

How are DAGs made?

A journey into modelling price elasticity

Paweł Olszewski

data scientist @ Allegro

27.05.2026

Meet **Tomek**,
a newly hired Data Scientist at Allegro.

His first assignment?

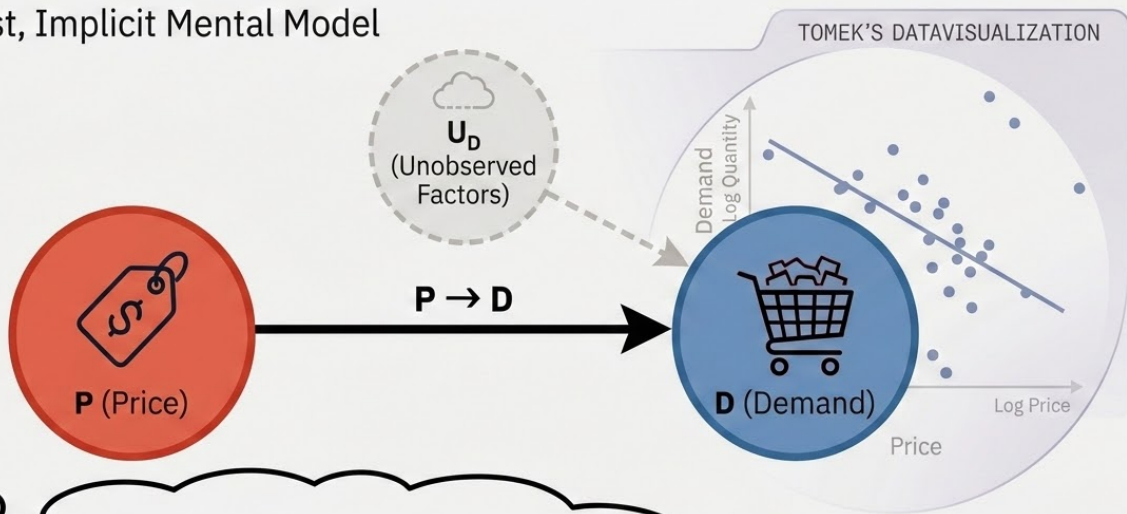
Automated price discounts!

Given a fixed monthly budget,
decide where to discount and how
much to maximize the total effect
on sales.



ACT 1: THE NAIVE OBSERVATIONAL MODEL

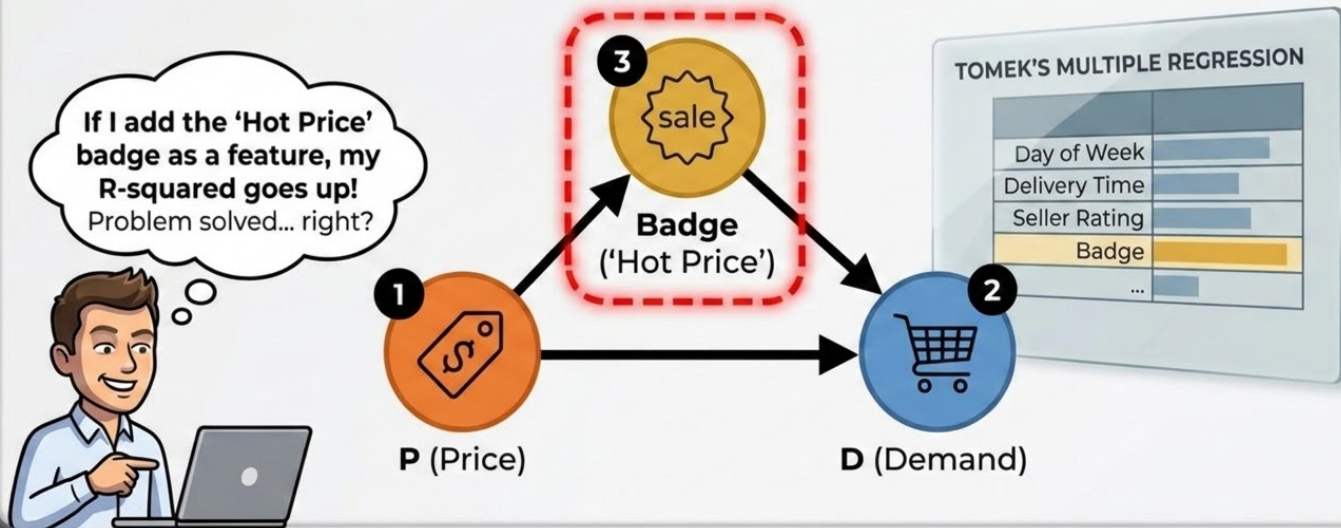
Tomek's First, Implicit Mental Model



So, the lower the price, the more we sell, right?
Simple regression, $\beta = -1.5$, I'm done!

