EERI
Economics and Econometrics Research Institute

A Test for Endogeneity in Regressions

Thomas B. Marvell

EERI Research Paper Series No 05/2025

ISSN: 2031-4892

EERI
Economics and Econometrics Research Institute
Avenue Louise
1050 Brussels
Belgium

Tel: +32 2271 9482
Fax: +32 2271 9480

A TEST FOR ENDOGENEITY IN REGRESSIONS

Thomas B. Marvell

ABSTRACT

Textbook theory predicts that *t*-ratios decline towards zero in regressions when there is increasing collinearity between two independent variables. This article shows that this rarely happens if the two variables are endogenous, and coefficients increase greatly with more collinearity. The purposes of this article are 1) to illustrate this bias and explain why it occurs, and 2) to use the phenomenon to develop a test for endogeneity. For the test, one creates a variable that is highly collinear with the independent variable of interest, and endogeneity is indicated if *t*-ratios do not decline with increasing collinearity. False negatives are possible, but not likely. The test is confirmed with algebraic examples and simulations. I give many empirical examples of the bias and the test, including testing exogeneity assumptions behind instrumental variables and Granger causality.

Keywords: endogeneity, collinearity, simultaneity, omitted variable bias, instrumental variables.

Thomas B. Marvell,[1] 155 Ridings Cove, Virginia, USA 23185. marvell@cox.net

[1] Thomas Marvell is Director of Justec Research, which specializes in criminal justice research.

A Test for Endogeneity in Regressions

Gauss-Markov assumptions require that independent variables (IVs) be exogenous, but there is no test for exogeneity that does not require an assumption of exogeneity, and researchers seldom know whether their regression results are biased. This article proposes a test for endogeneity that does not require exogeneity assumptions by making use of the relationship between collinearity and endogeneity. When endogeneity exists with collinear IVs, as the collinearity increases, with limited exceptions, their *t*-ratios do not decline towards zero. The absolute values of the coefficients and *t*-ratios either increase indefinitely or reach a plateau. Coefficients typically have opposite signs. I call this effect "whipsaw" because the two collinear IVs appear to play off against each other. Researchers regularly encounter these odd effects when regressions contain correlated IV lags or contain an IV and its square, producing coefficients with opposite signs and similar *t*-ratios. The latter is often interpreted as a curvilinear causal effect, but I show that this is at most partially true because increasing collinearity seldom reduces *t*-ratios. Under established theory whipsaw does not occur with exogeneity, so its presence indicates endogeneity. Thus, I develop a test for whether an IV is endogenous by creating a second IV that is highly collinear with the first (henceforth, the first IV is called the "key IV" and the second the "collinear IV"). If whipsaw occurs, there is endogeneity.

As used here, collinearity means high correlation between two variables, and correlations close to one (e.g., .99999) are called "extreme collinearity." Endogeneity is simultaneity or omitted variable bias (OVB). Simultaneity occurs when the dependent variable (DV) causes an IV. (Since the subject is regression analysis, "cause" simply means that a change in one variable leads a change in another.) It can be divided into three subtypes, which can coexist: 1) direct simultaneity where the DV causes the IV directly, 2) indirect simultaneity through other IVs in the regression, and 3) indirect simultaneity where the DV causes an omitted variable which causes the IV (called "OV simultaneity"). OVB occurs when an omitted variable causes both the IV and the DV. OVB and OV simultaneity can operate backwards in time. An omitted variable can cause the DV and cause an IV at a later time, and the DV can cause an omitted variable that causes an IV at a later time.

The outline of the article is as follows. The next section discusses relevant theory and research concerning collinearity with endogeneity. Section 3 gives algebraic examples of whipsaw. Section 4 illustrates whipsaw with simulations. Section 5 contains empirical examples of whipsaw and the whipsaw test: entering an IV and its square, using unit trends in panel regressions, and examining exogeneity assumptions when using instrumental variables and the Granger causation test. Section 6 concludes.

1. BACKGROUND

When simultaneity exists, regressions cannot separate the "forward" and "backward" causal impacts, mathematical calculations are impossible, and regression coefficients are biased (Harnack et al. 2017). Similarly, omitted variables prevent one from disentangling a regression coefficient from the impact of omitted variables, which of course are not entered in the regression equation, making mathematical calculations impossible. In both cases the error term is correlated with an IV, violating Gauss-Markov assumptions.

The Durbin–Wu–Hausman test, the only common test for endogeneity in regressions, entails comparing an OLS coefficient to that of a 2SLS regression in which instruments are used for the possibly endogenous IV. If the results differ there is endogeneity. The difficulty is that finding valid instruments is difficult because they must be exogenous to the DV (Hernam & Robins 2016; Wooldridge 2021). That is, they cannot be caused by the DV, there cannot be OV simultaneity, and there cannot be omitted variables that cause the DV and the instrument. These requirements can rarely be empirically verified. In contrast the proposed whipsaw test makes no such assumptions.

Likewise, strategies for curing or mitigating endogeneity in regressions require exogeneity assumptions, such that the strategies have limited utility if one cannot test the assumptions. Using lagged IVs, such as the Granger test, under the assumptions that causation cannot go backward in time (Granger 1969, Peters et al. 2017, Hoover 2001), encounters potential endogeneity problems due to OVB and OV simultaneity (Peters et al. 2017; Maziarz 2017). 2SLS uses instrumental variables that, among other problems, must be exogenous, and thus encounter the problems listed above. The required exogeneity is rarely proven, supported only by theory or general knowledge (Martens et al. 2006). The regression discontinuity design evaluates programs where there is artificial discontinuity in the treatment application. Simultaneity is possible because subjects can tailor their efforts to reach, or not reach, the threshold, and OVB is possible because other changes related to treatment effects might occur near the threshold (Imbens & Lemieux 2008). Using proxies for omitted variables is inexact and impossible if the omitted variables are unknown.

Traditional theory states that when two IVs become more collinear, standard errors grow and *t*-ratios decline towards zero (Belsley et al. 1980; Wooldridge 2021). Less well known, theory also predicts that collinearity biases coefficients; they increase with more collinearity, generally moving towards opposite signs (Hill & Atkins 2001). The supposition that collinearity always reduces *t*-ratios is often questioned. Spanos and McGuirk (2002) found through simulations that data characteristics can lead to erratic changes in coefficient and *t*-ratios as collinearity increases. They suggested that this uncertainty can lead to higher *t*-ratios with more collinearity. Mela and Kopalle (2002) argue that collinearity sometimes reduces coefficient variance when the two IVs are negatively correlated. Marvell (2010) suggests that coefficient sizes and *t*-ratios increase as collinearity increases when there is reciprocal causation, using crime rates and prison population as an example. Winship and Western (2016) show that the significance levels of collinear IVs can be exaggerated when the regression is misspecified. Goodhue et al (2017) and Kolmos (2020) argue that the combination of collinearity and errors-in-variables causes Type 1 errors. Chatelain and Ralf (2014) show that a coefficient can be significant even though the IV has no relationship with the DV if it is collinear with another IV, and the two IVs have outlier observations. Kalnins (2018, 2022), conducted a literature review and found that collinear IVs often have sizeable and significant coefficients of opposite signs, when theory would suggest they have the same sign. He argues that this can occur when collinear IVs share a common factor, and the effect is greater when the common factor accounts for more of the correlation between the two. None of these articles have taken the next step and used the failure of *t*-ratios to decline towards zero to test for endogeneity.

The following paragraphs show that with more extreme collinearity: 1) IV coefficients become larger, even with exogenous IVs. 2) Coefficients switch from the same sign to opposite signs, or from opposite signs to the same sign. 3) *T*-ratios decline towards zero without endogeneity. 4) *T*-ratios rarely decline towards zero with endogeneity because coefficients increase, and standard errors may decline or increase at a lower rate than with exogeneity. These are based on three equations.

First, using the model $y=\beta_1 x+\beta_2 z+r$ with Gaussian assumptions, the coefficient on x is:

$$\beta_1 = \left(\sum r_{xi} y_i\right) / \left(\sum r_{xi}^2\right) \quad \text{(Equation 1)}$$

where r_{xi} are the residuals when regressing x on z (plus any other IVs). Their absolute values decline with more collinearity.

