Alex Miller

Problem Set 2

1.

(a)

Given the moment generating condition of z’eiv = 0 and the formula eiv = y – xbiv, the instrumental variable estimator for β can be derived in the following way:

z’eiv = 0;

z’(y-xbiv) = 0;

z’y – z’xbiv = 0;

biv = (z’x)-1 \* (z’y);

where x is a matrix of the dimensions N x K+1, z is a matrix of the dimensions N x K+1, y is a matrix of the dimensions N x 1, and eiv is a matrix of the dimensions N x 1.

Our moment generating condition is conformable since z’ has the dimensions K+1 x N and eiv has the dimensions N x 1, so their product will be equal to K+1 x N. Likewise, our biv estimator is conformable because z’ has the dimensions K+1 x N and x has the dimensions N x K+1, so their product and its inverse will be of the dimensions K+1 x K+1; the z’y portion is also conformable and will be of the dimensions K+1 x 1, so the product of the two halves of our biv is conformable and the product will be K+1 x 1.

There is an implied relationship between each column of z and zero because we are using z to “fix” the non-zero relationship we believe one column of or x matrix has with the error term. The columns of z that are borrowed from the exogenous portions of x (x-k) have an implied relationship with zero because they do not have a casual relationship with the error matrix. The xk column of x does not have that same relationship with zero, which is why we substitute in an exogenous instrumental variable for it in zk.

(b)

The relevancy equation for an instrumental variable is as follows:

xk = x-kδ + zkθk + r

where xk is the endogenous variable, x-k are the exogenous variables of x, δ represents the set of coefficients on each column of x-k, zk is the instrumental variable such that E[z’ϵ]=0 and has no direct causal effect on y, θk is the coefficient on zk and represents the ability of the instrumental variable to explain exogenous variation in xk, and r is the error (and the endogenous portion of xk).

The hypothesis test used to see if zk is a relevant instrumental variable is an F-test of θk with the following hypotheses:

H0: θk = 0

H1: θk ≠ 0

If the p-value for the F-test is significant, then zk can explain exogenous variation in xk. The general rule of thumb for determining if it is a strong instrument is if the F-statistic is greater than 10.

(c)

var(biv) = E[(biv – E(biv))\*(biv – E(biv))’]

var(biv) = E[((z’x)-1(z’y)- β)\*((z’x)-1(z’y)- β)’]

Let a = (z’x)-1(z’y) - β

a = (z’x)-1(z’(xβ + ϵ)) - β

a = (z’x)-1(z’x)β + (z’x)-1(z’ϵ) - β

a = β + (z’x)-1(z’ϵ) - β

a = (z’x)-1(z’ϵ)

var(biv) = E[(z’x)-1z’ϵϵ’z(z’x)-1]; assuming (z’x)-1 is symmetric

var(biv) = (z’x)-1z’E[ϵϵ’z(z’x)-1]

Let VNxN = E[ϵϵ’] = Var(ϵ)

var(biv) = (z’x)-1z’Vz(z’x)-1

In order for our standard errors to scale with xk, we need to substitute xhat for z:

var(biv) = (xhat’x)-1xhat’Vxhat(xhat’x)-1

And STATA also utilizes a sample size adjustment:

var(biv) = (xhat’x)-1xhat’Vxhat(xhat’x)-1 \* (N/N-K)

2.

(a) Before estimating a model, all rows with missing data and rows for women with no education or who were never married were dropped from the sample. Missing data was dropped to simplify calculations in MATA, and unmarried women or women with no education were dropped to help normalize the variable distributions. This was accomplished with the following commands:

**. foreach v of var agefbrth educ knowmeth usemeth agefm idlnchld > electric radio tv urban age frsthalf heduc {**

**> drop if missing(`v')**

**>}**

**. drop if educ0==1**

**. drop if evermarr==0**

After those steps, the following model was used to estimate the relationship between the age a woman first gave birth and her level of education:

*(Age of first birth)= β0 + β1\*(education) + β2\*(know birth control dummy) + β3\*(use birth control dummy) + β4\*(age first married) + β5\*(ideal number of children) + β6\*(electricity dummy) + β7\*(radio dummy) + β8\*(TV dummy) + β9\*(Urban dummy) + β10\*(age)*

In other words, we would estimate that the age a woman first gives birth is going to be dependent on her education, knowledge of use of birth control, age when first married, her ideal number of children (as a measure of desire for children), an urban/rural dummy variable, her age (as a measure of generational culture change), and variables related to standard of living (electricity, radio, and television).

