Published in final edited form as:

Health Econ. 2019 May ; 28(5): 710–716. doi:10.1002/hec.3861.

Use of instrumental variables for endogenous treatment at the provider level

R. Tamara Konetzka^{1,2}, Fan Yang³, and Rachel M. Werner⁴

¹Department of Public Health Sciences, The University of Chicago, Chicago, Illinois

²Department of Medicine, The University of Chicago, Chicago, Illinois

³Department of Biostatistics and Informatics, Colorado School of Public Health, University of Colorado Denver, Aurora, Colorado

⁴Division of General Internal Medicine, University of Pennsylvania, Philadelphia, Pennsylvania

Abstract

Health economists are often interested in the effects of provider-level attributes (e.g., nonprofit status or quality rating) on patient outcomes, but estimation is subject to selection bias due to correlation with other omitted provider-level attributes that also affect patient outcomes. Recently, researchers have attempted to use patient-level instrumental variables, such as differential distance, to solve this problem of a provider-level endogenous treatment variable in settings where patients are nested within providers. However, to satisfy validity assumptions, an instrumental variable for a provider attribute must be at the provider level or a larger unit of aggregation, not at the patient level. A patient-level instrument cannot predict variation in a provider attribute separately from other, potentially unmeasured, provider attributes. In this paper, we explain this misapplication, review the extent of this problem in recent literature, and offer alternative approaches to avoid this misapplication of patient-level instrumental variables.

Keywords

instrumental variables; endogeneity

1 | INTRODUCTION

The instrumental variables (IV) approach has become commonplace in health economics and health services research. The popularity of the approach is rooted in its status as one of the few methods outside of randomized designs that can create treatment and control groups that are balanced across both observable and unobservable characteristics. When randomization is ethically or practically infeasible, the IV approach can provide a powerful alternative.

Limitations of the IV approach, even when properly implemented, have been debated (Garabedian, Chu, Toh, Zaslavsky, & Soumerai, 2014; Martens, Pestman, de Boer, Belitser, & Klungel, 2006). Our intent is not to revisit those limitations but to draw attention to a specific, fundamental yet increasingly common misapplication of the method: using an individual-level instrument to attempt to address an organization-level endogenous treatment variable when individuals are nested within organizations. Our primary contribution in this paper is to highlight this IV mismatch and review the extent of this problem in recent literature. We end with suggestions for addressing this issue.

2 | IV METHODS

IV methods are designed to address the problem of an endogenous treatment variable. Suppose that one is interested in estimating the effect of a patient-level treatment T on an outcome Y , controlling for patient i 's characteristics X :

$$Y_i = \alpha + \beta_1 T_i + \beta_2 X_i + \varepsilon_i \quad (1)$$

If T is randomly assigned, β_1 can be estimated consistently. However, if T is not randomly assigned, omitted variables may exist that are correlated with both T and Y , causing T to be correlated with the error term and precluding an unbiased estimate of β_1 . However, if a valid instrument W exists, it can be used to isolate exogenous variation in T , effectively “pseudo-randomizing” T and enabling unbiased estimation of β_1 .¹

The utility of the IV approach depends on the validity of the instrument. A valid IV must satisfy three key conditions: It must (a) be correlated with the potentially endogenous treatment variable; (b) have no direct effect on the outcome (other than through its impact on the treatment); and (c) be independent of the unmeasured confounders of the treatment–outcome relationship after conditioning on the measured confounders. The second and third conditions together are often referred to as the “exclusion restriction,” that is, the instrument can be validly excluded from the main estimation of interest.

When the treatment and the instrument are at the same level of aggregation (e.g., both at the patient level), addressing the endogeneity problem is relatively simple. In the well-known example of the effects of cardiac catheterization on patient mortality after a heart attack, differential distance, defined as the additional distance a patient would have to travel to be treated in a hospital offering cardiac catheterization versus the closest hospital, was used as an arguably valid instrument to predict receipt of cardiac catheterization (McClellan, McNeil, & Newhouse, 1994). Importantly, the endogenous treatment variable is at the patient level, defined as whether a patient gets catheterized, not whether the patient goes to a hospital that offers cardiac catheterization.

¹For a more comprehensive theoretical discussion of instrumental variables, see Wooldridge (2002). *Econometric analysis of cross section and panel data*. Cambridge, Massachusetts: The MIT Press.

