Causal Inference of Distressed Securities

Patrick Rogan

Columbia Engineering

May 4, 2016

Outline

Causal Inference of Distressed Assets

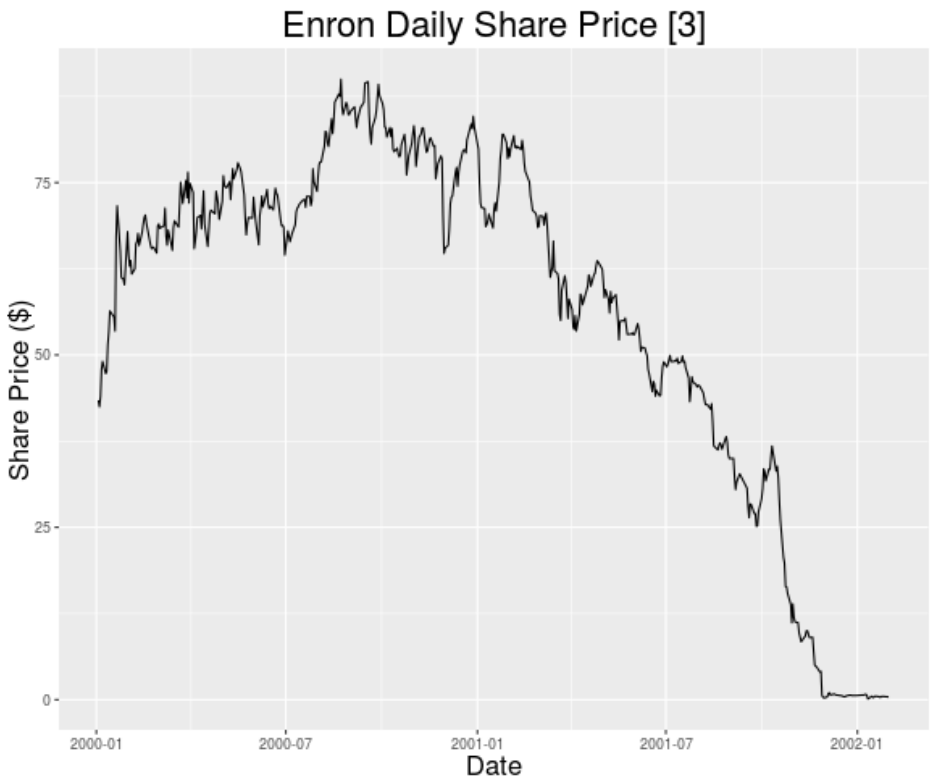
- Background/Motivation
- Problem Statement/Objective
- Previous Research
- Approach
- Experimental Design & Implementation
- Research Actions and Next Steps

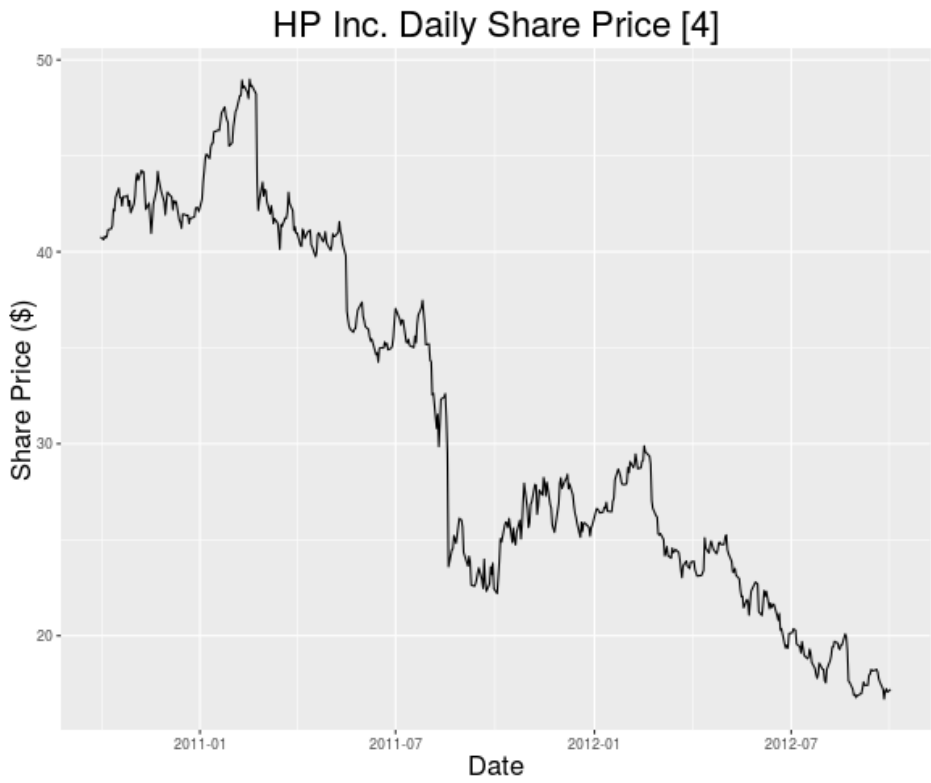
Background/Motivation

Understanding Distressed Assets
Over the past 30 years, bankruptcies of large publicly held American companies resulted in the loss of hundreds of billions of dollars in shareholder wealth [2].