Second, coefficients increase and move towards opposite signs as collinearity increases. The covariance of the coefficients is:

$$\mathrm{Cov}(\beta_1, \beta_2) = \frac{-\rho\sigma^2}{(1-\rho^2)\sqrt{\sum(x_i-\bar{x})^2}\sqrt{\sum(z_i-\bar{z})^2}} \quad \text{(Equation 2)}$$

where σ is the regression standard error, ρ is the correlation between x and z. The covariance is negative and increases indefinitely at an increasing rate with larger ρ. Coefficients move towards opposite signs if they have the same signs when entered individually. Their absolute sizes approach each other (but do not converge) and increase indefinitely (Hill & Atkins 2001, Spanos 2019, see table 1 below).

The third equation is textbook theory as to why standard errors increase with more collinearity. The standard error of β_1 is:

$$\sigma_{\beta_1} = \sqrt{\frac{\sum r_i^2}{\sum(x_i-\bar{x})^2(1-\rho^2)\,(n-2)}} \quad \text{(Equation 3)}$$

Where r_i are the regression errors, ρ is the correlation between x and z, and n is the sample size. The standard error increases as ρ increases. With non-OV simultaneity r_i decline because an IV now includes the DV, increasing the portion of DV variability accounted for (if the IV causes the DV). This counteracts the $(1-\rho^2)$ decline, and the standard error declines or increases less than if there were no simultaneity. With OVB and OV simultaneity, r_i include the effect of omitted variables and change in ways that are specific to the data.

2. ALGEBRAIC EXAMPLES

This section and the next (simulations) show how whipsaw works. The algebraic examples are approximations, not mathematical proofs, showing the impact of the first and second equations on coefficients due to endogeneity and extreme collinearity. The examples use the whipsaw test format. As will be seen later in the article, increasing coefficient size is usually the major reason why *t*-ratios fail to decline towards zero with endogeneity and increased collinerity. In the first example, there is reciprocal causation between the DV and IVs.

$$y = \mathrm{c} + \beta_1 x + \beta_2 z + e$$

$$x = c_x + \beta_3 y + e_x$$

$$z = c_z + \beta_4 y + e_z$$

$$y = c + \beta_1(c_x + \beta_3 y + e_x) + \beta_2(c_z + \beta_4 y + e_z) + e$$

$$y = c + \beta_1 c_x + \beta_1\beta_3 y + \beta_2 c_z + \beta_2\beta_4 y + (\beta_1 e_x + \beta_2 e_z + e)$$

$$y = (c + \beta_1 c_x + \beta_2 c_z) + \beta_1\beta_3(\beta_1 x + \beta_2 z + e) + \beta_2\beta_4(\beta_1 x + \beta_2 z + e) + (\mathrm{e} + \beta_1 e_x + \beta_2 e_z)$$

$$y = (c + \beta_1 c_x + \beta_2 c_z) + \beta_1^2\beta_3 x + \beta_1\beta_2\beta_3 z + \beta_1\beta_2\beta_4 x + \beta_2^2\beta_4 z + (e + \beta_1\beta_3 e + \beta_2\beta_4 e + \beta_1 e_x + \beta_2 e_z)$$

$$y = (c + \beta_1 c_x + \beta_2 c_z) + (\beta_1\beta_3 + \beta_2\beta_4)\beta_1 x + (\beta_1\beta_3 + \beta_2\beta_4)\beta_2 z + (e + \beta_1\beta_3 e + \beta_2\beta_4 e + \beta_1 e_x + \beta_2 e_z)$$

When ostensibly conducting the regression model in the first line, one is instead using the model in the last line, biasing c, β_1, β_2, and e. First, with respect to the coefficients:

From Equation 1:

$$(\beta_1\beta_3 + \beta_2\beta_4)\beta_1 = \left(\sum r_{x_i} y_i\right) / \left(\sum r_{x_i}{}^2\right)$$

$$\beta_1 = \frac{\left(\sum r_{x_i}\, y_i\right)}{\left(\sum r_{x_i}{}^2\right)(\beta_1\beta_3 + \beta_2\beta_4)}$$

Since β_1 is on both sides, there is an anomaly, and the equation cannot be solved. Equation 2, however, enables approximations. With more extreme collinearity $\beta_2 \rightarrow \sim -\beta_1$ (this is not a limit in the usual sense), and $\beta_4 \rightarrow \beta_3$. Thus, $(\beta_1\beta_3 + \beta_2\beta_4) \rightarrow \sim 0$, and β_1 increases. In addition, β_1 likely increases when r_{x_i} fall below one. If β_3 or $\beta_4 = 0$, there is no whipsaw even though there is simultaneity. That is, here whipsaw requires that both IVs be functions of *y*. The constant changes because $\beta_1 c_x$ and $\beta_2 c_z$ are added. The error term $[(1 + (\beta_1\beta_3 + \beta_2\beta_4)e + \beta_1 e_x + \beta_2 e_z]$ causes bias because it is correlated with the IVs through e_x and e_z.

In the second example, the DV causes one IV, which causes the second IV.

$y = c + \beta_1 x + \beta_2 z + e$

$x = c_x + \beta_3 y + e_x$

$z = c_z + \beta_4 x + e_z$

$y = c + \beta_1(c_x + \beta_3 y + e_x) + \beta_2(c_z + \beta_4 x + e_z) + e$

$y = (c + \beta_1 c_x + \beta_2 c_z) + \beta_1\beta_3 y + \beta_2\beta_4 x + e + \beta_1 e_x + \beta_2 e_z$

$y = (c + \beta_1 c_x + \beta_2 c_z) + \beta_1\beta_3(\beta_1 x + \beta_2 z + e) + \beta_2\beta_4 x + e + \beta_1 e_x + \beta_2 e_z$

$y = (c + \beta_1 c_x + \beta_2 c_z) + \beta_1\beta_3\beta_1 x + \beta_1\beta_3\beta_2 z + \beta_2\beta_4 x + (e + \beta_1\beta_3 e + \beta_1 e_x + \beta_2 e_z)$

$y = (c + \beta_1 c_x + \beta_2 c_z) + (\beta_1\beta_3\beta_1 + \beta_2\beta_4)x + \beta_1\beta_3\beta_2 z + (e + \beta_1\beta_3 e + \beta_1 e_x + \beta_2 e_z)$

$y = (c + \beta_1 c_x + \beta_2 c_z) + (\beta_1\beta_3 + \beta_2\beta_4/\beta_1)\beta_1 x + \beta_1\beta_3\beta_2 z + (e + \beta_1\beta_3 e + \beta_1 e_x + \beta_2 e_z)$

$$\beta_1 = \frac{(\sum r_{x_i} y_i)}{(\sum r_{x_i}^2)(\beta_1\beta_3 + \beta_2\beta_4/\beta_1)}$$

Applying Equation 2, with more extreme collinearity between x and z, $\beta_2 \rightarrow \sim -\beta_1$, and $\beta_4 \rightarrow 1$. Also, $\beta_1\beta_3=1$ because $(y=c+\beta_1 x+\beta_2 z+e) \rightarrow (y=c+\beta_1\beta_3 y+\beta_2 z+\beta_1 e_x+e)$. Thus, $\beta_1\beta_3+\beta_2\beta_4/\beta_1 \rightarrow \sim 0$. Whipsaw does not occur if $\beta_4=0$ even though there is simultaneity. The constant changes, and the error term is correlated with the IVs.

