Journal Pre-proof

The influence function of semiparametric two-step estimators with estimated control variables

Jinyong Hahn, Zhipeng Liao, Geert Ridder, Ruoyao Shi

PII: S0165-1765(23)00302-6
DOI: https://doi.org/10.1016/j.econlet.2023.111277
Reference: ECOLET 111277

To appear in: *Economics Letters*

Received date : 5 June 2023
Revised date : 21 July 2023
Accepted date : 27 July 2023

The Influence Function of Semiparametric Two-step Estimators with Estimated Control Variables

Jinyong Hahn* UCLA
Zhipeng Liao† UCLA
Geert Ridder‡ USC
Ruoyao Shi§ UC Riverside

July 21, 2023

Highlights

- Semiparametric two-step estimators with estimated control variables are considered.
- Common identification condition is relaxed to an index restriction.
- Influence function contains an extra term.

Abstract

We derive the influence function for two-step estimators with estimated control variables. We adopt a weaker identification assumption than is commonly used in the literature, and as a consequence, an extra term shows up in the influence function.

JEL Classification: C14, C31, C32

Keywords: Control Variable Approach, Index Restriction, Influence Function, Semiparametric Two-step Estimation.

*Department of Economics, UCLA, Los Angeles, CA 90095-1477, USA. Email: hahn@econ.ucla.edu.
†Department of Economics, UCLA, Los Angeles, CA 90095-1477, USA. Email: zhipeng.liao@econ.ucla.edu.
‡Department of Economics, University of Southern California, Los Angeles, CA 90089, USA. Email: ridder@usc.edu.
§Corresponding author. Department of Economics, UC Riverside, 900 University Ave., Riverside, CA 92521, USA. Tel: +1(951)827-1494. Email: ruoyao.shi@ucr.edu.

1 Introduction

When some regressors are endogenous in an econometric model, an attractive identification strategy is to use a moment restriction that conditions on control variables. These control variables typically are estimated in a first step as the residuals in a parametric or nonparametric relation between the endogenous regressors and instruments. In a second step, a conditional expectation of the dependent variable on the endogenous regressors and the control variables can be estimated nonparametrically as in Imbens and Newey (2009), and this conditional expectation can be averaged over the control variables to obtain the Average Structural Function (ASF). In applied work, however, parametric or semiparametric specifications along the line of Rivers and Vuong (1988) or Blundell and Powell (2004) are likely to be adopted, and it is of interest to understand how statistical inference about the estimated finite-dimensional parameters should be implemented.

The purpose of our paper is to develop a unified framework to understand inferential issues arising from such two-step estimation. We are interested in estimating a finite dimensional vector of parameters $\beta_* \in \mathbb{R}^{d_\beta}$, which is identified together with an unknown function $\lambda_*(\cdot)$ as the unique solution of a minimization problem. That is,

$$(\beta_*, \lambda_*) \equiv \arg\min_{\beta,\lambda} \mathbb{E}\left[\psi(Z, \beta, \lambda(v(\pi_*)))\right], \tag{1}$$

where Z is a vector of all observable variables, $v(\pi) \equiv v(Z, \pi)$ is the control variable known up to π_*, and π_* is a finite/infinite-dimensional parameter identified outside the model in a first step. The $v(\cdot)$ is determined by the procedure used to generate the control variable that enters as an argument of $\lambda_*(\cdot)$. The criterion function $\psi(z, \beta, \lambda)$ is known. An estimator $\hat{\beta}$ of β_* can be obtained using the sample analog of (1) with a first-step estimator of π_* and a nonparametric (e.g., the kernel or series) method estimating λ_*.

