Session Three:

Drivers of Noncompliance

2010 IRS Research Conference

Predicting Intentional and Inadvertent Non-Compliance

Prof. Kathleen M. Carley, Dawn Robertson,
Michael K. Martin, Ju-Sung Lee, Jesse L. St. Charles, Brian Hirshman

Internal Revenue Service
United States Department of the Treasury

POC: Kathleen M. Carley (PI)
Wean 5301
ISR, SCS, Carnegie Mellon
5000 Forbes Ave.
Pittsburgh, PA 15213 USA

412-268-6016
kathleen.carley@cs.cmu.edu

Carnegie Mellon

Center for Computational Analysis of Social and Organizational Systems
http://www.casos.cs.cmu.edu/

Intentional and Inadvertent Non-Compliance

- Goal: Develop models for characterizing and predicting intentional and inadvertent errors given tax returns
- Approach:
  - Team A: Develop theoretical “first principles” models using meta-analysis of information in the public literature
  - Team B: Develop empirical models using statistical machine learning techniques from fused EOAD and Preparer data
  - Team C: Combine these into unified error models
  - Analysis for entire return and for major line items
- Possible applications:
  - Improvement of services aimed at reducing error by customizing response by class of errors
  - Error model for use in simulations to enable more accurate forecasts of impact of other services on change in error rates
  - Improved training and support for examiners

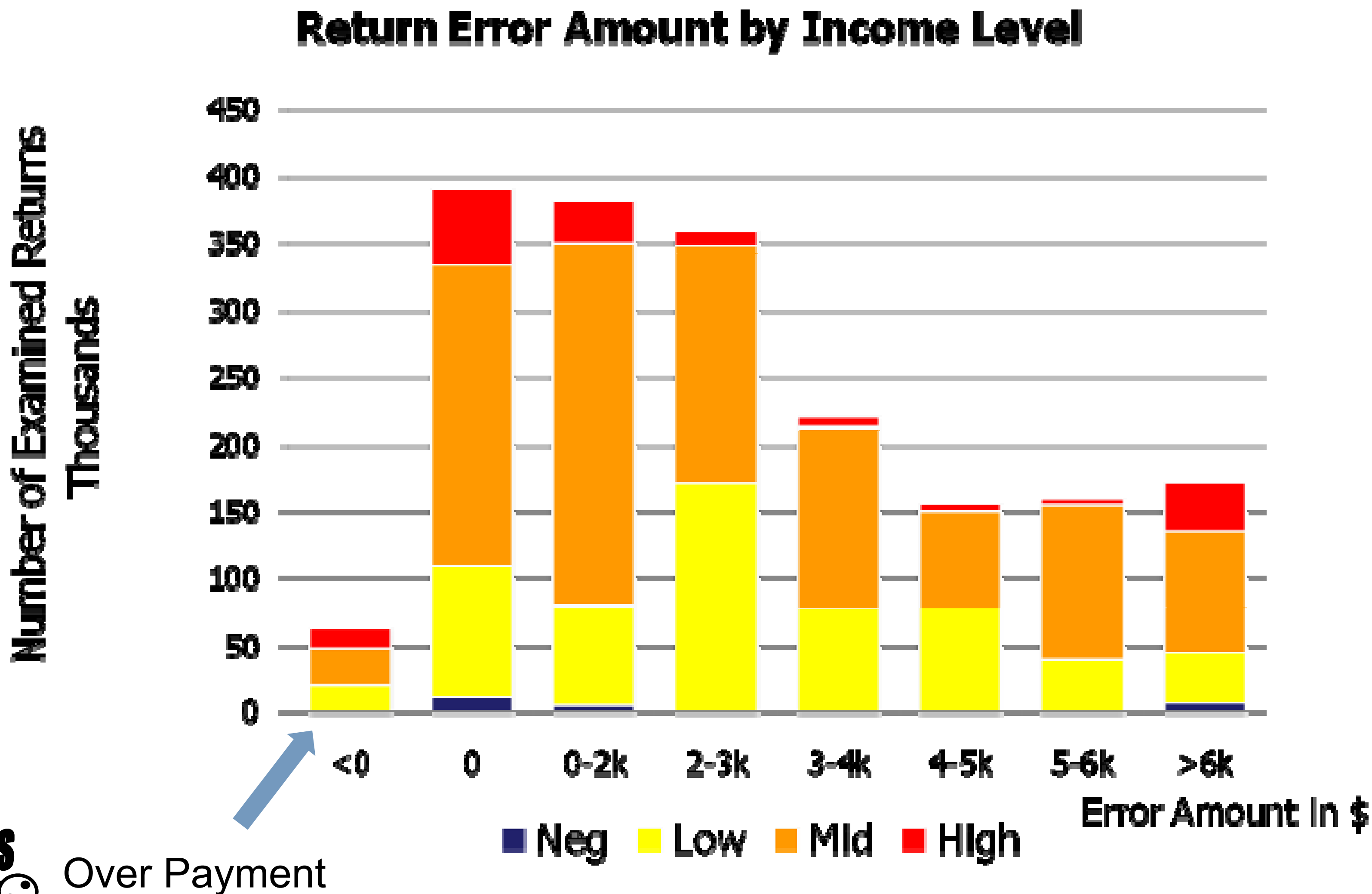
Categories for Variables Used in Models

- EITC: yes/no
- Age: <30, 30-60, >60
- Burden/complexity: from EOAD and IRTF: low, medium, high
- Late Code: On time, extension, late, no-file
- Filing Status: Single, Married-Filing Jointly, Married-Filing Separately, Head of Household
- Itemization: Yes/No
- Exemptions: 0-5, 6+
- Preparer: Self, Paid and IRS
- Error Amount
- Income

