

Article

On the Plausibility of the Latent Ignorability Assumption

Martin Huber ^{1,2} 

¹ Department of Economics, University of Fribourg, Bd. de Pérolles 90, 1700 Fribourg, Switzerland; martin.huber@unifr.ch

² Center for Econometrics and Business Analytics, St. Petersburg State University, 199034 St. Petersburg, Russia

Abstract: The estimation of the causal effect of an endogenous treatment based on an instrumental variable (IV) is often complicated by the non-observability of the outcome of interest due to attrition, sample selection, or survey non-response. To tackle the latter problem, the latent ignorability (LI) assumption imposes that attrition/sample selection is independent of the outcome conditional on the treatment compliance type (i.e., how the treatment behaves as a function of the instrument), the instrument, and possibly further observed covariates. As a word of caution, this note formally discusses the strong behavioral implications of LI in rather standard IV models. We also provide an empirical illustration based on the Job Corps experimental study, in which the sensitivity of the estimated program effect to LI and alternative assumptions about outcome attrition is investigated.

Keywords: instrument; non-response; attrition; sample selection; latent ignorability



Citation: Huber, Martin. 2021. On the Plausibility of the Latent Ignorability Assumption. *Econometrics* 9: 47. <https://doi.org/10.3390/econometrics9040047>

Academic Editors: Marc S. Paoletta and Ryo Okui

Received: 25 September 2021

Accepted: 3 December 2021

Published: 9 December 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

A frequently encountered complication when estimating the effect of a potentially endogenous treatment based on an instrumental variable (IV) methods is attrition/sample selection/non-response bias in the outcome. To account for this problem, the missing at random (MAR) assumption (e.g., [Rubin \(1976\)](#)), for instance, requires outcome attrition to only depend on observable variables. Alternatively, [Frangakis and Rubin \(1999\)](#) propose a latent ignorability (LI) restriction, which assumes attrition to be independent of the outcome conditional on the instrument and the treatment compliance type considered in [Angrist et al. \(1996\)](#). A compliance type is defined in terms of how the treatment (e.g., participation in a training) depends on or complies with the value of the instrument (e.g., random assignment to a training), such that the population generally consists of compliers (whose training participation always corresponds to the random assignment) and non-compliers. In the IV framework, we may even combine the MAR and LI assumptions to impose independence between attrition and outcomes conditional on the compliance type, the instrument, and further observed covariates.

We argue that LI is nevertheless quite restrictive, as attrition is not allowed to be related to unobservables affecting the outcome in a very general way. In fact, LI appears hard to justify in quite standard IV models with non-response and should therefore be cautiously scrutinized in applications. As an example, consider [Barnard et al. \(2003\)](#), who assess a randomized voucher program for private schooling with noncompliance (where the IV is the randomization and the treatment is private schooling) and attrition in the test score outcomes, because some children did not take the test. Unobservables as ability or motivation likely affect both test taking and test scores. LI (combined with MAR) requires that conditional on the compliance type (i.e. private schooling as a function of voucher receipt), voucher assignment, and observed covariates, test taking is not related to ability or motivation (and thus, test scores). Among compliers (only in private schooling when randomized in), those taking the test must thus have the same distribution of ability and motivation as those abstaining. However, even within compliers, heterogeneity in ability and motivation may be sufficiently high to selectively affect test taking such that LI fails.

As a second example, Mealli et al. (2004) as well as Mattei and Mealli (2007) consider a randomized trial on teaching breast self-examination (BSE), either based on mailed information (standard treatment) or on attendance in a course (new treatment), to investigate the impact on BSE practice (as a method of breast cancer prevention). However, a substantial share of women assigned to the course did not participate (noncompliance) and furthermore, not all study subjects did answer the follow-up survey on BSE practice (non-response). Unobservables likely affecting both survey response and BSE practice are interest in breast cancer prevention and risk awareness about breast cancer. Also within the subpopulation of compliers, differences in interest and risk awareness could systematically affect response behavior so that LI is violated, even conditional on observed covariates.