The STATA command to estimate this relationship (assuming exogeneity and homoskedasticity) is as follows:

**. reg agefbrth educ knowmeth usemeth agefm idlnchld electric radio tv urban age, robust**

which returns the following results:

**Linear regression Number of obs = 1257**

**F( 10, 1246) = 33.19**

**Prob > F = 0.0000**

**R-squared = 0.3227**

**Root MSE = 2.4949**

**----------------------------------------------------------------**

**| Robust**

**agefbrth | Coef. Std. Err. t P>|t|**

**-------------+--------------------------------------------------**

**educ | .2872645 .0321435 8.94 0.000**

**knowmeth | -.5889125 .5650983 -1.04 0.298**

**usemeth | -.4595086 .1849076 -2.49 0.013**

**agefm | .180416 .0219722 8.21 0.000**

**idlnchld | -.116089 .0364084 -3.19 0.001**

**electric | .4651771 .2201804 2.11 0.035**

**radio | -.4574288 .1731194 -2.64 0.008**

**tv | .2211023 .2950232 0.75 0.454**

**urban | -.2205582 .150136 -1.47 0.142**

**age | .0944838 .0117172 8.06 0.000**

**\_cons | 12.3325 .8151251 15.13 0.000**

**----------------------------------------------------------------**

Since this is an OLS estimation, the assumption E[x’ϵ] = 0 must hold true in order to have consistent estimators. As long as all of our x variables are exogenous, this will be true, but this assumption will be violated if there are unobserved variables that have a causal effect on y and are correlated with x or if the relationship between x and y cannot be disentangled due to simultaneity. Either the unobserved variables or the inverse equation, if not correctly accounted for, will be placed into the error structure of our model and will violate our core assumption since they’re correlated with x.

Interpreting the results of this regression, we can see that OLS is estimating the relationship between education and the age of first birth as strong, positive, and significant. For every additional year of education, our bOLS estimates that a woman will wait over 3 months more before having a child. Most of the control variables act as we would expect, but interestingly enough the variables for birth control were estimated to have a negative effect on the age of first birth.

(b)

In the model above, I hypothesize that education is an endogenous variable due to simultaneity. The level of education a woman receives can influence the age at which she deices to have a child, but having a child early can also negatively affect a woman’s ability to pursue education. This second relationship with education as the dependent variable and age of first birth being the independent variable would substitute into our first model, and the error term of the second model will enter the error structure of our first model. Since the error of the second model may contain exogenous factors that will affect both education and age of first birth or be correlated with the error of our first model, this induces bias.

(c)

The first instrumental variable candidate we will consider will be a dummy variable that is equal to 1 when the woman was born in the first six months of the year (from here on referred to as “first half”). In order to be a good instrumental variable, the first assumption is E[z’ϵ] = 0. I believe that the first half dummy fits this description because it is difficult to intuitively explain why a woman’s birth month would be correlated with age at first birth or with variables in the error structure.

The second condition is relevancy; we’ll test first half’s relevancy in part d, but for right now it will intuitively be relevant because a girl born in the first six months of the year will be younger than her classmates (if eligibility for school is determined around mid-year). This difference in age will lead girls born in the first 6 months of the year to perform worse, and ultimately seek less education.

The third condition is the exclusion restriction, or that z has a causal effect on y only though x. Again, there is no reason to believe that the month a woman was born will have a direct causal effect on the age a woman first gives birth.

(d)

An instrumental variable approach was used to estimate the relationship for our first model, with first half as our instrument for the endogenous variable education. This was done with the following command in STATA:

**. ivreg agefbrth knowmeth usemeth agefm idlnchld electric radio tv urban age (educ=frsthalf), robust**

which returned the following result:

