



IFAU – INSTITUTE FOR
LABOUR MARKET POLICY
EVALUATION

Sensitivity analysis of the unconfoundedness assumption in observational studies

Xavier de Luna
Mathias Lundin

WORKING PAPER 2009:12

The Institute for Labour Market Policy Evaluation (IFAU) is a research institute under the Swedish Ministry of Employment, situated in Uppsala. IFAU's objective is to promote, support and carry out scientific evaluations. The assignment includes: the effects of labour market policies, studies of the functioning of the labour market, the labour market effects of educational policies and the labour market effects of social insurance policies. IFAU shall also disseminate its results so that they become accessible to different interested parties in Sweden and abroad.

IFAU also provides funding for research projects within its areas of interest. The deadline for applications is October 1 each year. Since the researchers at IFAU are mainly economists, researchers from other disciplines are encouraged to apply for funding.

IFAU is run by a Director-General. The institute has a scientific council, consisting of a chairman, the Director-General and five other members. Among other things, the scientific council proposes a decision for the allocation of research grants. A reference group including representatives for employer organizations and trade unions, as well as the ministries and authorities concerned is also connected to the institute.

Postal address: P.O. Box 513, 751 20 Uppsala

Visiting address: Kyrkogårdsgatan 6, Uppsala

Phone: +46 18 471 70 70

Fax: +46 18 471 70 71

ifau@ifau.uu.se

www.ifau.se

Papers published in the Working Paper Series should, according to the IFAU policy, have been discussed at seminars held at IFAU and at least one other academic forum, and have been read by one external and one internal referee. They need not, however, have undergone the standard scrutiny for publication in a scientific journal. The purpose of the Working Paper Series is to provide a factual basis for public policy and the public policy discussion.

ISSN 1651-1166

Sensitivity analysis of the unconfoundedness assumption in observational studies*

XAVIER DE LUNA AND MATHIAS LUNDIN

*Department of Statistics
Umeå University, S-90187 Umeå, Sweden*

May 15, 2009

Abstract

In observational studies, the estimation of a treatment effect on an outcome of interest is often done by controlling on a set of pre-treatment characteristics (covariates). This yields an unbiased estimator of the treatment effect when the assumption of unconfoundedness holds, that is, there are no unobserved covariates affecting both the treatment assignment and the outcome. This is in general not realistically testable. It is, therefore, important to conduct an analysis about how sensitive the inference is with respect to the unconfoundedness assumption. In this paper we propose a procedure to conduct such a Bayesian sensitivity analysis, where the usual parameter uncertainty and the uncertainty due to the unconfoundedness assumption can be compared. To measure departures from the assumption we use a correlation coefficient which is intuitively comprehensible and ensures that the results of sensitivity analyses made on different evaluation studies are comparable. Our procedure is applied to the Lalonde data and to a study of the effect of college choice on income in Sweden.

Keywords: Causal inference, Effects of college choice, Propensity score, Register data.

JEL-codes: C11, C15

*We are grateful to Per Johansson, Maria Karlsson, Sune Karlsson and Thomas Laitila for helpful comments, and to Kent Eliasson and Olle Westerlund for our discussions on the effects of college choice on the labour market. The Swedish Institute for Growth Policy Studies is acknowledged for providing access to the register data used in Section 4 of this paper. The research presented in this article has been funded by the Swedish Council for Working Life and Social Research and the Institute of Labour Market Policy Evaluation, Uppsala.

1 Introduction

This paper proposes a standardised procedure to conduct sensitivity analysis of the assumption of unconfoundedness made in observational studies. We particularly focus on evaluation studies based on the Rubin model for causal inference, see Rubin (1974) and Holland (1986). In such studies the aim is to estimate the effect of a cause (e.g., a medical treatment or a labour market program) on an outcome of interest. Based on a model (a set of assumptions) a point estimate of the effect (a parameter of the model) is produced with a relevant estimation method. In association to such a point estimate, the following sequence of questions should be addressed:

- i) Scientific significance: What is an interesting effect? Is the estimated effect scientifically relevant?
- ii) Parameter uncertainty: What is the posterior/sampling variation? E.g., is the effect statistically significant?
- iii) Model uncertainty: Is the inference made in ii) sensitive to small departures from the model assumptions?

Point ii) is of interest only if the answer to point i) is affirmative, and similarly point iii) is of interest only if the answers to point i) and ii) are affirmative. While points i) and ii) are often addressed in evaluation studies, point iii) is seldom so. When the estimation is performed with small data sets, parameter uncertainty will typically dominate model uncertainty and therefore ignoring point iii) is often harmless. However, with large data sets, model uncertainty may be comparable or even dominate parameter uncertainty. This is an essential issue because effects tend to be statistically significant if you collect enough data and ignore model uncertainty. Pawitan (2001, Sec. 1.3), Greenland (2005), Copas and Eguchi (2005) have discussed the importance of model uncertainty.

In this paper, we focus on a major source of model uncertainty in studies evaluating the effect of a treatment with observational data. Namely, the uncertainty due to the assumption of unconfoundedness of the treatment assignment. In order to formalise this assumption, the Rubin model for causal inference is often used, where z is an indicator for treatment assignment ($z_i = 1$ when individual i is treated and $z_i = 0$ otherwise). Further, assume that the interest lies in studying the effect of the treatment on an outcome denoted $y(z)$, with $y_i(1)$ the outcome of individual i if she/he is treated and $y_i(0)$ if she/he is not. Then, in order to estimate the effect of the treatment, say $\tau = E(y(1) - y(0))$, it is most often assumed (either implicitly with parametric models or explicitly with non-parametric estimators such as matching) that $y(1)$ and $y(0)$ are independent of z when conditioning on a given set of covariates \mathbf{x} (characteristic of the individuals observed before treatment assignment). This is called in the sequel the unconfoundedness assumption (UA) because it holds if there is no unobserved confounding covariate u that is dependent of both the treatment assignment and the outcomes. Note that because the potential outcomes $y(z)$, $z = 0, 1$ cannot both

be observed for a given individual, there is in general too little information in the data to perform a statistical test of the unconfoundedness assumption. Because this assumption is essential in an observational study, a sensitivity analysis should be carried out where the consequences of the uncertainty attached to the assumption are investigated; see point iii) above.

Different strategies to address point iii) have been developed, see, e.g., Copas and Eguchi (2001, 2005), Rosenbaum (2002), Imbens (2003), and Greenland (2005). In this paper, we use a Bayesian approach (as advocated by Greenland, 2005) which has the advantage of making the two sources of variation ii) and iii)) conceptually easy to compare. We develop a sensitivity analysis based on the Rubin model for causal inference, by measuring the departure from the unconfoundedness assumption with a scalar measure of dependence –the correlation between the potential outcome $y(0)$ and the treatment assignment z given the covariates \mathbf{x} . The use of a standardised measure of dependence allows us to propose a standardised sensitivity analysis. Thus, we may compare the results of such analyses made for different evaluation studies –based on, different data, models, effects of interest, etc.– and say whether a given study is more or less sensitive than another. A correlation coefficient is, moreover, intuitively comprehensible by most empirical scientists, making the results of the analysis easy to communicate. Copas and Eguchi (2005) have also advocated the use of a scalar measure of dependence to perform a sensitivity analysis. However, while we compare parameter and model uncertainty by computing posterior distributions of the parameter of interest, Copas and Eguchi quantify analytically the bias due to a certain departure from the UA.

We have organised the paper as follows. In Section 2 we review how sensitivity analysis based on expanded models which parametrise the departure from the unconfoundedness assumption can be carried out within a Bayesian framework. We develop in this section our main contribution which consists in using a simple and standardised measure of dependence between the potential outcome and the treatment assignment to study sensitivity of the evaluation to the unconfoundedness assumption. The proposed procedure is then explicitly illustrated in Section 3 using linear regression to model the outcome and logistic regression to model the treatment assignment. We use for that purpose a toy application based on the Lalonde data (Lalonde, 1986). In Section 4, we perform a sensitivity analysis of an evaluation study presented in Lundin (2006), where the effect of college choice on income was evaluated with Swedish register data. Section 5 concludes the paper.

2 Sensitivity analysis: Models and methods

2.1 Modelling bias

In order to study departures from the UA it has been proposed in the literature to expand a null model (i.e. a model where the UA holds) by introducing parameters which model the departure from the assumption (Copas and Eguchi, 2001, 2005, Rosenbaum, 2002, Imbens,

2003, Greenland, 2005).

