

# Inference for high-dimensional nested regression

David Gold

Department of Statistics  
University of Washington

September 27, 2017

# Overview

# Endogeneity

- Regressors of interest in econometric research are often **endogenous**: they have non-trivial covariance with the noise term of the response  $Y$ .
- Endogeneity prevents us from discerning co-variation between  $X$  and  $Y$  due to a structural relationship from co-variation that is due to the correlation of random quantities that affect each or both variables.
- Three main sources of endogeneity are:
  - **confounding**, which is due to a discrepancy between the circumstances against which we specify an effect (usually causal) of interest and the circumstances in which we must conduct inference,
  - **errors in measurements** of predictor variables, and
  - the **interplay of mutually influential processes** that also exhibit random variation.

# Endogeneity: Example 1

# Endogeneity: Example 2

# Instrumental Variables

- Suppose that  $X$  is an endogenous variable and that

$$Y = X\beta + U$$

- Suppose also that  $Z$  is a third variable that satisfies ...
- Such a  $Z$  that satisfies ... is called an **instrumental variable**.

# Instrumental Variables: Overview

# Instrumental Variables: Overview (cont'd)



# Model

# Two-step estimation

# Inference for high-dimensional regression parameters

# One-step update

# Design

# Configurations

# Results

# Limitations



## Next steps

# Conclusion

# Acknowledgments