Revisiting "Causal Effects of Software and Systems Engineering Effort" Using New Causal Search Algorithms

Jim Alstad, Anandi Hira, Mike Konrad, A Winsor Brown

Alstad@acm.org, A.Hira@usc.edu, MDK@sei.cmu.edu, AWBrown@usc.edu

University of Southern California and/or Software Engineering Institute

BCSSE COCOMO® Forum

November 9-10, 2022

This material is based upon work funded and supported by the Department of Defense under Contract No. FA8702-15-D-0002 with Carnegie Mellon University for the operation of the Software Engineering Institute, a federally funded research and development center.

DM22-1086

Presentation Outline

- Motivation
- Causal Analysis in SW Cost

- Intro to Causal Inference
- Algorithms and Tool Used

- Datasets – COCOMO® II, COSYSMO 3.0

- Approaches and Results

- Conclusions and Questions

Motivation for this Line of Research

- Managers are frequently faced with issues of controlling project costs
  - My estimated cost is too high. What project aspects can I modify that would most likely reduce the cost?
  - I have some money to improve my organization's performance. Changing which organizational aspects would be most likely to improve cost performance?
  - As an acquirer, I need to add a new stakeholder and remove flexibility in modifying requirements. Is that likely to have a significant influence on project cost?
- Causal Analysis is a modern technique that analyzes datasets to determine causal relationships among its variables
- **The Goal of this Research:** Identify factors of software and systems engineering costs that are direct causes
  - To help manage real projects

Motivation of Revisit

- Causal Learning is an active research area
- New algorithms have been developed and incorporated into Tetrad since our previous analysis [1]
  - New search method (BOSS) to better find causal relationships
  - New validation method (Markov Checker) for causal search results

History of Causal Analyses for Effort

Boehm – COCOMO® Models [2]

- In-depth behavioral analyses for effort drivers
- Including COSYSMO models

Evidence-Based SW Engineering [3]

- Suggests running Experiments to identify causal relationships:
  - Cause precedes effect
  - Cause covaries with effect
  - Alternative explanations are implausible

Unified Code Count Maintenance [4]

- Software maintenance and upgrade data
- Project data has limited scope
  - Similar projects, from a single environment

Our difference: 2 calibration datasets (observational data) with varying values of cost drivers, application types, and project types

Presentation Outline

- Motivation
- Causal Analysis in SW Cost

- Intro to Causal Inference
- Algorithms and Tool Used

- Datasets – COCOMO® II, COSYSMO 3.0

- Approaches and Results

- Conclusions

We Employ Causal Search [5] as the Basis of Our Research

Causal Learning

- **Causal Search/Discovery** — Algorithms and domain knowledge on observational data
- **Causal Inference** — Algorithms to identify and quantify causal effects

Causal Search Algorithm Results

Result is a "causal graph", with each box representing a variable and each edge representing a causal relationship.

Here are the different possible types of edge:

X_1 directly causes X_2

X_1 directly causes X_2 or X_2 directly causes X_1

No directly causal relationship between X_1 and X_2

Tetrad Tool

- Implements causal search and inference algorithms
- Maintained by Carnegie Mellon University (CMU) and University of Pittsburgh
- For information, tutorials and tools:
  - https://sites.google.com/view/tetradcausal
  - https://www.ccd.pitt.edu

- Our 2020-2021 [1] results come from Tetrad versions 6.5.4 and 6.7.0
- Our 2022 results come from a pre-release version of Tetrad 7.1.3 that implements the BOSS algorithm [8]

New Causal Algorithms

BOSS

- Score-based search algorithm based on permutations
- Adaptation and optimization of the Greedy Relaxations of Sparsest Permutation (GRaSP) algorithm in [8]
- In testing, near-perfect adjacency and orientation of graph edges

Markov Checker

- The Causal Markov Condition (CMC) is "all graph-implied conditional independences (CIs) hold in the dataset"
- Many causal algorithms assume CMC
- The new Markov Checker algorithm (MC) tests for CMC
- MC reports what % of CIs implied by the graph hold in the dataset
- 1.0% is the expected rate of CIs failing due to chance; an actual rate near this suggests that the CMC holds, so the causal algorithm results can be trusted

Presentation Outline

- Motivation
- Causal Analysis in SW Cost

- Intro to Causal Inference
- Algorithms and Tool Used

- Datasets – COCOMO® II, COSYSMO 3.0

- Approaches and Results

- Conclusions

Datasets

COCOMO® II Calibration Dataset

- 16 organizations, various application types
- Variability in all 26 variables
- 161 projects
- See [6] for more details