In Example 3 an omitted variable (o) causes both IVs:

$y = c + \beta_1 x + \beta_2 z + \beta_3 o + e$

$x = c_x + \beta_4 o + e_x$

$z = c_z + \beta_5 o + e_z$

$o = (x - c_x - e_x)/\beta_4$

$y = c + \beta_1(c_x + \beta_4 o + e_x) + \beta_2(c_z + \beta_5 o + e_z) + \beta_3 o + e$

$y = c + \beta_1 c_x + \beta_1\beta_4 o + \beta_1 e_x + \beta_2 c_z + \beta_2\beta_5 o + \beta_2 e_z + \beta_3 o + e$

$y = c + \beta_1 c_x + \beta_1\beta_4(x - c_x - e_x)/\beta_4 + \beta_1 e_x + \beta_2 c_z + \beta_2\beta_5(x - c_x - e_x)/\beta_4 + \beta_2 e_z + \beta_3(x - c_x - e_x)/\beta_4$
$+ e$

$y = c + \beta_1 c_x + \beta_1 x - \beta_1 c_x - \beta_1 e_x + \beta_1 e_x + \beta_2 c_z + \beta_2\beta_5 x/\beta_4 - \beta_2\beta_5 c_x/\beta_4 - \beta_2\beta_5 e_x/\beta_4 + \beta_2 e_z + \beta_3 x/\beta_4$
$- \beta_3 c_x/\beta_4 - \beta_3 e_x/\beta_4 + e$

$y = (c + \beta_2 c_z - \beta_2\beta_5 c_x/\beta_4 - \beta_3 c_x/\beta_4) + \beta_1 x + \beta_2\beta_5 x/\beta_4 + \beta_3 x/\beta_4 + (e + \beta_2 e_z - \beta_1 e_x + \beta_1 e_x - \beta_3 e_x/\beta_4$
$- \beta_2\beta_5 e_x/\beta_4)$

$y = (c + \beta_2 c_z - \beta_2\beta_5 c_x/\beta_4 - \beta_3 c_x/\beta_4) + (1 + \beta_2\beta_5/\beta_1\beta_4 + \beta_3/\beta_1\beta_4)\beta_1 x + (e + \beta_2 e_z - \beta_2\beta_5 e_x/\beta_4 - \beta_3 e_x/\beta_4)$

With more extreme collinearity $\beta_2 \rightarrow\sim -\beta_1$, and $\beta_5 \rightarrow \beta_4$. Also, $\beta_3=0$, creating OVB. Thus $(1 + \beta_2\beta_5/\beta_1\beta_4 + \beta_3/\beta_1\beta_4) \rightarrow\sim (1 - \beta_5/\beta_4 + \beta_3/\beta_1\beta_4) \rightarrow\sim 0$, so β_1 increases. The constant changes, and the error term is correlated with the IVs.

In example 4 the omitted variable causes one IV, and that IV causes the second IV.

$$y = c + \beta_1 x + \beta_2 z + \beta_3 o + e$$

$$x = c_x + \beta_4 o + e_x$$

$$z = c_z + \beta_5 x + e_z$$

$$o = (x - c_x - e_x)/\beta_4$$

$$y = c + \beta_1(c_x + \beta_4 o + e_x) + \beta_2(c_z + \beta_5 x + e_z) + \beta_3(x - c_x - e_x)/\beta_4 + e$$

$$y = c + \beta_1 c_x + \beta_1\beta_4 o + \beta_2 c_z + \beta_2\beta_5 x + (\beta_3/\beta_4)x + (\beta_3/\beta_4)c_x + e + \beta_1 e_x - \beta_3 e_x/\beta_4 + \beta_2 e_z$$

$$y = (c + \beta_1 c_x + \beta_2 c_z + \beta_3 c_x/\beta_4) + \beta_1\beta_4(x - e_x)/\beta_4 + \beta_2\beta_5 x + \beta_3 x/\beta_4 + e - \beta_1 e_x - \beta_3 e_x/\beta_4 + \beta_1 e_x + \beta_2 e_z$$

$$y = (c + \beta_1 c_x + \beta_2 c_z + \beta_3 c_x/\beta_4) + \beta_1 x + \beta_2\beta_5 x + \beta_3 x/\beta_4 + (e - \beta_3 e_x/\beta_4 + \beta_2 e_z)$$

$$y = (c + \beta_1 c_x + \beta_2 c_z + \beta_3 c_x/\beta_4) + \beta_1(1 + \beta_2\beta_5/\beta_1 + \beta_3/\beta_1\beta_4)x + (e - \beta_3 e_x/\beta_4 + \beta_2 e_z)$$

With more extreme collinearity $\beta_2 \rightarrow\sim -\beta_1$, and $\beta_5 \rightarrow 1$. Also, $\beta_3=0$, creating OVB. $(1+ \beta_2\beta_5/\beta_1+\beta_3/\beta_1\beta_4) \rightarrow\sim 0$, so β_1 increases. If $\beta_5= 0$, there is no whipsaw. The constant changes, and the error term is correlated with the IVs.

Four algebraic examples do not establish a pattern, but they do show potential harmful consequences of endogeneity with collinearity between IVs. Coefficients increase greatly with more collinearity. The constants change, such that slopes change, and coefficients are biased. The error term is correlated with the IVs.

3. SIMULATIONS OF SIMULTAINITY AND OVB

The purposes of the simulations are to illustrate further why *t*-ratios fail to approach zero with endogeneity and extreme collinearity and to explain why the whipsaw test can fail to detect endogeneity. The simulations are OLS regressions with a sample size of 10,000, using STATA. The simulation equations are at the end of the tables. The DV is *y,* the key IV is *x*, and the collinear IV is *z*. Lower-case letters are random variables, normally distributed, with means of ten, and standard deviations are separate uniform random variables between one and two, except that *w*, which is used to create whipsaw, has a standard deviation of one.

I create collinearity two ways, the "K method" and the "exponent method." The first is used with simulations and the second with empirical data. In the K method, the collinear IV is the key IV plus K times a random variable (*w*), where K is an increasingly small number in successive regressions. The range here is K=10 to K=.001, and correlations range from near zero to more than .99999. This large range helps illustrate how whipsaw works. The disadvantage is that results vary greatly with different *w*'s unless the number of observations is very large.

In the exponent method the collinear IV is the key IV taken to a power slightly above one, here 1.02, producing correlations of about .99999. The exponent of 1.02 is chosen because results are virtually the same as those with smaller exponents in the empirical examples. Since even an IV and its square are highly correlated, the exponent method does not enable one to compare low collinearity to high collinearity situations.

Table 1 illustrates theory that *t*-ratios decline with more collinearity. Coefficients increase in opposite directions because covariances are negative and increase (Equation 2). The standard errors, proportional to the square root of $\sum r^2/(1-\rho^2)$ (Equation 3), increase because $\sum r^2$ does not change. *T*-ratios decline because standard errors increase more than coefficients increase.

Table 2 is an example of whipsaw where the DV causes one IV, and that IV causes the second. This is an unlikely situation empirically because the IVs do not cause the DV, but it illustrates well the workings of whipsaw. Coefficients and *t*-ratios increase greatly with more collinearity. Of course, the whipsaw effect is less pronounced with fewer observations and with less impact of *y* on *x*. With n=200, x=y*.1+b*.001, and K=.001, the coefficients on both x and z are 254, with standard deviations of 86.

Whipsaw does not occur without *b**.001 because it prevents r_x and r_z from falling towards zero with extreme collinearity. That is, when x=z+r_x, r_x decline to zero, but not when *x*=*z*+*b**.001+r_x, and the denominator in Equation 1 falls at a slower pace with more collinearity than r_xy when r_x fall below one. Whipsaw in table 2 is greater (coefficients and *t*-ratios increase) when .001 is changed to a larger fraction, and it still occurs if .0001 is substituted for .001. Table 2 results are similar when *y**.001 is substituted for *b**.001. Thus, here whipsaw requires that the collinear IV be caused directly or indirectly by the DV independently of the key IV. In empirical studies, one does not know whether this condition occurs, but the fact that the causal relationships need only be very small suggests that the whipsaw test is unlikely to have false negatives.

The main difference between tables 1 and 2 is the greater coefficient increase with more collinearity in table 2 because $\sum r_xy$/n and $\sum r_zy$/n decline less due to the lesser decline in r_x and r_z (again, due to the presence of *b**.001).

Table 3 simulates reciprocal causation between the DV and IVs.[1] The results are similar to table 2, and coefficients increase for the same reason they do in table 2.

Turning to OVB, in table 4 the omitted variable causes the key IV, which causes the collinear IV. Coefficients and *t*-ratios increase indefinitely. Whipsaw again results mainly from the large coefficient increases. There is no whipsaw if z does not include a*.001, but there is whipsaw if o*.001 or y*.001 are substituted for a*.001, analogous to similar findings in table 2. Table 4 results are similar if $o=x+a$, which is OV simultaneity.

OVB occurs when the constant is deleted (and the slope does not go through the origin), which biases coefficients. Table 5 is the regression in table 1 without the constant. Coefficients increase, but the *t*-ratios decline and then plateau at about 45. Whipsaw occurs without the additional term in the function for z (e.g., b*.001 in table 2). The lack of constant prevents r_x and r_z from falling towards zero with extreme collinearity because the error term contains the constant. Whipsaw is completely due to the relatively modest declines in $\sum r_x y$/n and $\sum r_z y$/n since $\sum r^2$ does not change, and standard errors are similar to those in table 1.

