Culture Remains Elusive: Comment on the Identification of Cultural Effects Using Instrumental Variables

Winston Chou*

April 6, 2016

Forthcoming in the American Sociological Review

Abstract

I evaluate a method recently proposed in this *Review* for estimating the causal effects of cultural values on behavior (Polavieja 2015). After deriving the assumptions of the method, I argue that they are unreasonably strong for most applications, even when ideal data are available. I conclude with some suggestions regarding the conceptualization and measurement of cultural effects in future research.

Keywords: Culture, migration, quantitative methods, instrumental variables

1 Introduction

Understanding the influence of cultural values on behavior remains a venerated if incomplete task in sociology. Progress in this domain has been marked by a wide gap between theory and practice. Sociologists benefit from a wealth of theoretical models of culture, many of which imply causal claims. For example, if a person were more benevolent, then she would donate more money (Miles 2015). If she were endowed with different cultural skills, then she would take greater advantage of structural opportunities (Swidler 1986).

*PhD candidate, Department of Politics, Princeton University. Email: wchou@princeton.edu. Website: http://princeton.edu/~wchou. Thanks to the *ASR* editors and three reviewers for comments.

Although claims like these are essential to cultural theorizing, researchers are typically wary of drawing similarly strong inferences from observational data. Consequently, theoretical elaboration has far outpaced empirical certitude; whereas many theories of culture posit causal relationships, researchers are often content to demonstrate the predictiveness of cultural variables, not estimate causal effects directly.

In an article recently published in this *Review*, Polavieja (2015) proposes a novel method for estimating the causal effects of cultural values. The intuition behind Polavieja's method is simple. Imagine that a researcher observes an employed migrant who also happens to be culturally traditional. A challenge to inferring that the migrant's traditionalism caused her employment status is that the two variables may be jointly affected by a third variable, such as the availability of work (Gerson 1985). However, suppose the researcher finds an observationally identical woman who did not migrate. Logically, this woman's traditionalism cannot be affected by the availability of work in the host country. To the extent that the two women are genuinely interchangeable, the non-migrant's traditionalism can be viewed as an unconfounded version of the migrant's traditionalism. Thus, it can be used to estimate the causal effect of traditionalism on labor force outcomes, free of bias from confounders such as the availability of work.

Polavieja mounts an interesting defense of this method, but fails to adequately elucidate its assumptions. This comment evaluates Polavieja's method in a sympathetic but ultimately critical light. After deriving "first-best" identification assumptions of the method (i.e., the necessary assumptions given ideal data), I explain why these are unlikely to hold. I then show that the "second-best" alternative proposed by Polavieja is not only unnecessary, but diminishes the transparency and validity of empirical research. Finally, I offer some suggestions regarding the conceptualization and measurement of cultural effects in future research.

2 Culture and causal inference with instrumental variables

Many prominent theories of culture imply causal claims. For example, in criticizing the "culture of poverty" thesis prominent in the 1970s, Swidler (1986, 281) argued,

> People do not readily take advantage of new structural opportunities which would require them to abandon established ways of life [...] not because they cling to cultural values, but because they are reluctant to abandon familiar strategies of action.

Not one but two causal claims are embedded in this hypothesis. A person would not pursue new opportunities if her only *values* were changed. Indeed, Swidler observed, many of those seemingly mired in the culture of poverty appeared to share middle-class values and aspirations. However, she would pursue these opportunities if endowed with the necessary *tools* – for instance, knowing "how to dress, talk in the appropriate style, or take a multiple-choice examination" (275).

"Tool kit" theory inspired an efflorescence of research, much of it examining the relative importance of these two forms of culture. However, scholars in this tradition have generally shied from claiming to measure their effects directly. For example, Vaisey (2009) argues that tool kit theory understates the importance of deep moral systems, such as "individualistic" systems prioritizing self-gratification and "theistic" systems motivated by faith. Using panel data, he shows that these systems are correlated with behavior years after being measured. However, despite statistically controlling for many variables, Vaisey concludes only that "the choice of moral [system] in 2002 is a very good overall predictor of behavior in 2005" (1703). As he observes, the challenge to viewing this predictiveness as causation is that moral systems are not chosen at random. Rather, because other factors may be "previously implicated in producing different moral schemas" (Ibid.), the observed correlation may provide a misleading indication of how an individual's behavior would vary if her moral system were somehow changed.

Of course, moral systems are hardly unique in this respect; most "treatments" of interest cannot be ethically or practically randomized by researchers. Thus, Polavieja identifies the *method of instrumental variables*, a popular tool for estimating causal effects from nonexperimental data, as a promising avenue for future research. The essence of the method of instrumental variables is to regard certained observed variables as if they had been randomized, e.g., through a "natural experiment" (Rosenzsweig and Wolpin 2000). Intuitively, if these instrumental variables (IVs) strongly affect the treatment, and if they affect the outcome only through the treatment, then they can be used to estimate causal effects even when the latter has not been randomly assigned. However, it must be stressed that the resulting estimates have a causal interpretation only if a set of strong (and often debatable) assumptions are met.