Let us specify a parametric null model in very general terms by denoting

$$p(y(0)|\mathbf{x}; \theta) \text{ and } p(y(1)|\mathbf{x}; \theta) \quad (1)$$

the density¹ functions of the potential outcomes $y(0)$ and $y(1)$ given a set of covariates \mathbf{x} , and a set of parameters θ . The parameter of interest, the program effect, is then $\tau(\theta)$ a function of θ . A model for program assignment is also needed to study the UA, i.e.

$$\Pr(z = 1|\mathbf{x}; \gamma) \quad (2)$$

must be specified.

The unconfoundedness assumption would not hold if there was an unobserved variable u (not included in \mathbf{x}) affecting both $y(z)$ and z . A sensitivity analysis can then be performed by considering the existence of such an unobserved confounding variable. This can be done with the following expanded model. A conditional model for the outcome is

$$p(y(0)|\mathbf{x}, u; \theta, \eta_1) \text{ and } p(y(1)|\mathbf{x}, u; \theta, \eta_1), \quad (3)$$

where the parameter η_1 describes how u affects $y(0)$ and $y(1)$, with a parametrisation assumed such that $\eta_1 = 0$ yields (1), and u has a known density function. Moreover, an assignment mechanism is specified by

$$\Pr(z = 1|\mathbf{x}, u; \gamma, \eta_2), \quad (4)$$

where the parameter η_2 describes how u affects z , and again the parametrisation is such that $\eta_2 = 0$ corresponds to (2).

We will use the following example throughout the paper to illustrate the different issues tackled.

Example 1 *[Normal/logistic model]* A simple example for a treatment evaluation model is obtained by considering a normal model for the potential outcomes (here $\theta = (\tau, \beta, \sigma)$)

$$y(0)|\mathbf{x}; \theta \sim N(\beta'\mathbf{x}, \sigma^2) \quad (5)$$

and

$$y(1)|\mathbf{x}; \theta \sim N(\beta'\mathbf{x} + \tau, \sigma^2)$$

Here a constant treatment effect τ is assumed. Further, a logistic regression model for the treatment assignment can be written as

$$\Pr(z = 1|\mathbf{x}; \gamma) = \exp(\gamma'\mathbf{x}) / (1 + \exp(\gamma'\mathbf{x})).$$

¹While using the terminology “density function”, the approach is not restricted to continuous valued outcomes

This normal-logistic null model can be expanded as follows²

$$y(0)|\mathbf{x}, u; \theta, \eta_1 \sim N(\beta'\mathbf{x} + \eta_1 u, \nu^2), \quad (6)$$

where u is Bernoulli with probability $1/2$ and $\nu = \nu(\theta, \eta_1)$, and

$$\Pr(z = 1|\mathbf{x}, u; \gamma, \eta_2) = \exp(\gamma'\mathbf{x} + \eta_2 u)/(1 + \exp(\gamma'\mathbf{x} + \eta_2 u)). \quad (7)$$

Greenland (2005) calls $\eta = (\eta_1, \eta_2)'$ the bias parameters. For $\eta_1\eta_2 = 0$ the UA holds, while otherwise a bias in the estimation of τ will be implied by not conditioning on u . Note that typically there is very little information in the data on $\eta_1\eta_2$ because u is not observed, thereby making the hypothesis $H_0 : \eta_1\eta_2 = 0$ difficult to test. Moreover, the result of such a test is bound to be very sensitive to the model assumptions made as noted by Imbens (2003). Therefore, an alternative often advocated is to perform a sensitivity analysis.

2.2 Bayesian sensitivity analysis

We describe here a Bayesian approach to sensitivity analysis, advocated recently by Greenland (2005). This approach allows us to quantify model uncertainty through a distribution function for the bias parameters η . The Bayesian sensitivity analysis yields a posterior distribution for the treatment effect that combines this model uncertainty with parameter uncertainty. For this purpose, we use probability distribution functions to describe our beliefs on parameters before and after the data has been observed. We use data observed on n individuals to update our prior $p(\theta, \gamma, \eta)$ (beliefs on the parameters before the data is observed) to obtain a posterior distribution

$$p(\theta, \gamma, \eta|\mathbf{y}, \mathbf{z}, X) \propto p(\mathbf{y}, \mathbf{z}|X; \theta, \gamma, \eta)p(\theta, \gamma, \eta)$$

where \mathbf{y} is the vector of n observed potential outcomes, \mathbf{z} the vector of n observed treatment assignments, and X is the matrix with element $X(i, j)$ being the observed value of covariate j for individual i . Moreover, the likelihood $p(\mathbf{y}, \mathbf{z}|X; \theta, \gamma, \eta) = L(\theta, \gamma, \eta)$ is obtained by integrating out the missing potential outcomes and \mathbf{u} from $p(\mathbf{y}, \mathbf{z}, \mathbf{y}_{miss}, \mathbf{u}|X; \theta, \gamma, \eta)$, where \mathbf{y}_{miss} is a vector containing the n unobserved potential outcomes (for each individual we have one observed and one unobserved outcome), and \mathbf{u} consists in the n unobserved values for u . For the parameter of interest, $\tau = \tau(\theta)$, the posterior distribution is

$$p(\tau|\mathbf{y}, \mathbf{z}, X) \propto \int_{\tau(\theta)=\tau} L(\theta, \gamma, \eta)p(\theta, \gamma, \eta)d(\theta, \gamma, \eta). \quad (8)$$

²Here we let the unobserved confounder affect $y(0)$. Alternative choices would have been to let u affect only $y(1)$, or both $y(0)$ and $y(1)$. If some knowledge exists on how a potential confounder would affect the outcome this should be used in the specification of the expanded model.

Example 2 [Normal/logistic cont'd] For the normal/logistic expanded model, we have (Imbens, 2003)

$$\begin{aligned}
L(\theta, \gamma, \eta) &= \prod_{i=1}^n \left[\frac{1}{2} \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right) \exp \left(-\frac{1}{2\sigma^2} (y_i - \tau z_i - \beta \mathbf{x}_i)^2 \right) \right. \\
&\quad \times \frac{(\exp(\gamma \mathbf{x}_i))^{z_i}}{1 + \exp(\gamma \mathbf{x}_i)} \\
&\quad + \frac{1}{2} \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right) \exp \left(-\frac{1}{2\sigma^2} (y_i - \tau z_i - \beta \mathbf{x}_i - \eta_1)^2 \right) \\
&\quad \left. \times \frac{(\exp(\gamma \mathbf{x}_i + \eta_2))^{z_i}}{1 + \exp(\gamma \mathbf{x}_i + \eta_2)} \right]. \tag{9}
\end{aligned}$$

A sensitivity analysis can then be performed by comparing two posteriors $p(\tau|\mathbf{y}, \mathbf{z}, X)$. The first is obtained by using the null model in conjunction with a prior for θ and γ . This is equivalent to using the expanded model with a prior $p(\theta, \gamma, \eta)$ which puts weight one on $\eta = 0$. The second posterior is obtained with a prior distribution $p(\theta, \gamma, \eta)$ describing how η is allowed to differ from the value zero. In particular, if the two posteriors do not differ much, then we may conclude that inference based on the null model is justified, because it is not sensitive to small departure from the UA.

The posterior distribution (8) is generally not straightforward to obtain and Markov Chain Monte Carlo Methods must often be used to obtain an empirical estimate (e.g., Gelman et al., 1995). A simpler approximate Bayesian sensitivity analysis can be used instead, by using asymptotic arguments. For this purpose, write

$$p(\tau|\mathbf{y}, \mathbf{z}, X, \eta) \propto \int_{\tau(\theta)=\tau} L(\theta, \gamma, \eta) p(\theta, \gamma|\eta) d(\theta, \gamma).$$

Using the prior $p(\theta, \gamma|\eta) \approx 1$, we obtain

$$p(\tau|\mathbf{y}, \mathbf{z}, X, \eta) \propto \int_{\tau(\theta)=\tau} L(\theta, \gamma, \eta) d(\theta, \gamma).$$

An asymptotic approximation of the latter is (Gelman et al., 1995, Chap. 4) the density function of the normal distribution with mean $\hat{\tau}_\eta$, the value maximising the likelihood $L(\theta, \gamma, \eta)$ for η fixed, and variance v_η , the usual asymptotic variance of the maximum likelihood estimator.