The LI or its combination with MAR is invoked in a range of further studies in the fields of medicine, (bio-)statistics, political science, and economics. O'Malley and Normand (2005) for instance suggest a maximum likelihood-based estimator and apply it to compare the relative effectiveness of two medical treatments among adults with refractory schizophrenia under treatment non-compliance and outcome attrition. Chen et al. (2015) use an LI approach to verify the robustness of their finding that high-calcium milk powder effectively reduces bone loss at the lumbar spine as well as height loss among treatment compliers consisting of postmenopausal women. Esterling et al. (2011) suggest a parametric estimator incorporating LI and apply it to measure the effect of participating in a deliberative session with U.S. politicians (about federal immigration and border control policies) on political impressions, e.g., whether public officials care about citizens' opinions. Frölich and Huber (2014) extend the LI framework to multiple outcome periods with increasing attrition across periods and evaluate the effect of a program aiming at increasing college achievement on students' grade point average in the first and second year. Adapting the LI approach to mediation analysis, Yamamoto (2013) disentangles the total treatment effect among compliers into its direct impact on the outcome and an indirect causal mechanism operating via an intermediate variable (or mediator), whose endogeneity is tackled by an LI assumption.

The remainder of this paper is organized as follows. Section 2 formally discusses the strong behavioral implications of LI in standard IV models with non-response. Section 3 provides an empirical illustration using the Job Corps experimental study, in which the estimated program effect under LI is compared to alternative assumptions about outcome attrition. Section 4 concludes.

2. IV Models with Nonresponse

Assume the following parametric IV model with nonresponse:

$$Y = \alpha_0 + D\alpha_1 + U, \quad D = 1(\beta_0 + Z\beta_1 \geq V), \quad R = 1(\gamma_0 + D\gamma_1 \geq W). \quad (1)$$

Y is the outcome of interest, D is the binary (and potentially endogenous) treatment, and R is the response indicator. Note that $1(\cdot)$ is the indicator function that is equal to one if its argument is satisfied and zero otherwise. Y is only observed if $R = 1$ and unknown if $R = 0$, implying non-response, sample selection, or attrition. Z is a randomly assigned instrument affecting D (but not directly Y or R) and assumed to be binary, e.g., the randomization indicator in an experiment. U, V, W denote arbitrarily associated unobservables, $\alpha_0, \alpha_1, \beta_0, \beta_1, \gamma_0, \gamma_1$ are coefficients.

Angrist et al. (1996) define four compliance types, denoted by T , based on how the potential treatment status depends on the instrument: An individual is a complier (defier) if her potential treatment state is one (zero) in the presence and zero (one) in the absence of the instrument and an always-taker (never-taker) if the potential treatment is always (never) one, independent of the instrument. Assume that β_1 is positive (a symmetric case could be made for a negative β_1). Then, an individual is a complier if $\beta_0 + \beta_1 \geq V > \beta_0$, an always taker if $\beta_0 \geq V$, and a never taker if $\beta_0 + \beta_1 < V$. Defiers do not exist due to the positive sign of β_1 .

We now impose the following latent ignorability (LI) assumption, see [Frangakis and Rubin \(1999\)](#), and critically assess it in the light of our standard IV model with attrition:

Assumption 1 (latent ignorability). $Y \perp R | Z, T$ (where ‘ \perp ’ denotes independence).

Which is equivalent to $Y \perp R | Z, D, T$ as Z and T perfectly determine D . Furthermore, we assume that the error term U is continuous, such that Y is continuous. Finally, for the moment we also impose that $U = V = W$ such that the same unobservable (e.g., motivation) affects the outcome (e.g., test score), treatment (e.g., private schooling), and response (e.g., test taking).

Note that Assumption 1 implies that the distribution of U among compliers is the same across response states given the instrument:

$$\begin{aligned} E(f(Y)|Z = 1, T = c, R = 1) &= E(f(Y)|Z = 1, T = c, R = 0) & (2) \\ \Leftrightarrow E(f(U)|Z = 1, \beta_0 + \beta_1 \geq U > \beta_0, \gamma_0 + \gamma_1 \geq U) & \\ = E(f(U)|Z = 1, \beta_0 + \beta_1 \geq U > \beta_0, \gamma_0 + \gamma_1 < U), & \end{aligned}$$

where $f(\cdot)$ denotes an arbitrary function with a finite expectation and the second line follows from the parametric model in (1). Obviously, the joint satisfaction of $U = V = W$ and (2) is impossible in this context, as the distribution of U conditional on $\gamma_0 + \gamma_1 \geq U$ and $\gamma_0 + \gamma_1 < U$, respectively, is non-overlapping. An analogous impossibility result holds for $E(f(Y)|Z = 0, T = c, R = 1) = E(f(Y)|Z = 0, T = c, R = 0)$, which is also implied by Assumption 1.