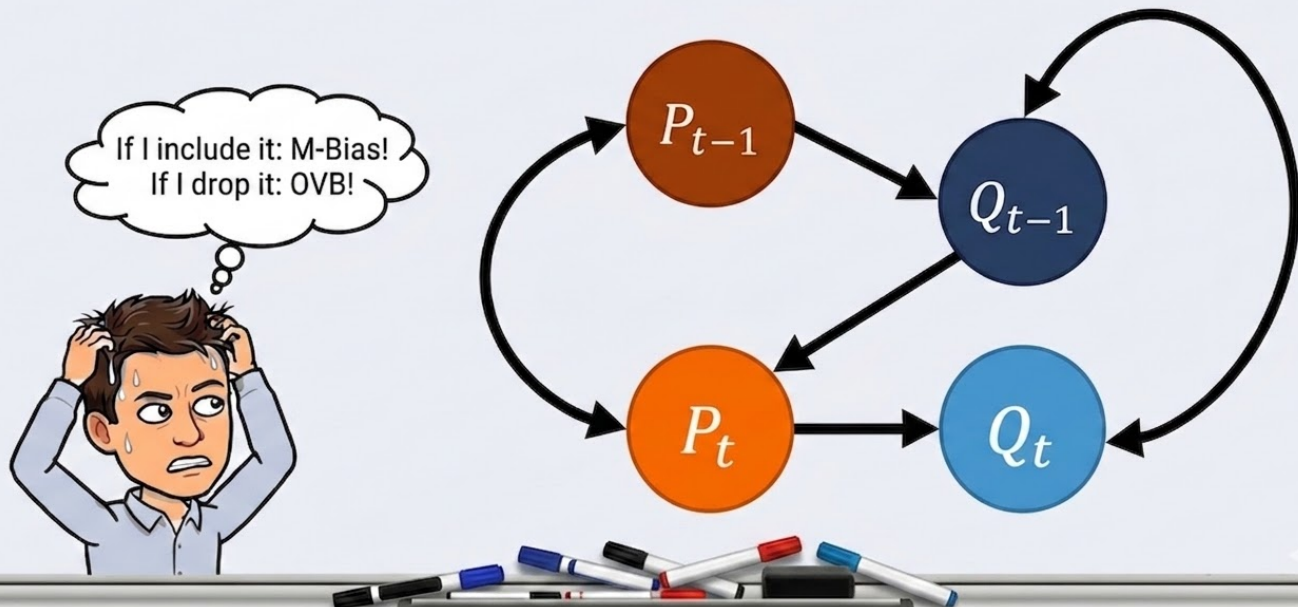
ACT 1: FEATURE STUFFING & THE MEDIATOR TRAP

Tomek Controls for a Post-Treatment Variable



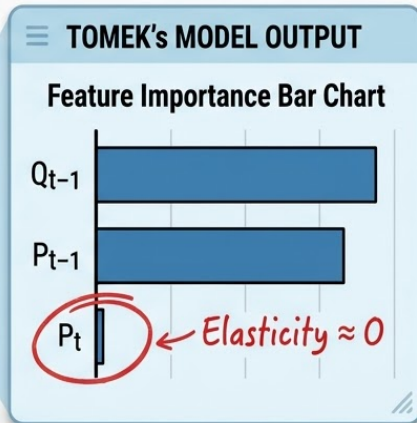
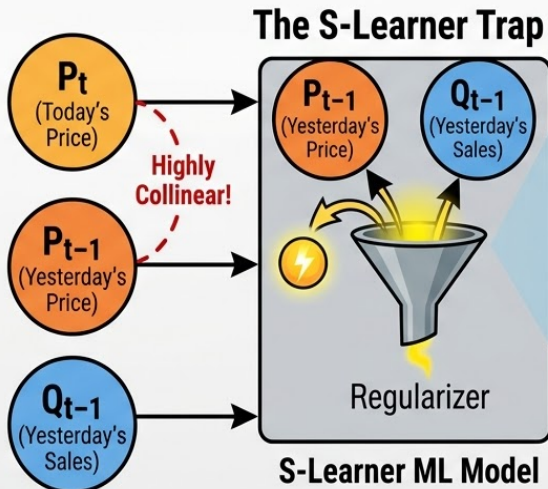
ACT 2: THE CONFOUNDER-COLLIDER DUALITY