OVB also occurs when the key IV has data errors or has an incorrect functional form, because the correct IV is an omitted variable. Simulations suggest that neither are likely to produce whipsaw. An example of data errors is the following exercise: x=a; error=b; digit=trunc(runiform(0,5)); errorx=0; errorx=errorx+error if digit=1; xx=x+errorx; z=x+r*K;

---

[1] The reciprocal causation model contains two steps because one cannot in STATA simulate one-step reciprocal causation, where the IVs directly cause the DV, and the DV directly causes the IVs.

y=xx+z+c; regress y on x and z. An example of misspecified IVs is: $x=a$; xx=x^2; $z=x+r$*K; y=$xx+z+c$; regress y on x and z. Neither produces whipsaw. Consistent with table 2, there is whipsaw in both cases if $z=a+xx$*.001+r*K, but empirically that seems an unlikely situation.

These simulations, which are limited to the whipsaw test format, cover only a small portion of likely endogeneity scenarios, but they suggest that whipsaw is caused primarily by coefficient growth with more collinearity due to lesser declines of $\sum r_x y/n$ and $\sum r_z y/n$ compared to declines in exogenous regressions. Based on the algebraic examples and simulations, false negatives do not occur if the collinear IVs has a causal connection with the DV independent of the key IV. This causal connection can be very weak. Other false negative situations cannot be ruled out due to the unavailability of mathematical proofs. If there is whipsaw, however, there must be endogeneity because of Equation 3.

4. EMPIRICAL EXAMPLES OF THE WHIPSAW EFFECT AND WHIPSAW TEST.

For reasons given earlier, I use the exponent method to create collinearity with empirical examples, using an exponent of 1.02. Correlation between the key IV and the collinear IV are almost always larger than .99999 in the examples. The exponent method cannot be used with dummy variables, and adjustments must be made if the key IV has negative values (see footnote 3). The standard for detecting endogeneity with the whipsaw test is whether t-ratios fail to decline (that is, increase or reach a plateau) with more extreme collinearity. A convenient procedure is to conduct the test with the key IV and its square, and then conduct it with the key IV and the key IV with a 1.02 exponent. Exogeneity is indicated if t-ratios (averaging the two) in the latter are not smaller than with an exponent of two. This leads to ambiguous results when the t-ratios are slightly smaller using the 1.02 exponent, which may indicate either plateauing or a decline towards zero. Hence, an additional test is conducted with a smaller exponent (here 1.0001) to determine whether there is plateauing. That is almost always the case in the empirical examples (I mention this issue only when *t*-ratios decline). In addition, a rule-of-thumb is that when *t*-ratios are less than one, endogeneity is likely inconsequential, unless the sample is small. Usual significance test criteria are irrelevant because the null is different; *t*-ratios below 1.95 do not mean that endogeneity is unlikely. If the key IV is a control variable, the danger of coefficient bias for IVs of interest caused by endogeneity is attenuated because it depends on the correlation between the IV of interest and the control. Thus, the rule of thumb is less restrictive, and I suggest a *t*-ratio of two.[2]

4.1. Adding Quadratic Terms

Innumerable regression studies enter an IV and its square, which are highly correlated. A random variable with a normal distribution, standard deviation of one, is correlated .9976 with its square. Wooldridge's (2021) use of empirical exercises provides a convenient large sample of regressions to illustrate this. OVB is likely because the regressions contain few IVs. Twenty-nine exercises contain 37 pairs of linear and quadratic IVs (table 6). Whipsaw and endogeneity are likely because *t*-ratios are similar for the two variables, and coefficients have opposite signs. *T*-ratios are below one in just one case, and they average less than two in only two more cases. The common explanation for the pattern in table 6 is a curvilinear relationship, but table 7 shows that

[2] The PcGets procedure usually uses a *t*-ratio of two when determining whether an INV should be dropped (see Krolzig and Hendry, 2001).

this can be only part of the explanation. Correlations between the two IVs increase to about .99999 when substituting IVs to a power of 1.02 for the quadratic IVs. The coefficients always increase greatly with more extreme collinearity. The *t*-ratios for $IV^{1.02}$ are about the same as those for IV^2, rather than decline towards zero. In all these regressions the sum of squares of residuals changes very little with more collinearity (not shown), which suggests that endogeneity is due to OVB and OV simultaneity.

4.2. Unit Trends in Panel Regressions

Another useful way to show many examples of whipsaw is to study unit time trends in a time-series cross-section regression. Researchers often use these trends as crude controls for unobserved variables, in addition to unit and time effects. The trends control for unobserved variables that in the aggregate have a linear impact on the DV net of other IVs. What the unobserved variables are is unknown; so whether they are endogenous is unknown. Endogenous trends bias results for IVs of interest if the trend coefficients are biased, and the trends are correlated with the IVs.

In table 8 the whipsaw test explores whether state trends are endogenous when state alcohol consumption is regressed on state trends and control variables (state effects, year effects, lagged DV, and age groups). When adding quadratic trends (Panel A), the coefficients on the trends and quadratic trends have opposite signs, and their *t*-ratios are similar, which further illustrates the bias seen in table 6. The whipsaw test with trends and trends to the 1.02 power (Panel B) strongly indicates endogeneity. The coefficients are much larger than in Panel A, and the *t*-ratios are larger except for five states. This whipsaw effect with alcohol consumption is unusually large. Using other DVs, I generally find lesser effects; *t*-ratios vary from state to state and are usually below two. I studied 37 DVs, a convenience sample. The greatest number of significant trends to the 1.02 power besides table 8 occur when DVs are crime rates and real welfare payments, and the least is prison population. The reason for these differences is unknown. The bottom line is that the researchers do not know whether trends are endogenous and should test for it.

4.3. Lagged Independent Variables and the Granger Test

The Granger test tests whether a key IV causes the DV with a lag. The right side contains lags of the DV and the IV of interest, and if a lag is significant, it "causes" the DV with a lag. For an empirical example, I conduct panel regressions with state per capita real personal income (RPI) and state per capita employment (EMP), selected because they are important economic indicators, and endogeneity seems likely. Simply regressing RPI on EMP, and vice versa, produces large *t*-ratios (coef.=.13, *t*=6.20 when RPI is DV, and .03 and 4.27 when EMP is DV, using the controls listed in table 9).

Table 9 illustrates the whipsaw test, as well as the effects of collinearity between IV lags. The correlation between RPI and its lag is .996, and the figure for EMP is .991. Using the Granger test, coefficients and *t*-ratios are similar for the two lags, but with opposite signs, which suggests whipsawing as in table 6. The whipsaw test, using lags taken to the 1.02 power, strongly suggests endogeneity because the coefficients and *t*-ratios increase with more collinearity, and the *t*-ratios are greater than one. As discussed earlier, this apparent causal relationship backwards in time is due to omitted variables. Table 10 is the same as table 9 but with differenced variables, and the pattern is similar to table 9 for the first lag, but there is no whipsaw for the second lag when the DV is EMP, consistent with the Granger test.

The main reason for whipsaw is the same as in the simulations: $\sum r_x y$ declines much less then $\sum r_x^2$ and (not shown) $\sum r_z y$ declines much less then $\sum r_z^2$ (table 11).

4.4. Testing For Endogeneity When Using Instruments.

There are two reasons to use the whipsaw test when using instrumental variables: to suggest whether the suspected endogenous IV actually is endogenous (table 12) and to test whether the instrument is exogenous (table 13). I use the convenient examples in Wooldridge (2021), Chapters 16 and 17, excluding dummy variable instruments.[3] A benefit of the whipsaw test for endogeneity of instrumented IVs is that it does not use instrumental variables, whose exogeneity may be suspect, as in the Housman (1978) test. The drawback is possible, but unlikely, false negatives. Table 12 contains 10 instrumented variables (excluding duplicates), and the whipsaw test suggests that all but two are endogenous, justifying the use of the instruments. The results for Example 16.7 are suspect because there are only 34 observations.