Imposing $U = V = W$ seems too extreme for most applications and was chosen for illustrative purposes. However, even if the unobserved terms in the various equations are not the same, but non-negligibly correlated as commonly assumed in IV models, identification may seem questionable. Suppose, for instance, that $W = \delta_1 V + \epsilon$, where ϵ is random noise and δ_1 is a coefficient. Then, Assumption 1 and the model in (1) imply that

$$\begin{aligned} E(f(U)|Z = 1, \beta_0 + \beta_1 \geq V > \beta_0, \gamma_0 + \gamma_1 \geq \delta_1 V + \epsilon) & (3) \\ = E(f(U)|Z = 1, \beta_0 + \beta_1 \geq V > \beta_0, \gamma_0 + \gamma_1 < \delta_1 V + \epsilon) & \\ \Leftrightarrow E\left(f(U)|Z = 1, \min\left(\beta_0 + \beta_1, \frac{\gamma_0 + \gamma_1 - \epsilon}{\delta_1}\right) \geq V > \beta_0\right) & \\ = E\left(f(U)|Z = 1, \beta_0 + \beta_1 \geq V > \max\left(\beta_0, \frac{\gamma_0 + \gamma_1 - \epsilon}{\delta_1}\right)\right). & \end{aligned}$$

If U is associated with either ϵ , V , or both, the latter equality does not hold in general, but only if the association of U, ϵ, V is of a very specific form, which raises concerns about Assumption 1.

Finally, we investigate an in terms of functional form assumptions more general IV model, where Y, D , and R are given by nonparametric functions denoted by ϕ, ψ , and η , respectively:

$$Y = \phi(D, U), \quad D = 1(\psi(Z, V) \geq 0), \quad R = 1(\eta(D, W) \geq 0). \tag{4}$$

Under this model, Assumption 1 implies that

$$\begin{aligned} E(f(U)|Z = 1, \psi(1, V) \geq 0, \psi(0, V) < 0, \eta(1, W) \geq 0) & (5) \\ = E(f(U)|Z = 1, \psi(1, V) \geq 0, \psi(0, V) < 0, \eta(1, W) < 0). & \end{aligned}$$

This can be satisfied in special cases, for instance if $U = \pi 1(\psi(1, V) \geq 0, \psi(0, V) < 0) + \epsilon$, with π denoting the (homogeneous) effect of being a complier and ϵ being random noise. Then, (5) simplifies to $E(f(\epsilon)|Z = 1, \psi(1, V) \geq 0, \psi(0, V) < 0, \eta(1, W) \geq 0) = E(f(\epsilon)|Z = 1, \psi(1, V) \geq 0, \psi(0, V) < 0, \eta(1, W) < 0)$, which holds because ϵ is independent of W . In general, identification requires that T is a sufficient statistic to control for the

endogeneity introduced by conditioning on R . This, however, implies that the association between U , V , and W is quite specific, otherwise Assumption 1 does not hold.

3. Empirical Illustration

As an illustration for treatment evaluation under LI and alternative assumptions about attrition, we consider the experimental evaluation of the U.S. Job Corps program (see for instance [Schochet et al. \(2001\)](#)), providing training and education for young disadvantaged individuals. We aim at estimating the effect of program participation (D) in the first or second year after randomization into Job Corps (Z) on log weekly wages of females in the third year (Y). Of the 4765 females in the experimental sample with observed treatment status, wages are only observed for 3682 individuals ($R = 1$), while 1083 do not report to work.

Reconsidering the IV model of (4), we assume that in each of ϕ , ψ , and η a vector of observed covariates, denoted by X , may enter as additional explanatory variables. Similar to [Frölich and Huber \(2014\)](#), Section 2.2, we assume that (i) Assumption 1 holds conditional on X (thus combining LI and MAR), (ii) $U \perp Z | X, T$ such that the instrument affects the outcome only through the treatment, (iii) $T \perp Z | X$ which is implied by random assignment, (iv) $\Pr(T = c) > 0$ and $\Pr(T = d) = 0$ so that compliers exist and defiers are ruled out, and (v) $0 < \Pr(Z = 1 | X) < 1$, ensuring common support in the covariates across instrument states. X (measured prior to randomization) includes education, ethnicity, age and its square, school and working status, and receipt of Aid to Families with Dependent Children (AFDC) and food stamps.

