AMERICAN SOCIOLOGICAL REVIEW

OFFICIAL JOURNAL OF THE AMERICAN SOCIOLOGICAL ASSOCIATION

ONLINE SUPPLEMENT
to article in
AMERICAN SOCIOLOGICAL REVIEW, 2017, VOL. 82

Culture as a Random Treatment: A Reply to Chou

Javier G. Polavieja
Universidad Carlos III de Madrid

Index

Part 1.

Having variables in the imputation regression that are also included in the structural equation can never harm the validity of the instrument. Here is the proof:

Definitions,

Let

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

be a population model where covariate T is endogenous, that is, $cov\ (T, \varepsilon) \neq 0$, and G is a vector of control variables that predict Y with E(ε |G)=0. We say Z is a valid instrument for T if $cov\ (Z, \varepsilon) = 0$ (exogeneity) and $cov\ (Z, T) \neq 0$ (relevance).

For simplicity, I introduce new notation for the imputation regression:

$$\hat{Z} = D\hat{\gamma}_1 + DG\hat{\gamma}_2 \tag{2}$$

where D is a collection of country of origin/ancestry dummies D={$d_1 \ldots, d_N$}, N is the number of origins, and DG is the interaction of the origin/ancestry dummies with all the covariates. Imputation regression 2 contains both intercepts and imputation variables G because

$$\sum_{s=1}^{N} d_s = 1 \text{ and } \sum_{s=1}^{N} d_s\, G = G$$

We assume that origin/ancestry is exogenous, that is, E(ε |D)=0.

Proposition 1: Having G variables in the imputation regression (i.e., variables that are also included in the structural equation) can never harm the instrument.

Proof of proposition 1

1. By the law of iterated expectations E(ε |D)=0 implies

$$cov\ (\varepsilon, D) = 0 \tag{3}$$

2. Similarly, E(ε |D, G)=0 implies

$$cov(\varepsilon, DG) = 0 \qquad (4)$$

3. By linearity of covariances

$$cov\ (\varepsilon, \hat{Z}) = cov\ (\varepsilon, D\hat{\gamma}_1 + DG\hat{\gamma}_2) = cov(\varepsilon, D\hat{\gamma}_1) + cov(\varepsilon, DG\hat{\gamma}_2)$$
$$= cov(\varepsilon, D)\hat{\gamma}_1 + cov(\varepsilon, DG)\hat{\gamma}_2 = 0 \qquad (5)$$

So, $\hat{Z}$ with G is a valid instrument.

Part 2.

Having variables in the imputation regression that are also included in the structural equation can add to the identification potential of SISTER at no (statistical or theoretical) cost. Adding further imputation predictors not included in the structural equation might also yield substantial statistical gains, but it always requires making the additional assumption that these predictors are exogenous to the outcome of interest. Let us see this in further detail.

Recall that imputation in SISTER is done in two steps. First, we model non-migrants' trait of interest (T_j) at country of origin using a set of multiple regressions of the form:

$$\hat{T}_j = \hat{\gamma}_{0,c_j} + \sum_{s=1}^{k} \hat{\gamma}_{s,c_j} X_{s,j} \qquad (1)$$

where c_j is the country of non-migrating individual j; $\hat{\gamma}_{0,c_j}$ is the intercept for c_j; and $\hat{\gamma}_{s,c_j}$ is a vector of multiple regression coefficients.

Second, we estimate the imputed value of the trait for migrants ($\hat{Z}_i$) by multiplying their own values of $\boldsymbol{X}$ by the coefficients estimated at their respective countries of origin/ancestry (c):

$$\widehat{T'}_i = \hat{\gamma}_{0,c_i} + \sum_{s=1}^{k} \hat{\gamma}_{s,c_i} X_{s,i} \qquad (2)$$