---

[3] In four cases the INVs tested for endogeneity have negative values, which become positive when squared and become missing variables when taken to a fractional power. When the INV is differenced, the undifferenced version is taken to the 2 or 1.02 power, and then differenced (table 12, 16.5#1). In other cases negative values of the INV are changed to positive, the INV is squared or taken to the 1.02 power, then signs are changed to negative whenever the INV is negative (table 13, 16.7). The salient feature of the collinear INV is its high correlation with the key INV, which these procedures accomplish.

The whipsaw test suggests that 9 of the 20 instruments are endogenous (table 13), although example 16.7 is suspect due to small sample size. As for example 15.6, Wooldridge (2021, p. 515) finds the results puzzling perhaps because the instrument KWW ("knowledge of the world of work") is endogenous.

5. CONCLUSION

Granger (1969) believes "there is little use in the practice of attempting to discuss causality without introducing time." I attempt to refute him. He also states that textbook approaches can make unrealistic assumptions, and one should not look on strange results as anomalies to be ignored, but as opportunities to develop new procedures (Granger 2012). Strange results - the whipsaw effect - occur when combining collinearity and endogeneity. With increasing collinearity, *t*-ratios rarely decline towards zero, and coefficients switch to opposite signs and increase to unrealistic levels. To test for endogeneity of an IV, I create a second IV highly collinear with the first. If whipsaw occurs, there is endogeneity. False negatives are possible, but unlikely. It is possible to distinguish between OMV and simultaneity (except simultaneity through an omitted variable) because simultaneity causes regression errors to decline. In the empirical examples, the regression errors decline only modestly; so the whipsaw effect is mainly due to OMV and simultaneity through omitted variables. Apparently, it is impossible to distinguish between these two. I illustrate whipsaw with algebraic models of endogeneity, simulations, and empirical examples. Perhaps the most useful application of the whipsaw test is testing the exogeneity assumptions behind instrumental variables and other strategies to deal with endogeneity. I also caution against using an IV and its square and against using successive IV lags when they are correlated.

REFERENCES

Belsley, D., E. Kun and R. Welsch. 1980. *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity.* New York: Wiley.

Chatelain, J.B. and K. Ralf. 2014. "Spurious regressions and near-multicollinearity, with an application to aid, policies and growth." *Journal of Macroeconomics* 39A:85-96.

Goodhue, D.L., W. Lewis and R. Thompson. 2017. "A multicollinearity and measurement error statistical blind spot: correcting for excessive false positives in regressions and PLS." *MIS Quarterly* 41:667-684.

Granger, C. 1969. "Investigating causal relations by econometric models and cross-spectral methods." *Econometrica*, 37:424–438.

Granger, C. 2012. "Useful conclusions from surprising results." *Journal of Econometrics* 169:142-146.

Harnack, D., E. Laminski, M. Schünemann, and K.R. Pawelzik. 2017. "Topological causality in dynamical systems." *Physical. Review Letters* 119:098301.

Hernam, M.A. and J.M. Robins. 2016. "Instruments for Causal Inference. An Epidemiologist's Dream?" *Epidemiology* 17:360-372.

Hill, R.C. and L.C. Adkins. 2001." Collinearity." In: B.H. Baltagi, Ed. *Companion to Theoretical Econometrics*. Oxford: Blackwell Publishers.

Hoover, K. 2001. *Causality in Macroeconomics.* Cambridge: Cambridge University Press.

Imbens, G.W. and T. Lemieux. 2008. "Regression discontinuity designs: a guide to practice." *Journal of Econometrics* 142:165-635.

Hausman, J.A. 1978. "Specification tests in econometrics." *Econometrica* 46:1251-1271.

Kalnins, A. 2018. “Multicollinearity: How common factors cause Type 1 errors in multivariate regression.” *Strategic Management Journal* 39:2362-2385.

Kalnins, A. 2022. “When does multicollinearity bias coefficients and cause type 1 errors? A reconciliation of Lindner, Puck, and Verbeke (2020) with Kalnins (2018).” *Journal of International Business Studies* 53:1536-1548.

Kielyte, J. 2008. "Estimating Panel Data Models in the Presence of Endogeneity and Selection," Journal of Economics and Econometrics, 51(2), 1-19.

Komlos, J. 2020. “Multicollinearity in the presence of errors-in-variables can increase the probability of Type-I error.” *Journal of Economics and Econometrics* 63:1-17.

Krolzig, H.M, and D.F. Hendry. 2001. “Computer automation of general-to-specific model selection procedures.” *Journal of Economic Dynamics and Control* 25:831-866.

Mela, C.F. and P.K. Kopalle. 2002. “The impact of collinearity on regression analysis: the asymmetric effect of negative and positive correlations.” *Applied Economics* 34:667-677.

Martens, E.P., W.R. Pestman, A. de Boer, S.V. Belitser and O.H. Klungel. 2006. “Instrumental Variables: Application and Limitations.” *Epidemiology* 17:260-267.

Marvell, T.B. 2010. “Prison population and crime.” In: Benson, B.L. & P.R. Zimmerman, eds. *Handbook on the Economics of Crime*. Northampton, MA: Edward Elgar.

Maziarz, M. 2017. “A review of the Granger-causality fallacy.” *The Journal of Philosophical Economics* 8:86-105.

Peters, J., D. Janzing and B. Schölkopf. 2017. *Elements of Causal Inference: Foundations and Systematic Learning Algorithms.* Cambridge, MA: MIT Press.

Spanos, A. 2019. “Near-collinearity in linear regression revisited: the numerical vs. the statistical perspective.” *Communications in Statistics – Theory and Methods* 48:5493-5516.

Spanos, A. and A. McGuirk. 2002. “The problem of near-multicollinearity revisited: erratic vs. systematic volatility.” *Journal of Econometrics* 108:365-393.

Winship, C. and B. Western. 2016. “Multicollinearity and model misspecification.“ *Sociological Science* 3:627-639.

Wooldridge, J. 2021. *Introductory Economics: A Modern Approach, Seventh Edition.* Boston, MA: Cengage Learning.

Table 1. Regression Simulations Without Endogeneity							
K	ρ *x&z*	β_x	σ_x	β_z	σ_z	$\sum r^2$	$1-\rho^2$
10	.13	.98	.016	.98	.0019	36112	.98
1	.77	.97	.025	1.01	.019	36112	.41
.1	.997	.86	.19	1.11	.19	36112	.0068
.01	.99997	-.16	1.93	2.14	1.93	36112	.000069
.001	.9999996	-10.38	19.34	12.35	19.34	36112	.00000072

Elements of Coefficients

K	$\sum r_x y/n$	$\sum r_x^2/n$	$\sum r_z y/n$	$\sum r_z^2/n$	$Cov(\beta_x,\beta_z)$
10	1.35	1.38	1.35	1.38	-.0000040
1	.55	.57	.98	.97	-.00038
.1	.0083	.0096	.011	.0097	-.037
.01	-.000015	.000097	.00021	.000097	-3.74
.001	-.000010	.00000097	.000012	.00000097	-374

$x=a$; $z=x+w*K$; $y=x+z+b$; regress *y* on *x* and *z*. n=10,000. ρ=correlation. β=coefficient. σ=standard error. $r_x y$, r_x^2, $r_z y$, and r_z^2 are from Equation 1. $Cov(\beta_x,\beta_z)$ is from Equation 2. β_x is $\sum r_x y$ divided by $\sum r_x^2$. The seed is 1313579.

Table 2. Reverse Causation

K=	ρ *x&z*	β_x	σ_x	β_z	σ_z	$\sum r^2$	$1-\rho^2$
10	.19	.39	.0050	-.000	.0009	8876	.96
1	.886	.39	.011	-.002	.0094	8876	.21
.1	.9990	.49	.095	-.10	.094	8875	.0027
.01	.999986	9.34	.936	- 8.94	.935	8795	.000028
.001	.999999	471	4.988	-471	4.999	4691	.00000048

Elements of Coefficients

K=	$\sum r_x y/n$	$\sum r_x^2/n$	$\sum r_z y/n$	$\sum r_z^2/n$	$Cov(\beta_x,\beta_z)$
10	1.36	3.50	-.014	99.78	-.00000091
1	.30	.78	-.0022	1.00	-.000089
.1	.0048	.0099	-.0010	.010	-.0089
.01	.00094	.00010	-.00090	.00010	-.87
.001	.00090	.0000019	-.00089	.0000019	-24

y=*a*; *x*=*y*+*b*; *z*=*x*+*b**.001+*w**K; regress *y* on *x* and z. See notes to table 1.