Greenland (2005) suggests to sample from the posterior distribution $p(\tau|\mathbf{y}, \mathbf{z}, X) = \int p(\tau|\mathbf{y}, \mathbf{z}, X, \eta) p(\eta|\mathbf{y}, \mathbf{z}, X) d\eta \approx \int p(\tau|\mathbf{y}, \mathbf{z}, X, \eta) p(\eta) d\eta$ with Algorithm 1. The latter approximation is obtained by noting that the data has little information on η . Repeating Step 1 to Step 3 of Algorithm 1 many times provides a sample which is used to estimate the posterior density $p(\tau|\mathbf{y}, \mathbf{z}, X)$. This density estimate is then compared with the approximate posterior $N(\hat{\tau}_0, v_0)$, i.e. the asymptotic distribution of $\hat{\tau}_0$, the maximum likelihood estimate based on $L(\theta, \gamma, \eta = 0)$.

Algorithm 1 Sampling from $p(\tau|\mathbf{y}, \mathbf{z}, X)$

Step 1: Sample η^s from $p(\eta)$.

Step 2: Maximising the likelihood $L(\theta, \gamma, \eta = \eta^s)$, thereby yielding $\hat{\tau}_{\eta^s}$ and its estimated variance v_{η^s} .

Step 3: Generate τ^s from a $N(\hat{\tau}_{\eta^s}, v_{\eta^s})$.

2.3 A standardised measure of departure from UA

2.3.1 Sensitivity analysis at a fixed level

Our aim is to study the sensitivity of an evaluation study to a small departure from the UA. A key issue is, therefore, how to define a small departure from $\eta_1\eta_2 = 0$. We argue that a measure of small departure must have the following characteristics: (i) be intuitively comprehensible by empirical scientists, and (ii) be standardised, in the sense that the values the measure can take is as far as possible independent from the context studied. Both these characteristics are important to make the analysis interpretable –in particular to be able to decide what is a small departure–, but also for results of sensitivity analyses of different evaluation studies to be comparable. A main purpose is here to make the results of the sensitivity analysis easy to interpret and report to third parties.

To fulfil these requirements we focus on the real issue of interest, which is the dependence between $y(0), y(1)$ and z given \mathbf{x} . Indeed, the bias parameters η have been introduced with the purpose to induce a dependence between the potential outcomes and the assignment mechanism, but η is itself not standardised and difficult to interpret.

To study the dependence (conditional on \mathbf{x}) between $y(0)$ and z in the expanded model³ (3-4) we consider the correlation $\rho(\mathbf{x}) = Corr(y(0), z|\mathbf{x})$. A correlation coefficient is a standardised measure, and is intuitively understandable by most empirical scientist, thereby fulfilling the requirements stated above.

Note that for Example 1 $\rho(\mathbf{x}) = 0, \forall \mathbf{x}$ if and only if $\eta_1\eta_2 = 0$. More generally, the use of the correlation coefficient $\rho(\mathbf{x})$ to measure a small departure from the null model is justified even if $\rho(\mathbf{x}) = 0, \forall \mathbf{x}$ does not imply UA ($\eta_1\eta_2 = 0$). Indeed, $\rho(\mathbf{x}) = 0, \forall \mathbf{x}$ implies that the average treatment effect can be estimated without bias:

$$\begin{aligned}\rho(\mathbf{x}) = 0, \forall \mathbf{x} &\Rightarrow E(y(0)z|\mathbf{x}) = E(y(0)|\mathbf{x})E(z|\mathbf{x}) \\ &\Rightarrow E(y(0)|\mathbf{x}, z = 1) = E(y(0)|\mathbf{x}, z = 0) = E(y(0)|\mathbf{x}).\end{aligned}$$

The last set of equalities ensures that conditional on \mathbf{x} , $y(0)$ observed for the controls are representative of the $y(0)$ not observed for the treated. To estimate the average treatment effect by conditioning on \mathbf{x} we also need that $Corr(y(1), z|\mathbf{x}) = 0, \forall \mathbf{x}$, see Footnote 3.

³The dependence between $y(1)$ and z does not need to be considered because the existence of a correlation between $y(0)$ and z is sufficient for the UA to be violated; see Footnote 2 however. Note also that if the average effect on the treated is of interest then only the dependence between $y(0)$ and z is relevant.

Typically $\rho(\mathbf{x})$ will vary with \mathbf{x} and, therefore, we work with $\rho_m = \rho(\arg \max |\rho(\mathbf{x})|)$. That is for a given application the maximum is taken over the observed values for \mathbf{x} .

There are two major issues that must be solved before a sensitivity analysis can be carried out. The first one is problem specific: the correlation $\rho(\mathbf{x})$ must be computed for a given bias expanded model. A concrete example on the normal/logistic expanded model of Example 2 is treated in Section 3. The second issue is of general character: What is a small departure from the null model? Or, in other words: What is a small correlation? We believe the latter question must be answered in general terms, that is without making reference to particular situations: models, sample size, etc. Copas and Eguchi (2005, Sec. 7) discuss this issue and list three possible strategies: (a) UA is assumed, (b) $\rho(\mathbf{x}) \neq 0$ but are small enough that if they would be estimable, they would not have been statistically significant, (c) $\rho(\mathbf{x})$ may be large.

The point of departure of a sensitivity analysis must be that the UA is reasonable approximation, and, therefore, potential departures are believed to be small, thereby making (c) uninteresting to entertain as pointed out by Copas and Eguchi (2005) as well.

Solution (b) has the disadvantage that it makes the sensitivity analysis depend on the sample size. In particular, an increase of the sample decreases the level of correlation that is not significant. However, such an increased sample size provides only little (if not at all) information on the UA. Only the observation of new covariates may provide such information.

Instead, we advocate the use of a given correlation coefficient, say $\rho_m = 0.01$ or 0.05 , or both as a standard measure of small departure from the UA implied correlation 0. In order to interpret these levels of correlation, let us consider the simple case where there is no covariates. Then, we can write $y(0) = bz + \nu$, with $E(\nu) = 0$, and a correlation of 1% and 5% corresponds to a $R = \frac{\sqrt{\text{Var}(bz)}}{\sqrt{\text{Var}(y(0))}}$ value of 1% and 5% respectively. That is bz explains 1%, respectively 5%, of the variability in $y(0)$. The use of such a standard (say $\rho_m = 0.05$ level sensitivity analysis) would help the evaluation discipline by making sensitivity analyses comparable.

2.3.2 Prior distributions for the bias parameters

We want to base the sensitivity analysis on ρ_m , i.e. we need to choose a prior $p(\rho_m)$. Typically, if we want to consider correlations less or equal to r in absolute value, then a uniform on the interval $(-r, r)$ may be chosen for $p(\rho_m)$. We could also have situations where the direction of the dependence is known and, e.g., only positive values are considered $(0, r)$. We have argued above for $r = 0.01$ and 0.05 as standards to perform a sensitivity analysis.

Because ρ_m is a function of η (and not the contrary), a choice for $p(\rho_m)$ does not correspond to a unique prior for η . In order to obtain a prior $p(\eta)$ yielding the desired prior $p(\rho_m)$ we need to use a rejection algorithm. This is described in Appendix A.1.

3 Normal/logistic expanded model: a toy application

3.1 Data and evaluation of a program effect

Dehejia and Wahba (1999) estimate the effect of a labour training program (National Supported Work Demonstration) on post-treatment earnings. The treatment group used is the 185 individuals of the Lalonde (1986) sample for which 1974 earnings is available and the control group, denoted CPS-3 by Dehejia and Wahba (1999), is 429 individuals that resemble the treatment group in terms of some of the pre-treatment variables. We replicate Dehejia and Wahba (1999) analysis for the CPS-3 subsample. The outcome variable is 1978 earnings, the treatment is program participation and the covariates used in the model are age, age squared, years of schooling, high school dropout status, race and 1974 earnings. The estimated effect of the program under the null model (5) is an earnings premium of $\hat{\tau} = 1270$ U.S. Dollars with a standard error of 798. Hence, in this example parameter uncertainty is too large to be able to conclude that we have evidence for a program effect. For instance, a 95% Bayesian credible interval (also a 95% confidence interval here) for τ is 1270 ± 1596 .