We make two technical contributions. First, we consider criterion functions $\psi(z, \beta, \lambda)$ that are general enough to nest many specific models as special cases, and provide a unified framework to understand the inferential problems. We follow Newey's (1994) path-derivative calculations to characterize the influence function that takes account of the estimation noise of the control variable, and therefore, our result is invariant to the specific nonparametric estimation method in the second step. Second, we consider moment conditions which are different from those imposed in the previous literature. The previous literature assumed that the "error" in the second step is (mean or quantile) independent of the endogenous regressor given a set of instruments, whereas we impose conditional independence given just the control variable. For example, in a model $Y = X\beta_* + \lambda_*(v) + \varepsilon$ with $X = W^\top \pi_* + v$, a common assumption is $\mathbb{E}[\varepsilon | X, W] = 0$, which is different from $\mathbb{E}[\varepsilon | X, v] = 0$. We adopt variants of the latter assumption when deriving the influence function, and show that this type of assumptions does make a difference in the asymptotic distribution. Although the condition $\mathbb{E}[\varepsilon | X, W] = 0$ is commonly used in early literature, recent applications of the control variable approach directly impose the conditional independence restriction given the control variable to achieve identification (see, e.g., Auerbach (2022) and Johnsson and Moon (2021)). Therefore, our results can be applied to derive the asymptotic distribution of the two-step estimators in these recent works.

The rest of the paper is organized as follows. Section 2 derives the influence function of the semiparametric two-step estimator when π_* is estimated in a parametric first step. Section 3 extends the result in Section 2 to the case where π_* is estimated in a nonparametric first step. Section 4 concludes. Proofs and demonstrative examples of the main results of this paper are included in a supplemental appendix.

2 Two-step Estimation with a Parametric First Step

In this section, we derive the influence function of the semiparametric two-step estimator $\hat{\beta}$ when π_* is parametrically specified. Since the focus is on β_*, we profile out the nonparametric component λ by solving

$$h\left(v\left(\pi\right);\beta,\pi\right)\equiv\arg\min_{\lambda}\mathbb{E}\left[\psi(Z,\beta,\lambda(v(\pi)))\right] \tag{2}$$

for any β and any π. The properties of $h\left(v\left(\pi\right);\beta,\pi\right)$ are characterized by the optimality condition of the above minimization problem, which together with the optimality condition of β_* enables us to derive the influence function.

Assumption 1 (Optimality Condition). *(i) For any β, any π and any square integrable functions $\lambda\left(\cdot\right)$ and $\lambda_1\left(\cdot\right)$ of $v\left(\pi\right)$, there exist functions $\psi_\lambda(\cdot)$, $\psi_\beta(\cdot)$, $\psi_{\beta,\lambda}(\cdot)$, $\psi_{\beta,\beta}(\cdot)$, $\psi_{\lambda,\beta}(\cdot)$ and $\psi_{\lambda,\lambda}(\cdot)$ of z, β and λ such that*

$$\begin{aligned}
\left.\frac{\partial\mathbb{E}\left[\psi(Z,\beta,\lambda\left(v\left(\pi\right)\right)+\tau\lambda_1\left(v\left(\pi\right)\right))\right]}{\partial\tau}\right|_{\tau=0} &= \mathbb{E}\left[\psi_\lambda(Z,\beta,\lambda\left(v\left(\pi\right)\right))\lambda_1\left(v\left(\pi\right)\right)\right],\\
\frac{\partial\mathbb{E}\left[\psi(Z,\beta,\lambda(v(\pi)))\right]}{\partial\beta} &= \mathbb{E}\left[\psi_\beta(Z,\beta,\lambda(v(\pi)))\right],\\
\left.\frac{\partial\mathbb{E}\left[\psi_\beta\left(Z,\beta,\lambda\left(v\left(\pi\right)\right)+\tau\lambda_1\left(v\left(\pi\right)\right)\right)\right]}{\partial\tau}\right|_{\tau=0} &= \mathbb{E}\left[\psi_{\beta,\lambda}\left(Z,\beta,\lambda\left(v\left(\pi\right)\right)\right)\lambda_1\left(v\left(\pi\right)\right)\right],\\
\frac{\partial\mathbb{E}\left[\psi_\beta\left(Z,\beta,\lambda\left(v\left(\pi\right)\right)\right)\right]}{\partial\beta^\top} &= \mathbb{E}\left[\psi_{\beta,\beta}\left(Z,\beta,\lambda\left(v\left(\pi\right)\right)\right)\right],\\
\frac{\partial\mathbb{E}\left[\psi_\lambda\left(Z,\beta,\lambda\left(v\left(\pi\right)\right)\right)\right]}{\partial\beta} &= \mathbb{E}\left[\psi_{\lambda,\beta}\left(Z,\beta,\lambda\left(v\left(\pi\right)\right)\right)\right],\ \textit{and}\\
\left.\frac{\partial\mathbb{E}\left[\psi_\lambda\left(Z,\beta,\lambda\left(v\left(\pi\right)\right)+\tau\lambda_1\left(v\left(\pi\right)\right)\right)\right]}{\partial\tau}\right|_{\tau=0} &= \mathbb{E}\left[\psi_{\lambda,\lambda}\left(Z,\beta,\lambda\left(v\left(\pi\right)\right)\right)\lambda_1\left(v\left(\pi\right)\right)\right];
\end{aligned}$$