COSYSMO 3.0 Calibration Dataset

- Covers various types of systems
  - > 2 orders of magnitude size variation
- Variability in all 18 variables
- 68 projects
- See [7] for more details

Each dataset is reasonably representative of projects of its type

Presentation Outline

- Motivation
- Causal Analysis in SW Cost

- Intro to Causal Inference
- Algorithms and Tool Used

- Datasets – COCOMO® II, COSYSMO 3.0

- Approaches and Results

- Conclusions

Direct Causes of Effort Found in 2021 Analysis [1]
(FGES Algorithm with Bootstrapping and Weak Signal Analysis #4)

COCOMO® II - Effort

- Size (SLOC)
- Team Cohesion (TEAM)
- Platform Volatility (PVOL)
- Reliability (RELY)
- Storage Constraints (STOR)
- Time Constraints (TIME)
- Product Complexity (CPLX)
- Process Maturity (PMAT)
- Risk and Architecture Resolution (RESL)

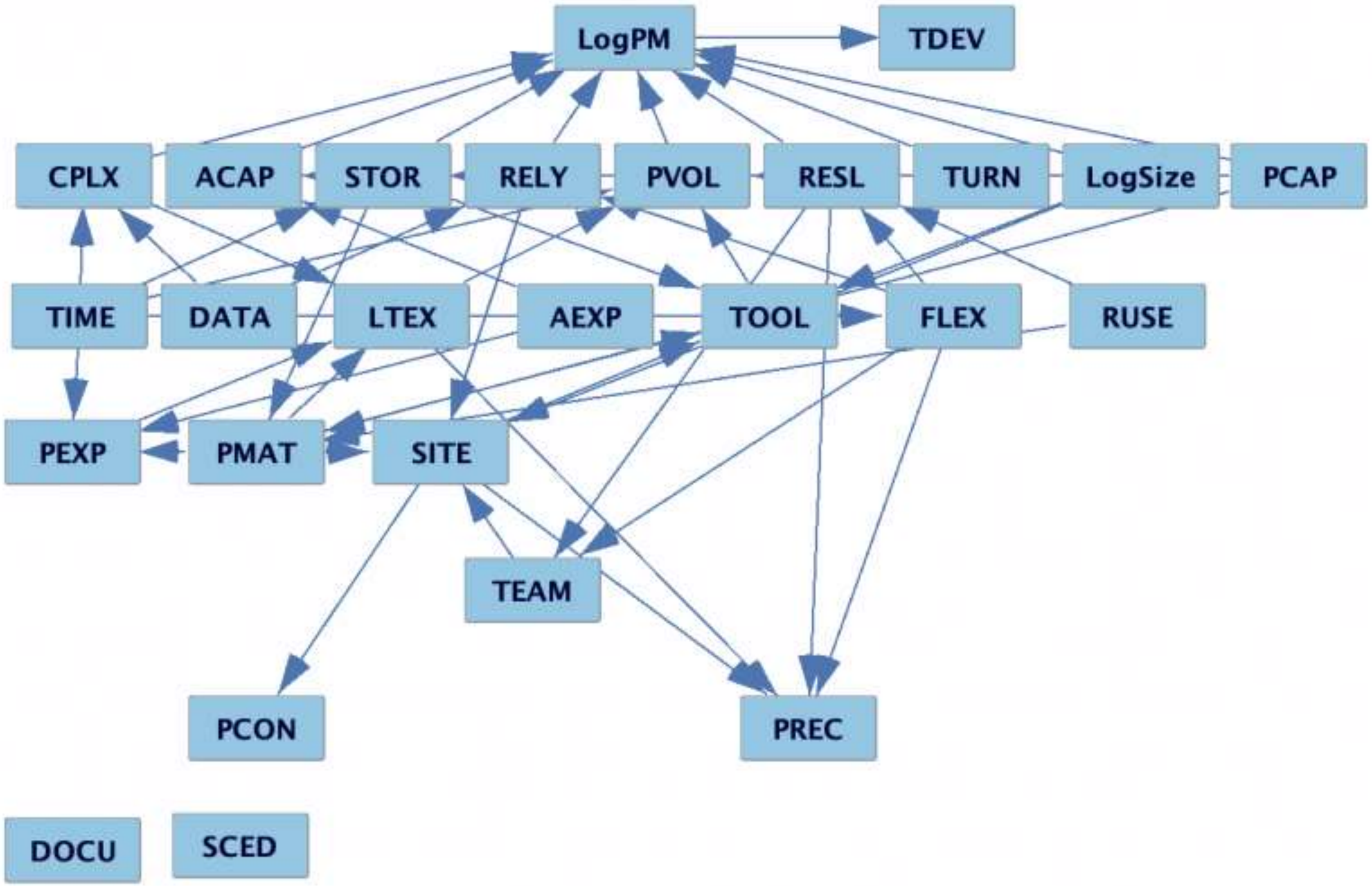
COCOMO® II - Schedule

- Size (SLOC)
- Platform Experience (PLEX)
- Schedule Constraint (SCED)
- Effort (LogPM)

COSYSMO 3.0 - Effort

- Size
- Level of Service Requirements (LSVC)

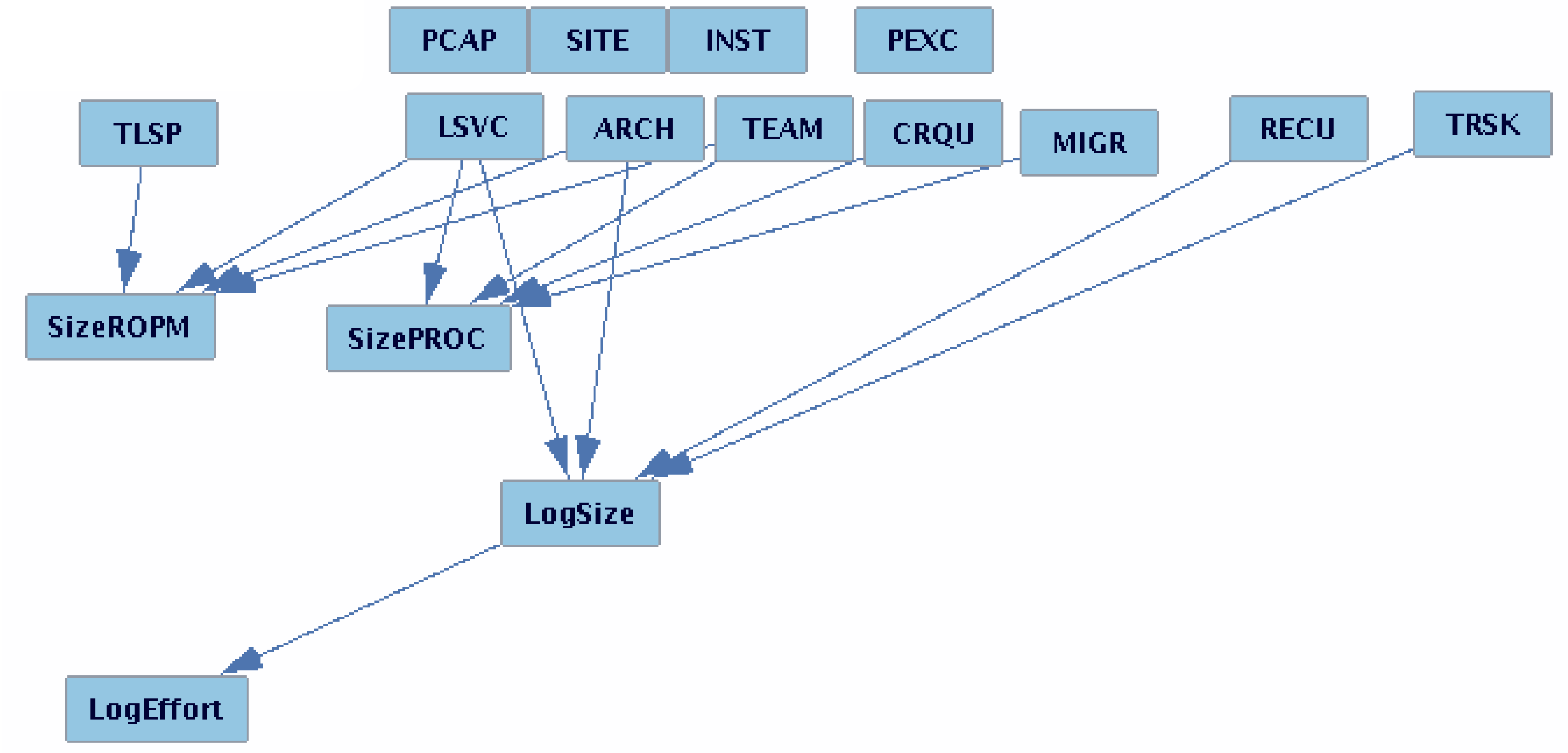
COCOMO® II BOSS Results



Markov Check

- % dependent = 4.08%
- Reasonably close to the expected 1%
- Suggests strong agreement between graph and dataset

