Taylor & Francis
Taylor & Francis Group

Check for updates

Comment: Individualized Treatment Rules Under Endogeneity

Sukjin Han

Department of Economics, University of Bristol, Bristol, UK

1. Heterogeneity, Individualized Treatments, and Endogeneity

For the past few decades, scholars in statistical and social sciences have acknowledged that heterogeneity is prevalent in the subjects of their studies. The literature on individualized treatments is a great example that embraces heterogeneity and develops fruitful policies that benefit members of targeted populations. I congratulate Yifan Cui, Eric Tchetgen Tchetgen, Hongxiang Qiu, Marco Carone, Ekaterina Sadikova, Maria Petukhova, Ronald C. Kessler, and Alex Luedtke for successfully pursuing this literature's endeavor. Their works—Cui and Tchetgen Tchetgen (2020) and Qiu et al. (2020)—especially stand out as they relax the assumption of "no unmeasured confounders" that this entire literature has relied on. In experimental and observational studies, there are ample examples where endogeneity in treatment decisions cannot be solely captured by observables, that is, measured confounders. In a sense, allowing for treatment endogeneity becomes relevant in a world of heterogeneity, because typically, unobserved heterogeneity is a core factor of individuals decisions to receive treatments. Therefore, in my view, considering endogeneity in developing individualized treatments is quite natural.

I appreciate what Cui and Tchetgen Tchetgen (2020) and Qiu et al. (2020) have achieved. Both use instrumental variable (IV) methods, but each of them makes unique contributions. Under a set of identifying conditions, Cui and Tchetgen Tchetgen (2020) derive a simple, closed-form expression for the counterfactual mean based on which the (non-stochastic) optimal regime is identified. The most notably result is Theorem 2.2, in which the optimal regime is identified even without observing the endogenous treatment. This is especially relevant in experimental settings where partial compliance is suspected but the information about treatment decisions is missing, for example, due to the cost of fidelity assessments. For example, the experimental dataset in Murphy et al. (2001) has information from fidelity checks, but only aggregated information (i.e., the compliance rate) is available. Qiu et al. (2020) consider regimes that are allowed to be stochastic, which is pertinent to the existence of budget constraints. They target not only optimal treatments but also optimal encouragements, which is a sensible idea given the use of the IV framework. With these notable features, they proceed by employing the main

identifying conditions similar to Cui and Tchetgen Tchetgen (2020). I appreciate that they develop the asymptotic distributions for the estimated gains from optimal regimes, which enables inference.

I read the two papers with great pleasure. In this note, I discuss the main identifying conditions commonly employed in the papers. Inspired by their analyses, I propose alternative identifying conditions. Then, motivated from these discussions, I turn the focus from the point identification to partial identification. Along the way, I mention two closely related and complementing papers on optimal dynamic treatment regimes in longitudinal settings written by myself. I conclude by making remarks on possible future directions for research.

2. About Main Identifying Conditions

In the presence of heterogeneity and endogeneity, identifying the average treatment effects (ATEs), and thus optimal treatment regimes, is a challenging task. To overcome this challenge, Cui and Tchetgen Tchetgen (2020) and Qiu et al. (2020) utilize identifying conditions (Assumptions 7 and 8 in the former, Conditions A5b-1(b) and 2(b) in the latter) that are introduced in Wang and Tchetgen Tchetgen (2018). Since the identification, estimation, and inference of the two papers rely on these conditions, I would like to discuss them in depth here. In the next section, I propose alternative identifying conditions.

For convenience, I follow the notation and labels of Cui and Tchetgen Tchetgen (2020); in the parenthesis, I indicate the labels of Qiu et al. (2020). I attempt to interpret Assumptions 7 and 8 (Conditions A5b-1(b) and 2(b)) and understand their implications, starting from Assumption 8 (Condition A5b-2(b)). To facilitate the discussion, we maintain Assumption 10 (Condition A6a) that $Z \perp (A_z, Y_a)|L$ (causal IV). Then, Assumption 8, which the authors call "independent compliance type," can be expressed as

$$E[A_1 - A_{-1}|L, U] = E[A_1 - A_{-1}|L].$$

This expression reveals that the compliance type is defined by the response to IV, that is, $A_1 - A_{-1}$, and that the assumption concerns conditional mean independence between $A_1 - A_{-1}$ and U. Since $A_1 - A_{-1}$ is not binary, $A_1 - A_{-1}$ and U can be associated through higher order moments. Nonetheless, $A_1 - A_{-1}$ takes only three values ($A_1 - A_{-1} \in \{-1, 0, 1\}$), and thus

the mean independence would still significantly restrict the joint distribution, yielding a specific form of treatment endogeneity. At least to an econometrician like myself, this is an unfortunate feature. The dependence of $A_1 - A_{-1}$ and U is the source of what the econometrics literature calls "compliance heterogeneity." For example, the literature on the local ATE (LATE) and the marginal treatment effect (MTE) essentially builds on this notion. I believe this issue can be partly mitigated by treating L to be high-dimensional. In fact, both papers' estimation methods allow for high-dimensional explanatory covariates.[1]