X covariates can be of two types: (1) imputation regressors included in the structural equation, $\boldsymbol{G}$; and (2) imputation regressors not included in the structural equation, which I denote Z^*. Because $\boldsymbol{G}$ covariates are also predictors of the outcome in the structural equation, they are by construction orthogonal to the error term and thus can never harm the validity of the instrument (see Part 1). If the imputation regression includes an additional instrument Z^* that is not included in $\boldsymbol{G}$, then this instrument must be itself exogenous for $\widehat{T'}_i$ to be valid. Hence, $cov\,(Z^*, \varepsilon) = 0$. We can therefore distinguish between three different possible types of imputation regressions in SISTER, which I call, respectively, (1) intercept imputation; (2) intercept $+\,\boldsymbol{G}$ imputation; and (3) intercept $+\,\boldsymbol{G} + Z^*$ imputation. Intercept imputation uses variation in the trait at country of origin/ancestry ($\hat{\gamma}_{0,c_j}$) as the only source of identification; intercept $+\,\boldsymbol{G}$ imputation adds imputation regressors that are covariates included in the structural equation; and intercept $+\,\boldsymbol{G} + Z^*$ imputation includes (at least) one extra exogenous parameter Z^*not included in the structural equation.

Table S1 shows SISTER estimates for synthetic traits ($\widehat{T'}_i$) imputed using these three different types of imputation regressions. Model 1 imputes on country of origin/ancestry alone; Model 2 adds age, schooling, and parental education, which are also covariates in the structural equation; and Model 3 adds religious denomination, which is assumed exogenous to FLFP. All models are fitted to the European Social Survey data originally used in my 2015 article. The last rows of the table present a simple comparison of SISTER estimates across these three imputation regressions, taking the intercept model (Model 1) as the benchmark.

Note that Model 2 uses variation in the intercepts and variation in the slopes of $\boldsymbol{G}$ as sources of identification. Model 2 yields an 8 percent increase in instrument relevance (*z* score of $\widehat{T'}_i$ in the 1st stage), and a 17 percent increase in the *z* score of the SISTER estimate in the 2nd stage (with a 13 percent reduction in its standard error). The Wu-Hausman test for endogeneity also improves: imputing on the intercepts alone yields a *z* score of 1.78 ($p > |z| = .075$); whereas adding $\boldsymbol{G}$ in the imputation of $\widehat{T'}$ yields a *z* score of 2.33 ($p > |z| = .020$). This is a 31 percent increase.[1] In summary, Model 2 shows that using variation in the slopes of $\boldsymbol{G}$ across countries of origin can add to the identification potential of the SISTER method by improving the statistical efficiency of the estimate. Because our instrument is scalar, these improvements come at no cost in terms of degrees of freedom. Moreover, Model 2 does not require making any further exogeneity

[1] Because the Wu-Hausman test is only as good as the instruments used, comparing the results of this test across instruments is a good way of judging instrument performance (see, e.g., Hausman 1978).

assumptions, because, as noted above, covariates $\boldsymbol{G}$ are controls in the structural equation and thus orthogonal to the error term. Also, as shown in Part 1, adding $\boldsymbol{G}$ covariates to the imputation regression can never harm the validity of the instrument. Hence, Model 2 is unquestionably preferable to Model 1.

Table S1. SISTER Estimates for Three Different Imputation Regressions

Variables	MODEL 0 Probit	MODEL 1 SISTER	MODEL 2 SISTER	MODEL 3 SISTER
Traditionalism	-.056*	-.376**	-.381**	-.261**
	(.023)	(.142)	(.123)	(.070)
Age	-.037**	-.029**	-.029**	-.033**
	(.003)	(.005)	(.005)	(.003)
Age^2	-.002**	-.002**	-.002**	-.002**
	(.000)	(.000)	(.000)	(.000)
Schooling	.047**	.030**	.030**	.038**
	(.007)	(.009)	(.009)	(.007)
Constant	-.001	1.347*	1.370**	.855**
Observations	2,915	2,915	2,915	2,915
Imputation step predictors		$\hat{T}_1$=origin	$\hat{T}_2$=origin + age + schooling + parental education	$\hat{T}_3$= origin + age + schooling + parental education + religious denomination
1st stage $\hat{T}$'on T (z)		7.34**	7.91**	17.85**
Wald test (Chi2[1])		3.43+	4.52*	8.64**
Wu-Hausman test of endogeneity (z)		1.78+	2.33*	2.45*
Improvement on Model 1:				
% Increase in z (SISTER estimate)			17.0	40
% Decrease in Std. Err (SISTER estimate)			13.1	48.8
% Increase in instrument relevance (z)			7.8	125.7
% Increase in Wu-Hausman test (z)			30.9	37.6