We compare semiparametric LATE estimation based on the latter assumptions (see Theorem 1 in [Frölich and Huber \(2014\)](#)) to (i) MAR-based LATE estimation as in Section 2.3 of [Frölich and Huber \(2014\)](#) (assumptions: $Y \perp R | X, Z, D$, $(U, T) \perp Z | X$, $\Pr(T = c) > 0$, $\Pr(T = d) = 0$, $0 < \Pr(Z = 1 | X) < 1$), (ii) the so-called Wald estimator among those with $R = 1$ which ignores sample selection, and (iii) the method of [Fricke et al. \(2020\)](#), which tackles sample selection and treatment endogeneity by two distinct instruments. In the latter approach, which allows for non-ignorable selection related to U in a more general way than LI, we use the number of kids younger than 6 in the household 2.5 years after random assignment as instrument for R . We apply a semiparametric version of the estimator outlined in Equation (23) of [Fricke et al. \(2020\)](#) along with the weighting function in their expression (21).

Table 1 provides descriptive statistics for the covariates, the treatment, and the instruments in the total sample and for working and not working females. Across the latter groups for instance education, aid receipt, previous job status, and Job Corps participation differ importantly, pointing to non-random selection into employment. In the case that such socioeconomic characteristics also affect the wage outcome, then systematic differences in these variables across the employment states of females generally entail a bias in treatment effect estimation if one does not control for them. Table 2 presents the effect estimates, standard errors, and p -values based on 1999 bootstraps using the quantile method. The effect under LI + MAR (based on Theorem 1 of [Frölich and Huber \(2014\)](#)) of 0.12 log points virtually identical to the Wald estimator which ignores sample selection bias, and both are statistically significantly different from zero. The MAR-based estimate is one third higher, but not significantly differently so. The method of [Fricke et al. \(2020\)](#) based on two instruments (2 IVs) yields virtually the same effect as MAR and is neither statistically significantly different from any other estimator, nor from zero at any conventional level.

It seems important to understand the differences in the behavioral assumptions of the estimators. LI + MAR, for instance, assumes that given the covariates and program assignment, unobservables like ability and motivation do not jointly affect employment and wages among compliers. In contrast, the method of [Fricke et al. \(2020\)](#) does not rely on this restriction and allows for more general forms of sample selection, at the cost of also requiring a valid instrument for employment. In our illustration, the results persistently point to a positive wage effect and are therefore rather robust to the different assumptions

considered. The fact that LI + MAR, MAR, the approach based on two IVs, and even the Wald estimator (which ignores the selection problem) all yield qualitatively similar estimates may give some confidence to our findings, as the latter are not sensitive to the kind of model imposed on the sample selection process. However, such an agreement among different methods controlling for sample selection or outcome attrition need not necessarily occur in other empirical contexts. For this reason, the plausibility of the alternative sets of assumptions needs to be thoroughly scrutinized in the evaluation problem at hand.

Table 1. Descriptive statistics.

	Total Sample		Working		Not Working	
	Mean	Std.Dev	mean	Std.Dev	Mean	Std.Dev
education: 12 years	0.23	0.42	0.25	0.44	0.17	0.37
education: 13 or more years	0.03	0.18	0.04	0.19	0.01	0.10
race: black	0.54	0.50	0.53	0.50	0.56	0.50
race: Hispanic	0.19	0.39	0.18	0.38	0.21	0.40
age	18.59	2.18	18.66	2.19	18.37	2.14
in school prior to randomization	0.63	0.48	0.63	0.48	0.61	0.49
school information missing	0.02	0.14	0.02	0.13	0.03	0.17
in job prior to randomization	0.61	0.49	0.65	0.48	0.47	0.50
received AFDC	0.41	0.49	0.40	0.49	0.45	0.50
received food stamps	0.54	0.50	0.52	0.50	0.60	0.49
treatment: Job Corps participation	0.45	0.50	0.46	0.50	0.41	0.49
instrument: randomization	0.64	0.48	0.66	0.48	0.60	0.49
instrument: kids under 6	0.77	0.90	0.73	0.88	0.88	0.95
instrument kids under 15	1.15	1.26	1.12	1.23	1.25	1.34

Table 2. Effect estimates.