Cui and Tchetgen Tchetgen (2020) do relax Assumption 8 by introducing Assumption 7. As noted in their paper, Assumption 8 implies Assumption 7. Also, Condition A5b-1(b) in Qiu et al. (2020) implies Assumption 7. Condition A5b-1(b) posits that $E[Y_1 - Y_{-1}|L, U] = E[Y_1 - Y_{-1}|L]$, that is, the treatment effect $Y_1 - Y_{-1}$ and the confounder U are conditionally mean independent. This limits the treatment effect heterogeneity. We definitely want to avoid the additive structure $Y_a = g_a(L) + U$, because it significantly restricts *unobserved* heterogeneity. Also, with binary Y as in the empirical examples of both papers, it is more natural to assume a nonseparable structure of binary choice models.

Therefore, I would like to investigate Assumption 7 as it is. Again, under the causal IV assumption, Assumption 7 can be expressed as

$$\text{cov}\{E[Y_1 - Y_{-1}|L, U], E[A_1 - A_{-1}|L, U]|L\} = 0.$$

Let us consider a more generic problem after suppressing L: for some generic scalar-valued functions δ and γ of U, consider the assertion that

$$\text{cov}(\delta(U), \gamma(U)) = 0.$$

This assertion requires a specific relationship between the distribution of U and the shape of conditional mean functions $\delta(\cdot)$ and $\gamma(\cdot)$. This is most clearly seen in the following simple example. Suppose $U \in \{u_1, u_2\}$ is binary (e.g., high and low unobserved health types) with $\Pr[U = u_1] = p_1 > 0$. Then, it should be that either δ or γ is a constant function. To see this, let $\delta_1 \equiv \delta(u_1)$, $\delta_2 \equiv \delta(u_2)$, $\gamma_1 \equiv \gamma(u_1)$, and $\gamma_2 \equiv \gamma(u_2)$. Then,

$$\begin{aligned}
&\text{cov}(\delta(U), \gamma(U)) \\
&= p_1\delta_1\gamma_1 + (1 - p_1)\delta_2\gamma_2 - \{p_1\delta_1 + (1 - p_1)\delta_2\} \\
&\quad \{p_1\gamma_2 + (1 - p_1)\gamma_2\} \\
&= p_1(1 - p_1)\delta_1(\gamma_1 - \gamma_2) - p_1(1 - p_1)\delta_2(\gamma_1 - \gamma_2).
\end{aligned}$$

Therefore, $\text{cov}(\delta(U), \gamma(U)) = 0$ if and only if $\delta_1(\gamma_1 - \gamma_2) = \delta_2(\gamma_1 - \gamma_2)$, or equivalently, $\delta_1 = \delta_2$ or $\gamma_1 = \gamma_2$, which proves the claim. This result shows that when U is distributed as Bernoulli, Assumption 7 is equivalent to Assumption 8 or Condition A5b-1(b). More generally, $\delta(\cdot)$ and $\gamma(\cdot)$ need to follow specific shapes given the particular distribution of U. It would be worth investigating how sensitive the estimation and inference results of Cui and Tchetgen Tchetgen

(2020) and Qiu et al. (2020) are when these identifying conditions fail or nearly fail. For example, one can ask how much suboptimal the estimated regime would become under the misspecification.

3. Alternative Identifying Condition for Optimal Regimes

Inspired by the identification analyses of Cui and Tchetgen Tchetgen (2020) and Qiu et al. (2020), I would like to propose a simple identifying condition for optimal dynamic regimes. This condition can be used as an alternative to Assumptions 7 or 8 in Cui and Tchetgen Tchetgen (2020) and Conditions A5b-1(b) or 2(b) in Qiu et al. (2020). It can also be an alternative to the identifying conditions in Han (2020), which proposes the use of extra exogenous variables besides binary IVs to identify the dynamic treatment effects and optimal sequential treatment regimes in longitudinal settings.

To state the condition, I follow the notation in Cui and Tchetgen Tchetgen (2020).