Source: Calculated by the author from the European Social Survey, rounds 1, 2, and 3 combined, restricted migrants sample (1st-, 1.5-, and 2nd-generation immigrant women age 16 to 65 years from ESS-sampled origins and not in education).
Note: S.E. clustered at destination. SISTER = IV probit models. All models control for migrant generation, host language proficiency, and type of destination location. Models 1, 2, and 3 also control for parental education. Robust standard errors are in parentheses.
+p < .1; *p < .05; **p < .01 (two-tailed tests).

Model 3 also yields sizeable statistical gains over Model 1 (the *z*-score of the SISTER estimate increases as much as 40 percent with a 49 percent reduction in the standard errors and a huge increase in instrument relevance) with the same degrees of freedom. Yet, Model 3 requires making the additional assumption that religious denomination is exogenous to FLFP. Only if we believe this assumption should we choose Model 3 over Model 2.

Part 3.

SISTER rests on the assumption that culture of birth is exogenous. But if this is the case, why not use country of origin/ancestry directly as an instrument in standard IV (SIV) estimation? The answer is because this would lead to the too-many-too-weak instruments problem.

To see this, Table S2 compares the results of fitting two IV probit models: Model 1 uses countries of origin/ancestry as instruments for immigrant women's traditionalism. Model 2 is the SISTER intercept-imputation model described in Part 2. Hence, Model 1 is the SIV equivalent to Model 2. There is, however, one very noticeable difference between the two: Model 1 requires using 23 – 1 = 22 different origin dummies as instruments, whereas Model 2 uses only one.

Table S2. Comparing Standard IV and SISTER Estimates using Country of Origin/Ancestry as Instruments

Variables	MODEL 0 Probit	MODEL 1 Standard IV	MODEL 2 SISTER
Traditionalism	-.056* (.023)	-.588** (.086)	-.376** (.142)
Constant	-.001	2.280**	1.347*
Observations	2,915	2,915	2,915
IV description		*Countries of origin/ancestry* (22 instruments)	*Synthetic traits (1 instrument)*
Imputation predictors		No imputation	$\hat{T}$=origin/ ancestry
F test for IV relevance		3.99**	53.70**

Source: Calculated by the author from the European Social Survey, rounds 1, 2, and 3 combined, restricted migrants sample (1st-, 1.5-, and 2nd-generation immigrant women age 16 to 65 years from ESS-sampled origins and not in education).
Note: S.E. clustered at destination. SISTER = IV probit models. All models control for age, age squared, schooling, migrant generation, host language proficiency, and type of destination location. Models 1 and 2 also control for parental education. Robust standard errors are in parentheses.
+p < .1; *p < .05; **p < .01 (two-tailed tests).

The last row of Table S2 shows the F-statistic for instrument strength. This is a joint test of whether all excluded instruments are significantly different from zero in the first stage. Stock and Yogo (2005) show that for an instrument to be sufficiently strong, the F-statistic should be bigger than 10 in case of a single endogenous regressor. Values below 10 imply the instrument(s) is (are) too weak to provide reliable estimates. The F-statistic for model 1 is only 3.99, revealing SIV estimates suffer from the well-known too-many-too-weak instruments problem (see, e.g., Andrews and Stock 2005). In stark contrast, the F-statistic for Model 2 is 53.7. This is way above Stock and Yogo's threshold for instrument relevance. Comparing Models 1 and 2 thus reveals how imputing on the intercepts alone can boost instrument strength. By taking a large number of potential instruments and condensing them in a scalar using external data, SISTER provides a parsimonious way of modelling the first stage in IV estimation. This is a great advantage over SIV, because it helps us circumvent the too-many-too-weak instruments problem.

Part 4. SIMULATION STUDY

I now address Chou's concern about selection bias (SB). I use simulation experiments to directly address this concern because Chou draws on simulation experiments to develop his argument.[2]

Chou's concern about SB is both legitimate and important. Yet in posing the problem of non-random selection as a special case of split-sample IV estimation, Chou assumes there is a valid IV for people who do not migrate. This assumption is puzzling, for if we had such valid IV we would not have an endogenous preference problem to solve.