These assumptions are commonly introduced in terms of structural diagrams or models. For example, consider a standard linear regression:

$$Y = \alpha + \tau T + \varepsilon \tag{1}$$

where Y is the probability of employment, T is a binary variable indicating traditionalism, and ε is an error term that captures other causes of labor force participation. The quantity of interest in this model is τ, the effect of traditionalism on labor force participation. It is well-known that τ cannot be consistently estimated when T is correlated with ε, as when it is affected by a "lurking" variable that also affects Y.

The IV solution involves an additional variable, Z, which is strongly correlated with T but not ε. More precisely, the necessary conditions for Z to be a valid instrument are, first, that Z must not covary with other causes of Y. This exogeneity condition is expressed by the following equation:

$$\text{Cov}(Z, \varepsilon) = 0. \tag{2}$$

Exogeneity is sometimes taken to mean that the instrument is as good as randomly assigned, since it implies that the instrument is uncorrelated with unobserved causes of the outcome. The second condition is the exclusion restriction, which states that the instrument affects the outcome only through its effect on the treatment, and can therefore be excluded from a model like equation (1). Lastly, the instrument must covary with the treatment, or

$$\text{Cov}(T, Z) \neq 0 \tag{3}$$

Importantly, the first two IV assumptions, exogeneity and the exclusion restriction, cannot be tested from the data (Morgan and Winship 2007). Even when the instrument is known to be randomly assigned, and hence exogenous, the exclusion restriction can only ever be established in theory (Rosenzweig and Wolpin 2000). Of course, an issue arising in most nonexperimental work is that the instrument is not known to be randomly assigned, and so the assumption of exogeneity is also rarely *prima facie* credible. I consider this aspect of IV methods in greater detail below.

3 The proposed methodology

Polavieja extends the method of IV to answer the following question: How does a woman's traditionalism affect her labor force outcome? The problem with simply comparing the labor force outcomes of highly and lowly traditional women is that assignment to traditionalism is nonrandom. Consequently, highly traditional women will likely differ from lowly traditional women along many unobservable or hard-to-measure dimensions. For example, women may "self-select" into or out of traditionalism on the basis of variables correlated with labor force participation, such as the availability of work or endowment of abilities. Gerson (1985) describes many women whose expanded work opportunities and discovery of latent talents led to greater ambivalence towards traditional feminine roles. These women would not be a good comparison group for highly traditional women, since modifying the traditionalist views of the latter group would not endow them with the same distributions of work opportunities or talents.

To address this problem, Polavieja introduces the Survey-based Imputation of Synthetic Traits used as Exogenous Regressors (SISTER) method, consisting of the following steps:

1. **Sampling step.** First, the researcher draws a sample of migrating and non-migrating women born in the same country. For migrants, only traditionalism at destination is observed, whereas only traditionalism at origin is observed for non-migrants.

2. **Imputation step.** Second, observationally equivalent non-migrating women are used to impute fictional values of traditionalism at origin for migrants. These values represent the levels of traditionalism that would have been observed for migrants, had they chosen not to migrate.

3. **IV step.** Lastly, these imputed ("synthetic") values are used as an instrument for the actually observed values among migrants, ostensibly to identify the causal effect of traditionalism on labor force outcomes.

The intuition behind this procedure is simple, and derives from the supposed unconfoundedness of the traits of non-migrating women. As Polavieja (2015, 184) reasons,

> The fundamental property of synthetic traits is that they are *by construction* exogenous to the destination environment, because they are imputed using information from observational-equivalent women who did not migrate.

While compelling, this logic has several flaws. These are best illustrated by considering the imputation and IV steps of the SISTER method separately. Thus, the next section examines some "first-best" assumptions of SISTER, which assume that the researcher does not need to impute any data. In particular, I assume the researcher has collected data on migrants before and after migration, and believes that the pre-migration values would have remained constant without migration. This eliminates the need to impute values based on *observationally* equivalent non-migrants, since the values of *exactly* equivalent non-migrants are observed. As these values are logically unaffected by the context of reception, and are therefore "exogenous" (in Polavieja's sense) to the destination environment, the researcher may be tempted to use these pre-migration values as an instrument. Unfortunately, as I show below, even these ideal data will rarely meet the IV assumptions.

3.1 First-best assumptions

The first-best assumptions of SISTER are as follows. First, the exogeneity assumption states that cultural traits measured before migration are unrelated to unobserved causes of labor force outcomes after migration. In other words, pre-migration cultural traits are as good as randomly assigned with respect to post-migration outcomes. Second, the exclusion restriction states that pre-migration traits do not affect post-migration labor force outcomes except through post-migration traits. Lastly, pre-migration traits must covary with post-migration traits. For space, I address the first two assumptions only.

Even in the first-best case, the exogeneity assumption will rarely be credible. This is because women with high and low traditionalism before migration may differ in ways that affect their labor force outcomes even after migration. For example, a conventional source of confounding bias is unobserved ability. Clearly, latent attributes like ability are not discarded in the process of migration (see figure 1). SISTER is motivated by the reasoning that pre-migration traditionalism will not be confounded by variables in the receiving country, such as the state of the destination economy. However, prior traditionalism may still be confounded due to attributes that follow migrants across borders, such as ability.