Assumption A. The following two conditions hold for every $L = l$: (a) either $E[Y_1|L, U] \geq E[Y_{-1}|L, U]$ a.s. $[F_{U|L}]$ or $E[Y_1|L, U] \leq E[Y_{-1}|L, U]$ a.s. $[F_{U|L}]$; (b) either $E[A_1|L, U] \geq E[A_{-1}|L, U]$ a.s. $[F_{U|L}]$ or $E[A_1|L, U] \leq E[A_{-1}|L, U]$ a.s. $[F_{U|L}]$.

Assumption A(a) posits that, conditional on L, the sign of the ATE is maintained across individuals defined by unobserved type U, that is, $\text{sign}\{E[Y_1 - Y_{-1}|L, U]\} = \text{sign}\{E[Y_1 - Y_{-1}|L]\}$ a.s. $[F_{U|L}]$. For example, Cui and Tchetgen Tchetgen (2020) mention in their empirical example (Section 5) that "one might expect that having a third child would generally reduce a mother's labor participation even if the effects are heterogeneous." Assumption A(a) can be consistent with this story. This assumption even allows the following: the effect of a third child on a mother's labor participation to be positive for less educated women ($L \leq l_0$) regardless of their level of aspiration (U), and be negative for highly educated women ($L > l_0$) regardless of their U.

A similar interpretation can be made for Assumption A(b) in terms of the effect of the IV on the treatment decision. This part is weaker than the LATE monotonicity assumption that $\Pr(A_1 \geq A_{-1}) = 1$ (which also appears as Assumption 9 in Cui and Tchetgen Tchetgen (2020)). This is because when $A_1 \geq A_{-1}$ a.s. then $E[A_1|L, U] \geq E[A_{-1}|L, U]$ a.s. $[F_{U|L}]$, but the converse is not necessarily true.

Theorem 3.1. Under Assumptions 2–6 (Cui and Tchetgen Tchetgen 2020) and Assumption A, $\arg\max_{\mathcal{D}} E[Y_{\mathcal{D}(L)}]$ is nonparametrically identified by

$$\arg\max_{\mathcal{D}} E[Y_{\mathcal{D}(L)}] = \arg\max_{\mathcal{D}} E\left[\frac{ZI\{A = \mathcal{D}(L)\}YA}{\delta(L)f(Z|L)}\right]. \quad (1)$$

Furthermore,

$$\arg\max_{\mathcal{D}} E[Y_{\mathcal{D}(L)}] = \arg\max_{\mathcal{D}} E\left[\frac{I\{Z = \mathcal{D}(L)\}Y}{\delta(L)f(Z|L)}\right]. \quad (2)$$

---

[1] It is worth noting that even under Assumption 8, A_z itself does not have to be mean independent of U. However, U seems hard to be cancelled out in $E[A_1|L, U] - E[A_{-1}|L, U]$ when A_z has structure of a binary choice model.

Proof. To prove (1), by the first and second derivations in the proof of Theorem 2.1 in Cui and Tchetgen Tchetgen (2020),

$$E\left[\frac{ZI\{A=\mathcal{D}(L)\}YA}{\delta(L)f(Z|L)}\right] = E\left[\tilde{\gamma}(L,U)\tilde{\delta}(L,U)\frac{I\{\mathcal{D}(L)=1\}}{\delta(L)}\right]$$

$$+ E\left[\kappa(L,U)\right], \qquad (3)$$

where $\kappa(L,U) \equiv \tilde{\delta}(L,U)\frac{E[Y_{-1}|L,U]}{\delta(L)}$, $\tilde{\delta}(L,U) \equiv \Pr(A=1|Z=1,L,U) - \Pr(A=1|Z=-1,L,U)$, and $\delta(L) \equiv \Pr(A=1|Z=1,L) - \Pr(A=1|Z=-1,L)$. Recall that, with $\Delta(L) \equiv E[Y_1 - Y_{-1}|L]$,

$$\arg\max_{\mathcal{D}} E[Y_{\mathcal{D}(L)}] = \arg\max_{\mathcal{D}} E\left[\Delta(L)I\{\mathcal{D}(L)=1\}\right]. \quad (4)$$

We prove that maximizing the right side of (3) is equivalent to maximizing $E\left[\Delta(L)I\{\mathcal{D}(L)=1\}\right]$ under Assumption A. Without loss of generality, assume $\delta(L) > 0$. Under Assumption A, for any given l,

$$\text{sign}\{\tilde{\gamma}(l,U)\} = \text{sign}\{\Delta(l)\} \text{ a.s.}$$

and

$$\text{sign}\{\tilde{\delta}(l,U)\} = \text{sign}\{\delta(l)\} = 1 \text{ a.s.}$$

