The problem of endogeneity

John Antonakis
Faculty of Business and Economics,
University of Lausanne

AOM Meeting
Chicago, 2018

The “typical” management study

Measure some attitude, behavior, style, choice (*x*), at some level of analysis and model that as a cause of *y*. If *x* is measured and not manipulated (random) it is probably endogenous with respect to *y*.

Are there un-modeled variables that cause *x* and *y*? E.g.,

- manager-level variables (e.g., ability, personality)
- subordinate-level variables (e.g., ability, personality)
- organizational variables (e.g., resources, training, culture)

If so, then then *x* correlates with *e* of *y.*

Why is endogeneity a problem?

Suppose the true model is (Antonakis et al., 2010, *LQ*):

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 z_i + e_i \quad \text{Eq. 1}$$

Instead of Eq. 1, we estimate:

$$y_i = \varphi_0 + \varphi_1 x_i + v_i \quad \text{Eq. 2}$$

In the presence of endogeneity is: $\varphi_1 = \beta_1$?

If *z* and *x* are correlated (irrespective of the direction), we can note:

$$z_i = \gamma_1 x_i + u_i \quad \text{Eq. 3}$$

The endogeneity problem is evident when substituting Eq. 3 into Eq. 1:

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 (\gamma_1 x_i + u_i) + e_i, \quad \text{Eq. 4}$$

Multiplying out gives (notice, the error term v_i, which is the error term of Eq. 2):

$$y_i = \beta_0 + \beta_1 x_i + \underbrace{(\beta_2 \gamma_1 x_i + \beta_2 u_i + e_i)}_{v_i} \quad \text{Eq. 5}$$

Or, rearranging as a function of *x* gives

$$y_i = \beta_0 + (\beta_1 + \beta_2 \gamma_1) x_i + (\beta_2 u_i + e_i) \quad \text{Eq. 6}$$

In the presence of endogeneity, the OLS estimate will be inconsistent:

$$\beta_1 (from\ Eq.\ 1: y_i = \beta_0 + \beta_1 x_i + \beta_2 z_i + e_i) \neq$$
$$\varphi_1 (from\ Eq.\ 2: y_i = \varphi_0 + \varphi_1 x_i + v_i)$$

$$\varphi_1 = \frac{Cov(y,x)}{Var(x)} = \beta_1 + \beta_2 \gamma_1 \neq \beta_1 \quad \text{Eq. 7}$$

unless: $\beta_2 = 0$ or $\gamma_1 = 0$ (i.e., *x* and *z* are orthogonal and *x* is exogenous); these conditions are achieved if *x* is manipulated.

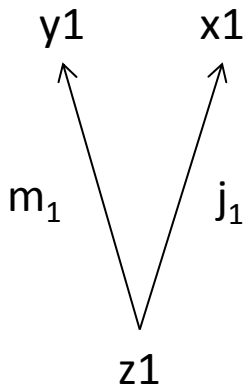
Researchers think they have an ACE up their sleeve; temporal ordering!