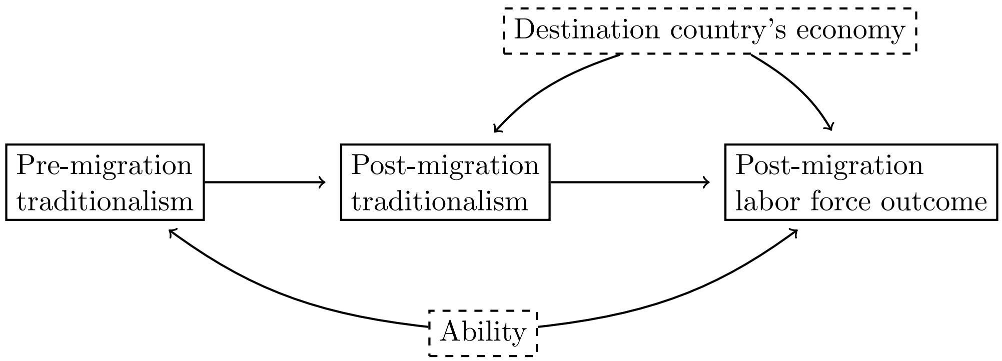

Figure 1: Pre-Migration Traditionalism is Not a Valid Instrument

The second assumption, or exclusion restriction, is also hard to justify. Assuming it is reasonable to posit a causal effect of traditionalism on employment outcomes, it is also reasonable to posit an effect on other behaviors known to affect employment independently of migration status, such as childbearing and educational attainment. For example, consider a woman who would have a child if she were highly traditional prior to migration, and whose child would continue to affect her post-migration employment status. The exclusion restriction would not be satisfied, since pre-migration traditionalism affects post-migration labor force participation through a channel other than post-migration traditionalism. Unfortunately, there is no fix for this problem, as controlling for prior childbearing will induce post-treatment bias (Rosenbaum 1984).

An even stronger assumption that has been left implicit is that individuals migrate at random. This assumption will not be satisfied if some women migrate because they are sure to be employed in the receiving country, nor will it be satisfied if some women do not migrate because they are highly traditional. Polavieja states that the potential bias from selection into migration is unknown (187). In fact, nonrandom migration has the startling implication that, even if the exogeneity assumption is satisfied in the population, it may not be satisfied when migrants and non-migrants are analyzed separately, as in SISTER.

The following example illustrates. Suppose pre-migration traditionalism (Z) is randomly assigned in the sending country, so that Z is exogenous in the population. Further assume that, among individuals observed to have high pre-migration traditionalism ($Z = 1$), 25% would be employed after migration. By contrast, 50% of individuals with low pre-migration traditionalism ($Z = 0$) would be employed after migration. If a random sample of the population were to migrate, revealing their post-migration employment outcomes, a simple comparison of the average employment rate among high and low traditionalism migrants would be unbiased for the average effect of high traditionalism on employment, $\tau = 0.25$.

Random migration is contrasted with selective migration in tables 1 and 2. Table 1 represents the expected employment rate among migrants when they represent a *random* draw from the sending country. However, suppose that individuals migrate only when they have low traditionalism or are sure to be employed in the receiving country. This is represented by asterisks in the corresponding cells in Table 1. This leaves Table 2. The same comparison of low- and high-traditionalism migrants results in a starkly biased estimate of the average causal effect, $\hat{\tau} = 0.5$, despite random assignment! This is due to collider bias, which arises from the dual selection into migration on the basis of employment and traditionalism (Elwert and Winship 2014).[1]

		Employed at destination	
		Yes	*No*
Traditionalism at origin	*High*	25%*	75%
	Low	50%*	50%*

*: Migrating cells when migration is non-random.

Table 1: Expected Distribution of Migrants' Observed Outcomes when Migration is Random

		Employed at destination	
		Yes	*No*
Traditionalism at origin	*High*	100%	–
	Low	50%	50%

Table 2: Expected Distribution of Migrants' Observed Outcomes when Migration is Non-Random

[1] Thanks to a reviewer for this point. A detailed simulation study demonstrating the effects of collider bias on SISTER is available from author's website.

These examples illustrate how, even with ideal data, SISTER entails excessively strong assumptions. In effect, the researcher assumes that pre-migration values are as good as randomly assigned, that they do not affect post-migration outcomes through alternative channels, and that the decision to migrate is random.

3.2 The second-best alternative

When researchers do not repeatedly observe the same individuals, they must resort to Polavieja's proposed version of SISTER. This version is "second-best" because it relies on missing data imputation. In fact, this section shows that the imputation step is unnecessary, as it nearly always assumes that the researcher has a valid, non-imputed instrument. However, while imputation is generally unnecessary, it is not always innocuous. Not only does the imputation step reduce the transparency of the relevant assumptions, it can also inadvertently invalidate a valid instrument.