3.2 Sensitivity analysis of the evaluation

We now conduct a sensitivity analysis of the above evaluation based on the theory presented in the previous sections. We need first to compute the correlation $\rho(\mathbf{x})$ for the Normal/logistic expanded model (3-4). We obtain (see Appendix A.2)

$$\rho(\mathbf{x}) = \frac{\frac{1}{4}\eta_1(a-b)}{\sqrt{\left(\frac{1}{2}(a+b) - \frac{1}{4}(a+b)^2\right) \cdot (\sigma^2 + \frac{1}{4}\eta_1^2)}}, \quad (10)$$

where

$$a = \frac{\exp(\mathbf{x}\gamma + \eta_2)}{1 + \exp(\mathbf{x}\gamma + \eta_2)}$$

and

$$b = \frac{\exp(\mathbf{x}\gamma)}{1 + \exp(\mathbf{x}\gamma)}.$$

Based on this correlation Algorithms 2 and 3 are used. We set $p_0(\eta) = p_0(\eta_1)p_0(\eta_2)$ with $p_0(\eta_1)$, the density of a uniform on $(-0.52, 0.52)$ and $p_0(\eta_2)$ the density of a normal distribution function with mean zero and variance equal to nine. A uniform is chosen for η_1 because this parameter is known to be bounded as $-0.52 \leq \eta_1 \leq 0.52$ (see Appendix A.3). A normal with large variance is used for η_2 to represent lack of information. In Algorithm 2 0.5 million replicates are drawn in Step 1 and in Step 2 we use the function `density` with parameter `kernel="cosine"` in R to estimate the density $p_0(\rho_m)$. The large amount of replicates was necessary to obtain a good approximation of $p_0(\rho_m)$ on the support of $p(\rho_m)$, which is important in Step 3 of Algorithm 3. The latter algorithm is run until 10'000

replicates are obtained. The outputs of both algorithms for ρ_m uniform on $(-0.05, 0.05)$ and on $(-0.01, 0.01)$ are displayed in Figures 3 and 4 respectively (Appendix A.4).

The sensitivity analysis can then be implemented with Algorithm 1. We run the algorithm 1000 times by sampling η with replacement from the output of Algorithm 3. The likelihood was given in Example 2.

The outputs of Algorithm 1 are displayed in Figure 1 together with the inference under the null model, i.e. a normal density with mean $\hat{\tau} = 1270$ and variance 798^2 . We note that the posteriors taking into account model uncertainty (plain lines) are not dramatically different from the inference under UA (dashed). We therefore conclude that at the 1% and 5% correlation level (as defined in Section 2.3) the inference in this evaluation is not sensitive to small departures from UA. Because parameter uncertainty was large (see above), it would have not have been necessary to look at model uncertainty in a practical situation. However, this exercise allows us to show a situation where sensitivity is low, which we can put in contrast with the next application.

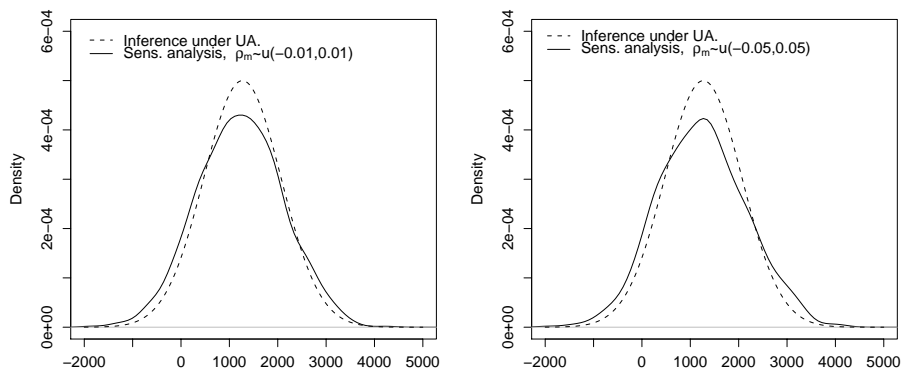


Figure 1: Posteriors for τ under the null model (dashed) and the expanded model (plain) with $|\rho_m| \leq 0.01$ (left panel) and $|\rho_m| \leq 0.05$ (right panel).

4 Application to a study of the effect of college choice on income

Lundin (2006) presented an evaluation of the effect of college choice on income performed with Swedish register data. In the evaluation, graduates in business/economics from old and new Swedish universities are compared. The universities classified as old are the Universities of Stockholm, Gothenburg, Lund, Uppsala, and Umeå as well as the Stockholm School of Economics. The new universities/colleges are those installed after 1965. The cohorts included in the sample are the ones that graduated from senior high school between the years 1990-1996 and then graduated from university with a degree in business or economics

before 2001. The old university group ($z = 1$) consists of 3343 individuals and the new university group ($z = 0$) consists of 3413 individuals.

A generalised additive model approach (Hastie and Tibshirani, 1990), controlling for the propensity score, $\Pr(z = 1|\mathbf{x})$, was used to estimate the effect on income of graduating from an old university ($z = 1$) instead of a new university ($z = 0$). The estimated earnings premium (log scale) 3 years after graduation was estimated to 0.061 with a standard error of 0.01. In order to perform a sensitivity analysis based on the Normal/logistic expanded model, we approximate the non-parametric functions in the generalised additive models used in Lundin (2006) with polynomial functions. The covariates included in the logistic model are gender, country of birth (Sweden/abroad), parents country of birth (Sweden/abroad), parents level of education (7 levels from elementary school to graduate education), whether the individual lived in a county with an old university prior to college choice, age at college start, duration of college studies, average grade in senior high school (polynomial of order 2) and the family’s disposable income (polynomial of order 2). The covariates are introduced in the normal model through a polynomial of order 3 of the estimated propensity score. Using these polynomial regression models we obtain an earnings premium of 0.062 with 0.01 as standard error. Here parameter uncertainty is low and there is strong evidence of college choice effect under UA.

We perform a sensitivity analysis of Lundin’s (2006) evaluation using the same setting as in Section 3.2, with the differences that a uniform on the interval $(-0.66, 0.66)$ is used for $p_0(\eta_1)$ and 300 replicates are computed in Algorithm 1. The outputs of Algorithms 2 and 3 for ρ_m uniform on $(-0.05, 0.05)$ and on $(-0.01, 0.01)$ are displayed in Figures 6 and 5 respectively (Appendix A.4).

The outputs of Algorithm 1 are displayed in Figure 2 together with the inference under the null model, i.e. a normal density with mean $\hat{\tau} = 0.06$ and variance 0.01^2 . We can see that this evaluation is sensitive to departures from the UA. The sensitivity is most obvious at the 5% correlation level, where, for instance, $\Pr(\tau \leq 0|\mathbf{y}, \mathbf{z}, X) = 0.01$ when $p(\rho_m)$ is uniform on $(-0.05, 0.05)$, while the same probability is equal to $3.8e-9$ under UA. The larger sensitivity in this application is not surprising since we have a much larger sample compared to the toy application, thereby increasing the importance of model uncertainty compared to parameter uncertainty. Note, however, that the observed sensitivity of the posterior distribution does not seem to invalidate the conclusion that there is evidence for a college choice effect on income. In other words, there is sensitivity if the inference of interest is to measure the size of the effect, while the decision about whether there is effect or not is not that sensitive.

5 Discussion

We have argued in this paper for a standardised sensitivity analysis to the unconfoundedness assumption. A purpose is to make comparable the results of sensitivity analyses made on different evaluation studies. Obviously our approach stands on several arbitrary model choices, such as, for instance, the distribution assumption on the confounder u . However,

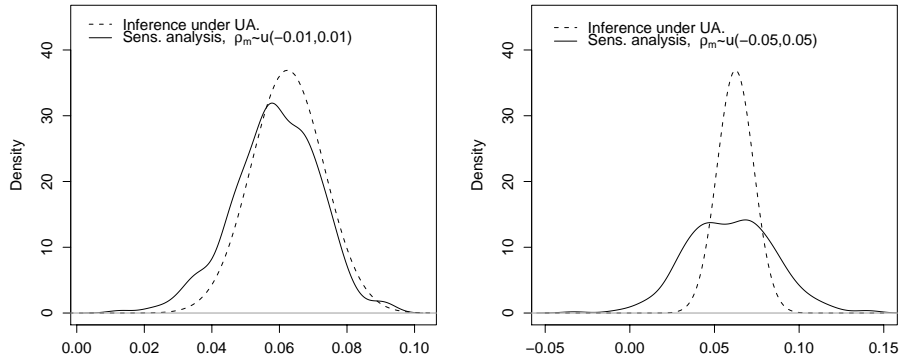


Figure 2: Posteriors for τ under the null model (dashed) and the expanded model (plain) with $|\rho_m| \leq 0.01$ (left panel) and $|\rho_m| \leq 0.05$ (right panel).

we do not believe that these choices have a large impact on the conclusions of the analysis because we focus the analysis on the correlation between the potential outcome and the treatment assignment. The modelling process is there only to induce such a correlation. In any case, if sensitivity is observed then the inference made under the null model should be questioned, this independently of the arbitrary choices made to conduct the analysis.