- Measure *x* at Time 1
- Measure *y* at Time 2
- Regress y on *x*

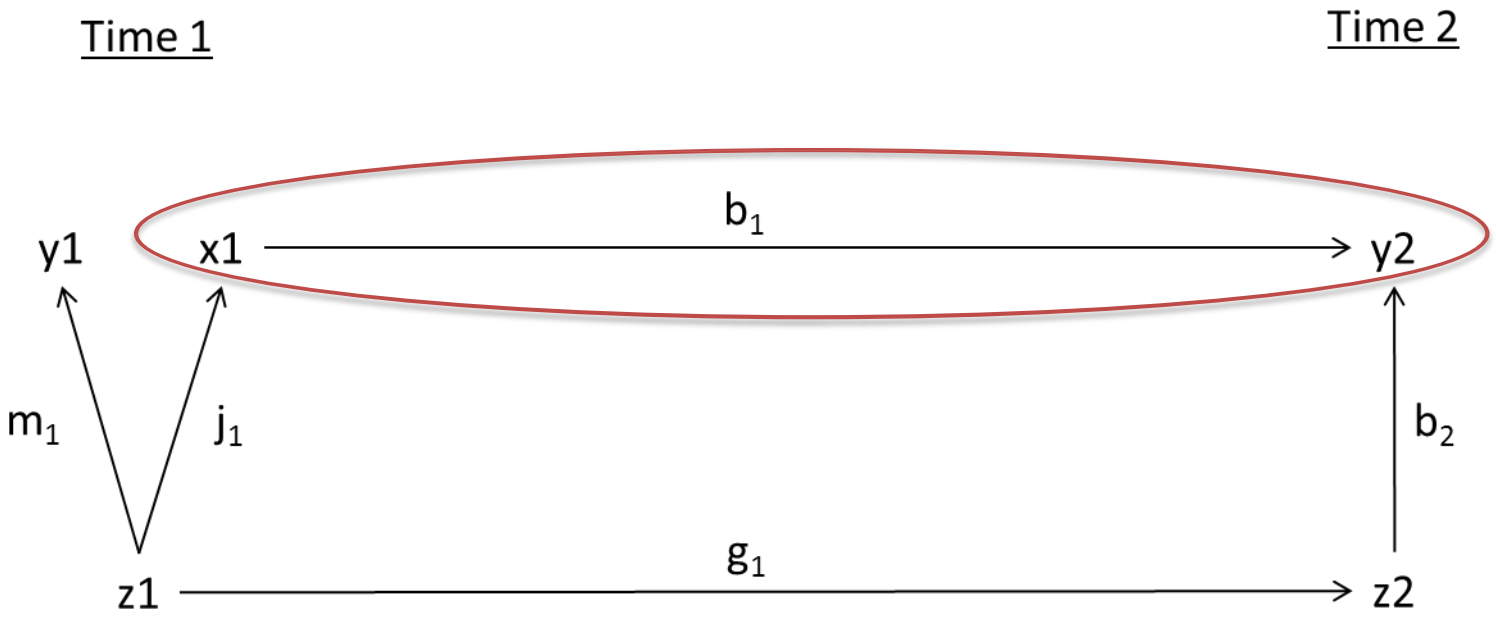
The above is still problematic if *x* is not exogenous; this design is the tranquilizing drug of naiveté.

Researchers repeatedly make the *post-hoc ergo propter hoc* fallacy! Many editors and reviewers ignore this problem!

Suppose the following (at time 1)



Thus, we cannot estimate the effect of x1→y1 without controlling for z1; z1 will correlate with itself over time.



The only time b_1 will give you a true estimate of the effect on y is if either j_1, or unusually, g_1, or b_2 are zero.

Unaddressed endogeneity is….

- A failsafe way to get desk rejected from some top journals nowadays (e.g., *SMJ, JOpsM, LQ*)
- Is a rampant problem
  - researchers fail to address at least 66% and up to 90% of design and estimation conditions that make causal claims invalid (Antonakis et al., 2010, *LQ*).
  - 79.73% of LQ articles (between 1990-2014) had one or endogeneity threats in testing mediation models (Antonakis et al., 2014, *LQ*)
  - In estimating mediation models, 74% of articles had endogenous predictors and 98% used wrong estimator (i.e., OLS-type instead of instrumental-variable type, Fischer et al., 2017, *JOM*)
- It will cost you in terms of citations (Antonakis et al., 2014, *LQ*)

There are ways to deal with it

- Isolate *x* from the disturbance using an appropriate design (*x* is then "as-if" randomized).
- It takes longer to think of the *right design* (than modeling all omitted causes), but at the end, it pays!
- Researchers must do "slower science"; take their time, be creative, and do their homework, instead of spewing out lots of bad-quality research.
- …and then go back and re-examine some apparent "knowns" and "received wisdom", e.g., the Queen Bee phenomenon:

The Leadership Quarterly xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

The Leadership Quarterly

journal homepage: www.elsevier.com/locate/leaqua

The queen bee: A myth? The effect of top-level female leadership on subordinate females

Paulo Roberto Arvate[a,*], Gisele Walczak Galilea[b], Isabela Todescat[c]

[a] São Paulo School of Business Administration and Center of Studies on Microeconomic Applied, Getulio Vargas Foundation, Rua Itapeva, 474, 01332-010 São Paulo, SP, Brazil