1. **SISTER is unnecessary.** In the standard application of SISTER, Polavieja advises that the instrument be imputed based on observationally equivalent non-migrants using multiple regression (175). Specifically, suppose that, for each migrant i and non-migrant j, the researcher observes a set of covariates, such as race, gender, age, etc. Denote these covariates (along with an intercept) by $\mathbf{X}_i$ for migrants and $\mathbf{X}_j$ for non-migrants. The proposed imputation methodology involves first modeling non-migrants' traditionalism (T_j) using regression, or

$$\hat{T}_j = \mathbf{X}_j^\top \hat{\boldsymbol{\gamma}} \tag{4}$$

where $\hat{\boldsymbol{\gamma}}$ are estimated regression coefficients. Then the imputed value of traditionalism for migrants is estimated by multiplying their values of X by the estimated coefficients:

$$Z_i = \mathbf{X}_i^\top \hat{\boldsymbol{\gamma}}. \tag{5}$$

For simplicity, suppose the researcher only observes one covariate. Then, for each migrant, the value of the instrument is $Z_i = \hat{\gamma}_0 + \hat{\gamma}_1 X_i$. However, due to the linearity of covariances, it is readily shown that the validity of Z as an instrument implies the validity of X. First, if Z is to covary with T, it must be that $\hat{\gamma}_1 \mathrm{Cov}(X, T) \neq 0$. But this implies that both

$$\text{Cov}(X, T) \neq 0 \tag{6}$$

and

$$0 = \text{Cov}(Z, \varepsilon) = \text{Cov}(\hat{\gamma_0}, \varepsilon) + \hat{\gamma_1}\text{Cov}(X, \varepsilon) \implies \text{Cov}(X, \varepsilon) = 0. \tag{7}$$

Equation 7 establishes the exogeneity of X. Lastly, X must satisfy the exclusion restriction, since it is collinear with Z and cannot be included alongside it in a regression. This verifies that the validity of Z as an instrument implies the validity of X.

In general, if the imputed instrument is valid, the researcher will nearly always have a valid, non-imputed instrument. A formal proof of this statement, generalizing to multiple imputation regressors, is provided in the appendix (Proposition 1). Therefore, the imputation step is superfluous, as one can simply use the non-imputed variables as the instruments.

2. **SISTER threatens transparency and validity.** Although the imputation step is rarely necessary, it is not necessarily innocuous. By transforming the non-imputed imputation regressors into a single variable, the imputation step hides the dependence of SISTER on all of imputation regressors. For example, Polavieja uses "age, years of schooling, parental education, and religious denomination as predictors of traditionalism in the imputation regression" (175). Equation 8 below shows that all of these variables must satisfy the exogeneity condition; it is not enough, as Polavieja suggests, to include "at least one regressor that is (arguably) orthogonal to the error term in the structural equation of interest" (Ibid.). This misconception is damaging given that IV methods rely on assumptions that must be argued theoretically rather than demonstrated from the data (Morgan and Winship 2007). While Polavieja claims that regression diagnostics can test whether the instrument is exogenous, this is a misinterpretation of the Wald exogeneity test, which assesses whether the instrumented variable is endogenous on the assumption that the instrument is exogenous (182).

   Unfortunately, the imputation step is not just obscurant, it can also reduce the validity of empirical estimates. To see this, suppose the instrument is now imputed with two variables, only one of which is exogenous, so that $\text{Cov}(X_1, \varepsilon) = 0$ but $\hat{\gamma}_2 \text{Cov}(X_2, \varepsilon) \neq 0$. But then

$$\text{Cov}(Z, \varepsilon) = \text{Cov}(\hat{\gamma}_0, \varepsilon) + \hat{\gamma}_1 \text{Cov}(X, \varepsilon) + \hat{\gamma}_2 \text{Cov}(X_2, \varepsilon) \neq 0, \tag{8}$$

This demonstrates that, even when the researcher has a genuinely valid instrument, the imputed variable is not necessarily valid.

4 Recommendations for future research

This comment evaluates a recently proposed method for estimating the causal effects of culture. I argue that Polavieja's SISTER method makes assumptions that are unreasonably strong for most applications, even when the data are ideal. When these data are not available, SISTER is an unnecessary and obscurant alternative to standard IV methods. Thus, it should not be used.

Are there more promising avenues for causal analysis in cultural research? An important task for future research will be to sharpen the conceptualization of cultural effects. It has become increasingly popular to view causation as a comparison between potential states of the world, one in which a cause is applied to a unit and another in which it is not (Holland 1986). For instance,

> If we say, "This boy has grown tall because he has been well fed," we are not merely tracing out the cause and effect in an individual instance; we are suggesting that he might quite probably have been worse fed, and that in this case he would have been shorter.
>
> R.A. Fisher, quoted in Rubin (2005, 323)

As Holland (1986) argues, thought experiments of this kind presuppose manipulability: in principle, one could design an experiment in which some boys were randomly chosen to be better fed than others. It is far less clear how one might similarly manipulate a person's values without distorting other fundamental characteristics relevant to outcomes of interest. This is not just a philosophical quibble, but bears, for instance, on the specification of statistical models that purportedly identify the effect of cultural values while "holding constant" variables such as parentage and context of upbringing. These models have no obvious interpretation, as no realistic intervention could induce such an effect (cf. Rosenbaum 1984).

A fruitful direction may therefore be to focus on the more manipulable aspects of culture, such as the habits and skills highlighted in tool kit theory. For example, cognitive and dual-process theories of culture posit that individuals can be led to act in ways that are more or less motivated by their underlying moral systems (Vaisey 2009). Thus, Miles (2015) conducts an experiment in which individuals are randomly assigned numbers to memorize, increasing their cognitive load and thus the probability of relying on deep-seated values. This design facilitates inference by manipulating the salience, rather than content, of values. Along similar lines, Sen and Wasow (2016) discuss research designs for estimating the influence of race, another seemingly immutable characteristic, by disaggregating it into its more flexible parts. Of course, in turning from Parsonsian "unmoved movers" to the more mutable aspects of culture, these designs arguably sacrifice some of classical theory's grand ambitions for enhanced credibility and interpretation. How to manage these tradeoffs will be an important question as the causal analysis of culture progresses.