Sometimes model uncertainty can be incorporated in the final inference through Bayesian model averaging (Draper, 1995). However, this is reasonable only if you have grounds to specify prior probabilities on all alternative models. If you cannot, you may use a uniform prior on all alternative. When the outcome variable has finite support, the choice of a uniform prior for the bias parameters would yield an analysis similar to Manski's (1990), where bounds on the estimated effect are computed by not using the UA. In this paper, we instead argue that when the unconfoundedness assumption is a reasonable approximation a sensitivity analysis should be carried. For this purpose, we propose the use of an informative prior to measure small departures from the assumption. However, the information contained in the prior is arbitrary and is used only for the sake of the sensitivity analysis, and not to perform inference on the effect of interest. If the analysis shows no sensitivity, then the inference in point ii) (see Introduction) is performed conditional on the unconfoundedness assumption. On the other hand, if sensitivity to the UA is observed then we have no means to obtain a correct inference, otherwise than using a pessimistic totally uninformative prior on η .

References

- Copas, J. and Eguchi, S. (2001). Local sensitivity approximations for selectivity bias. *Journal of the Royal Statistical Society, Ser. B* 63 871-895.
- Copas, J. and Eguchi, S. (2005). Local model uncertainty and incomplete data bias. *Journal of the Royal Statistical Society, Ser. B* 67 459-513.
- Dehejia, R. and Wahba, S. (1999). Causal effects in non-experimental studies: Re-evaluating the evaluation of training programs. *Journal of the American Statistical Association*, 94, 1053-1062.
- Draper, D. (1995). Assessment and propagation of model uncertainty. *Journal of the Royal Statistical Society, Ser. B* 57 45-97.
- Gelman, A., Carlin, J., Stern, H. and Rubin, D. (1995). *Bayesian Data Analysis*, Chapman & Hall, London.
- Greenland, S. (2005). Multiple-bias modelling for analysis of observational data. *Journal of the Royal Statistical Society, Ser. A* 168, 267-306.
- Hastie, T.J. and Tibshirani, R.J. (1990). *Generalized Additive Models*, Chapman & Hall, London.
- Holland, P. (1986). Statistics and causal inference. *Journal of the American Statistical Association*, 81, 945-960.
- Imbens, G. (2003). Sensitivity to exogeneity assumptions in program evaluation. *The American Economic Review*, 93, 126-132.
- Lalonde, R. (1986). Evaluating the econometric evaluations of training programs with experimental data. *The American Economic Review*, 76, 604-620.
- Lundin, M. (2006). Effects of college choice on incomes in Sweden. *Working paper*.
- Manski, C.F. (1990). Nonparametric bounds on treatment effects. *American Economic Review Papers and Proceedings*, 80, 319-323.
- Morgan, B.J.T. (1984). *Elements of Simulation*. Chapman & Hall, London.
- Pawitan, Y. (2001). *In All Likelihood: Statistical modelling and Inference Using Likelihood*. Oxford University Press.
- Rosenbaum, P. (2002). *Observational Studies*, Springer, Berlin.
- Rubin, D.B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66, 688-701.

A Implementation details

A.1 Prior distributions for the bias parameters: rejection algorithm

Because ρ_m is a function of η (and not the contrary), a choice for $p(\rho_m)$ does not correspond to a unique prior for η . In order to obtain a prior $p(\eta)$ yielding the desired prior $p(\rho_m)$ we need to use a rejection algorithm. We operate in two stages. In a first stage (Algorithm 2) we choose an non-informative prior for η , $p_0(\eta)$. The latter implies a prior for ρ_m , $p_0(\rho_m)$ which we obtain by simulation.

Algorithm 2 Density function of $p_0(\rho_m)$

Step 1: Generate many η values from $p_0(\eta)$ and compute their corresponding ρ_m value.

Step 2: Estimate the density function $p_0(\rho_m)$ based on the simulated values in Step 1.

The output of Algorithm 2 is a set of replicates of the vector (η_1, η_2, ρ_m) and a corresponding empirical density function $p_0(\rho_m)$. Both are used in a second stage (Algorithm 3) which consists of a rejection algorithm (Morgan, 1984, p. 100) with envelope $p_0(\rho_m)$.

Algorithm 3 Replicates from $p(\eta)$ for a given $p(\rho_m)$

Step 1: Sample (η_1, η_2, ρ_m) with replacement from the output of Algorithm 2. (i.e. generate ρ_m from $p_0(\rho_m)$)

Step 2: Generate a uniform number, v , on the interval $(0, p_0(\rho_m))$.

Step 3: Accept (η_1, η_2, ρ_m) if $\rho_m \in S(p(\rho_m))$ and $v < Kp(\rho_m)$, where $S(p(\rho_m))$ is the support of $p(\rho_m)$, and K is a constant such that $Kp(\rho_m) < p_0(\rho_m)$ on $S(p(\rho_m))$.

Step 4: Repeat Steps 1-3 until the desired number of triplets (η_1, η_2, ρ_m) has been accepted.

The output of Algorithm 3 is a set of replicates of the vector (η_1, η_2, ρ_m) such that the ρ_m 's are from the desired prior $p(\rho_m)$, and thereby the η from a density $p(\eta)$ actually yielding $p(\rho_m)$. In our applications we use $p(\rho_m) = 1/(2r)$ on the support $S(p(\rho_m)) = (-r, r)$ with $r = 0.01$ and 0.05 .

A.2 Correlation for the normal/logistic model

We give here the details of the computation of $\rho(\mathbf{x}) = \text{Corr}(y(0), z|\mathbf{x})$ for the model of Example 1 and 2.

$$\begin{aligned}\rho(\mathbf{x}) &= \frac{\text{Cov}(z, y(0)|\mathbf{x})}{\sqrt{\text{Var}(z|\mathbf{x})\text{Var}(y(0)|\mathbf{x})}} \\ &= \frac{E(zy(0)|\mathbf{x}) - E(z|\mathbf{X})E(y(0)|\mathbf{x})}{\sqrt{\text{Var}(z|\mathbf{x})\text{Var}(y(0)|\mathbf{x})}},\end{aligned}$$

where

$$E(z|\mathbf{x}) = \frac{1}{2} \left(\frac{\exp(\mathbf{x}\gamma + \eta_2 u)}{1 + \exp(\mathbf{x}\gamma + \eta_2 u)} + \frac{\exp(\mathbf{x}\gamma)}{1 + \exp(\mathbf{x}\gamma)} \right),$$

$$E(y(0)|\mathbf{x}) = E(E(y(0)|\mathbf{x}, u)|\mathbf{x}) = E(\mathbf{x}\beta + \eta_1 u|\mathbf{x}) = \mathbf{x}\beta + \frac{1}{2}\eta_1,$$

$$\begin{aligned} E(zy(0)|\mathbf{x}) &= E(E(zy(0)|\mathbf{x}, u)|\mathbf{x}) \\ &= E(E(z|\mathbf{x}, u)E(y(0)|\mathbf{x}, u)|\mathbf{x}) \\ &= E\left(\frac{\exp(\mathbf{x}\gamma + \eta_2 u)}{1 + \exp(\mathbf{x}\gamma + \eta_2 u)}(\mathbf{x}\beta + \eta_1 u)|\mathbf{x}\right) \\ &= \frac{1}{2}(\mathbf{x}\beta + \eta_1) \left(\frac{\exp(\mathbf{x}\gamma + \eta_2)}{1 + \exp(\mathbf{x}\gamma + \eta_2)} \right) + \frac{1}{2}\mathbf{x}\beta \left(\frac{\exp(\mathbf{x}\gamma)}{1 + \exp(\mathbf{x}\gamma)} \right), \end{aligned}$$