To address the question of non-random selection in a way that is useful for the discussion of the SISTER method, we must acknowledge the problem of endogenous preferences. This implies recognizing that social embeddedness makes it virtually impossible for us to observe a valid instrument for non-migrating respondents. In other words, we cannot have an exogenous instrument for women's traditionalism unless they migrate. This can be expressed formally as follows:

[2] Chou has provided two different versions of his experiments: A first preprint version (Chou 2016), which supplemented the comment I was originally asked to respond to, and a second version (Chou 2017), which Chou developed while my reply was under review. This simulation study addresses the first version of Chou's experiments more directly but my results are relevant to both.

$$T_i = T_i^* + \eta_i \tag{1}$$

Expression 1 decomposes observed traditionalism into two theoretical components: a purely exogenous component T_i^* and an endogenous component, which I call social embeddedness and denote η_i. This latter component would capture all the contextual effects that exert a joint influence on both observed traditionalism and the outcome of interest (FLFP), net of the standard individual-level predictors (η_i would thus capture societal influences such as the role of the welfare state, labor-market institutions, the educational system, and the macro-economic environment). By introducing this latter component we account for the problem of endogenous preferences, which is the very problem we want to address.

We observe T_i^* only for migrants. This we can call the *epidemiological principle* (i.e., the idea that migration allows us to identify the exogenous component of culture because it removes people from their original embedding social environments). Because this is an ideal setup, we can assume, following Chou (2016), that we actually observe this exogenous component (T_i^*) fully for migrants, so we do not need to impute it.

As in Chou's own setup, the selection problem occurs if (1) the probability to migrate $\Pr(M = 1)$ is potentially influenced by women's traditionalism (e.g., if more traditional women are less likely to migrate) and/or (2) it is potentially influenced by a set of factors that are not correlated with T_i but can affect the outcome of interest at destination, Y_i. This is what Chou calls *colliders*, denoted μ_i. One such collider could indeed be, as he argues, the availability and density of co-ethnic networks at country of destination (see Chou 2016). Formally the selection problem can be expressed as follows:

$$\Pr(M = 1) = \Pr(\beta_0 T_i + \beta_1 \mu_i + V_i') \tag{2}$$

$$Y_i = f(\eta_i, \mu_i) \tag{3}$$

where V_i' is a random component that captures all other unobserved determinants of migration that are unrelated to the outcome of interest, Y_i (FLFP in this case). As in Chou's formalization, the intensity of selection bias is captured by the beta parameters. Yet note that, because I am not disregarding the problem of endogenous preferences, this time the outcome variable is potentially affected not only by colliders (μ_i) but also by embedding factors (η_i). The latter must thus be regarded as an additional potential source of selection bias.

Simulation Experiment

To illustrate the implications of the selection problem thus posed, I conduct a Monte Carlo simulation experiment based on 10,000 trials using the following generative model:

1. The true instrument (T_i^*) (only observed for migrants) has the standard normal distribution.

2. Both destination networks, μ_i, and all other unobserved determinants of migration, V_i', have the standard normal distribution (as in Chou's experiment).

3. Embeddedness, η_i, also has the standard normal distribution.

4. Propensity to migrate is generated by the model represented in Equation 2 ($\Pr(M = 1) = \Pr(\beta_0 T_i + \beta_1 \mu_i + V_i' > 0)$), where the beta parameters are fixed at the exact means provided by Chou ($\beta_0 = -2$; $\beta_1 = 2$), assuming, as he does, that more traditional women are less likely to migrate, whereas women with high access to networks at destination are more likely to migrate.[3]

5. Finally, labor force participation is a deterministic function of both networks and embeddedness as shown in Equation 3. Specifically, $Y_i = 1$ when $\eta_i + \mu_i > 0$.