3 | ENDOGENEITY AT MULTIPLE LEVELS

In many applications, the “treatment” of interest is an attribute of a provider or an organization and does not vary for patients within a provider—for example, a hospital’s profit status or its performance on a report card. These provider-level attributes are rarely randomly assigned, so estimating their effect on patient outcomes is subject to selection bias. However, unlike an endogenous patient-level treatment, there is potential for selection bias at two different levels: the hospital level (h) and the patient level (i).

$$Y_{ih} = \pi + d_1 T_h + d_2 X_i + d_3 Z_h + \eta_{ih} \quad (2)$$

In Equation (2), selection bias will exist if there exists omitted patient characteristics *or* omitted hospital characteristics that are correlated with treatment T_h . Either way, d_1 will be biased. In other words, condition (c) above—that the instrument must be independent of the unmeasured confounders of the treatment–outcome relationship after conditioning on the measured confounders—needs to be satisfied at both patient and organization levels.

Take the example of estimating the effect of for-profit status of a hospital on patient outcomes. Patients treated at for-profit hospitals may be different in unmeasured ways from patients treated at other hospitals—in preferences for preventive care, level of social support, or health status. To solve this problem, an instrument W_i can be used. A logical choice may be the differential distance instrument, using the additional distance a patient would have to travel to reach a for-profit hospital versus the nearest hospital. The use of this instrument could result in groups of patients that are balanced on observed and unobserved characteristics, minimizing confounding by patient selection.

However, the organizational selection problem remains: Hospitals that choose to be for-profit are likely to be different in unobserved ways from hospitals that do not, perhaps in resources, mission, or management—factors that may be correlated with for-profit status and also patient outcomes. A logical choice to address this bias, again, may be IV. An instrument W_h that exogenously pseudo-randomizes hospitals to be for-profit—say, an instrument based on what competitor hospitals in the market are doing, or a change in laws about for-profit status—could solve the problem, creating treatment and comparison groups of hospitals that differ only in for-profit status.

It is a mistake, however, to use a patient-level instrument W_i to attempt to solve the endogeneity of an organization-level treatment variable or to imply that it does. If the goal of the analysis is to estimate a causal effect of for-profit status, an erroneous approach might be to use patient-level differential distance as an IV. Although this would likely result in treatment and comparison groups balanced on patient-level characteristics, it does nothing to solve the critical provider selection issue. When the unmeasured hospital-level confounders are not independent of the proposed patient-level instrument, condition (c) will be violated. Put another way, the patient-level instrument does not create exogenous variation in for-profit status separate from other hospital attributes.

This issue is well established in the literature on quasi-experimental design. In Cook and Campbell's classic text on quasi-experimental design and their advice on the level of randomization, "The research question also determines at which level of aggregation units should be randomized ... The unit of randomization must reflect the aggregate-level question and cannot be at a level of aggregation less than itself" (Cook & Campbell, 1979). The same applies to pseudo-randomization using IV.

4 | EVIDENCE FROM THE LITERATURE

To assess the extent of this problem, we conducted a literature search of major health economics and health services research journals² from 2010 through June 2018³ using the search term "instrumental variable." Of 99 papers that used IV, 10 used an individual-level instrument (in all cases derived from distance) for an organization-level endogenous treatment variable. These papers are listed in Table 1 with selected details of the approach for each.⁴

The papers varied in the extent to which the inability to address organization-level endogeneity with the chosen instrument was acknowledged or partially remedied. Most papers framed the conceptual problem as one of provider behavior and discussed the policy implication in those terms, thus implying a causal effect of a provider attribute on patient outcomes. However, the majority of the 10 papers empirically addressed only patient selection associated with a provider attribute, ignoring the provider selection issue. A few of the papers indirectly acknowledged the provider selection issue by discussing and controlling for selected facility-level confounders (Joyce, McGuire, Bartels, Mitchell, & Grabowski, 2018; Lee, Chertow, & Zenios, 2010). Finally, a paper studying the effects of a hospital having a catheterization lab argued that cardiac interventions are multifaceted and that allowing the catheterization treatment to be confounded by other hospital-level correlates is an advantage (Sanwald & Schober, 2017) but ignored that some hospital-level correlates could be considered part of the intervention (like the skill of cardiac surgeons), whereas others would not (such as hospital payer mix).