COYSYMO 3.0 BOSS Results



Markov Check

- % dependent = 9.52%
- Not as close to the expected 1%
- Suggests agreement between graph and dataset

Changes to the **Direct** Causes of Effort/Schedule

COCOMO® II - Effort

- Size (SLOC) (LogSize)
- ~~Team Cohesion (TEAM)~~
- Platform Volatility (PVOL)
- Reliability (RELY)
- Storage Constraints (STOR)
- ~~Time Constraints (TIME)~~
- Product Complexity (CPLX)
- ~~Process Maturity (PMAT)~~
- Risk and Architecture Resolution (RESL)
- **Programmer Capability (PCAP)**
- **Analyst Capability (ACAP)**
- **Computer Turnaround Time (TURN)**

And 12 of the COCOMO® II factors influence Effort indirectly

COCOMO® II - Schedule

- ~~Size (SLOC) (LogSize)~~
- ~~Platform Experience (PLEX)~~
- ~~Schedule Constraint (SCED)~~
- Effort (LogPM)

COSYSMO 3.0 - Effort

- Size
- ~~Level of Service Requirements (LSVC)~~

And 3 of the COSYSMO 3.0 factors influence Effort indirectly

Presentation Outline

- Motivation
- Causal Analysis in SW Cost

- Intro to Causal Inference
- Algorithms and Tool Used

- Datasets – COCOMO® II, COSYSMO 3.0

- Approaches and Results

- Conclusions

Conclusion – Causal Search

- How do BOSS results compare to previous results [1]?
  - Our previous results used the algorithms WSA#4 (custom script outside of Tetrad), and FGES with bootstrapping. This time, we did no bootstrapping and used BOSS.
  - COCOMO® II Results
    - Gives similar # of causes without the extra data analyst effort
    - Provided additional controls for project managers but fewer controls for organizations' upper management or acquirers
  - COSYSMO 3.0 Results
    - Discovered additional causes
- Does Markov Checker seem to provide useful info?
  - Idea and its implementation by Elias Bareinboim (Columbia), Joe Ramsey (CMU), and Mike
    - This presentation may be the first formal reporting of such a validity check anywhere in the literature.
  - Evaluated validity of causal structures obtained from causal search
  - Results: We think a reasonably low "% dependent" was returned on both COCOMO® and COSYSMO datasets
  - Initial evaluation of Markov Checker: seems useful but difficult to interpret "% dependent"

Future Work

- Compute new estimation “mini-models” based on our new searches
- Continue to check out new algorithms as they become available in Tetrad

Bibliography

1. Alstad, J., Hira, A., Brown, A.W., Konrad, M. (2021). " Investigating Causal Effects of Software and Systems Engineering Effort", ICEAA 2021 Online Workshop.
2. Boehm, B. W. (1981). *Software Engineering Economics.* Upper Saddle River, NJ: Prentice-Hall, Inc.
3. Kitchenham, B. A., Dyba, T., & Jorgensen, M. (2004, May). Evidence-based software engineering. In *Proceedings. 26th International Conference on Software Engineering* (pp. 273-281). IEEE.
4. Hira, A., Boehm, B., Stoddard, R., & Konrad, M. (2018). Further Causal Search Analyses With UCC's Effort Estimation Data.
5. Spirtes, P. (2010). Introduction to causal inference. *Journal of Machine Learning Research*. 11, May (2010), 1643–1662.
6. Boehm, B. W., et al. (2000). *Software Cost Estimation with COCOMO II.* Upper Saddle River, NJ: Prentice-Hall, Inc.
7. Alstad, JP. (2019). "Development of COSYSMO 3.0", *Procedia Computer Science 153*
8. Lam, W. Y., Andrews, B., & Ramsey, J. (2022, February). Greedy Relaxations of the Sparsest Permutation Algorithm. In *The 38th Conference on Uncertainty in Artificial Intelligence*.