The results of this simulation are generated using Ox version 7.00 (see Doornik 2012) and shown in Figure S1, where I present false positive frequencies (type-I error at 5 percent level *t*-tests) under sample sizes 500, 1,000, and 5,000 for four different scenarios.[4] The upper-left panel of Figure S1 plots type-I error rates for standard OLS estimates of traditionalism (T_i) over the full sample. Note that in sharp contrast to Chou's results, my experiment yields a rejection frequency of 100 percent for standard OLS

---

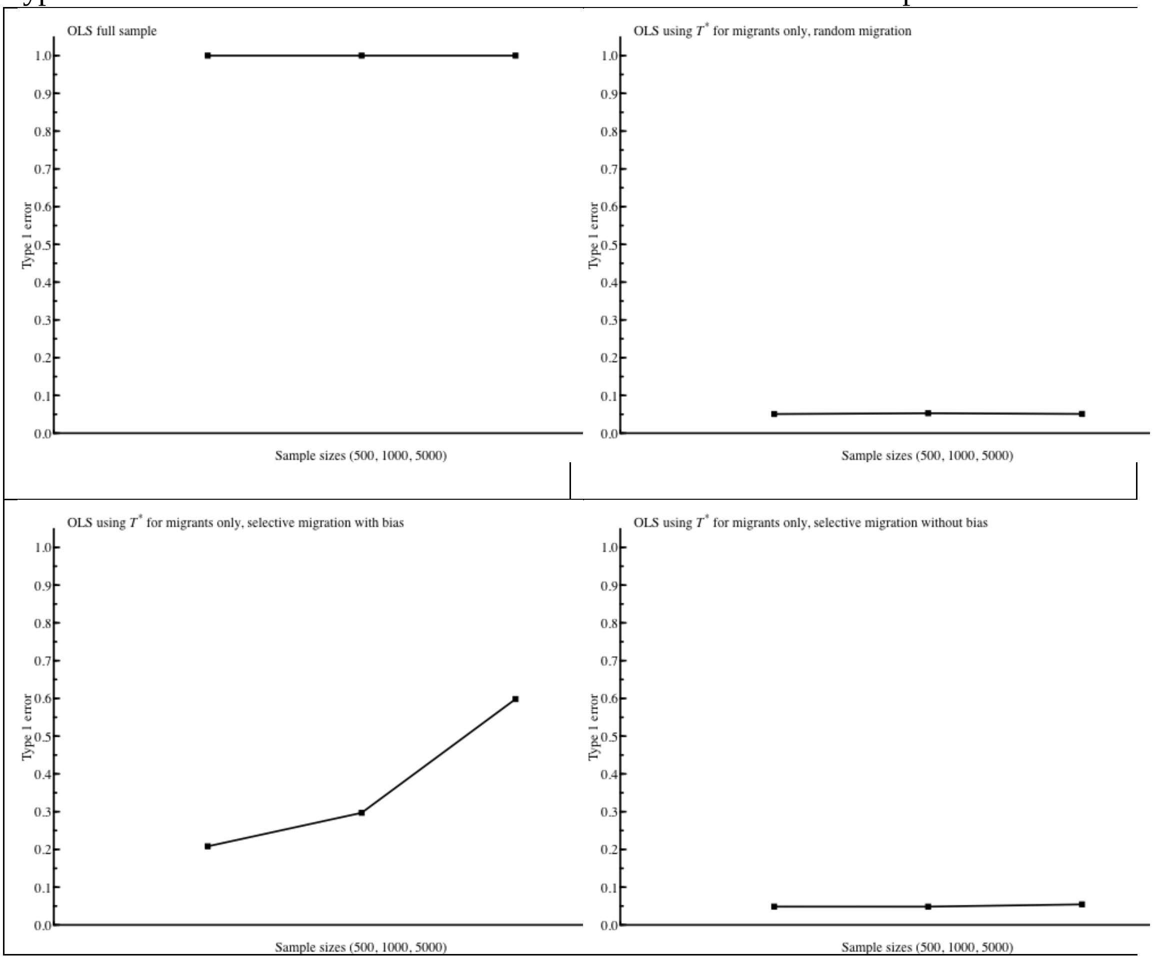
[3] Note that both of these assumptions are questionable. As explained in Polavieja (2015), more traditional women might actually be more likely to migrate than non-traditional women if migration is driven by family reunification. Likewise, dense networks at destination could hinder (not favor) FLFP if they are composed of traditional co-ethnics. The direction of potential selection bias is thus far from obvious.

