

An Expected Value Model for Maximizing Revenues from Sales Tax Audits

Michael Sherman, Nicole White & Ying Du | The University of Texas at Austin | McCombs School of Business | Department of Information, Risk & Operations Management | MSIROM: Business Analytics

Project Goal

Develop a simple, accurate, and deployable predictive model for determining which sales tax returns to audit.

Motivating Problems

- Tax evasion exceeds \$3.1 trillion globally per year.
- ~20% of US tax returns have inaccuracies.
- In some developing countries, up to 50% of taxes are evaded.
- Audits are expensive. Only ~1% of tax returns are audited.

Traditional Methods

- Audit companies that pay out the most in wages.
- Audit companies with the highest gross incomes.
- Audit companies with the most past instances of fraud.
- Audit companies deducting the greatest % of their income.

Prior Research

- Very little published research on predicting fraud from tax returns.
- Of the published research, primary focus is the *presence* of fraud rather than the *extent* of fraud.
- Real-world auditing operations are concerned with both finding fraud and maximizing the collection of owed taxes.

Our Method

Expected Value of Audit =
Probability Claim is Fraudulent

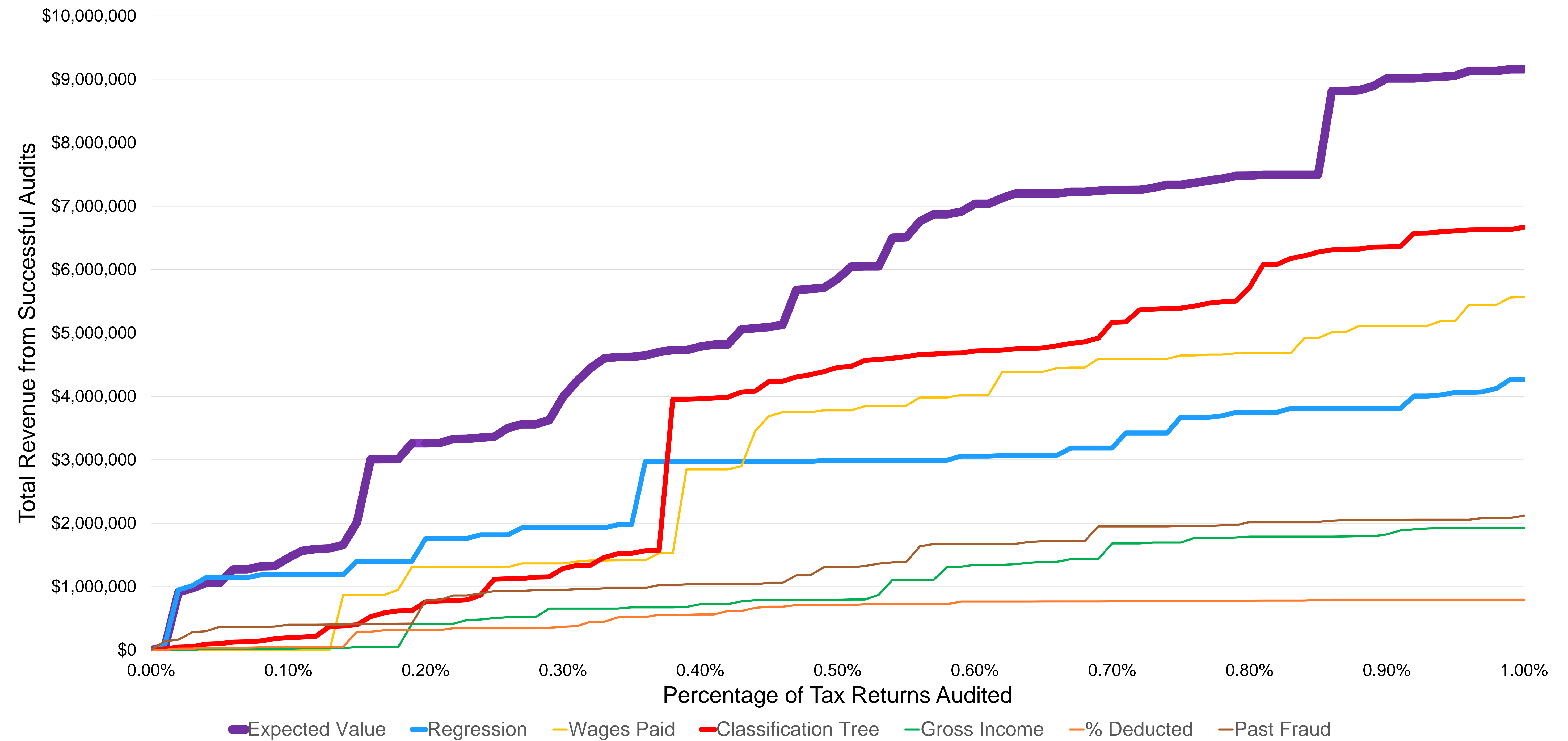
×

Predicted Amount of Underpaid Tax

- For each return, create two simple models:
 - A classification tree assigns each return a probability of being fraudulent.
 - A multiple linear regression assigns each return a predicted amount of underpayment.
- Multiply the results of the two models to get an expected value.
- Rank returns by expected value and audit from highest to lowest.

Results

Revenue from Different Tax Audit Selection Methods



Model	Lift
Expected Value	17.58
Classification Tree	12.80
Wages Paid	10.69
Multiple Linear Regression	8.19
Past Frequency of Fraud	4.07
Gross Income	3.69
% Income Claimed as Deductions	1.52
Random	1.00

The expected value model generates more revenue than established audit selection methods and alternative models while remaining easy to construct and interpret.

Acknowledgements

- Dr. Maytal Saar-Tsechansky
- Dr. Rajiv Garg
- Department of Information, Risk & Operations Management
- McCombs School of Business
- MSIROM: Business Analytics Program