Why does software cost so much?
Towards a Causal Model

How can we better control costs in software development and sustainment? This LENS is collaborating with the larger community of cost estimators in applying causal learning to program datasets to better understand which factors should be adjusted to reduce software lifecycle costs. The resulting causal model will be incorporated into a program dashboard to enable more effective program control.

Problem
• DoD leadership continues to ask “Why does software cost so much?”
• DoD program offices need to know where to intervene to control software costs

Solution
An actionable, full causal model of software cost factors immediately useful to DoD programs and contract negotiators

Actionable intelligence
• Enhance program control of software cost throughout the development and sustainment lifecycles
• Inform “could/should cost” analysis and price negotiations
• Improve contract incentives for software-intensive programs
• Increase competition using effective criteria related to software cost

Technical Approach
Working with collaborators, we will identify and prepare datasets for causal learning to establish key cause-effect relationships among project factors and outcomes, segmented by software lifecycle and software super-domain. For example, for Quality, we might have this causal graph:

The resulting causal models will then be “stitched” using CMU algorithms to create a universal causal model, but estimated and calibrated for lifecycle and super-domain. These estimated models will be the basis for improved program management. (See figure at left.)

Collaborative Approach
First, we train each collaborator so that they are capable of running causal analyses on their own proprietary datasets. Then, they derive causal models and only share with the SEI general information about the dataset and search and the resulting causal graph, sufficient for integrating into a universal model.

Summary
Causal learning has come of age from both a theoretical and tooling standpoint and provides better basis for program control than models based on correlation. Application to cost estimation requires large amounts of quality data. Now is the time to engage the larger community of cost estimators in deriving improved cost models that enable improved program control.

Who would use the tool?
A program's key stakeholders, as part of their regular progress reviews, will examine progress to date and evaluate proposed corrective actions, according to their respective roles:

- Contract Manager: Implications for contract commitments?
- Acquisition Manager: Implications for end user and dependencies on other programs?
- Program Manager: Change priorities for product content and reset expectations?
- Chief Engineer: What changes to plans, resources, and tools are needed to implement action?
Why does software cost so much?