Therefore,

$$\text{sign}\{E[\tilde{\gamma}(L,U)\tilde{\delta}(L,U)/\delta(L)|L=l]\} = \text{sign}\{\Delta(l)\}. \quad (5)$$

Note that choosing $\mathcal{D}^*$ that maximizes $E[\Delta(L)I\{\mathcal{D}(L)=1\}]$ in (4) is equivalent to choosing $\mathcal{D}^*$ such that $\mathcal{D}^*(L) = 1$ whenever $\Delta(L) > 0$ and $\mathcal{D}^*(L) = -1$ otherwise. But by (5), the latter is equivalent to choosing $\mathcal{D}^*$ such that $\mathcal{D}^*(L) = 1$ whenever $E[\tilde{\gamma}(L,U)\tilde{\delta}(L,U)/\delta(L)|L] > 0$ and $\mathcal{D}^*(L) = -1$ otherwise. Finally, this is equivalent to choosing $\mathcal{D}^*$ that maximizes

$$E\left[E[\tilde{\gamma}(L,U)\tilde{\delta}(L,U)/\delta(L)|L]I\{\mathcal{D}(L)=1\}\right]$$

$$= E\left[\tilde{\gamma}(L,U)\tilde{\delta}(L,U)\frac{I\{\mathcal{D}(L)=1\}}{\delta(L)}\right],$$

which completes the proof as $E\left[\kappa(L,U)\right]$ in (3) does not depend on $\mathcal{D}$.

To prove (2), by the first derivation in the proof of Theorem 2.2 in Cui and Tchetgen Tchetgen (2020) (except the last equality),

$$E\left[\frac{I\{Z=\mathcal{D}(L)\}Y}{\delta(L)f(Z|L)}\right] = E\left[\tilde{\gamma}(L,U)\tilde{\delta}(L,U)\frac{I\{\mathcal{D}(L)=1\}}{\delta(L)}\right]$$

$$+ E\left[\kappa(L,U)\right], \qquad (6)$$

where $\kappa(L,U)$ is defined in Cui and Tchetgen Tchetgen (2020) and does not depend on $\mathcal{D}$. Then, by the same argument as in the previous case, we have the desired result. $\square$

Although Assumption A is helpful to identify the optimal dynamic regime, it cannot be directly used to identify the value function, whereas Assumption 8 in Cui and Tchetgen Tchetgen (2020) and Conditions A5b-1(b) or 2(b) in Qiu et al. (2020) are powerful enough to identify it.[2]

4. Partial Identification Approach

Although the point identification approach permits powerful inference on treatment effects and optimal regimes, as is shown in Cui and Tchetgen Tchetgen (2020) and Qiu et al. (2020), it has to rely on some versions of extrapolative assumptions. This is inevitable, because when the instrument is binary, it is known to have no identifying power for general non-compliers, for example, always-takers and never-takers, but only identifies the effect for compliers as the LATE. Therefore, to identify the ATE, which is the effect for the entire population that includes always-takers and never-takers, and which is a relevant parameter to recover the optimal regime, it requires means of extrapolation. Broadly speaking, the assumptions used in Cui and Tchetgen Tchetgen (2020), Qiu et al. (2020), and Han (2020) operate in a way that extrapolates the LATE to the effects for different subpopulations.

Extrapolation inevitably demands prior beliefs from the researcher or policy maker. Sometimes, one may be interested in knowing how much she can learn in the absence of extrapolative assumptions and how sensitive her point identification results are to these assumptions. This view is shared by Qiu et al. (2020) who mention that "[i]n future work, it would be interesting to develop a framework for sensitivity analyses for Condition A5b." In this regard, partial identification can be a fruitful approach. Cui and Tchetgen Tchetgen (2020) briefly touch upon this approach. In the Appendix, they derive bounds on the average potential outcome $E[Y_{\mathcal{D}(L)}]$ for the case of binary Y using Balke and Pearl's (1997) bounds and find regimes that maximize corresponding lower bounds. What could have been done further is a related sensitivity analysis, for example, how the estimated regime under Assumption 7 is consistent with the estimated regime under their bound analysis.

The partial identification approach to learn optimal dynamic regimes and welfare is also considered in Han (2019). In this article, I characterize the identified set of the optimal dynamic sequential regime that maximizes a general form of welfare that nests the average potential outcome. From data that are generated from multi-period settings where unmeasured confounders exist, I use an IV method to establish the sharp partial ordering of welfares (in terms of possible regimes), the identified set of the optimal dynamic regime, and sharp bounds on the optimized welfare (i.e., the value function). I also propose a wide range of identifying assumptions that can tighten the results and can be easily incorporated within the paper's framework.