	LI + MAR	MAR	Wald	2 IVs
effect	0.12	0.16	0.12	0.16
standard error	0.06	0.06	0.05	0.33
bootstrap <i>p</i> -values (quantile-based)	0.05	0.00	0.03	0.65

4. Conclusions

The latent ignorability assumption (LI) has been applied in a range of instrumental variable-based evaluations of treatment effects for tackling outcome attrition. By means of standard IV models, we demonstrated that LI is in fact a rather restrictive assumption: it requires the statistical association of outcome attrition and unobservables affecting the outcome to be fully captured by the treatment compliance type, at least conditional on covariates. In other words, the treatment compliance type is assumably a sufficient statistic to control for the endogeneity of the attrition process. This appears disputable in many empirical applications. For this reason, we recommend complementing LI approaches by alternative strategies for modelling outcome attrition whenever feasible in the data at hand, in order to investigate the sensitivity of the treatment effect estimates as we did in our empirical application. Even though the true attrition model is typically unknown in any application, obtaining rather robust results across alternative assumptions on the attrition process (including LI) may increase the confidence in the estimated effects.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin. 1996. Identification of causal effects using instrumental variables. *Journal of American Statistical Association* 91: 444–72. [\[CrossRef\]](#)
- Barnard, John, Constantine E. Frangakis, Jennifer L. Hill, and Donald B. Rubin. 2003. A principal stratification approach to broken randomized experiments: A case study of school choice vouchers in new york city. *Journal of the American Statistical Association* 98: 299–323. [\[CrossRef\]](#)
- Chen, Y., Q. Zhang, Y. Wang, Y. Xiao, R. Fu, H. Bao, and M. Liu. 2015. Estimating the causal effect of milk powder supplementation on bone mineral density: A randomized controlled trial with both non-compliance and loss to follow-up. *European Journal of Clinical Nutrition* 69: 824–30. [\[CrossRef\]](#)
- Esterling, Kevin M., Michael A. Neblo, and David M. J. Lazer. 2011. Estimating treatment effects in the presence of noncompliance and nonresponse: The generalized endogenous treatment model. *Political Analysis* 19: 205–26. [\[CrossRef\]](#)
- Frangakis, Constantine E., and Donald B. Rubin. 1999. Addressing complications of intention-to-treat analysis in the combined presence of all-or-none treatment-noncompliance and subsequent missing outcomes. *Biometrika* 86: 365–79. [\[CrossRef\]](#)
- Fricke, Hans, Markus Frölich, Martin Huber, and Michael Lechner. 2020. Endogeneity and non-response bias in treatment evaluation—Nonparametric identification of causal effects by instruments. *Journal of Applied Econometrics* 35: 481–504. [\[CrossRef\]](#)
- Frölich, Markus, and Martin Huber. 2014. Treatment evaluation with multiple outcome periods under endogeneity and attrition. *Journal of the American Statistical Association* 109: 1697–711. [\[CrossRef\]](#)
- Mattei, Alessandra, and Fabrizia Mealli. 2007. Application of the principal stratification approach to the faenza randomized experiment on breast self-examination. *Biometrics* 63: 437–46. [\[CrossRef\]](#) [\[PubMed\]](#)
- Mealli, Fabrizia, Guido W. Imbens, Salvatore Ferro, and Annibale Biggeri. 2004. Analyzing a randomized trial on breast self-examination with noncompliance and missing outcomes. *Biostatistics* 5: 207–22. [\[CrossRef\]](#) [\[PubMed\]](#)
- O'Malley, A. James, and Sharon-Lise T. Normand. 2005. Likelihood methods for treatment noncompliance and subsequent nonresponse in randomized trials. *Biometrics* 61: 325–34. [\[CrossRef\]](#) [\[PubMed\]](#)
- Rubin, Donald B. 1976. Inference and missing data. *Biometrika* 63: 581–92. [\[CrossRef\]](#)
- Schochet, Peter Z., John Burghardt, and Steven Glazerman. 2001. *National Job Corps Study: The Impacts of Job Corps on Participants' Employment and Related Outcomes*. Washington, DC: Mathematica Policy Research, Inc.
- Yamamoto, Teppei. 2013. *Identification and Estimation of Causal Mediation Effects with Treatment Noncompliance*. Working Paper. Cambridge, MA: MIT Department of Political Science.