$$\begin{aligned} Cov(z, y(0)|\mathbf{x}) &= \frac{1}{2}(\mathbf{x}\beta + \eta_1) \left(\frac{\exp(\mathbf{x}\gamma + \eta_2)}{1 + \exp(\mathbf{x}\gamma + \eta_2)} \right) + \frac{1}{2}\mathbf{x}\beta \left(\frac{\exp(\mathbf{x}\gamma)}{1 + \exp(\mathbf{x}\gamma)} \right) \\ &\quad - \frac{1}{2} \left(\frac{\exp(\mathbf{x}\gamma + \eta_2)}{1 + \exp(\mathbf{x}\gamma + \eta_2)} + \frac{\exp(\mathbf{x}\gamma)}{1 + \exp(\mathbf{x}\gamma)} \right) (\mathbf{x}\beta + \frac{1}{2}\eta_1) \\ &= \frac{1}{2}\eta_1 \left(\frac{\exp(\mathbf{x}\gamma + \eta_2)}{1 + \exp(\mathbf{x}\gamma + \eta_2)} \right) - \frac{1}{4}\eta_1 \left(\frac{\exp(\mathbf{x}\gamma + \eta_2)}{1 + \exp(\mathbf{x}\gamma + \eta_2)} + \frac{\exp(\mathbf{x}\gamma)}{1 + \exp(\mathbf{x}\gamma)} \right) \\ &= \frac{1}{4}\eta_1 \left(\frac{\exp(\mathbf{x}\gamma + \eta_2)}{1 + \exp(\mathbf{x}\gamma + \eta_2)} - \frac{\exp(\mathbf{x}\gamma)}{1 + \exp(\mathbf{x}\gamma)} \right). \end{aligned}$$

$$\begin{aligned} Var(y(0)|\mathbf{x}) &= E(Var(y(0)|\mathbf{x}, u)|\mathbf{x}) + Var(E(y(0)|\mathbf{x}, u)|\mathbf{x}) \\ &= \sigma^2 + Var(\mathbf{x}\beta + \eta_1 u|\mathbf{x}) = \sigma^2 + \frac{1}{4}\eta_1, \end{aligned}$$

and

$$\begin{aligned}
\text{Var}(z|\mathbf{x}) &= E(\text{Var}(z|\mathbf{x}, u)|\mathbf{x}) + \text{Var}(E(z|\mathbf{x}, u)|\mathbf{x}) \\
&= \frac{1}{2} \left(\frac{\exp(\mathbf{x}\gamma + \eta_2)}{1 + \exp(\mathbf{x}\gamma + \eta_2)} - \left(\frac{\exp(\mathbf{x}\gamma + \eta_2)}{1 + \exp(\mathbf{x}\gamma + \eta_2)} \right)^2 \right) \\
&\quad + \frac{1}{2} \left(\frac{\exp(\mathbf{x}\gamma)}{1 + \exp(\mathbf{x}\gamma)} - \left(\frac{\exp(\mathbf{x}\gamma)}{1 + \exp(\mathbf{x}\gamma)} \right)^2 \right) \\
&\quad + \frac{1}{2} \left(\left(\frac{\exp(\mathbf{x}\gamma + \eta_2)}{1 + \exp(\mathbf{x}\gamma + \eta_2)} \right)^2 + \left(\frac{\exp(\mathbf{x}\gamma)}{1 + \exp(\mathbf{x}\gamma)} \right)^2 \right) \\
&\quad - \frac{1}{4} \left(\frac{\exp(\mathbf{x}\gamma + \eta_2)}{1 + \exp(\mathbf{x}\gamma + \eta_2)} + \frac{\exp(\mathbf{x}\gamma)}{1 + \exp(\mathbf{x}\gamma)} \right)^2 \\
&= \frac{1}{2} \left(\frac{\exp(\mathbf{x}\gamma + \eta_2)}{1 + \exp(\mathbf{x}\gamma + \eta_2)} + \frac{\exp(\mathbf{x}\gamma)}{1 + \exp(\mathbf{x}\gamma)} \right) \\
&\quad - \frac{1}{4} \left(\frac{\exp(\mathbf{x}\gamma + \eta_2)}{1 + \exp(\mathbf{x}\gamma + \eta_2)} + \frac{\exp(\mathbf{x}\gamma)}{1 + \exp(\mathbf{x}\gamma)} \right)^2.
\end{aligned}$$

The resulting value for $\rho(\mathbf{x})$ was given in (10). This correlation has to be computed many times (in our applications 0.5 million) using Algorithm 2 in order to obtain a good estimate of $p(\rho_m)$ as needed in Algorithm 3. The optimal option would have been to estimate the vector γ for a given value of η with the likelihood (9); see below on the maximisation of the likelihood. However, this would have been extremely computer intensive and we chose to compute approximate values for γ as follows. For $\eta_1 \approx 0$ we have that γ maximise the likelihood:

$$L(\gamma; \eta_2) = \prod_{i=1}^n \left[\frac{1}{2} \frac{(\exp(\gamma \mathbf{x}_i))^{z_i}}{1 + \exp(\gamma \mathbf{x}_i)} + \frac{1}{2} \frac{(\exp(\gamma \mathbf{x}_i + \eta_2))^{z_i}}{1 + \exp(\gamma \mathbf{x}_i + \eta_2)} \right].$$

Instead of maximising this likelihood 0.5 million times we retrieve $\gamma(\eta_2)$ from $\hat{\gamma}$ the maximum likelihood estimator of $L(\gamma; \eta_2 = 0)$. Integrating out u from (7) we have

$$\begin{aligned}
\Pr(z = 1|\mathbf{x}; \gamma, \eta_2) &= \frac{1}{2} \exp(\gamma' \mathbf{x}) / (1 + \exp(\gamma' \mathbf{x})) \\
&\quad + \frac{1}{2} \exp(\gamma' \mathbf{x} + \eta_2) / (1 + \exp(\gamma' \mathbf{x} + \eta_2)). \tag{11}
\end{aligned}$$

Denote $X = x_1 + \gamma_2/\gamma_1 x_2 + s + \gamma_q/\gamma_1 x_q$, where x_1, \dots, x_q are the q covariates of the q dimensional \mathbf{x} vector. Then, (11) is a cumulative distribution function $F(x)$ (mixture of two logistic) and can be rewritten

$$F(x) = \frac{1}{2} \frac{\exp((x - m)b)}{1 + \exp((x - m)b)} + \frac{1}{2} \frac{\exp((x - m - \eta_1/b)b)}{1 + \exp((x - m - \eta_1/b)b)},$$

where $m = -\gamma_0/\gamma_1$ and $b = \gamma_1$.

When we are assuming that $\eta_2 = 0$, we are considering instead the logistic cumulative distribution function

$$G(x) = \frac{\exp((x - n)c)}{1 + \exp((x - n)c)},$$

In the case η_2 is actually equal to zero then $n = m$ and $c = b$. However, if $\eta_2 \neq 0$ (misspecification) then we can retrieve n and c by equating the two first moments implied by $F(x)$ and $G(x)$. We have

$$\begin{aligned}\tilde{\mu} &= E_F(X) = m + \eta_1/2b, \\ E_f(X) &= \frac{1}{2} \left(2\frac{\pi}{3b} + m/2 + (m + \eta_1/b)/2 \right), \\ \tilde{\sigma} &= Var_F(X) = \frac{\pi}{3b} + \frac{m}{2} + \frac{(m + \eta_1/b)}{2} - \frac{(2m + \eta_1/b)}{4}.\end{aligned}$$

The expectation of X^2 is obtained by noting that $F(x)$ is a mixture of two density functions. Moreover,

$$\begin{aligned}\mu &= E_G(X) = n, \\ \sigma &= Var_G(X) = \frac{\pi}{3c}.\end{aligned}$$

Equating $\tilde{\mu} = \mu$ and $\tilde{\sigma} = \sigma$ yields

$$\begin{aligned}m &= n - \eta_1/2b, \\ \frac{1}{b} &= \frac{3}{\pi} \left(\frac{\pi}{3c} - \frac{m}{2} - \frac{(m + \eta_1/b)}{2} + \frac{(2m + \eta_1/b)}{4} \right),\end{aligned}$$

which we solve for m and b :

$$\begin{aligned}m &= n - \frac{\eta_1}{c\alpha}, \\ b &= \frac{c\alpha}{2}, \\ \alpha &= \sqrt{4\pi + 3\eta_1/\pi}.\end{aligned}$$

Thus, assuming $\eta_2 = 0$ we obtain estimates of

$$\gamma'_0 = cn, \quad \gamma'_1 = c, \quad \gamma'_2 = \frac{c\gamma_2}{b}, \dots, \gamma'_p = \frac{c\gamma_p}{b},$$

by maximising the likelihood $L(\gamma; \eta_2 = 0)$ given above, thereby yielding the estimates

$$\begin{aligned}\hat{c} &= \hat{\gamma}'_1, \\ \hat{n} &= \hat{\gamma}'_0/\hat{c}, \\ \hat{m} &= \hat{n} - \eta_1/(\hat{c}\alpha), \\ \hat{b} &= \hat{c}\alpha/2.\end{aligned}$$

Finally, we retrieve estimates of the correctly specified model

$$\begin{aligned}\hat{\gamma}_0 &= \hat{m}\hat{b}, \\ \hat{\gamma}_1 &= \hat{b}, \\ \hat{\gamma}_j &= \frac{\hat{\gamma}'_j\hat{b}}{c}, \quad j = 2, \dots, p.\end{aligned}$$

As noted above, this heuristic holds for $\eta_1 \approx 0$. When η_1 is “large” then η_2 will tend to be small for the small correlations of interest and therefore the calibration made above will typically be of little effect.