5 References

Elwert, Felix and Christopher Winship. 2014. "Endogenous selection bias: The problem of conditioning on a collider variable." *Annual Review of Sociology*, 40:31-53.

Gerson, Kathleen. 1985. *Hard choices: How women decide about work, career, and motherhood.* UC Press.

Holland, Paul W. 1986. "Statistics and causal inference." *Journal of the American Statistical Association*, 81(396):945-960.

Miles, Andrew. 2015. "The (re)genesis of values: Examining the importance of values for action." *American Sociological Review*, 80(4):680-704.

Morgan, Stephen and Christopher Winship. 2007. "Counterfactuals and causal inference." Cambridge University Press.

Polavieja, Javier. 2015. "Capturing culture: A new method to estimate exogenous cultural effects using migrant populations." *American Sociological Review*, 80(1):166-191.

Rosenbaum, Paul. 1984. "The consequences of adjustment for a concomitant variable that has been affected by the treatment." *Journal of the Royal Statistical Society, Series A*, 656-666.

Rosenszweig, Mark R. and Kenneth I. Wolpin. 2000. "Natural 'natural experiments' in economics." *Journal of Economic Literature*, 38:827-874.

Rubin, Donald. 2005. "Causal inference using potential outcomes." *Journal of the American Statistical Association*, 100(469):322-331.

Sen, Maya and Omar Wasow. 2016. "Race as a 'bundle of sticks': Designs that estimate effects of seemingly immutable characteristics." *Annual Review of Political Science*, 19.

Swidler, Ann. 1986. "Culture in action: Symbols and strategies." *American Sociological Review*, 51(2):273-286.

Vaisey, Stephen. 2009. "Motivation and justification: A dual-process model of culture in action." *American Journal of Sociology*, 114(6):1675-1715.

6 Appendix

Definitions. Let

$$Y = \alpha + \tau T + \varepsilon \tag{1}$$

be a linear population model for Y with $\mathrm{Cov}(T, \varepsilon) \neq 0$. We say that a variable Z is a valid instrument for T or that Z satisfies the instrumental variables (IV) assumptions if $\mathrm{Cov}(Z, \varepsilon) = 0$ and $\mathrm{Cov}(Z, T) \neq 0$. Excludability is implied by the omission of Z from equation 1.

Proposition 1. Let $X_1, X_2, \ldots, X_k$ be variables in a multiple regression imputation function:

$$Z = \hat{\gamma}_0 + \sum_{j=1}^{k} \hat{\gamma}_j X_j \tag{2}$$

We further assume that

$$\hat{\gamma}_j \mathrm{Cov}(X_j, \varepsilon) \neq 0 \implies \hat{\gamma}_j \mathrm{Cov}(X_j, \varepsilon) \neq -\sum_{j'} \hat{\gamma}_{j'} \mathrm{Cov}(X_{j'}, \varepsilon), j \neq j'. \tag{3}$$

This assumption rules out the exceptional event that the scaled covariances cancel each other out.

Then Z satisfies the IV assumptions only if X_j satisfies the IV assumptions for at least one j.

Proof of Proposition 1.

1. Because Z is valid by assumption, $\mathrm{Cov}(Z, T) \neq 0$. By definition of Z, this implies that

$$\mathrm{Cov}(\hat{\gamma}_0 + \sum_{j=1}^{k} \hat{\gamma}_j X_j, T) \neq 0. \tag{4}$$

2. By linearity of covariances, equation 3 implies that

$$\hat{\gamma}_j \mathrm{Cov}(X_j, T) \neq 0 \tag{5}$$

for at least one j.

3. Because Z is valid by assumption, $\text{Cov}(Z, \varepsilon) = 0$. By definition of Z, this is the same as
$$\text{Cov}(\hat{\gamma_0} + \sum_{j=1}^{k} \hat{\gamma_j} X_j, \varepsilon) = 0. \tag{6}$$

4. By linearity of covariances, this can be rewritten as
$$\sum_{j=1}^{k} \text{Cov}(\hat{\gamma_j} X_j, \varepsilon) = 0. \tag{7}$$

5. By equation 5, $\hat{\gamma_j} \neq 0$ for at least one j. By our assumption and exogeneity of Z,
$$\text{Cov}(X_j, \varepsilon) = 0 \tag{8}$$
for this j.

6. Equations 5 and 8 together imply that this X_j is a valid instrument for T. *Q.E.D.*

7 Simulation Study

This simulation study examines the effect of collider bias on instrumental variables (IV) methods that analyze migrants separately from non-migrants. I especially focus on the SISTER method introduced by Polavieja (2015) in the *American Sociological Review*. Several scenarios are analyzed; however, across all scenarios, I assume the following model for an outcome Y

$$Y_i = -T_i + \epsilon_i, \tag{1}$$

with

$$\epsilon \sim \mathcal{N}(0, 1) \tag{2}$$

and

$$T_i \sim B(\pi_i), \tag{3}$$

where the definition of π_i varies depending on the scenario.