**Instrumental variables (2SLS) regression Number of obs = 1257**

**F( 10, 1246) = 19.23**

**Prob > F = 0.0000**

**R-squared = .**

**Root MSE = 3.068**

**----------------------------------------------------------------**

**| Robust**

**agefbrth | Coef. Std. Err. t P>|t|**

**-------------+--------------------------------------------------**

**educ | -.4061488 .5222373 -0.78 0.437**

**knowmeth | -.5724511 .5791272 -0.99 0.323**

**usemeth | .184628 .5212286 0.35 0.723**

**agefm | .2339232 .045602 5.13 0.000**

**idlnchld | -.2914252 .1404728 -2.07 0.038**

**electric | 1.374126 .7428649 1.85 0.065**

**radio | .2064404 .5356958 0.39 0.700**

**tv | 2.443559 1.716088 1.42 0.155**

**urban | -.2090493 .1842827 -1.13 0.257**

**age | .029316 .051578 0.57 0.570**

**\_cons | 17.44933 4.016713 4.34 0.000**

**----------------------------------------------------------------**

Additionally, the relevance of our instrumental variable was tested with the following commands:

**. reg educ frsthalf knowmeth usemeth agefm idlnchld electric radio tv urban age**

**. test frsthalf=0**

Which returned the following results:

**Source | SS df MS Number of obs = 1257**

**-------------+--------------------------- F( 10, 1246) = 86.71**

**Model | 5724.98514 10 572.498 Prob > F = 0.0000**

**Residual | 8226.26705 1246 6.60214 R-squared = 0.4104**

**-------------+--------------------------- Adj R-squared = 0.4056**

**Total | 13951.2522 1256 11.1076 Root MSE = 2.5695**

**----------------------------------------------------------------**

**educ | Coef. Std. Err. t P>|t|**

**-------------+--------------------------------------------------**

**frsthalf | -.3408922 .1470156 -2.32 0.021**

**knowmeth | .1166217 .8806162 0.13 0.895**

**usemeth | .9401642 .1752231 5.37 0.000**

**agefm | .0792706 .0162945 4.86 0.000**

**idlnchld | -.2488002 .0385749 -6.45 0.000**

**electric | 1.291074 .2244566 5.75 0.000**

**radio | .9365029 .1796187 5.21 0.000**

**tv | 3.183781 .2551467 12.48 0.000**

**urban | .0157911 .1629464 0.10 0.923**

**age | -.0938992 .0109514 -8.57 0.000**

**\_cons | 7.420528 .9853358 7.53 0.000**

**----------------------------------------------------------------**

**( 1) frsthalf = 0**

**F( 1, 1246) = 5.38**

**Prob > F = 0.0206**

Starting our interpretation of the relevancy test first, the F-statistic testing whether the coefficient on first half is actually equal to zero was calculated as 5.38 with an associated p-value of about 0.02. This is a statistically significant result, and while it’s not as strong as recommended (an F-stat of 10 or greater), it’s strong enough for our purposes.

Next, interpreting the differences in parameters from OLS to our IV regression, the biggest notable difference is the sign change on the education coefficient. The education coefficient was a highly significant 0.287, but now it has been reduced statistically to zero (a not significant -0.406). Besides that, the use of birth control had a relatively significant negative effect on age of first birth (which runs counter to the general idea of birth control), but now has an appropriate positive sign (not significant, however). The coefficient on the radio dummy also now makes more sense intuitively for the same reasons. Otherwise, most other variables kept their sign, but lost significance, reducing the explanatory power of these given independent variables vastly compared to OLS.

(e)

Replication of the canned OLS and IV regression parameters is possible through the following commands in MATA:

**. mata**

**: mata clear**

**: function inv(X)**

**{**

**return(issymmetric(X)?invsym(X):luinv(X))**

**}**

A function called “inv” was defined to avoid confusion between symmetric and asymmetric inverses.

**: y=st\_data(.,("agefbrth"))**

**: x=st\_data(.,("educ", "knowmeth", "usemeth", "agefm", "idlnchld", "electric", "radio", "tv", "urban", "age"))**

**: x=(J(rows(x),1,1),x);**

**: n = rows(x)**

**: k = cols(x)**

The data were imported from STATA, assigned to matrices, and our N and K values were defined.

**/\*OLS\*/**

**: b = inv(x'x)\*(x'y);**

**: b**

**: yhat = x\*b;**

**: ehat = y-yhat;**

**: e2 = ehat\*ehat'**

**: vhat=diag(e2)**

**: v\_robust = inv(x'x)\*x'\*vhat\*x\*inv(x'x)\*(n/(n-k))**

**: se\_robust = sqrt(diagonal(v\_robust))**

**: se\_robust**

A normal OLS regression was performed in MATA, assuming that there is heteroskedasticity and exogenous x’s.