Tomek's Econometric Nightmare



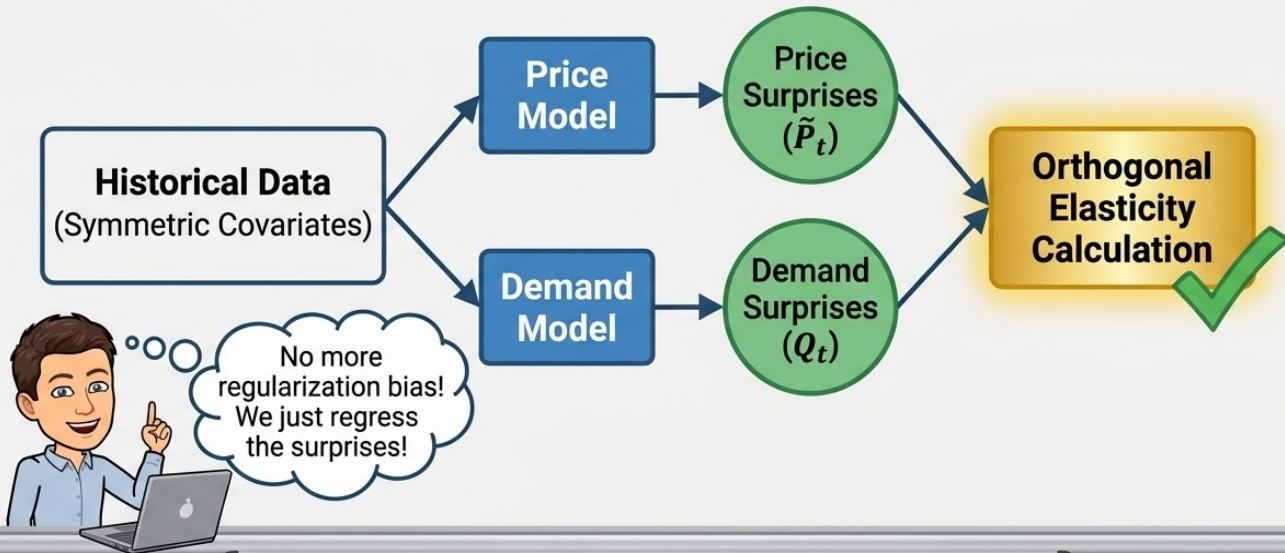
ACT 3: THE S-LEARNER TRAP & REGULARIZATION BIAS

Tomek Discovers Weight Stealing



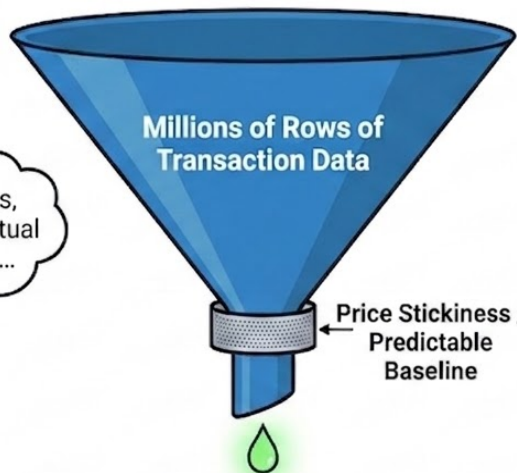
ACT 4: DOUBLE MACHINE LEARNING

Tomek Isolates the Unexpected Jumps

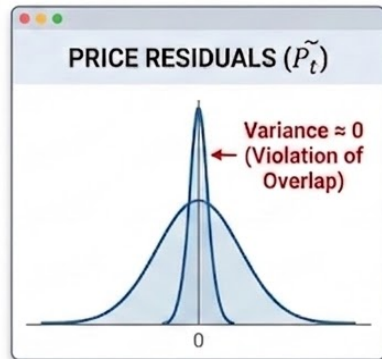


ACT 4: THE POSITIVITY ASSUMPTION

The Small Data Reality of Sticky Prices



Millions of rows,
but almost no actual
price changes...