Table 3. Reciprocal Causation

K =	ρ $x\&z$	β_x	σ_x	β_z	σ_z	$\sum r^2$	$1-\rho^2$
10	.43	.54	.0033	-.0018	.0014	18641	.82
1	.978	.56	.014	-.017	.014	18642	.043
.1	.9998	.54	.14	.0034	.14	18644	.00046
.01	.999998	-16	1.35	17	1.34	18362	.0000047
.001	.9999999	-656	4.82	656	4.82	6536	.00000012

Elements of Coefficients

K	$\sum r_x y/n$	$\sum r_x^2/n$	$\sum r_z y/n$	$\sum r_z^2/n$	$Cov(\beta_x,\beta_z)$
10	9.51	17	-.18	99	-.0000019
1	.52	.95	-.016	.99	-.00019
.1	.0053	.0098	.000034	.0099	-.019
.01	-.0016	.00010	.0017	.00010	-1.83
.001	-.0018	.0000028	.0019	.0000028	-23

xx=*a*; *zz*=*b*; *y*=*xx*+*zz*+*c*; *x*=*xx*+*y*+*d*; z=*x*+y*.001+w*K; regress y on *x* and *z*. See notes to table 1.

Table 4. Omitted Variable Bias							
K	ρ *x&z*	β_x	σ_x	β_z	σ_z	$\sum r^2$	$1-\rho^2$
10	.24	1.71	.0047	1.00	.0011	11516	.94
1	.92	1.71	.012	1.00	.011	11516	.15
.1	.991	1.60	.11	1.11	.11	11514	.0018
.01	.99991	-9.98	1.08	13	1.08	11382	.000018
.001	.9999998	-542	4.99	545	4.99	5257	.00000016

Elements of Coefficients

K	$\sum r_x y/n$	$\sum r_x^2/n$	$\sum r_z y/n$	$\sum r_z^2/n$	Cov(β_x,β_z)
10	8.80	5.14	96	96	-.0000012
1	1.39	.81	.97	.97	-.00012
.1	.015	.0096	.011	.0097	-.012
.01	-.00097	.000098	.0012	.000098	-1.17
.001	-.0011	.0000021	.0012	.0000021	-25

x=*o*+*a*; [*o*=omitted variable]; *z*=*x*+*a**.001+*w**K; *y*=*x*+*z*+*o*; regress *y* on x and *z*.

See notes to table 1. The RESET test is negative.

Table 5. Regressions Without Constants

K	ρ *x&z*	β_x	σ_x	β_z	σ_z	$\sum r^2$	$1-\rho^2$
10	.13	1.33	.015	1.06	.003	41948	.98
1	.77	.78	.020	1.61	.013	41948	.41
.1	.997	-4.68	.15	7.06	.13	41948	.0068
.01	.99997	-59	1.36	67	1.34	41948	.000069
.001	.9999996	-605	13.53	607	13.42	41948	.00000072

Elements of Coefficients

K	$\sum r_x y/n$	$\sum r_x^2/n$	$\sum r_z y/n$	$\sum r_z^2/n$	$Cov(\beta_x,\beta_z)$
10	2.57	1.94	247	233	- .000020
1	.46	.59	3.74	3.33	-.00036
.1	-.090	.019	.16	.023	-.0020
.01	-.014	.00023	.014	.00023	-1.82
.001	-.0014	.0000023	.0014	.0000023	-180

These regressions are the same as table 1, but without constants. Constants are highly significant in table 1 (not shown).

Table 6. Wooldridge's Uses of IVs and Their Squares							
Pg.	Ex.	DV	IV	IV		IV^2	
				Coef	*t*	Coef	*t*
189	-	wage	experience	.298	7.27	-.0061	6.78
190	6.2	ln(price)	rooms	-.545	3.24	.062	4.86
193	6.3	exam result	GPA	-1.629	3.39	.293	2.93
"	"	exam result	ACT score	-.128	1.30	.0045	2.08
202	6.5	GPA	class size	-.0608	3.69	.00546	2.41
211	-	ln(salary)	Roe	.0215	1.61	-.00008	3.08
"	-	r&d exp.	sales	.00030	2.14	.0000000070	1.89
213	-	Prate	totemp	-.00043	4.78	.0000000039	3.90
227	7.5	ln(wage)	experience	.029	5.92	-.00058	5.43
"	"	ln(wage)	tenure	.032	4.63	-.00059	2.49
235	7.10	ln(wage)	experience	.029	5.40	-.00058	4.91
"	"	ln(wage)	tenure	.032	4.65	-.00059	2.51
240	-	inlf.	experience	.039	6.50	-.00060	3.33
251	-	Sleep	age	-8.70	.78	.128	.96
252	-	SATscore	class size	19.30	5.04	-2.19	4.13
265	8.1	ln(wage)	experience	.0268	5.22	-.00054	5.03
"	"	ln(wage)	tenure	.0291	4.19	-.00053	2.19
268	8.3	arrests	sentence ln	.0178	1.76	-.00052	2.49

280	8.7	cigarettes	age	.771	4.81	-.0090	5.18
285	8.8	inlf.	experience	.039	6.96	-.00060	3.23
288	-	smokes	age	.020	3.33	-.00026	4.33
296	9.1	arrests	Pcnv	.553	3.58	-.730	4.68
“	“	arrests	prison time	.287	6.49	-.0296	7.66
“	“	arrests	income	-.0034	4.25	.0000072	2.81
355	10.8	fertility	time	-2.531	6.50	.0196	3.95
428	13.1	# kids	age	.532	3.85	-.0058	3.71
430	13.2	ln(wage)	experience	.0296	8.29	-.00040	5.15
472	14.4	ln(wage)	experience	.106	6.80	-.0047	6.85
508	15.4	ln(wage)	experience	.107	4.89	-.0022	6.64
511	15.5	ln(wage)	experience	.044	3.29	-.00090	2.24
544	16.5	ln(wage)	experience	.035	1.77	-.00071	1.55
568	17.1	inlf.	experience	.206	6.42	-.0032	3.10
576	17.2	work hours	experience	131.564	7.61	-1.864	3.47
592	17.5	wageoffer	experience	.044	2.70	-.00086	1.96
658	-	particip.	Mrate	.239	5.69	-.087	2.02
“	-	particip.	ln(emp)	-.112	8.00	.0057	6.33
“	-	particip.	age	.0059	5.90	-.00007	3.50
From Wooldridge (2021). I exclude examples that are very similar to earlier examples.							

Table 7. Comparing Results in Table 6 to Results When IVs Have a Smaller Power.											
Ex.	IV	IV alone		IV and IV^2				IV and $IV^{1.02}$			
				IV		IV^2		IV		$IV^{1.01}$	
		Coef	*t*	. Coef	*t*	Coef	*t*	Coef	*t*	Coef	*t*
6.2	rooms	.255	13.74	-.545	3.29	.062	4.86	-40.331	5.13	38.353	5.16
6.3	GPA	-.554	1.69	-1.629	3.39	.293	2.93	-67.412	2.74	64.741	2.72
"	ACT	.082	7.33	-.128	1.30	.005	2.08	-10.028	2.07	9.312	2.08
6.5	Hsize	-.023	4.59	-.061	3.69	.005	2.41	-2.095	3.37	1.987	3.34
7.5	exper	.005	2.85	.029	5.92	-.001	5.43	1.134	6.08	-1.045	6.06
"	tenure	.017	5.84	.032	4.63	-.001	2.49	.605	2.68	-.551	2.61
7.6	exper	.005	2.85	.029	5.92	-.001	5.43	1.078	5.53	-.993	5.52
"	tenure	.015	5.37	.029	4.63	-.001	2.31	.545	2.45	-.496	2.38
7.10	exper	.005	2.85	.029	5.89	-.001	5.40	1.129	6.05	-1.401	6.03
"	tenure	.017	5.85	.032	4.65	-.001	2.59	.613	2.71	-.559	2.64
8.1	exper	.003	1.96	.027	5.22	.001	5.03	1.078	5.63	-.993	5.61
"	tenure	.016	4.56	.029	4.19	.001	2.19	.545	2.44	-.496	2.37
8.3	avsen	.003	.59	.018	1.76	-.001	2.49	.618	2.04	-.577	2.05
8.7	age	-.045	1.58	.771	4.81	-.009	5.18	40.351	5.31	- 36.734	5.32
8.8	exper	.022	10.36	.039	6.96	-.001	3.23	.705	3.21	-.368	3.11
9.1	Pcnv	.133	3.30	.553	3.58	-.730	4.68	24.116	4.54	-24.289	4.56
"	ptime	-.041	4.63	-.287	6.49	-.030	7.66	12.964	8.89	-12.400	8.93