To assess the effect of collider bias, I further assume that migration is correlated with Y and T. In particular,

$$M_i \sim B(\mu_i), \tag{4}$$

where

$$\mu_i = \Phi(\epsilon_i - \pi_i), \tag{5}$$

and Φ is the standard normal distribution function.

The behavioral interpretations for these equations are as follows. Equation (1) states that "traditional" individuals ($T_i = 1$) will have lower outcomes than non-traditional individuals. Equations (1, 3-5) state that, *ceteris paribus*, individuals are more likely to migrate when they have higher values of the outcome (e.g., post-migration labor force participation) and are not traditional. In other words, individuals select into migration on the basis of post-migration outcomes and their pre-migration values. According to Polavieja, selection on cultural values can be diagnosed by low instrument relevance. In scenario 3, I demonstrate that this is not generally the case.

Scenario 1

In the baseline scenario, I assume that π_i is a function of ϵ_i, so that T is endogenous. However, the researcher has a valid instrument Z (i.e., is exogenous, satisfies the exclusion restriction, and affects T_i). In particular,

$$Z \sim \mathcal{N}(0,1), \tag{6}$$

and

$$\pi_i = \Phi(Z_i - \epsilon_i). \tag{7}$$

I estimate four models in all scenarios. First, equation 1 is estimated with ϵ_i omitted by ordinary least squares (OLS):

$$\hat{Y}_i = \hat{\tau} T_i. \tag{8}$$

Because of the omitted variable, the OLS estimate is inconsistent for $\tau = -1$.

Second, I estimate τ with two-stage least squares using Z as the instrument. The whole sample is used. This is the standard IV approach, and the IV estimate is consistent for τ. Third, *and for migrants only*, I estimate τ with two-stage least squares using Z as the instrument. Due to collider bias, this IV estimate is inconsistent for τ. To illustrate this, fourth, I reestimate this model but allow M be exogenously determined, i.e. by setting $\mu_i = 1/2$ for all i. In other words, I resimulate the data but allow individuals to migrate completely at random. The IV estimate for τ from this model is also consistent.

The results, based on 4 × 5,000 Monte Carlo simulations, are shown in figure 1, which plots the root-mean-squared error (RMSE) the estimates of τ as the sample size grows from 500 to 10,000. Figure 1 also shows the coverage of the 95% confidence interval (CI) at these sample sizes. As the sample size becomes infinitely large, the RMSE should converge to 0. The true value of the parameter should be covered by the 95% CI in .95 of simulations.

As figure 1 shows, the IV with migrants only under non-random migration is both inconsistent and has poor coverage even as the sample size grows. Note that these issues are not specific to SISTER, as I do not use the imputation step of SISTER until scenario 3.

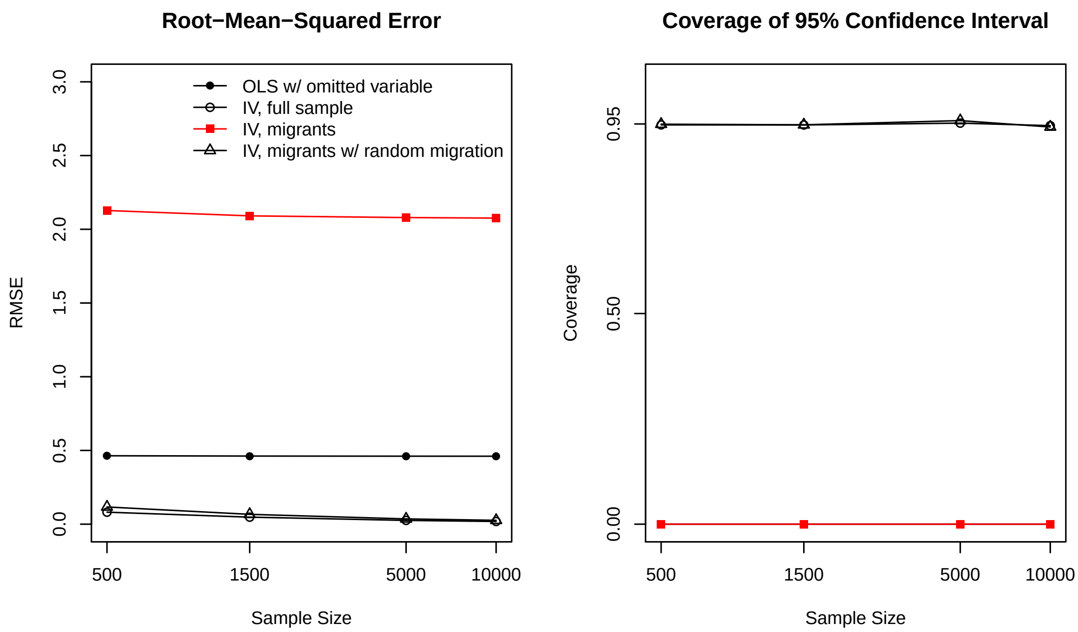

Figure 1: Scenario 1 Simulation Results

As such, these results illustrate some general implications of collider bias when analyzing migrants and non-migrants separately. Also note that, consistent with expectations, OLS with an omitted variable performs poorly, having high RMSE even as $n \to \infty$ and low coverage proportion. On the other hand, IV with the full sample and under random migration is consistent, and the estimated confidence intervals cover the true value of $\tau = -1$ in the correct proportion of trials. The latter has higher variance at each sample size since half of the sample (the non-migrants) is discarded.