Only two papers attempted to address endogeneity due to unmeasured facility-level attributes. Rahman, Foster, Grabowski, Zinn, and Mor (2013), in studying the effect of the concentration of hospital referrals to skilled nursing facilities, included both hospital and nursing facility fixed effects in some specifications, thereby identifying the effects of referral concentration through changes over time. In the most thorough example, Joyce et al. (2018) conducted robustness checks using a dose-response strategy (the percent of patients in a dementia Special Care Unit) and comparisons to patients without dementia to isolate the effect of Special Care Units from unmeasured facility quality.

²Our search included the following journals: *Health Economics*, *Journal of Health Economics*, *Health Services Research*, *Medical Care*, and *Medical Care Research and Review*

³We began the search with 2010 because this misapplication seems a recent phenomenon; indeed, we found just one relevant paper in 2010–2012 with the majority published in 2013 or later.

⁴Sometimes, a research question can reasonably be framed as a question about where individuals get care and not about the causal effect of the organization-level attribute. In these cases, if no attempts were made to imply a causal effect of the organization-level attribute, we excluded the paper before arriving at our total.

The consequences of using a patient-level IV in cases of endogenous provider-level treatment can be serious. The estimated treatment effect is subject to omitted provider-level variables bias, which renders it untrustworthy, just as if no IV had been used. In other words, the estimated effect may be incorrect in direction, magnitude, and/or significance. To understand the extent to which this omission matters, we found several examples of papers seeking to answer similar questions to those in Table 1, while addressing provider-level endogeneity. In a paper on profit status of nursing homes (Grabowski, Feng, Hirth, Rahman, & Mor, 2013) in which nonprofit homes were found to provide higher quality using the mismatched instrument, the authors note that a prior paper, employing a longitudinal analysis of changes in profit status, found little effect (Grabowski & Stevenson, 2008). Similarly, Rahman, Norton, and Grabowski (2016) found that vertical integration led to shorter length of stay and lower spending in skilled nursing facilities, whereas papers that addressed facility-level endogeneity found the opposite: Vertical integration led to longer length of stay and higher spending (David, Rawley, & Polsky, 2013; Konetzka, Stuart, & Werner, 2018). Although these studies are subject to other caveats and differed in other ways, making direct comparisons difficult, they demonstrate that the potential consequences of a mismatched instrument can be substantial.

5 | DISCUSSION

To address this error in the use of IV, researchers must be aware of the multiple levels of potential selection bias in any analysis of organizational-level treatment and must explicitly acknowledge that a patient-level IV will not solve the provider selection problem. Bypassing the issue by framing a research question as one of patient selection into providers with and without an attribute of interest will be an unsatisfactory solution in most cases. Misattributing endogeneity to patient-level selection in the presence of organization-level selection is subtle, seldom acknowledged, and unlikely to be clear to readers who are not exceptionally well versed in IV.

Reasonable (if not ideal) alternatives exist. First, if a valid provider-level IV can be found, one potential solution is to aggregate all data to the provider level and use only a provider-level instrument. This assumes that the focus on patient-level outcomes is not critical to the study. One might also separately estimate the model using patient-level and provider-level IVs to assess the importance of selection at each level.

If researchers want to retain the focus on patient-level outcomes and cannot identify more than one valid instrument, a practical alternative is to use multiple forms of quasi-experimental design simultaneously. For example, if longitudinal data are available, one might retain the patient-level IV but incorporate provider fixed effects or a difference-in-differences model to address provider-level selection, which accounts for selection on time-invariant unobservable provider characteristics that are correlated with both the treatment and the outcome, though still subject to time-varying confounding. In the absence of longitudinal data, authors could find other ways to isolate plausibly exogenous variation in provider-level treatment, such as the dose-response approach used in Joyce et al. (2018) wherein some residents were hypothesized to be more affected by the presence of a dementia special care unit than others. Although their approach does not rule out all possible

omitted variables, it narrows the scope of the potential confounders. As a last resort, provider-level matching on observable characteristics may help if the unobservable confounders are correlated with the observables. In the absence of these or similar approaches, researchers that address only patient-level selection must at a minimum avoid causal language about provider-level effects and explicitly acknowledge that the provider selection issue has not been solved.