(ii) $\mathrm{Var}\left(\psi_\beta\left(Z,\beta_*,\lambda_*(v)\right)\right)$ *and* $\mathrm{Var}\left(\psi_\lambda\left(Z,\beta_*,\lambda_*(v)\right)\right)$ *are non-singular.*

From the optimality of $h\left(v\left(\pi\right);\beta,\pi\right)$, we have

$$\mathbb{E}\left[\psi_\lambda\left(Z,\beta,h\left(v\left(\pi\right);\beta,\pi\right)\right)\lambda\left(v\left(\pi\right)\right)\right]=0 \tag{3}$$

for any function $\lambda(v(\pi))$ of $v(\pi)$ and any β. The profiled version of the minimization problem (1) becomes

$$\min_{\beta} \mathbb{E}\left[\psi(Z, \beta, h(v(\pi_*); \beta, \pi_*))\right]. \tag{4}$$

Therefore, β_* satisfies the following first-order condition

$$\mathbb{E}\left[J(Z, \beta_*, \pi_*)\right] = 0, \tag{5}$$

where

$$J(Z, \beta, \pi) \equiv \psi_\beta(Z, \beta, h(v(\pi); \beta, \pi)) + \psi_\lambda(Z, \beta, h(v(\pi); \beta, \pi)) \frac{\partial h(v(\pi); \beta, \pi)}{\partial \beta}, \tag{6}$$

and the derivative $\partial h(v(\pi); \beta, \pi) / \partial \beta$ exists by Assumption 1 and Assumption 2 below.

The influence function of $\hat{\beta}$ is calculated using the arguments in Newey (1994), which shows that the function $J(Z, \beta, \pi)$ is the key for the calculation, because: *(i)* $J(Z, \beta_*, \pi_*)$ is the score of $\hat{\beta}$ when π_* is known; *(ii)* the impact of estimating π_* on the score function of $\hat{\beta}$ is the derivative $\partial \mathbb{E}\left[J(Z, \beta_*, \pi_*)\right] / \partial \pi^\top$ times the influence function of $\hat{\pi}$; *(iii)* the Hessian matrix of $\hat{\beta}$ is given by $\partial \mathbb{E}\left[J(Z, \beta_*, \pi_*)\right] / \partial \beta^\top$.

For ease of notation, we suppress the dependence of the derivatives of $\psi(z, \beta, \lambda)$ on the parameters when they are evaluated at the true parameter values. Therefore $v \equiv v(\pi_*)$, $\psi_\beta(Z) \equiv \psi_\beta(Z, \beta_*, \lambda_*(v))$, $\psi_\lambda(Z) \equiv \psi_\lambda(Z, \beta_*, \lambda_*(v))$ and the other notations are understood similarly. Define

$$g_*(v) \equiv \frac{\mathbb{E}\left[\psi_{\beta,\lambda}(Z) | v\right]}{\mathbb{E}\left[\psi_{\lambda,\lambda}(Z) | v\right]} \text{ and } \Psi_{\beta,\beta} \equiv -\mathbb{E}\left[\psi_{\beta,\beta}(Z) - \psi_{\lambda,\lambda}(Z)\, g_*(v) g_*(v)^\top\right]. \tag{7}$$

Assumption 2 (Mild Regularity Conditions). *(i) Differentiation under expectation of $\psi_\lambda(\cdot)$ and $\psi_\beta(\cdot)$ is allowed; (ii) the influence function of $\hat{\pi}$ is $\varphi_\pi(z)$; (iii) $\Psi_{\beta,\beta}$ is non-singular; (iv) $\psi_{\lambda,\beta}(Z) = \psi_{\beta,\lambda}(Z)$ almost surely; (v) $h(v(\pi); \beta, \pi)$ is continuously differentiable in β and π, and $v(\pi)$ is continuously differentiable in π for any β and π.*