Although the partial identification approach may not deliver informative recommendations for optimal regimes, I believe that it still has value for policy making for at least three reasons. First, as mentioned above, one may conduct sensitivity analyses for priors used in point and partial identification. Second, even with uninformative results, suboptimal regimes can be easily detected and removed from the policy menu. Third, the lack of informativeness may guide the researcher and the policy

---

[2] Assumption A may still have identifying power for the value function if one is willing to take a partial identification approach.

maker toward better data collection, so that the informativeness in policy suggestions is not driven from potentially arbitrary assumptions but from more informative data.

5. Concluding Remarks

It is exciting to see works like Cui and Tchetgen Tchetgen (2020) and Qiu et al. (2020) in the literature on individualized treatments that was pioneered by Murphy et al. (2001), Murphy (2003), and Robins (2004) and has grown rapidly since then. The literature seems to have entered a new territory as it begins to concern the problem of treatment endogeneity. Many interesting questions can emerge along the way. I would like to list a few here. First, it would be interesting to consider the issues of policy invariance in the context of individualized treatments. For example, it is well known that different IVs induce different individuals to respond to the policy. Qiu et al. (2020) acknowledge this point when they define the local average encouragement effect for compliers. Indeed, the definition of this local parameter raises the question of who the compliers are, especially because the parameter is defined by contrasting two different IVs. More generally, it would be important to define an appropriate—that is, policy invariant—target population relevant to the optimized policy, which is inevitably defined by comparing multiple hypothetical policies. Second, as a way of overcoming the identification challenge under endogeneity, it would be interesting to continue developing frameworks that are designed to incorporate rich data structures, for example, panels, multiple IVs, and high-dimensional covariates. Third, related to the second question, extending the online approach (e.g., reinforcement learning (Shortreed et al. 2011) to the context of treatment endogeneity and noncompliance would be interesting for future work. Finally, the theory of estimation and inference is worth further exploration. For example, for their asymptotic theory, Qiu et al. (2020) introduce a convergence rate requirement for the estimated optimal regime that does not seem innocuous. It would be interesting to investigate whether it is possible to develop asymptotic theory for the treatment and encouragement effects directly from the optimization problems (Equations (2) and (3) in their paper), which may bypass the intermediate estimation of the optimizers (the optimal regimes).

References

Balke, A., and Pearl, J. (1997), "Bounds on Treatment Effects From Studies With Imperfect Compliance," *Journal of the American Statistical Association*, 92, 1171–1176. [194]

Cui, Y., and Tchetgen Tchetgen, E. (2020), "A Semiparametric Instrumental Variable Approach to Optimal Treatment Regimes Under Endogeneity," *Journal of the American Statistical Association* (just-accepted). [192,193,194,195]

Han, S. (2019), "Optimal Dynamic Treatment Regimes and Partial Welfare Ordering," arXiv no. 1912.10014. [194]

——— (2020), "Optimal Dynamic Treatment Regimes and Partial Welfare Ordering," arXiv no. 1912.10014. [193,194]

Murphy, S. A. (2003), "Optimal Dynamic Treatment Regimes," *Journal of the Royal Statistical Society*, Series B, 65, 331–355. [195]

Murphy, S. A., van der Laan, M. J., Robins, J. M., and Conduct Problems Prevention Research Group (2001), "Marginal Mean Models for Dynamic Regimes," *Journal of the American Statistical Association*, 96, 1410–1423. [192,195]

Qiu, H., Carone, M., Sadikova, E., Petukhova, M., Kessler, R. C., and Luedtke, A. (2020), "Optimal Individualized Decision Rules Using Instrumental Variable Methods," *Journal of the American Statistical Association* (just-accepted). [192,193,194,195]

Robins, J. M. (2004), "Optimal Structural Nested Models for Optimal Sequential Decisions," in *Proceedings of the Second Seattle Symposium in Biostatistics*, Springer, pp. 189–326. [195]

Shortreed, S. M., Laber, E., Lizotte, D. J., Stroup, T. S., Pineau, J., and Murphy, S. A. (2011), "Informing Sequential Clinical Decision-Making Through Reinforcement Learning: An Empirical Study," *Machine Learning*, 84, 109–136. [195]

Wang, L., and Tchetgen Tchetgen, E. (2018), "Bounded, Efficient and Multiply Robust Estimation of Average Treatment Effects Using Instrumental Variables," *Journal of the Royal Statistical Society*, Series B, 80, 531. [192]