A.3 Bound for η_1

Fitting the null model (5) to the data provides estimates of β and σ^2 . The latter estimated residual variance defines bounds for η_1 in the expanded model (3) because of the constraint $\nu^2 > 0$. We have

$$\sigma^2 = \text{Var}(y_0|\mathbf{x}) = \eta_1^2 \text{Var}(u) + \text{Var}(y_0|\mathbf{x}, u) = \frac{\eta_1^2}{4} + \nu^2,$$

where we have used the fact that u is assumed independent of \mathbf{x} . Hence, $\nu > 0$ implies

$$|\eta_1| < 2\sigma.$$

A.4 More figures

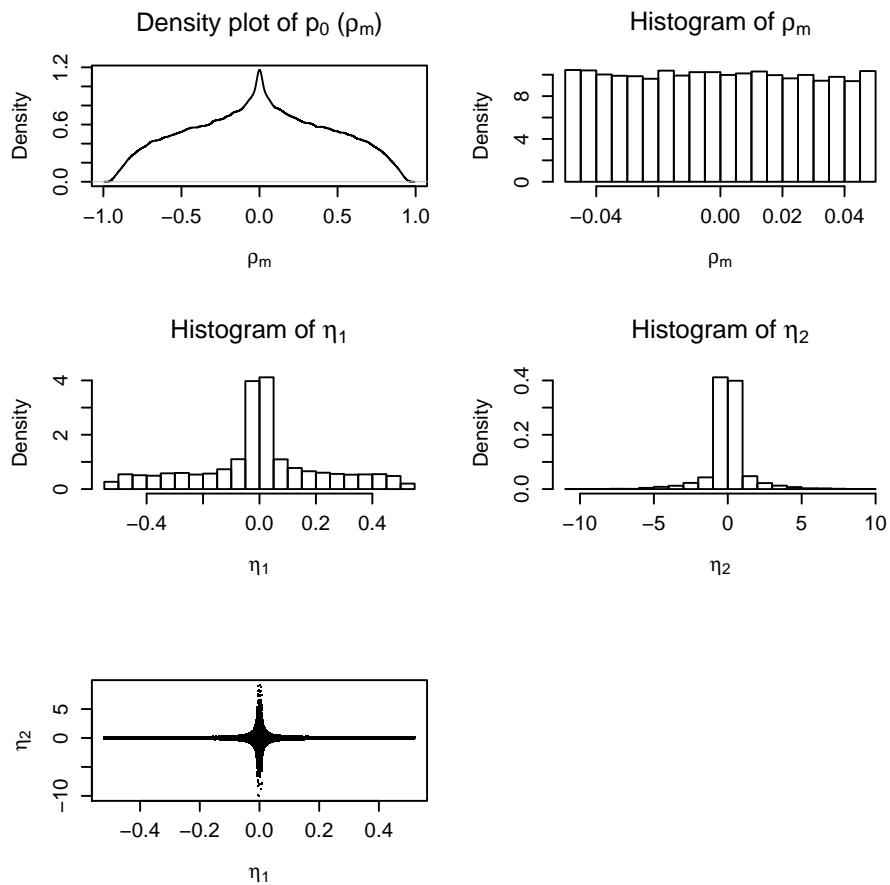


Figure 3: Outputs of Algorithms 2 and 3 for $|\rho_m| \leq 0.05$. From left to right and top to bottom: the estimated density $p_0(\rho_m)$ (Algorithm 2); the histogram of the replicates of ρ_m , η_1 and η_2 from Algorithm 3; a scatter plot of the latter η_1 and η_2 against each others.

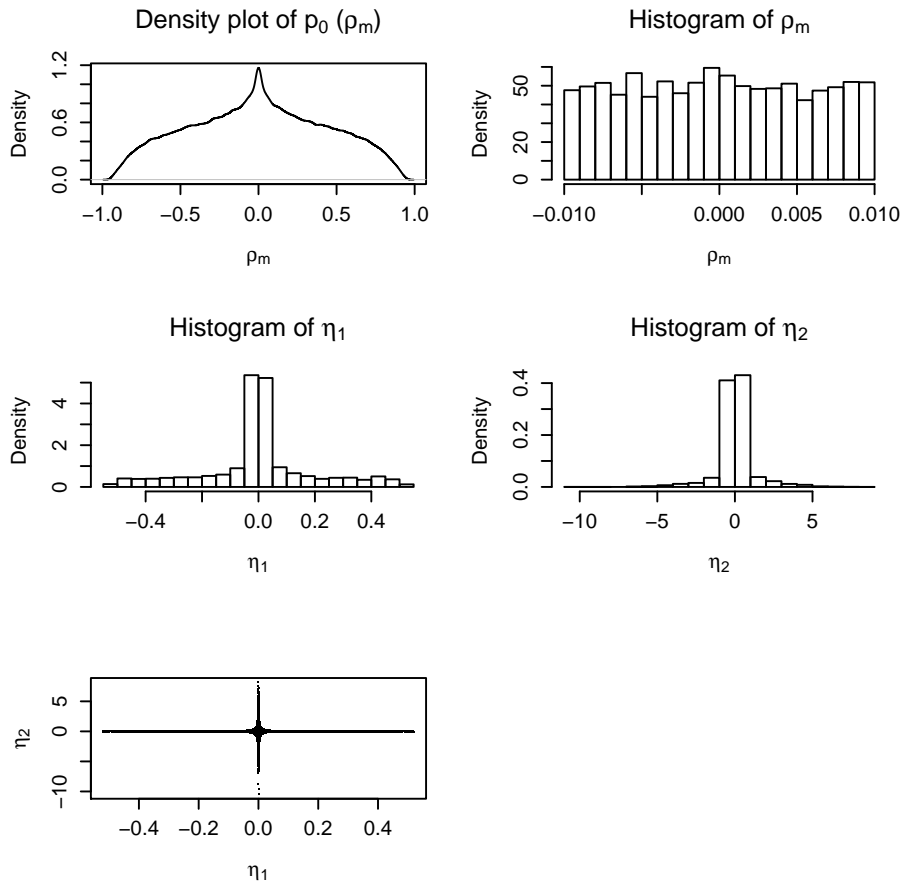


Figure 4: Outputs of Algorithms 2 and 3 ($|\rho_m| \leq 0.01$) for the toy application of Section 3. From left to right and top to bottom: the estimated density $p_0(\rho_m)$ (Algorithm 2); the histogram of the replicates of ρ_m , η_1 and η_2 from Algorithm 3; a scatter plot of the latter η_1 and η_2 against each others.