[4] I wish to thank Sophocles Mavroeidis for his invaluable help in programming these simulations. Simulation commands are presented in Part 4.1.

estimates of traditionalism. This is because, contrary to Chou's setup, my formalization does not ignore the problem of endogenous preferences (i.e., because T_i is endogenous for non-migrating observations).

The upper-right panel of Figure S1 plots simulation results under purely random migration. Logically, if there is no selection bias, the rate of type-I error for instrument T_i^* estimates is 5 percent (i.e., regression estimates should recover the true effect of the instrument in 95 percent of trials). Note that T_i^* is the formal equivalent to SISTER estimates in this setup.

Figure S1. Simulation Experiments Accounting for Social Embeddedness: Probability of Type-I Error under Different Theoretical Scenarios and Different Sample Sizes



But what happens under non-random migration? This crucially depends on the migration rule. The lower panels of Figure S1 plot false positive rates for T_i^* under two different scenarios. In the first scenario, presented in the lower-left panel, the migration decision is determined not only by networks and traditionalism, as in Chou, but also by embeddedness. All three effects add to selection bias. In this case, the rate of type-I error for T_i^* is very large (even larger than the one obtained in Chou's experiment) but still considerably lower than the one obtained for standard OLS estimates for the full sample.

In the second scenario, the migration decision is influenced by the same three factors (networks, traditionalism, and embeddedness), but in this case their effects do not add to selection bias. This latter scenario reveals an interesting implication: namely, the fact that there is selection bias does not necessarily invalidate the instrument. In other words, we might have selective migration but no estimation bias. How is this possible? Simply because accounting for embeddedness also opens up the (theoretical) possibility that η_i interacts with the selection parameters, β_0 and β_1, in such a way that their potentially biasing effects cancel each other out. In this case, estimates would be identical to those obtained under purely random migration, as shown in Figure S1.

In summary, this simulation experiment reveals that (1) even under extreme forms of non-random selection, SISTER estimates can be less biased than standard OLS estimates (because the latter are endogenous); and (2) the potentially biasing impact of selective migration is theoretically undetermined as it depends on how embedding factors, colliding factors, and women's own traditionalism interact with each other in influencing the migration decision. In other words, once we account for endogeneity in the full (non-selected) sample, it is possible to have situations in which endogeneity and selection bias work in opposite directions, thus leading to very different results from those obtained in Chou's simulation experiment. This means simulation experiments cannot shed definitive light on the potential biasing impact of selective migration.