"	Inc	-.001	4.37	-.003	4.25	.000	2.81	-.113	3.41	.099	3.37
10.8	time	-1.150	6.12	-2.531	6.50	.020	3.95	-64.514	5.27	58.119	5.18
13.1	age	. 020	2.40	.532	3.85	-.006	3.71	25.617	3.73	-23.207	3.73
13.2	exper	.012	11.27	.030	8.29	-.000	5.15	.812	5.91	-.740	5.82
14.4	exper	.033	2.98	.106	6.88	-.005	6.85	3.506	7.78	-3.274	7.71
15.4	exper	.062	3.16	.107	4.89	-.002	6.64	2.140	6.07	-1.949	6.09
15.5	exper	.015	3.89	.044	3.29	-.001	2.24	1.476	2.79	-1.359	2.76
16.5	exper	.010	1.23	.035	1.77	-.001	1.55	1.209	2.07	-1.111	2.06
17.1	exper	.120	8.79	.206	6.42	-.003	3.10	3.940	3.06	-3.569	2.97
17.2	exper	76.779	11.69	131.564	7.61	-1.864	3.47	2424.730	3.58	- 2189.349	3.47
17.5	exper	.015	2.96	.044	2.70	-.001	1.96	1.428	2.46	-1.313	2.44
See notes to table 6. Only examples with accompanying data sets are included. In example 15.5 lowering the exponent to 1.0001 led to *t*-ratios of .83, indicating exogeneity.											

Table 8. Regressing Alcohol Consumption on State Specific Trends								
	A. Trend & Trend Squared				B. Trend & Trend to 1.02 Power			
	Trend		$Trend^2$		Trend		$Trend^{1.02}$	
	Coef	*t*	Coef	*t*	Coef	*t*	Coef	*t*
Alabama*	-.007	3.59	.00013	4.32	.578	10.84	-.533	10.84
Alaska#	-.021	7.58	.00027	6.56	.332	6.57	-.311	6.69
Arizona	-.007	4.57	.00008	3.25	.680	11.79	-.630	11.81
Arkansas*	-.005	2.64	.00009	3.18	.685	12.08	-.632	12.07
California#	-.013	9.09	.00018	7.60	.494	7.28	-.459	7.31
Colorado	-.016	7.06	.00024	7.05	.346	9.31	-.321	9.37
Connecticut*	-.010	5.44	.00016	5.25	.506	11.55	-.467	11.56
Delaware	-.009	5.12	.00021	6.53	.391	9.49	-.359	9.47
Florida	-.009	5.46	.00013	5.34	.562	10.18	-.521	10.17
Georgia	-.006	2.88	.00005	1.65	.730	14.00	-.675	14.02
Hawaii#	-.013	7.80	.00021	8.02	.471	7.87	-.436	7.92
Idaho*	-.011	6.02	.00015	5.35	.540	10.06	-.501	10.06
Illinois	-.009	5.82	.00013	5.09	.573	10.21	-.531	10.20
Indiana*	-.007	3.88	.00013	4.46	.601	11.71	-.554	11.71
Iowa*	-.010	5.26	.00010	6.44	.459	10.12	-.423	10.09
Kansas*	-.008	4.50	.00014	5.06	.550	10.89	-.508	10.87
Kentucky*	-.012	5.58	.00021	6.16	.401	11.45	-.371	11.44

Louisiana*	-.008	4.01	.00011	4.00	.615	12.07	-.568	12.04
Maine*	-.016	6.23	.00028	6.66	.229	12.65	-.218	12.64
Maryland	-.014	7.71	.00020	6.54	.419	10.04	-.390	10.11
Massachusetts	-.011	5.51	.00017	5.08	.486	11.46	-.452	11.62
Michigan	-.014	6.91	.00022	6.67	.378	10.13	-.350	10.15
Minnesota*	-.013	5.30	.00021	5.63	.406	13.53	-.376	13.45
Mississippi*	-.006	2.84	.00010	3.29	.636	12.10	-.587	11.06
Missouri*	-.007	4.01	.00014	4.82	.567	12.34	-.523	12.29
Montana*	-.014	6.85	.00027	7.34	.332	8.42	-.307	8.45
Nebraska	-.009	4.79	.00013	4.50	.562	12.23	-.521	12.22
Nevada*	-.016	11.59	.00021	9.75	.410	6.56	-.382	6.65
New Hampshire	-.017	6.35	.00027	6.03	.250	14.79	-.233	14.71
New Jersey*	-.009	5.70	.00015	5.55	.524	10.04	-.484	10.04
New Mexico	-.011	5.88	.00016	5.06	.541	10.38	-.502	10.44
New York#	-.016	8.24	.00024	8.11	.359	6.94	-.333	6.97
North Carolina*	-.009	4.55	.00016	5.27	.511	10.80	-.471	10.82
North Dakota*	-.012	6.03	.00023	7.58	.385	8.55	-.356	8.50
Ohio*	-.008	4.78	.00014	4.82	.563	11.53	-.520	11.53
Oklahoma*	-.008	5.02	.00014	5.53	.567	10.61	-.523	10.60
Oregon*	-.012	6.21	.00020	6.48	.443	10.08	-.410	10.05
Pennsylvania*	-.008	5.24	.00016	5.93	.514	9.24	-.474	9.22

Rhode Island*	-.011	5.77	.00016	5.48	.504	9.47	-.467	9.46
South Carolina	-.006	2.71	.00008	2.29	.678	12.91	-.626	12.98
South Dakota*	-.005	2.84	.00010	3.29	.631	12.64	-.582	12.59
Tennessee*	-.010	5.06	.00018	6.06	.474	10.09	-.438	10.09
Texas	-.008	4.83	.00010	4.28	.682	10.81	-.613	10.80
Utah	-.009	6.54	.00010	5.15	.672	6.19	-.623	8.17
Vermont	-.024	8.02	.00040	7.78	--	--	--	--
Virginia*	-.011	5.53	.00017	5.50	.509	11.95	-.471	11.98
Washington	-.012	6.08	.00016	5.20	.521	11.67	-.484	11.66
West Virginia*	-.009	4.56	.00016	4.95	.508	11.73	-.469	11.73
Wisconsin	-.013	5.70	.00019	5.49	.431	12.28	-.400	12.24
Wyoming #	-.016	8.00	.00027	7.87	.332	7.32	-.309	7.37

These are two pooled time-series cross-section regressions, with annual state data for 1970 to 2020. The DV is logged alcohol consumption per capita. There are 50 separate state-specific trends (which are zero except for the state, where they are a counter). Additional IVs are state and year effects, a lagged DV, and nine age categories (percent 20-24 to 60-64) also logged. The regression is weighted by the square root of state population and is clustered by state. *The trend is not significant when entered alone, without a powered trend. # The average *t*-ratio with $trend^2$ is higher than the average with $trend^{1.02}$. -- Dropped by the regression program because of collinearity. The alcohol data are from https://www.niaaa.nih.gov/sites/default/files/pcyr1970-2021.txt.

The population data are from https://www2.census.gov/programs-surveys/popest/datasets/

Table 9. Granger Test: Real Personal Income and Employment (Levels)

A. DV=real personal income. IV=employment.

	First IV lag				Second IV lag			
	Lag		Lag Exponent		Lag		Lag Exponent	
	Coef	*t*	Coef	*t*	Coef	*t*	Coef	*t*
One lag	.05	2.83						
Granger	.40	6.69			-.37	6.03		
Whipsaw sqr.	-4.84	3.31	.42	3.56	3.75	2.75	-.33	3.00
Whipsaw 1.02	-263.33	3.52	249.33	3.52	207.68	2.98	-196.65	2.98

B. DV=employment. IV=real personal income.

	First IV lag				Second IV lag			
	Lag		Lag Exponent		Lag		Lag Exponent	
	Coef	*t*	Coef	*t*	Coef	*t*	Coef	*t*
One lag	.00	.98						
Granger	.05	2.95			-.06	3.46		
Whipsaw sqr.	-.21	1.37	.03	1.69	.17	1.25	-.05	1.66
Whipsaw 1.02	-13.74	1.83	13.26	1.86	12.26	1.82	-11.84	1.82

These are time series-cross section regressions with data for 50 states, years 1970-2020.