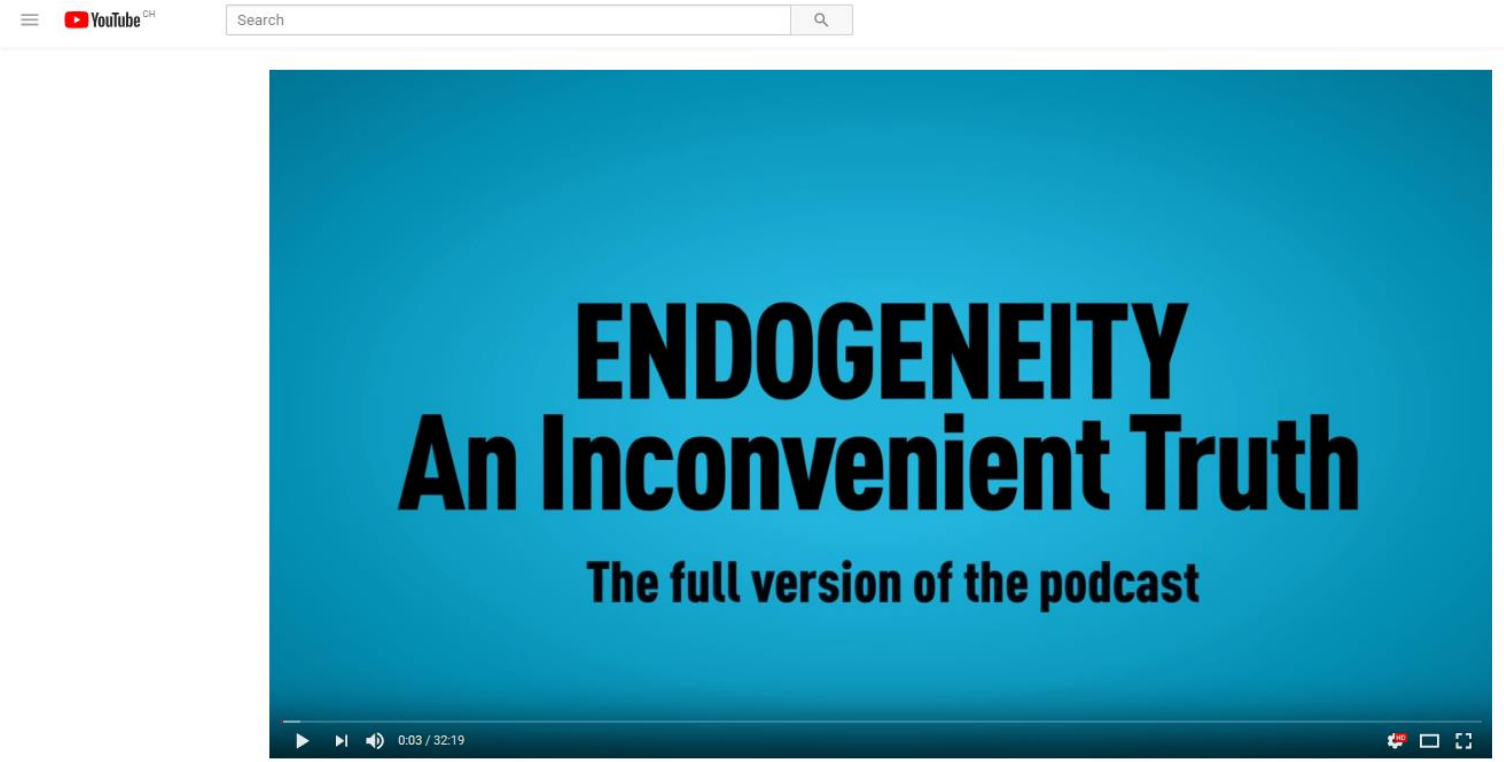
[b] São Paulo School of Business Administration, Getulio Vargas Foundation and Institute of Education and Research, INSPER, Brazil

[c] São Paulo School of Business Administration, Getulio Vargas Foundation, Brazil

ABSTRACT

We investigate the effect of female leadership on gender differences in public and private organizations. Female leadership impact was constructed using a quasi-experiment involving mayoral elections, and our research used a sample of 8.3 million organizations distributed over 5600 Brazilian municipalities. Our main results show that when municipalities in which a woman was elected leader (treatment group) are compared with municipalities in which a male was elected leader (control group) there was an increase in the number of top and middle female managers in public organizations. Two aspects contribute to the results: time and command/role model. The time effect is important because our results are obtained with reelected women – in their second term – and the command/role model (the queen bee phenomenon is either small, or non-existent) is important because of the institutional characteristics of public organizations: female leaders (mayor) have much asymmetrical power and decision-making discretion, i.e., she chooses the top managers. These top managers then choose middle managers influenced by female leadership (a role model). We obtained no significant results for private organizations. Our work contributes to the literature on leadership by addressing some specific issues: an empirical investigation with a causal effect between the variables (regression-discontinuity design – a non-parametric estimation), the importance of role models, and how the observed effects are time-dependent. Insofar as public organizations are concerned, the evidence from our large-scale study suggests that the queen bee phenomenon may be a myth; instead, of keeping subordinate women at bay, our results show that women leaders who are afforded much managerial discretion behave in a benevolent manner toward subordinate women. The term "Regal Leader" instead of "Queen Bee" is thus a more appropriate characterization of women in top positions of power.

For an intuitive introduction to the problem see my video on YouTube.



Endogeneity: An inconvenient truth (full version), by John Antonakis

You'll get to see what Endogeneity looks like!