4.1. SIMULATION COMMANDS

This do-file is for Ox, version 7.00 (see Doornik 2012).

```
#include <oxstd.oxh>
#include <oxdraw.oxh>

main()
{
decl vn = <500;1000;5000>;              // sample sizes
decl cn = sizerc(vn);                   // # of different sample sizes

// The following variables, T to M, are of dimention n x 1
decl T;                                 // Traditionalism
decl Tstar;                                         // Exogenous component of traditionalism
decl eta;                               // factors that affect both outcome and traditionalism
decl U;                                 // factors that affect both outcome and migration decision
decl v;                                 // factors that only affect the migration decision
decl Y;                                 // outcome
decl M;                                 // migration indicator variable: M= 1 if migrant 0 otherwise, random migration
decl Mwithout;                          // migration indicator variable, selective migration without bias
decl Mwith;                             // migration indicator variable, selective migration with bias

            // The following are parameters (coefficients, non-random)
decl beta0          = -2;                           // determines effect of traditionalism on prob to migrate
decl beta1          =  2;                           // determines effect of U on prob to migrate
            // variables that hold simulation output
decl crep = 10000;                      // # of Monte Carlo replications
decl mtfull   = new matrix[cn][crep];   // t statistic for the null hypothesis beta=0 in regression of Y on T using all data
decl mtmigr   = new matrix[cn][crep];   // t statistic for the null hypothesis beta=0 in regression of Y on Tstar using migrants,
random migration
decl mtmigrwithout   = new matrix[cn][crep];        // t statistic for the null hypothesis beta=0 in regression of Y on Tstar using
migrants, selective migration without bias
decl mtmigrwith   = new matrix[cn][crep];                       // t statistic for the null hypothesis beta=0 in regression of Y on Tstar
using migrants, selective migration with bias
            parallel for(decl i = 0; i < crep; ++i)         // run crep simulations
{           // For each replication i,
            // first, generate the data:
Tstar       = 12et(max(vn), 1);         // draw data over maximum sample
eta         = rann(max(vn), 1);         // >>
U           = rann(max(vn), 1);
v           = rann(max(vn), 1);
T           = Tstar + eta;
Y           = eta + U .> 0;             // labor 12eter participation 12eterosc: Y=1 if eta+U > 0 or Y=0 otherise.
M           = v .> 0;                   // migration decision: M=1 if beta0*T+beta1*U+v > 0 or M=0 otherise.
Mwithout= beta0*T+beta1*U+v .> 0;       // migration decision: M=1 if beta0*T+beta1*U+v > 0 or M=0 otherise.
Decl b0 = beta0+12et(1,1);              // random draw of coefficient on T in migration decition equation
decl b1 = beta1+12et(1,1);              // random draw of coefficient on U in migration decition equation
Mwith       = b0*T+b1*U+v .> 0;                     // migration decision: M=1 if beta0*T+beta1*U+v > 0 or M=0 otherise.

For(decl j = 0; j < cn; ++j)            // produce results for different sample sizes
{
            // select sample size
decl n = vn[j];                         // sample size j from vn
decl Yn = Y[:n-1][];                    // select the first n rows of Y
decl Tn = T[:n-1][];
decl Mn = M[:n-1][];
```

```
decl Mwithoutn = Mwithout[:n-1][];
decl Mwithn = Mwith[:n-1][];
decl Tstarn = Tstar[:n-1][];
                                        // first, compute the OLS estimator using T over the full sample
decl coefs, xxinv, sterr;               // auxiliary variable to compute regression output
olsc(Yn, 1~Tn, &coefs, &xxinv);         // run regression of Y on T, obtain estimates coefs, and sum of squares of regressor
sterr = sqrt(varc(Y)*xxinv[1][1]);                // standard error (under the null hypothesis that coef = 0, so no
13eteroscedasticity)
mtfull[j][i] = coefs[1]/sterr;          // store t statistic
                                        // next, compute OLS estimators using Tstar for migrants. This depends on migration rule
        // start with random migration
decl Ym = selectifr(Yn, Mn);            // outcome of migrants only
decl Tstarm = selectifr(Tstarn, Mn);    // Tstar of migrants only

olsc(Ym, 1~Tstarm, &coefs, &xxinv);     // run regression of Ym on Tstarm, obtain estimates coefs, and sum of squares of
regressor
sterr = sqrt(varc(Ym)*xxinv[1][1]);     // standard error (under the null hypothesis that coef = 0, so no 13eteroscedasticity)
mtmigr[j][i] = coefs[1]/sterr;          // t statistic
        // then, selective migration without selection bias
Ym = selectifr(Yn, Mwithoutn);          // outcome of migrants only
Tstarm = selectifr(Tstarn, Mwithoutn);  // Tstar of migrants only

olsc(Ym, 1~Tstarm, &coefs, &xxinv);     // run regression of Ym on Tstarm, obtain estimates coefs, and sum of squares of
regressor
sterr = sqrt(varc(Ym)*xxinv[1][1]);     // standard error (under the null hypothesis that coef = 0, so no 13eteroscedasticity)
mtmigrwithout[j][i] = coefs[1]/sterr;   // t statistic
        // finally, selective migration with selection bias
Ym = selectifr(Yn, Mwithn);             // outcome of migrants only
Tstarm = selectifr(Tstarn, Mwithn);     // Tstar of migrants only
olsc(Ym, 1~Tstarm, &coefs, &xxinv);     // run regression of Ym on Tstarm, obtain estimates coefs, and sum of squares of
regressor
sterr = sqrt(varc(Ym)*xxinv[1][1]);     // standard error (under the null hypothesis that coef = 0, so no 13eteroscedasticity)
mtmigrwith[j][i] = coefs[1]/sterr;                // t statistic

}
}
decl RejFreqFull = meanr(fabs(mtfull) .> 1.96);             // rejection frequency of 5% level t test in full sample
decl RejFreqMigr = meanr(fabs(mtmigr) .> 1.96);             // rejection frequency (type I error) of 5% level t test for migrants,
random migration
decl RejFreqMigrWithout = meanr(fabs(mtmigrwithout) .> 1.96);       // rejection frequency (type I error) of 5% level t test for
migrants, selective migration without bias
decl RejFreqMigrWith = meanr(fabs(mtmigrwith) .> 1.96);     // rejection frequency (type I error) of 5% level t test for migrants,
selective migration with bias
        // print results in table
println("%r", {"n=500", "n=1000", "n=5000"}, "%c", {"OLS full", "Tstar random", "Tstar without", "Tstar with"},
RejFreqFull~RejFreqMigr~RejFreqMigrWithout~RejFreqMigrWith);
        // plot results in separate graphs
SetDraw(SET_LINE, 1, 0, 30);
SetDraw(SET_SYMBOL, 1, 0, 100);
for(decl i = 0; i < 4; ++i)                                 // plot each of the four graphs separately
{
decl sname = i==0 ? "OLS_full_sample" : i==1 ? "Tstar_random" :
        i == 2 ? "Tstar_without" : "Tstar_with";            // file name to store each graph
        decl stitle = i==0 ? "OLS full sample" :
        i==1 ? "OLS using $T^*$ for migrants only, random migration" :
        i==2 ? "OLS using $T^*$ for migrants only, selective migration without bias"
        : "OLS using $T^*$ for migrants only, selective migration with bias";     // title for each graph
decl result = i==0 ? RejFreqFull : i==1 ? RejFreqMigr : i==2 ? RejFreqMigrWithout : RejFreqMigrWith;
```

```
DrawTitle(0, stitle);
DrawText(0, “Type 1 error”, -1, -1,
  -1, -1, TEXT_YLABEL, 90,-1);
DrawText(0, “Sample sizes (500, 1000, 5000)”, -1, -1,
   -1, -1, TEXT_XLABEL, 0,-1);
DrawAxis(0, /*iIsXaxis*/0, /*dAnchor*/0, /*dAxmin*/0, /*dAxmax*/1.049, /*dFirstLarge*/0,
   /*dLargeStep*/0.1, /*dSmallStep*/.05, /*iFreq*/0);
DrawAxis(0, /*iIsXaxis*/1, /*dAnchor*/0, /*dAxmin*/0, /*dAxmax*/4, /*dFirstLarge*/5,
   /*dLargeStep*/5, /*dSmallStep*/5, /*iFreq*/0);
//                      DrawAdjust(ADJ_PAPERSCALE, 150);
                        DrawMatrix(0, result’, “", 1, 1, 2, 1);
                        ShowDrawWindow();
                        SaveDrawWindow(sprint(sname,".png"));            // save graph as eps (use .eps, .pdf for different formats)
```