Scenario 2

In the second scenario, traditionalism is only endogenous for migrants. For precision, I introduce some additional notation. Let T_i^0 denote pre-migration traditionalism. It is distributed according to the Bernoulli distribution with parameter π_i^0. Then, as before, Z

is exogenous,

$$Z \sim \mathcal{N}(0,1), \tag{9}$$

and π_i is determined by

$$\pi_i^0 = \Phi(Z_i). \tag{10}$$

However, after M_i is realized, the value of traditionalism observed, denoted T_i, is affected by ϵ_i for migrants only. Specifically,

$$T_i \sim B(\pi_i), \tag{11}$$

and

$$\pi_i = \Phi(Z_i - \epsilon_i), \tag{12}$$

for all i for whom $M_i = 1$. However, migrants' pre-migration values (T_i^0) are observed, and can be used as an instrument for T_i.

For this scenario, I reestimate all the models but replace Z with T_i^0. As figure 2 shows, with the exception of the IV regression that (a) uses migrants only when (b) migration is endogenous, the other estimators have lower RMSE relative to scenario 1. This is because the treatment variable is only endogenous for migrants. Indeed, the OLS estimator has slightly better coverage when the sample size is small, which is intuitive since the confidence intervals are larger. However, the problematic IV estimator does not improve; indeed, it has a slightly higher RMSE relative to scenario 1. This is because the instrument is now a mapping from Z to $\{0,1\}$, and not Z itself.

Scenario 3

Lastly, I examine the scenario in which T is only endogenous for migrants, but their pre-migration values must be estimated using observationally equivalent non-migrants. This is the scenario closest to SISTER. The only observed covariate is Z, which satisfies the instrumental variables assumptions. As recommended by Polavieja (2015), migrants' pre-