Theorem 1 (Parametric First Step). *Suppose that Assumptions 1 and 2 hold. Then the influence function of $\hat{\beta}$ is*

$$\Psi_{\beta,\beta}^{-1}\left(\varphi_\beta(Z) + \Psi_{\beta,\pi}\varphi_\pi(Z)\right), \tag{8}$$

where

$$\varphi_\beta(Z) \equiv \psi_\beta(Z) - g_*(v)\psi_\lambda(Z), \tag{9}$$

$$\Psi_{\beta,\pi} \equiv \mathbb{E}\left[\left(\delta_\beta(Z) - \delta_g(Z)\right) \frac{\partial v(\pi_*)}{\partial \pi^\top}\right], \tag{10}$$

$$\delta_\beta(Z) \equiv \left[\psi_{\lambda,\beta}(Z) - g_*(v)\psi_{\lambda,\lambda}(Z)\right] \frac{\partial \lambda_*(v)}{\partial v}, \tag{11}$$

$$\delta_g(Z) \equiv \psi_\lambda(Z) \frac{\partial g_*(v)}{\partial v}. \tag{12}$$

The proof of Theorem 1 is included in the supplemental appendix.

Remark 1 (Index Restriction)**.** *The adjustment in the score function of $\hat{\beta}$ can be simplified under an extra assumption*

$$\mathbb{E}\left[\psi_\lambda(Z)\middle| v(\pi_*), \frac{\partial v(\pi_*)}{\partial \pi^\top}\right] = 0, \tag{13}$$

because in this case,

$$\Psi_{\beta,\pi} = \mathbb{E}\left[\delta_\beta(Z)\frac{\partial v(\pi_*)}{\partial \pi^\top}\right].$$

Condition (13) is further implied by

$$\mathbb{E}\left[\psi_\lambda(Z)\,|X, W\right] = 0 \tag{14}$$

since $v(\pi_*) \equiv v(X, W, \pi_*)$ is a function of X and W. Condition (14) becomes the commonly used identification condition in the literature of the control variable approach. On the other hand, in view of (3) the influence function of $\hat{\beta}$ derived here only uses

$$\mathbb{E}\left[\psi_\lambda(Z)\,|v(\pi_*)\right] = 0 \tag{15}$$

and (5), which is different from (14). Although condition (14) is popular in early literature, recent applications of the control variable approach such as Auerbach (2022) and Johnsson and Moon (2021) use variants of (15), which are imposed on the control variables directly. Under the weaker condition (15), Theorem 1 shows that the extra term $\mathbb{E}\left[\delta_g(Z)\frac{\partial v(\pi_*)}{\partial \pi^\top}\right]$ in the influence function of $\hat{\beta}$ may not be negligible, when assumption (13) does not hold.

3 Two-step Estimation with a Nonparametric First Step

In this section, we extend the influence function formula of $\hat{\beta}$ obtained in the previous section to the case where π_* is nonparametrically specified in the first step. Suppose that there are L functions $\pi_{*,l}(w_l)$ $(l = 1, \ldots, L)$ estimated separately in the first step.

Assumption 3 (Identification of $\pi_{*,l}$)**.** *For each $l = 1, \ldots, L$, $\pi_{*,l}$ is identified by the following conditional moment condition*

$$\mathbb{E}\left[\mu_l(Z_l, \pi_{*,l})|\, W_l\right] = 0,$$

where $\mu_l(z_l, \pi_l)$ is a first step residual function, Z_l is a sub-vector of Z and W_l is a sub-vector (of exogenous variables) of Z_l.

We follow Newey (1994) and consider any one-dimensional path of densities of Z indexed by $\tau \in \mathbb{R}$ such that the path hits the true density at $\tau = 0$. Let $\pi_{*,l,\tau}$ denote the counterpart of $\pi_{*,l}$ under the path τ, i.e., $\pi_{*,l,\tau}$ satisfies

$$\mathbb{E}_\tau\left[\mu_l(Z_l, \pi_{*,l,\tau})\pi_l(W_l)\right] = 0 \tag{16}$$

for any square integrable function $\pi_l(\cdot)$, where $\mathbb{E}_\tau[\cdot]$ denotes the conditional expectation taken under the path density indexed by τ.