ACT 5: THE CAUSAL FOREST

Grouping by Elasticity, Not by Demand

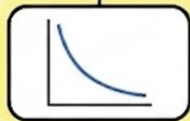
Just import
CausalForestDML...
let the library handle
the variance penalty math!



Standard ML: Groups
by similar sales levels.

Product
Embeddings (X)

Algorithm maximizes
the difference in
ELASTICITY



Highly Elastic Products



Inelastic Products



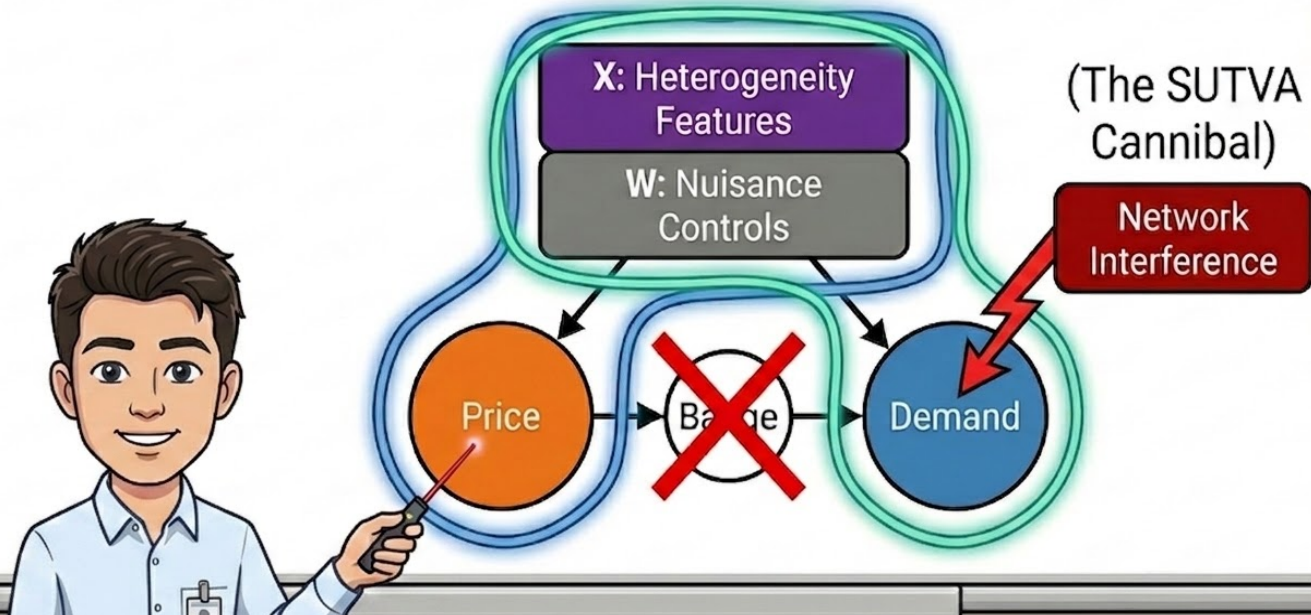
ACT 6: SUTVA & CANNIBALIZATION

The Monster Under the Bed



IT'S ABOUT DAGS WE MAKE ALONG THE WAY

Tomek's Final Causal Architecture



Tips for price elasticity modelling

1. Identify and exclude mediators (if you cannot model the $P \rightarrow M$ effect separately).
2. You probably should include the history of the outcome in your features (autoregression). But don't forget to include history of the treatment as well.
3. An ML S-learner will regularize some of your treatment effect towards zero.
4. Model the treatment to expose problems with positivity.
5. Estimating heterogeneity of the effect (rather than just ATE) is another level of difficulty. Consider pairing DML with CausalForest (note: the final stage of DML is *non-causal*).

Chances are no one else is going to want to see your DAG. Draw it anyway :)