REFERENCES

- Benatar S, Garrett AB, Howell E, & Palmer A (2013). Midwifery care at a freestanding birth center: A safe and effective alternative to conventional maternity care. *Health Services Research*, 48(5), 1750–1768. 10.1111/1475-6773.12061 [PubMed: 23586867]
- Bowblis JR, & McHone HS (2013). An instrumental variables approach to post-acute care nursing home quality: Is there a dime's worth of evidence that continuing care retirement communities provide higher quality? *Journal of Health Economics*, 32(5), 980–996. 10.1016/j.jhealeco.2013.06.007 [PubMed: 23999575]
- Cook TD, & Campbell DT (1979). *Quasi-experimentation: Design and analysis issues for field settings* Boston, MA: Houghton Mifflin.
- David G, Rawley E, & Polsky D (2013). Integration and task allocation: Evidence from patient care. *Journal of Economics and Management Strategy*, 22(3), 617–639. 10.1111/jems.12023
- Garabedian LF, Chu P, Toh S, Zaslavsky AM, & Soumerai SB (2014). Potential bias of instrumental variable analyses for observational comparative effectiveness research. *Annals of Internal Medicine*, 161(2), 131–138. 10.7326/M13-1887 [PubMed: 25023252]
- Grabowski DC, Feng Z, Hirth R, Rahman M, & Mor V (2013). Effect of nursing home ownership on the quality of post-acute care: An instrumental variables approach. *Journal of Health Economics*, 32(1), 12–21. 10.1016/j.jhealeco.2012.08.007 [PubMed: 23202253]
- Grabowski DC, & Stevenson DG (2008). Ownership conversions and nursing home performance. *Health Services Research*, 43, 1184–1203. 10.1111/j.1475-6773.2008.00841.x [PubMed: 18355255]
- Joyce NR, McGuire TG, Bartels SJ, Mitchell SL, & Grabowski DC (2018). The impact of dementia special care units on quality of care: An instrumental variables analysis. *Health Services Research*, 53(5), 3657–3679. 10.1111/1475-6773.12867 [PubMed: 29736944]
- Konetzka RT, Stuart EA, & Werner RM (2018). The effect of integration of hospitals and post-acute care providers on Medicare payment and patient outcomes. *Journal of Health Economics*, 61, 244–258. 10.1016/j.jhealeco.2018.01.005 [PubMed: 29428772]
- Lee DK, Chertow GM, & Zenios SA (2010). Reexploring differences among for-profit and nonprofit dialysis providers. *Health Services Research*, 45(3), 633–646. 10.1111/j.1475-6773.2010.01103.x [PubMed: 20403066]
- Martens EP, Pestman WR, de Boer A, Belitser SV, & Klungel OH (2006). Instrumental variables: Application and limitations. *Epidemiology*, 17(3), 260–267. 10.1097/01.ede.0000215160.88317.cb [PubMed: 16617274]
- McClellan M, McNeil BJ, & Newhouse JP (1994). Does more intensive treatment of acute myocardial infarction in the elderly reduce mortality? Analysis Using Instrumental Variables. *JAMA*, 272(11), 859–866. 10.1001/jama.1994.03520110039026 [PubMed: 8078163]
- Rahman M, Foster AD, Grabowski DC, Zinn JS, & Mor V (2013). Effect of hospital-SNF referral linkages on rehospitalization. *Health Services Research*, 48(6 Pt 1), 1898–1919. 10.1111/1475-6773.12112 [PubMed: 24134773]
- Rahman M, Norton EC, & Grabowski DC (2016). Do hospital-owned skilled nursing facilities provide better post-acute care quality? *Journal of Health Economics*, 50, 36–46. 10.1016/j.jhealeco.2016.08.004 [PubMed: 27661738]
- Sanwald A, & Schober T (2017). Follow your heart: Survival chances and costs after heart attacks-an instrumental variable approach. *Health Services Research*, 52(1), 16–34. 10.1111/1475-6773.12509 [PubMed: 27444099]

- Short PF, Moran JR, Yang TC, Camacho F, Gusani NJ, Mackley HB, Anderson RT (2016). Effects of hospital type and distance on lymph node assessment for colon cancer among metropolitan and nonmetropolitan patients in appalachia. *Medical Care Research and Review*, 73(5), 546–564. 10.1177/1077558715619052 [PubMed: 26613702]
- Wooldridge JM (2002). *Econometric analysis of cross section and panel data* Cambridge, Massachusetts: The MIT Press.
- Xu J, Liu G, Deng G, Li L, Xiong X, & Basu K (2015). A comparison of outpatient healthcare expenditures between public and private medical institutions in urban China: An instrumental variable approach. *Health Economics*, 24(3), 270–279. 10.1002/hec.3015 [PubMed: 24327571]

TABLE 1

Published papers using patient-level instrumental variables with provider-level endogenous treatment, 2010–2018