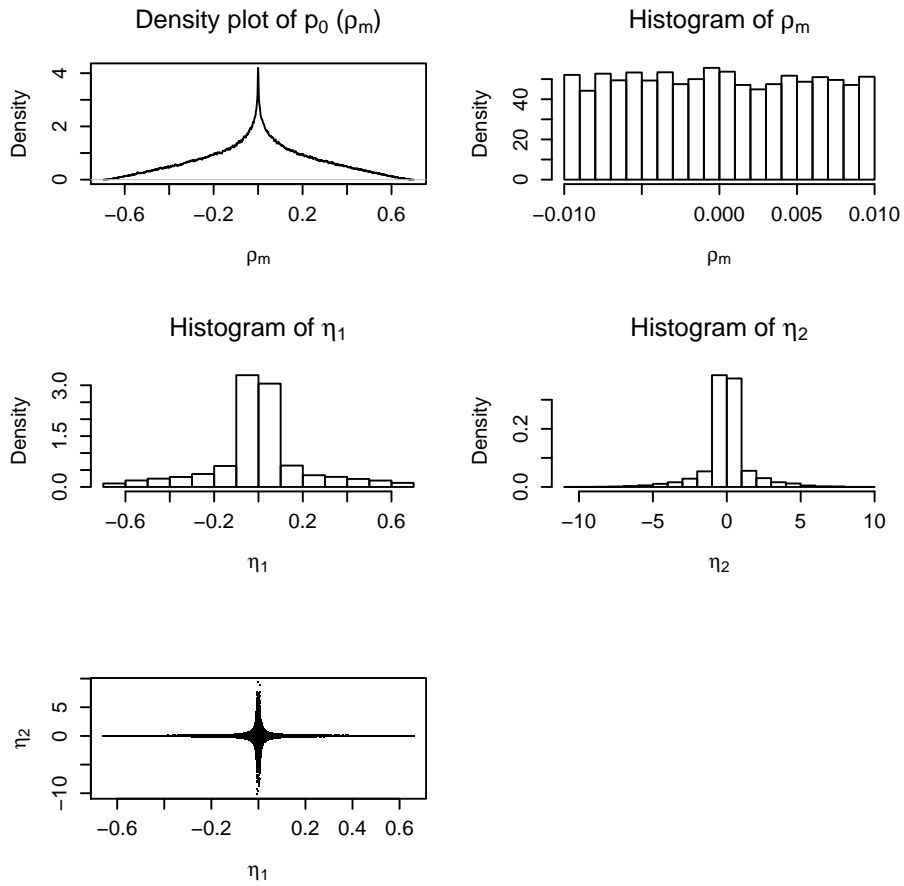


Figure 5: Outputs of Algorithms 2 and 3 ($|\rho_m| \leq 0.01$) for the application of Section 4. From left to right and top to bottom: the estimated density $p_0(\rho_m)$ (Algorithm 2); the histogram of the replicates of ρ_m , η_1 and η_2 from Algorithm 3; a scatter plot of the latter η_1 and η_2 against each others.

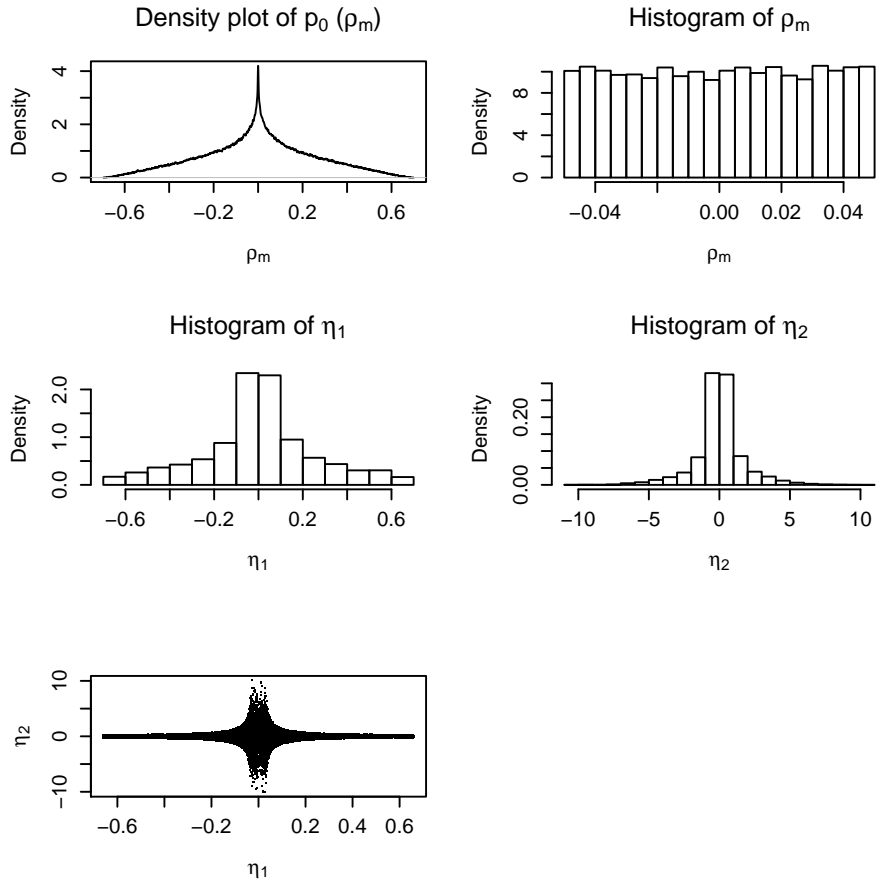


Figure 6: Outputs of Algorithms 2 and 3 for $|\rho_m| \leq 0.05$. From left to right and top to bottom: the estimated density $p_0(\rho_m)$ (Algorithm 2); the histogram of the replicates of ρ_m , η_1 and η_2 from Algorithm 3; a scatter plot of the latter η_1 and η_2 against each others.

Publication series published by the Institute for Labour Market Policy Evaluation (IFAU) – latest issues

Rapporter/Reports

- 2009:1** Hartman Laura, Per Johansson, Staffan Khan and Erica Lindahl, "Uppföljning och utvärdering av Sjukvårdsmiljarden"
- 2009:2** Chirico Gabriella and Martin Nilsson "Samverkan för att minska sjukskrivningar – en studie av åtgärder inom Sjukvårdsmiljarden"
- 2009:3** Rantakeisu Ulla "Klass, kön och platsanvisning. Om ungdomars och arbetsförmedlars möte på arbetsförmedlingen"
- 2009:4** Dahlberg Matz, Karin Edmark, Jörgen Hansen and Eva Mörk "Fattigdom i folkhemmet – från socialbidrag till självförsörjning"
- 2009:5** Pettersson-Lidbom Per and Peter Skogman Thoursie "Kan täta födelseintervaller mellan syskon försämra deras chanser till utbildning?"
- 2009:6** Grönqvist Hans "Effekter av att subventionera p-piller för tonåringar på barnafödande, utbildning och arbetsmarknad"
- 2009:7** Hall Caroline "Förlängningen av yrkesutbildningarna på gymnasiet: effekter på utbildningsavhopp, utbildningsnivå och inkomster"
- 2009:8** Gartell Marie "Har arbetslöshet i samband med examen från högskolan långsiktiga effekter?"
- 2009:9** Kennerberg Louise "Hur försörjer sig nyanlända invandrare som inte deltar i sfi?"
- 2009:10** Lindvall Lars "Bostadsområde, ekonomiska incitament och gymnasieval"

Working papers

- 2009:1** Crépon Bruno, Marc Ferracci, Grégory Jolivet and Gerard J. van den Berg "Active labor market policy effects in a dynamic setting"
- 2009:2** Hesselius Patrik, Per Johansson and Peter Nilsson "Sick of your colleagues' absence?"
- 2009:3** Engström Per, Patrik Hesselius and Bertil Holmlund "Vacancy referrals, job search and the duration of unemployment: a randomized experiment"
- 2009:4** Horny Guillaume, Rute Mendes and Gerard J. van den Berg "Job durations with worker and firm specific effects: MCMC estimation with longitudinal employer-employee data"
- 2009:5** Bergemann Annette and Regina T. Riphahn "Female labor supply and parental leave benefits – the causal effect of paying higher transfers for a shorter period of time"
- 2009:6** Pekkarinen Tuomas, Roope Uusitalo and Sari Kerr "School tracking and development of cognitive skills"
- 2009:7** Pettersson-Lidbom Per and Peter Skogman Thoursie "Does child spacing affect children's outcomes? Evidence from a Swedish reform"
- 2009:8** Grönqvist Hans "Putting teenagers on the pill: the consequences of subsidized contraception"

- 2009:9** Hall Caroline "Does making upper secondary school more comprehensive affect dropout rates, educational attainment and earnings? Evidence from a Swedish pilot scheme"
- 2009:10** Gartell Marie "Unemployment and subsequent earnings for Swedish college graduates: a study of scarring effects"
- 2009:11** Lindvall Lars "Neighbourhoods, economic incentives and post compulsory education choices"
- 2009:12** de Luna Xavier and Mathias Lundin "Sensitivity analysis of the unconfoundedness assumption in observational studies"

Dissertation series

- 2009:1** Lindahl Erica "Empirical studies of public policies within the primary school and the sickness insurance"
- 2009:2** Grönqvist Hans "Essays in labor and demographic economics"