Other IVs in all regressions are two DV lags, state and year effects, and nine age categories (percent 20-24 to 60-64). RPI is per capita, and EMP is per 1000 capita. The continuous variables are logged. The regressions are clustered. The lag exponent is the lag squared (the Whipsaw sqr. line) or the lag taken to the 1.02 power (the Whipsaw 1.02 line). Correlations between lagged IVs and their squares are about .99. The corresponding correlations for IVs taken to the 1.02 power are about .999999. Third lags are far from significant. RPI and EMP are I(1) and cointegrated. RPI data are from https://www.bea.gov/data/income-saving/personal-income-by-state EMP data are from https://www.bea.gov/data/income-saving/personal-income-by-state.

Table 10. Granger Test: Real Personal Income and Employment (Differences)								
A. DV=real personal income. IV=employment.								
	First IV lag				Second IV lag			
	Lag		Lag Exponent		Lag		Lag Exponent	
	Coef	*t*	Coef	*t*	Coef	*t*	Coef	*t*
One lag	.37	4.80						
Granger	.34	3.71			.10	1.28		
Whipsaw sqr.	-5.56	3.00	.47	3.20	3.38	1.89	-.26	1.81
Whipsaw 1.02	-296.13	3.19	280.18	3.19	163.59	1.82	-154.50	1.82
B. DV=employment. IV=real personal income.								
	First IV lag				Second IV lag			
	Lag		Lag Exponent		Lag		Lag Exponent	
	Coef	*t*	Coef	*t*	Coef	*t*	Coef	*t*
One lag	.05	2.93						
Granger	.05	3.02			.00	.09		
Whipsaw sqr.	-.16	1.54	.04	1.99	.06	.63	-.01	.63
Whipsaw 1.02	-11.42	2.27	11.03	2.29	2.53	.51	-2.43	.51

This is the same as table 9 except that the variables are differenced, and state dummies are dropped. The lagged IV^2 is about .99 correlated with the lagged IV. The corresponding correlation for IVs taken to the 1.02 power is about .9999.

Table 11. Elements of Coefficients in Tables 9 and 10				
Regressions in table 9 (levels)				
DV=real personal income				
	First IV lag		Second IV lag	
	$\sum r_x y$	$\sum r_x^2$	$\sum r_x y$	$\sum r_x^2$
exp=2	-.00084	.00017	.00068	.000015
exp=1.02	-.000018	.000000069	.000015	.000000072
DV = employment				
	First IV lag		Second IV lag	
	$\sum r_x y$	$\sum r_x^2$	$\sum r_x y$	$\sum r_x^2$
exp=2	-.0022	.010	.0019	.011
exp=1.02	-.000058	.0000042	.000055	.0000045
Regressions in table 10 (differences)				
DV = real personal income				
	First IV lag		Second IV lag	
	$\sum r_x y$	$\sum r_x^2$	$\sum r_z y$	$\sum r_z^2$
exp=2	-.00083	.00015	.00052	.000016
exp=1.02	-.000018	.000000060	.000010	.000000062
DV = employment				
	First IV lag		Second IV lag	

	$\sum r_x y$	$\sum r_x^2$	$\sum r_z y$	$\sum r_z^2$
exp=2	-.0018	.011	.00071	.012
exp=1.02	-.000052	.0000046	.000012	.0000048
r_x is the residual when regressing the particular IV on all other IVs.				

Table 12. Test of Endogeneity of Instrumented Independent Variables

Exam-ple	DV	Instrum-ented IV	Instrumented IV & IV^2				Instrumented IV & $IV^{1.02}$			
			Coef	*t*	Coef	*t*	Coef	*t*	Coef	*t*
15.1	Lwage	**educ**	-.073	.69	.007	1.73	-7.408	1.56	7.005	1.59
15.2	Lwage	educ	.133	1.78	-.003	.90	3.245	.87	-2.962	.85
15.3	Lbwght	**packs**	-.149	3.83	.047	1.69	-3.275	2.01	3.177	1.96
15.5	Lwage	**educ**	-.078	.75	.007	1.80	-7.504	1.61	7.094	1.64
15.6	Lwage	**IQ**	-.005	.60	.000	1.06	-.418	1.10	.377	1.11
15.8	Lwage	**educ**	-.078	.75	.007	1.80	-7.504	1.61	7.094	1.64
15.9	Kids	**educ**	.120	1.27	-.010	2.68	8.957	2.53	-8.47	2.58
16.5#1	hours	**Lwage#**	.370*	2.76	-.165*	3.04	9.250*	1.89	-9.099*	1.89
16.5#2	Lwage	**hours**	.000	4.00	-.000	4.95	.018	3.31	-.015	3.72
16.6	inf	**open**	-.520	2.09	.002	1.33	-18.222	1.66	16.303	1.64
16.7	Gc	r3#	-.000	.04	.000	.10	.001	.01	-.001	.02

The instrumented IV is suspected to be endogenous. There are two regressions in each line, with the instrumented IV and its squared and, second, with the instrumented IV and it taken to the 1.02 power. Other IVs are not shown. The correlation between the IV and the IV taken to the 1.02 power is .9999 or higher. Bold faced instrumented IV names are endogenous. L in a DV name means the DV is logged. * Divided by 1000. # The IV has negatIVe values before adjustments (see note 10).

Table 13. Test of Exogeneity of Instruments.										
			Instrument & Its Square				Instrument & It to the 1.02 Power			
Ex. #	DV	Instrument	Instrument		Instrument^2		Instrument		$\text{Instrument}^{1.02}$	
			Coef	*t*	Coef	*t*	Coef	*t*	Coef	*t*
15.1	Lwage	**fatheduc**	-.080	1.94	.005	2.39	-3.469	2.52	3.441	2.52
15.2	Lwage	Sibs	-.047	3.07	.002	1.38	-.506	.86	.456	.81
15.3	Lbwght	cigprice	-.006	.67	.000	.77	-.321	.73	.287	.73
15.5	Lwage	motheduc	.003	.06	-.001	.39	.827	.48	-.790	.48
"		**fatheduc**	-.043	1.04	.002	.95	-1.834	1.19	1.717	1.19
15.6	Lwage	**KWW**	-.026	2.27	.000	2.74	-1.512	2.83	1.386	2.84
15.8	Lwage	motheduc	.001	.02	-.000	.03	.400	.22	-.375	.22
"		**fatheduc**	-.072	1.64	.004	1.93	-3.218	2.00	3.030	2.01
"		hueduc	.047	.62	-.001	.17	.564	.17	-.494	.16
15.9	kids	meduc	.015	.30	-.000	.21	.657	.39	-.639	.40
"		Feduc	-.016	.28	.001	.34	.668	.34	-.662	.35
16.5#1	hours	**exper**	68.867	6.72	-.706	2.17	662.411	1.71	-574.722	1.59
"		**expersq**	2.602	10.13	-.001	5.71	64.236	6.84	- 54.827	6.70
16.5#2	Lwage	Age	.023	.47	-.000	.53	1.286	.52	-1.173	.52
"		kidslt6	-.035	.14	.005	.03	.565	.05	- .598	.05
"		**nwifeinc**	.011	1.38	-.000	.78	.400	1.11	-.362	1.09

16.6	inf	Lland	-4.563	.63	.346	1.00	-318.672	.95	300.723	.96
16.7	Gc	lag Lgc#	1.727	.14	-.063	.09	-142.122	.23	133.733	.23
"		**lag Lgy**#	9.031	.86	-.486	.88	607.142	1.20	-569.186	1.20
"		**lag r3#**	-.004	1.55	.001	1.97	-.174	1.76	.171	1.77

See notes to table 12. Where more than one instrument is shown they are in the same regression, except that they are in separate regressions in 16.5#1 to avoid perfect collinearity.