

Carnegie Mellon University

AutoMerge

Francisco Fonseca Octavio Mesner

March 17, 2016

Goals

Discuss

- ▶ Project Overview
- ▶ Data
- ▶ Describe Lasso
- ▶ Possible causal pathways

MOMS Data

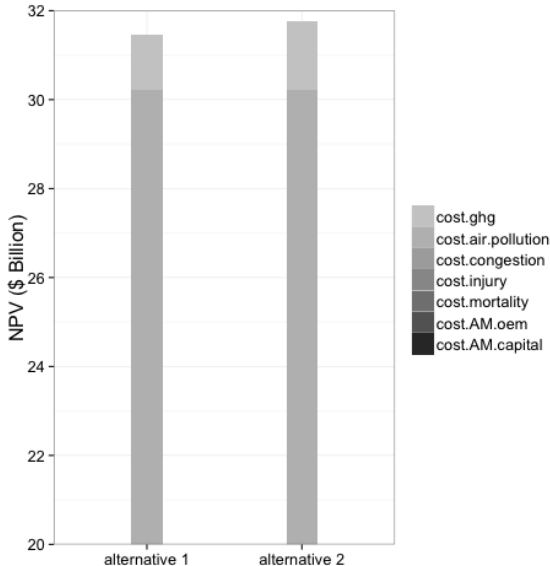
Overview

MOMS Data

Methods

Evaluation
Lasso

- ▶ 744 subject observations
- ▶ 352 covariates collected
- ▶ 121 variables with $>70\%$ missing (from substudy?)
- ▶ 57 (8%) preterm births/629(92%) full term (58 missing)



Cross Validation

MOMS Data

Methods

Evaluation

Lasso

How well does a model perform on new data?

What is Lasso?

- ▶ Builds on multiple linear regression
- ▶ Includes a penalty term for size of parameters