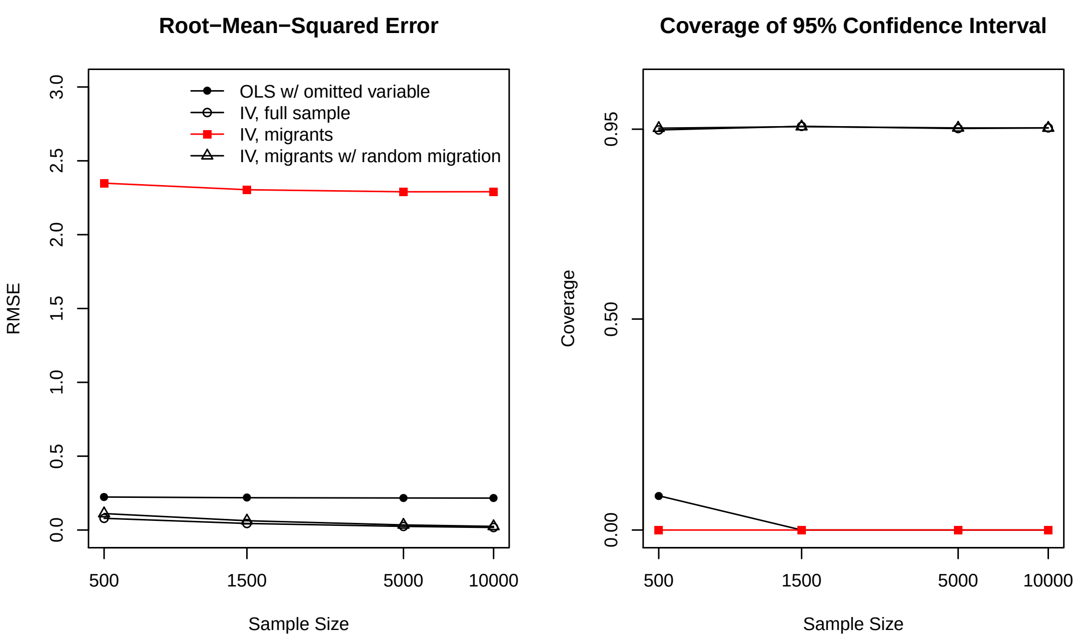

Figure 2: Scenario 2 Simulation Results

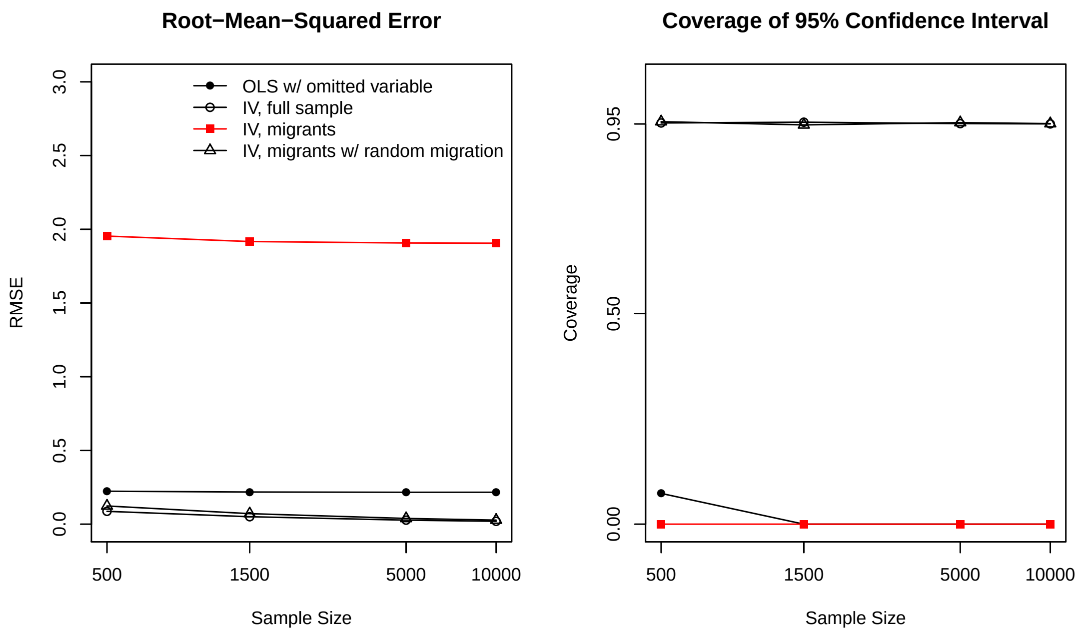

Figure 3: Scenario 3 Simulation Results

migration values are estimated using linear regression, i.e.,

$$\hat{T_i^0} = \hat{\gamma} Z_i, \tag{13}$$

where $\hat{\gamma}$ is estimated from an linear regression of traditionalism on Z for migrants.

As with Scenarios 1 and 2, I estimate the following four models: a linear regression model, an IV regression with $\hat{T_i^0}$ as the instrument for the full sample, an IV regresion with $\hat{T_i^0}$ as the instrument for migrants only when migration is endogenous, and an IV regression with $\hat{T_i^0}$ as the instrument for migrants when migration is exogenous.

As figure 3 shows, the SISTER estimator does not perform well. It has a slightly better RMSE relative to scenario 2, as the instrument is now a bijective function of Z_i rather than a mapping to $\{0, 1\}$. However, its RMSE is even higher than OLS and its coverage remains poor.

7.1 Can selection be diagnosed from low instrument relevance? No.

Polavieja conjectures that selection into migration can be diagnosed by low correlation between the imputed instrument and the endogenous variable (174). The intuition is that, if migrants and non-migrants are sufficiently different, the imputation regressors (estimated from non-migrants) will generate poor predictions for migrants. The problem with this argument, as with SISTER more generally, is that it fails to recognize that the instrument is just a function of the migrants' own covariates. When the imputation regressors are nonzero, the instrument will contain information from migrants.

Figure 4 shows a histogram of 5,000 pairwise correlations between the imputed variable and the endogenous variable for migrants from my simulations. The red dashed line represents the overall correlation between the imputed variable and the endogenous variable from Polavieja's sample (180). It is obvious that the simulation correlations are higher on average despite nonrandom selection into migration.

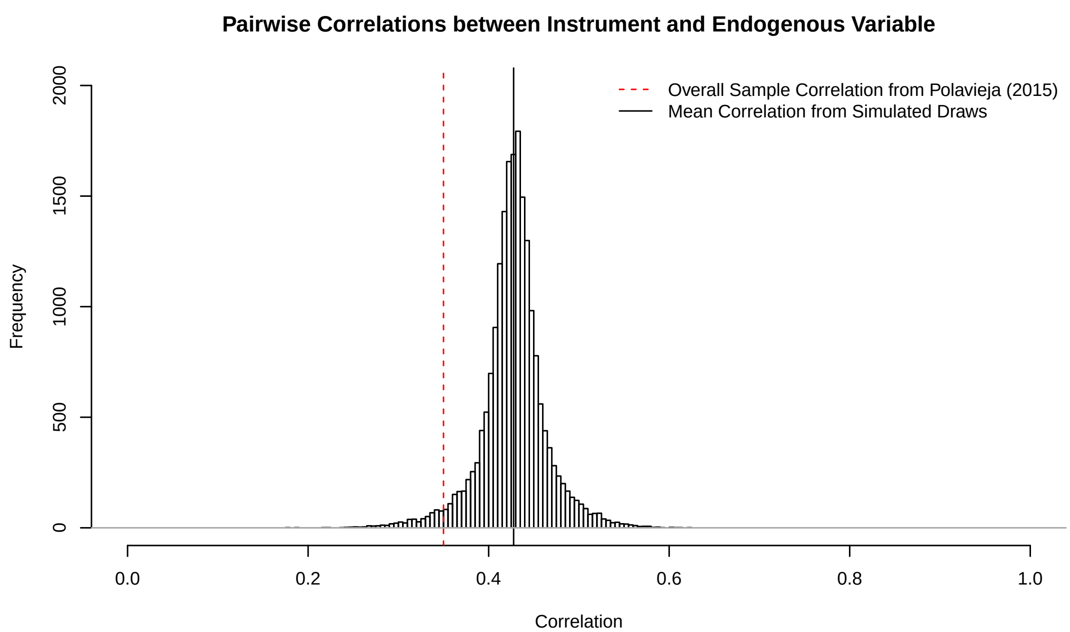

Figure 4: Histogram of Correlations between Imputed Instrument and Endogenous Variable

References

Polavieja, Javier G. 2015. “Capturing Culture: A New Method to Estimate Exogenous Causal Effects using Migration Populations.” *American Sociological Review*, 80(1):166-191.