**/\*IVREG\*/**

**: xsansk=st\_data(.,("knowmeth", "usemeth", "agefm", "idlnchld", "electric", "radio", "tv", "urban", "age"))**

**: z=(J(rows(xsansk),1,1),st\_data(.,("frsthalf")),xsansk);**

Here x-k and z are defined based upon the data in x and our instrumental variable.

**: biv = inv(z'x)\*(z'y);**

**: biv**

**: yhativ = x\*biv;**

**: ehativ = y-yhativ;**

**: e2iv = ehativ\*ehativ'**

**: vhativ=diag(e2iv)**

**: xk = x[.,2]**

**: theta = inv(z'z)\*(z'xk)**

**: xkhat = z\*theta**

**: xhat = (J(rows(x),1,1),xkhat,xsansk);**

**: v\_robustiv = inv(xhat'x)\*xhat'\*vhativ\*xhat\*inv(xhat'x)\*(n/(n-k))**

**: se\_robustiv = sqrt(diagonal(v\_robustiv))**

**: se\_robustiv**

The OLS estimation in MATA returned a b matrix consisting of:

**+----------------+**

**1 | 12.33250167 |**

**2 | .2872645038 |**

**3 | -.5889124654 |**

**4 | -.4595085907 |**

**5 | .1804160116 |**

**6 | -.1160889879 |**

**7 | .4651771453 |**

**8 | -.4574288098 |**

**9 | .2211022531 |**

**10 | -.2205581876 |**

**11 | .0944838111 |**

**+----------------+**

with robust standard errors:

**+---------------+**

**1 | .8151251257 |**

**2 | .0321434972 |**

**3 | .5650982889 |**

**4 | .1849075668 |**

**5 | .0219721935 |**

**6 | .0364083872 |**

**7 | .2201803805 |**

**8 | .1731194257 |**

**9 | .2950231905 |**

**10 | .1501359769 |**

**11 | .0117171964 |**

**+---------------+**

The IV regression estimation in MATA returned a biv matrix consisting of:

**+----------------+**

**1 | 17.44933436 |**

**2 | -.4061488024 |**

**3 | -.5724510632 |**

**4 | .1846280052 |**

**5 | .233923229 |**

**6 | -.2914252214 |**

**7 | 1.374125736 |**

**8 | .2064404333 |**

**9 | 2.443559462 |**

**10 | -.2090492941 |**

**11 | .0293160484 |**

**+----------------+**

with robust standard errors:

**+---------------+**

**1 | 4.016713185 |**

**2 | .5222373347 |**

**3 | .5791271969 |**

**4 | .5212286156 |**

**5 | .0456020101 |**

**6 | .1404727778 |**

**7 | .7428649389 |**

**8 | .5356958408 |**

**9 | 1.71608823 |**

**10 | .1842827232 |**

**11 | .0515780222 |**

**+---------------+**

Both sets of b’s and standard errors matched the STATA results exactly.

(f)

In order to test for endogeneity of xk, a Hausman-Wu test was performed. The Hausman-Wu test recovers the errors from the relevancy equation (which can be seen as the endogenous portion of x), and places those errors in the original regression. Then, an F-test is performed to see if μ (the coefficient on the relevancy errors) is equal to zero. The hypothesis test is as follows:

H0: μ = 0 => Exogeneity

H1: μ≠ 0 => Endogeneity

If μ=0, then what we believe to be the endogenous portion of x is uncorrelated with y, so there is no endogeneity problem. This test was performed in STATA on our data with the following commands:

**. reg educ frsthalf knowmeth usemeth agefm idlnchld electric radio tv urban age**

**. predict r, resid**

**. reg agefbrth educ knowmeth usemeth agefm idlnchld electric radio tv urban age r, robust**

**. test r=0**

which returned the following results:

**Source | SS df MS Number of obs = 1257**

**-------------+-------------------------- F( 10, 1246) = 86.71**

**Model | 5724.98514 10 572.50 Prob > F = 0.0000**

**Residual | 8226.26705 1246 6.6021 R-squared = 0.4104**

**-------------+-------------------------- Adj R-squared = 0.4056**

**Total | 13951.2522 1256 11.108 Root MSE = 2.5695**

**----------------------------------------------------------------**

**educ | Coef. Std. Err. t P>|t|**

**-------------+--------------------------------------------------**

**frsthalf | -.3408922 .1470156 -2.32 0.021**

**knowmeth | .1166217 .8806162 0.13 0.895**

**usemeth | .9401642 .1752231 5.37 0.000**

**agefm | .0792706 .0162945 4.86 0.000**

**idlnchld | -.2488002 .0385749 -6.45 0.000**

**electric | 1.291074 .2244566 5.75 0.000**

**radio | .9365029 .1796187 5.21 0.000**

**tv | 3.183781 .2551467 12.48 0.000**

**urban | .0157911 .1629464 0.10 0.923**

**age | -.0938992 .0109514 -8.57 0.000**

**\_cons | 7.420528 .9853358 7.53 0.000**

**----------------------------------------------------------------**

**Linear regression Number of obs = 1257**

**F( 11, 1245) = 30.85**

**Prob > F = 0.0000**

**R-squared = 0.3242**

**----------------------------------------------------------------**

**| Robust**

**agefbrth | Coef. Std. Err. t P>|t|**

**-------------+--------------------------------------------------**

**educ | -.4061489 .4236948 -0.96 0.338**

**knowmeth | -.5724511 .574847 -1.00 0.320**

**usemeth | .1846281 .4307069 0.43 0.668**

**agefm | .2339232 .0374021 6.25 0.000**

**idlnchld | -.2914253 .1111937 -2.62 0.009**

**electric | 1.374126 .5965466 2.30 0.021**

**radio | .2064406 .4412309 0.47 0.640**

**tv | 2.44356 1.397291 1.75 0.081**

**urban | -.2090493 .1501056 -1.39 0.164**

**age | .029316 .0415861 0.70 0.481**

**r | .6964056 .4246121 1.64 0.101**

**\_cons | 17.44934 3.260704 5.35 0.000**

**----------------------------------------------------------------**

**( 1) r = 0**

**F( 1, 1245) = 2.69**

**Prob > F = 0.1012**

In a strict interpretation, we would fail to reject the null hypothesis of exogeneity, but there still may be endogeneity present since it is so close.

Generally speaking, we cannot test our instrumental variable for exogeneity. If it is relevant and our xk is correlated with y, then it will have some correlation with the error term through that relationship. We typically have to infer the exogeneity of our instrumental variable through intuitive logic. The only exception here is when the model is overidentified, we can see if the instrument violates the orthogonality condition (E[z’ϵ]=0) and is therefore endogenous.

(g)

Although the chances that it does not violate the orthogonality condition are less than that of the first half dummy, I propose that the number of children a woman ultimately has is a potential instrumental variable candidate. We can infer that the more children a woman has, the more her ability to seek further education is hindered, but presumably the number of children that a woman ultimately has is not going to have a causal effect on the age in which she had her first child. One worry here is that there is correlation outside of the relationship with x, since women that bear children early have more time to give birth to further children, but just because a woman can have more children does not mean she will, and hopefully this reverse causal story will not enter our error structure.

Relevancy was tested with the following commands:

**. reg educ frsthalf ceb knowmeth usemeth agefm idlnchld electric radio tv urban age**

**. test frsthalf=0**

**. test ceb=0**

**. test frsthalf ceb**

which returned the following results:

**Source | SS df MS Number of obs = 1257**

**-------------+-------------------------- F( 11, 1245) = 92.80**

**Model | 6285.54127 11 571.41 Prob > F = 0.0000**

**Residual | 7665.71091 1245 6.1572 R-squared = 0.4505**

**-------------+-------------------------- Adj R-squared = 0.4457**

**Total | 13951.2522 1256 11.108 Root MSE = 2.4814**

**----------------------------------------------------------------**

**educ | Coef. Std. Err. t P>|t|**

**-------------+--------------------------------------------------**

**frsthalf | -.2772549 .1421318 -1.95 0.051**

**ceb | -.4490822 .0470661 -9.54 0.000**

**knowmeth | 1.015149 .8556227 1.19 0.236**

**usemeth | 1.172097 .1709526 6.86 0.000**

**agefm | .0484158 .0160647 3.01 0.003**

**idlnchld | -.1459455 .0387806 -3.76 0.000**

**electric | 1.100722 .2176773 5.06 0.000**

**radio | .8279265 .1738334 4.76 0.000**

**tv | 2.936676 .2477563 11.85 0.000**

**urban | -.1146981 .157953 -0.73 0.468**

**age | -.0074368 .0139271 -0.53 0.593**

**\_cons | 5.599283 .9705094 5.77 0.000**

**----------------------------------------------------------------**

**( 1) frsthalf = 0**

**F( 1, 1245) = 3.81**

**Prob > F = 0.0513**

**( 1) ceb = 0**

**F( 1, 1245) = 91.04**

**Prob > F = 0.0000**

**( 1) frsthalf = 0**

**( 2) ceb = 0**

**F( 2, 1245) = 48.40**

**Prob > F = 0.0000**

Where first half was a relatively weak instrumental variable by itself, the number of children ever born proves to be a very strong and very relevant instrumental variable; far surpassing the suggested F-stat of 10. With these two instrumental variables, our new model was estimated with:

**. ivreg agefbrth knowmeth usemeth agefm idlnchld electric radio tv urban age (educ=frsthalf ceb), robust**

and returned:

**Instrumental variables (2SLS) regression Number of obs = 1257**

**F( 10, 1246) = 23.72**

**Prob > F = 0.0000**

**R-squared = .**

**Root MSE = 3.6044**

**----------------------------------------------------------------**

**| Robust**

**agefbrth | Coef. Std. Err. t P>|t|**

**-------------+--------------------------------------------------**

**educ | 1.297538 .1520357 8.53 0.000**

**knowmeth | -.612896 .7887489 -0.78 0.437**

**usemeth | -1.397988 .2888629 -4.84 0.000**

**agefm | .1024582 .0300775 3.41 0.001**

**idlnchld | .1393685 .0622647 2.24 0.025**

**electric | -.8591227 .3729448 -2.30 0.021**

**radio | -1.424658 .2760376 -5.16 0.000**

**tv | -3.016924 .6443323 -4.68 0.000**

**urban | -.2373262 .221269 -1.07 0.284**

**age | .1894305 .0229925 8.24 0.000**

**\_cons | 4.877492 1.567532 3.11 0.002**

**----------------------------------------------------------------**

Interestingly enough, all of the signs of our coefficients that changed from our first IV estimation have returned to the signs of the OLS estimated coefficients. As I’ll demonstrate later, this is probably because we have violated the orthogonality condition and overidentified our model. The second IV has, however, increased the significance of our education coefficient.

(h)

Endogeneity was tested with another Hausman-Wu test conducted with the following commands:

**. reg educ frsthalf ceb knowmeth usemeth agefm idlnchld electric radio tv urban age**

**. predict r, resid**

**. reg agefbrth educ knowmeth usemeth agefm idlnchld electric radio tv urban age r, robust**

**. test r=0**

which returned an F statistic of 89.12 and an associated p-value of 0.0000. Adding the second instrumental variable has perhaps better identified the exogenous portion of x, leaving us with an endogenous residual that is very correlated with y. The test for endogeneity comes out overwhelmingly positive as we reject then null hypothesis of exogeneity.

(i)

The Sargon test for over identification was performed using the ivreg2 command courtesy of Baum, Schaffer, and Stillman. This test is otherwise performed by recovering the residuals from an IV regression, and running a second regression with the residuals as the dependent variable and x-k and z as the independent variables. The test statistic is calculated from the R2 of this regression with the following hypothesis test:

H0: R2 = 0 => Identified

H1: R2≠ 0 => Overidentified

The intuition of this test is that if there is a small R2 value, the exogenous variables in our model cannot explain the variation in the error term and the model is well identified. Otherwise, large R2 values mean that the exogenous variables can explain variation in the error, and the orthogonality condition has been violated. The following command was used to do this test:

**. ivreg2 agefbrth knowmeth usemeth agefm idlnchld electric radio tv urban age (educ=frsthalf ceb)**

The Sargon test portion of these results reported a test statistic of 8.507 and an associated P-value of 0.0035. Here we can reject the null hypothesis and confirm that adding one more instrumental variable does indeed push the model to overidentification.

(j)

The rational for using more than one instrument in an instrumental variable regression is to improve the model efficiency. Adding more data, where appropriate, to a model will always result in more efficient estimators. Here, more instrumental variables will be able to better estimate the exogenous variation in x. But, as we add more instrumental variables, we also increase the risk of violating the E[z’ϵ]=0 condition. In this particular case, adding the number of children ever born did increase the strength of our z, but at the cost of violating a basic assumption.