Assumption 4. *(i) Suppose that there exists a function $\mu_{l,\pi}(z_l, \pi_l)$ such that*

$$\left.\frac{\partial \mathbb{E}\left[\mu_l(Z_l, \pi_{*,l,\tau})\pi_l(W_l)\right]}{\partial \tau}\right|_{\tau=0} = \left.\mathbb{E}\left[\mu_{l,\pi}(Z_l, \pi_{*,l})\pi_l(W_l)\frac{\partial \pi_{*,l,\tau}(W_l)}{\partial \tau}\right]\right|_{\tau=0}$$

for any square integrable function $\pi_l(\cdot)$ and $l = 1, \ldots, L$; (ii) $v(z, \pi)$ is differentiable in $\pi \equiv (\pi_1, \ldots, \pi_L)^\top$ and it depends on π only through its value $\pi(w)$; (iii) $|\mathbb{E}[\mu_{l,\pi}(Z_l, \pi_{*,l})| W_l]| > 0$ almost surely for $l = 1, \ldots, L$.

Assumption 4 is mainly used to derive the effect of the first step nonparametric estimator on the influence function of $\hat{\beta}$. Condition (ii) imposes smoothness on the control variable in terms of π. Condition (iii) is a local identification condition of the unknown parameters $\pi_{*,l}$ $(l = 1, \ldots, L)$.

Theorem 2 (Nonparametric First Step)**.** *Suppose that Assumptions 1, 2(i, iii, iv, v), 3 and 4 hold. Then the influence function of $\hat{\beta}$ is*

$$\Psi_{\beta,\beta}^{-1}\left(\varphi_\beta(Z) + \delta_\pi(W)^\top \varphi_\pi(Z)\right),$$

where $\delta_\pi(W) \equiv (\delta_{1,\pi}(W_l), \ldots, \delta_{L,\pi}(W_L))^\top$,

$$\delta_{l,\pi}(W_l) \equiv \mathbb{E}\left[\left[\delta_\beta(Z) - \delta_g(Z)\right]\frac{\partial v(\pi_*)}{\partial \pi_l}\right| W_l\Bigg], \text{ and} \tag{17}$$

$$\varphi_\pi(Z) \equiv -\left(\frac{\mu_1(Z_1, \pi_{*,1})}{\mathbb{E}\left[\mu_{1,\pi}(Z_1, \pi_{*,1})| W_1\right]}, \cdots, \frac{\mu_L(Z_L, \pi_{*,L})}{\mathbb{E}\left[\mu_{L,\pi}(Z_L, \pi_{*,L})| W_L\right]}\right)^\top, \tag{18}$$

$\Psi_{\beta,\beta}$, $\varphi_\beta(\cdot)$, $\delta_\beta(\cdot)$ and $\delta_g(\cdot)$ are defined in (7), (9), (11) and (12), respectively.

The proof of Theorem 2 is included in the supplemental appendix.

4 Conclusion

In this paper, we derive the influence function of semiparametric two-step estimators where an unknown function/control variable is estimated in a first step, which can be parametric, semiparametric or fully nonparametric. The influence function is derived under an index restriction that is different from the common identification condition in the literature. The supplemental appendix provides illustrative examples of the influence function formula.

References

Auerbach, Eric (2022): "Identification and estimation of a partially linear regression model using network data," Econometrica, 90 (1), 347–365.

Blundell, Richard W and James L Powell (2004): "Endogeneity in semiparametric binary response models," The Review of Economic Studies, 71 (3), 655–679.

Imbens, Guido W and Whitney K Newey (2009): "Identification and estimation of triangular simultaneous equations models without additivity," Econometrica, 77 (5), 1481–1512.

Johnsson, Ida and Hyungsik Roger Moon (2021): "Estimation of peer effects in endogenous social networks: Control function approach," Review of Economics and Statistics, 103 (2), 328–345.

Newey, Whitney K (1994): "The asymptotic variance of semiparametric estimators," Econometrica, 62 (6), 1349–1382.

Rivers, Douglas and Quang H Vuong (1988): "Limited information estimators and exogeneity tests for simultaneous probit models," Journal of Econometrics, 39 (3), 347–366.

Economics Letters MS#EL57080

Highlights

* Semiparametric two-step estimators with estimated control variables are considered.
* Common identification condition is relaxed to an index restriction.
* Influence function contains an extra term.