[1] Images courtesy of http://www.brandsoftheworld.com
[2] http://www.bankruptcydata.com/Research/Largest_Overall_All-Time.pdf

Background/Motivation

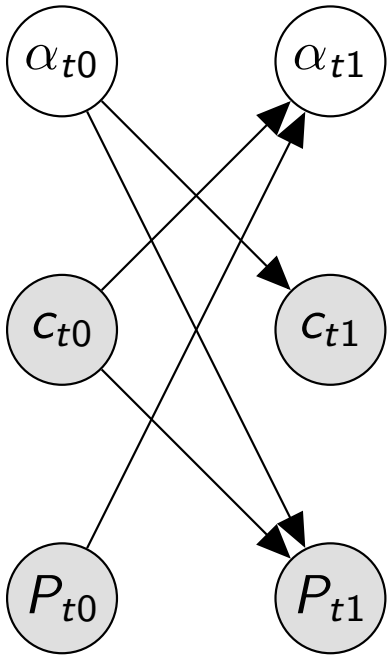




Can we use causal inference to differentiate outcomes?

[3] Data Retrieved From: http://www.gilardi.com/pdf/enro13ptable.pdf
[4] Data Retrieved From: http://finance.yahoo.com/

Problem Statement/Objective



Research Question Can commonly reported financial indicators be causally linked to future prices performance in the case of assets that have recently undergone a substantial decrease in share price [5]?

We assume the following situation: latent variables (eg. changing market conditions, fraud, macroeconomic forces) inform reported corporate financials. We attempt to find a causal relationship between these variables and share price.

[5] P. Rogan (2016). EECS 6898: Project Proposal. Unpublished Manuscript.

Previous Research

Granger Causality

- Examines all relationships between predictors across time.
- Empirical research has found this model tends to identify causal relationships from simple correlations.
- Popular in the fields of economics and finance.

Causal Inference of Stock Prices

- S. Kleinberg investigated the causal relationship between individual stock prices from 2000 - 2007 [6].
- Investigation found short term relationships between certain assets (eg. the prices of financial companies appear to be influenced by changing prices in the technology sector).

[6] S. Kleinberg. Causality, Probability, and Time. Cambridge University Press, 2012. Cambridge Books Online.