References

Andrews, Donald W. K., and James H. Stock 2005. “Inference with Weak Instruments.” NBER Technical Working Paper No. 313, August (http://www.nber.org/papers/t0313.pdf).

Chou, Wiston. 2016. “Culture Remains Elusive: Comment on Polavieja, *ASR*, February 2015.” Preprint. Retrieved July 20, 2016 (http://www.princeton.edu/~wchou/asr-comment/preprint.pdf).

Chou, Wiston. 2017. Online supplement to “Culture Remains Elusive: Comment on Identification of Cultural Effects using Instrumental Variables.” *American Sociological Review* 82(2) (http://asr.sagepub.com/supplemental).

Doornik, Jurgen A. 2012. *Object-Oriented Matrix Programming Using Ox*, 3rd ed. London: Timberlake Consultants Press and Oxford (http://www.doornik.com).

Hausman, Jerry A. 1978. “Specification Tests in Econometrics.” *Econometrica* 46(6):1251–71.

Polavieja, Javier G. 2015. “Capturing Culture: A New Method to Estimate Exogenous Cultural Effects using Migrant Populations.” *American Sociological Review* 80(1):166–91.

Stock, James H., and Motohiro Yogo. 2005. “Testing for Weak Instruments in Linear IV Regression.” Pp. 80–108 in *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, edited by D. W. K. Andrews and J. H. Stock. Cambridge, UK: Cambridge University Press.