Paper ^a	Endogenous provider-level treatment variable	Instrument	Limitation discussed?	Remaining endogeneity concern
Midwifery care at a freestanding birth center: A safe and effective alternative to conventional maternity care (Benatar, Garrett, Howell, & Palmer, 2013)	Freestanding birth center	Cube root of distance to birth center	No	Choice to be a freestanding birth center is endogenous. Other features of birth centers—such as location or quality of staffing—present sources of omitted variables bias.
An instrumental variables approach to post-acute care nursing home quality: Is there a dime's worth of evidence that continuing care retirement communities provide higher quality? (Bowblis & McHone, 2013)	Whether a nursing home is part of a Continuing Care Retirement Community (CCRC)	Differential distance from patient home to the nearest CCRC nursing home and the nearest traditional nursing home	No	Choice to be a CCRC is endogenous. Sometimes results are phrased in the context of patient admission to a CCRC, but title and discussion imply causal effect of being a CCRC.
Effect of nursing home ownership on the quality of post-acute care: An instrumental variables approach (Grabowski et al., 2013)	Nonprofit nursing home ownership	Differential distance from patient home to closest nonprofit and closest for-profit nursing home	No	Choice to be a nonprofit is endogenous. Authors refer to “instrumenting for ownership status” though the patient-level instrument does not do this. Conceptual model and literature review focus on provider behavior.
The impact of dementia special care units on quality of care: An instrumental variables analysis (Joyce et al., 2018)	Whether a nursing home has a dementia Special Care Unit (SCU)	Differential distance from patient home to closest nursing home and nursing home with SCU	Yes	Choice to have an SCU is endogenous. Facility-level quality recognized as unmeasured confounder, but ignored more general problem of omitted facility characteristics. Appropriately used complementary methods to try to identify SCU effect.
Re-exploring differences among for-profit and nonprofit dialysis providers (Lee et al., 2010)	Profit status of dialysis facilities	Differential distance from patient home to dialysis facilities	Yes	Choice of profit status is endogenous. Some potential facility-level confounders acknowledged, but discussion and conceptual framework focus on effect of profit status.
Effect of hospital-SNF referral linkages on rehospitalization (Rahman et al., 2013)	Hospital-nursing home referral concentration	Predicted referrals based on patient distance to nursing homes	Yes	Choice of referral linkages is endogenous. Limited/indirect acknowledgment of the provider selection problem. Appropriately incorporated nursing home and hospital fixed effects in some specifications, identifying off changes over time to partially address the issue.
Do hospital-owned skilled nursing facilities provide better post-acute care quality? (Rahman et al., 2016)	Hospital-nursing home integration	Differential distance from a patient's home to the nearest hospital with and without a	No	Choice to be hospital-owned is endogenous. Treatment sometimes phrased as where patients get care, but conceptual model and literature review focus on provider behavior.

Paper ^a	Endogenous provider-level treatment variable	Instrument	Limitation discussed?	Remaining endogeneity concern
Follow your heart: survival chances and costs after heart attacks—an instrumental variable approach (Sanwald & Schober, 2017)	Hospital with catheterization lab	skilled nursing facility Differential distance from patient home to the closest hospital and the closest hospital with catheterization lab	Yes	Choice to have a catheterization lab is endogenous. Authors argue that the hospital-level correlates of having a catheterization lab are part of the intervention and advocate for more hospitals with this capability. However, it is impossible to identify which hospital characteristics are causing the outcome—some can be considered part of the outcome but others are confounders.
Effects of hospital type and distance on lymph node assessment for colon cancer among metropolitan and nonmetropolitan patients in Appalachia (Short et al., 2016)	Hospital type (accredited or high-volume)	Differential distance from patient home to the nearest hospital and the nearest accredited or high-volume hospital	No	Accreditation status and volume are endogenous. Goal of the paper in part is to look at where patients get care, but causal language used with respect to hospital type.
A comparison of outpatient healthcare expenditures between public and private medical institutions in urban China: An instrumental variable approach (Xu et al., 2015)	Public versus private hospital in China	Indicator of whether the patient's closest hospital is private	No	Choice to be public or private is endogenous. Goal of the paper is to compare costs, but discussion implies the causal effect of public versus private institutions.

^aFrom a search of the following major health economics and health services research journals: *Health Economics*, *Journal of Health Economics*, *Health Services Research*, *Medical Care*, and *Medical Care Research and Review*.