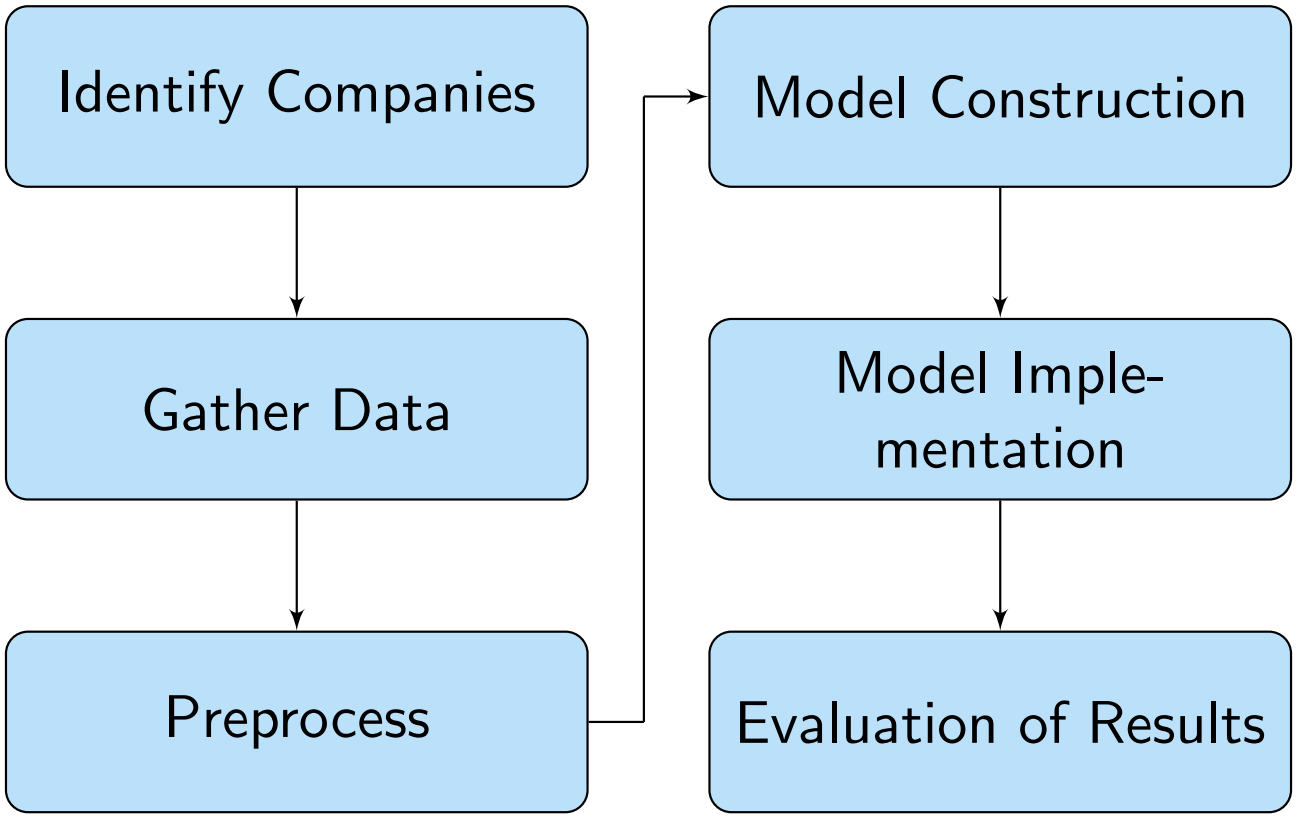
Approach

What we're looking for: We seek to infer relationships of the form "Drop in revenue and cash causes bankruptcy in 6-12 months with a probability of 0.4." To do this we need:

- Regular reports on the financial status of publicly traded companies.
- Time series data on the price of assets.
- A temporal-logical model.

Experimental Design & Implementation

Overview



Straightforward desktop-scale implementation.

Experimental Design & Implementation

Specifics

Identification of Companies: Manual search of sources. Identify companies that are "stable" and those that undergo significant financial declines resulting in bankruptcy or recovery.

Gather Data: Stock price data can be easily obtained through an R API for Yahoo Finance (quantmod) [7]. Financial reports can be obtained through the Security Exchange Commission's Electronic Data Gathering, Analysis, and Retrieval (EDGAR) system [8].

Preprocess Data: Ability to process data retrieved from Yahoo Finance already exists. SEC 10-Q forms are text based, processing techniques under construction.

[7] quantmod: Quantitative Financial Modelling Framework. http://www.quantmod.com/.
[8] SEC.gov. https://www.sec.gov/edgar/searchedgar/companysearch.html.

Experimental Design & Implementation

Specifics (Continued)

Model Construction/Implementation: Create temporal-logical based model. As part *Causality, Probability, and Time*, S. Kleinberg has provided an analysis and data. Use this assist in the creation of a viable model.

Model Evaluation: Devise a method for evaluating the performance of the model. This will likely be a holdout dataset or splitting timeseries data into in sample and out of sample sections.

Research Actions and Next Steps

Research Actions

- Perform FTP retrieval of financial statements for identified companies.
- Process records into a usable format.
- Investigate automated implementations of temporal-logical models.
- Evaluate results against hold out dataset.

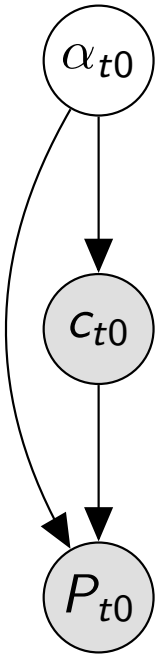
Next Steps

- Pull in other data sources (what we had previously deemed latent).
- Investigate the effects of "interventions" using Rubin causal model.

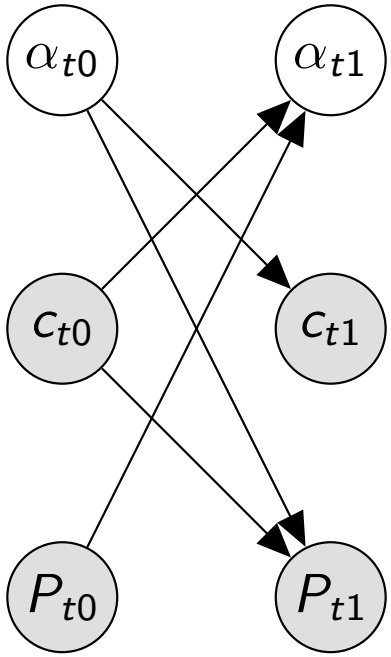
Questions

Questions?

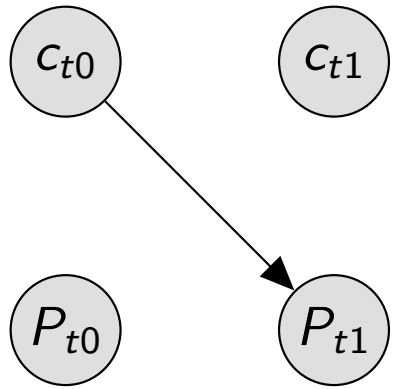
Backup: Various DBNs



(a) Initial State



(b) Time Evolution (All Variables)



(c) Time Evolution (Observed Variables)