, Cheitlin MD, Eagle KA, Gardner TJ, Gibbons RJ, Lewis RP, Orourke RA, Ryan TJ. Guidelines for the evaluation and management of heart failure. *Journal of the American College of Cardiology* 1995; **26**(5):1376–1398.
19. Hunt SA, Baker DW, Chin MH, Cinquegrani MP, Feldman AM, Francis GS *et al.* ACC/AHA guidelines for the evaluation and management of chronic heart failure in the adult: executive summary. *Journal of the American College of Cardiology* 2001; **38**(7):2101–2113.
20. Levy A, Johnston K, Briggs A. Pharmacoepidemiological methods for assessing safety and effectiveness, In *Safe and Effective*, NM (ed.). Canadian Pharmacists Association, Ottawa, Ontario, Canada, 2007.
21. Heiat A, Gross CP, Krumholz HM. Representation of the elderly, women, and minorities in heart failure clinical trials. *Archives of Internal Medicine* 2002; **162**(15):1682–1688.
22. D'Agostino Jr RB, D'Agostino Sr RB. Estimating treatment effects using observational data. *Journal of the American Medical Association* 2007; **297**(3):314–316.
23. Wright JH. Detecting lack of identification in GMM. *Econometric Theory* 2003; **19**:322–330.
24. Sheps SB, Reid RJ, Barer ML, Krueger H, McGrail KM, Green B *et al.* Hospital downsizing and trends in health care use among elderly people in British Columbia. *Canadian Medical Association Journal* 2000; **163**(4):397–401.
25. Polanczyk CA, Rohde LE, Philbin EA, Salvo TGD. A new casemix adjustment index for hospital mortality among patients with congestive heart failure. *Medical Care* 1998; **36**:1489–1499.
26. Hardin JW, Hilbe JM. *Generalized Estimating Equations*. Chapman & Hall: London, 2003.
27. Martens EP, Pestman WR, de Boer A, Belitser SV, Klungel OH. Instrumental variables: application and limitations. *Epidemiology* 2006; **17**(3):260–267.
28. Newhouse JP, McClellan M. Econometrics in outcomes research: the use of instrumental variables. *Annual Review of Public Health* 1998; **19**:17–34.

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/sim