Error Adjustment	Bin
<$0	0
$0	1
>$0 , <$2K	2
>$2K , <$3k	3
>$3K , <$4k	4
>$4K , <$5k	5
>$5K , <$6k	6
>$6k	7

Income	Bin
AGI < $0	Negative
AGI = 0	Low
$0 < AGI <$15k	Low
$15K < AGI < $30k	Middle
$30K < AGI < $50k	Middle
$50k < AGI < $80k	Middle
$80k < AGI < $120k	Middle
AGI > $120k	High

Most Errors Lead to Under-Reporting



Inadvertent and Intentional Errors for tax returns as a Whole

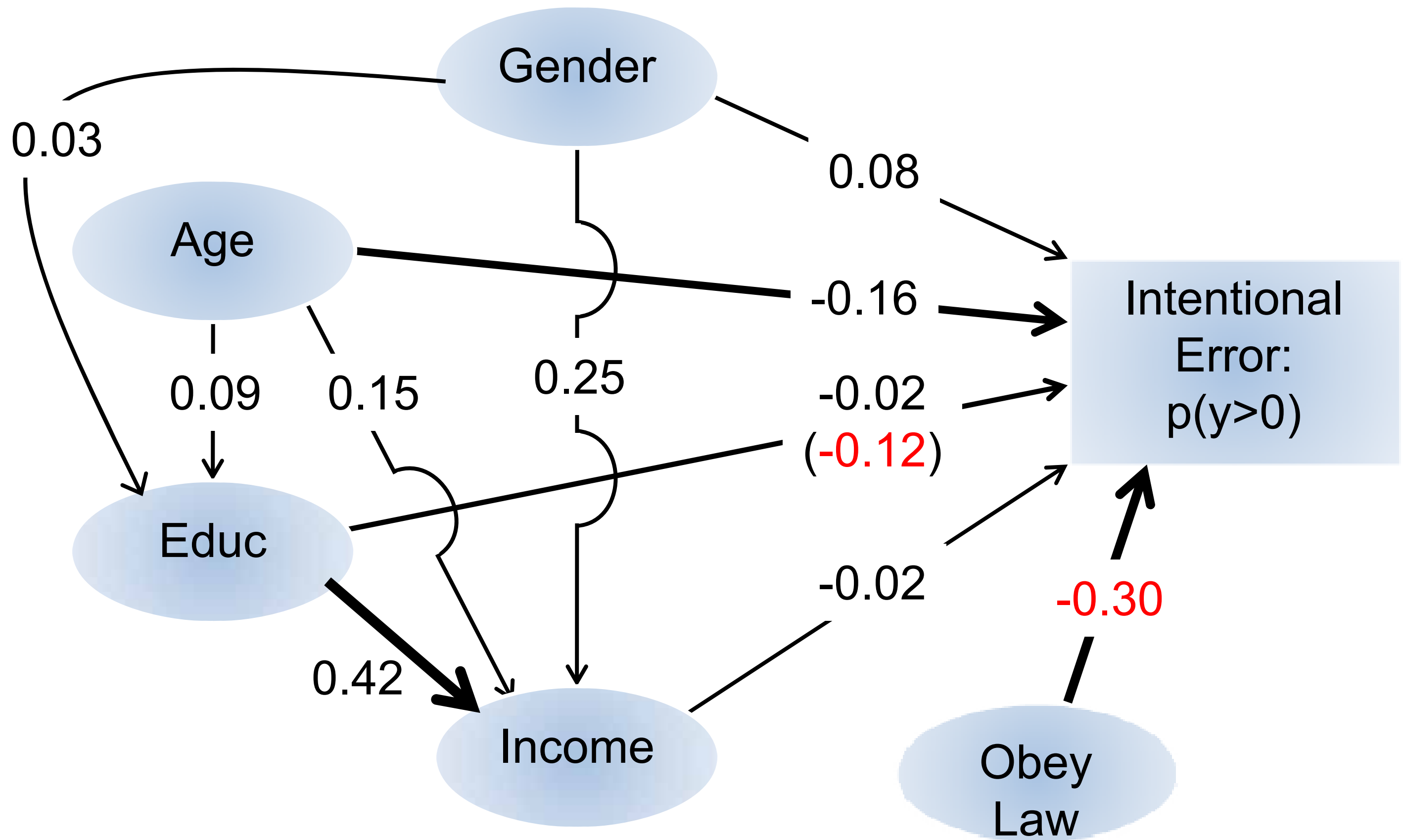
Distribution of Errors as marked by examiners per income group

	Negative	Low	Middle	High	Total
Inadvertent	23,498	270,356	630,648	117,910	1,042,412
Not Inadvertent	5,133	308,807	424,153	46,392	784,485
Intentional	4,671	10,290	46,396	14,110	75,467
Not Intentional	23,960	568,873	1,008,405	150,192	1,751,430
Total	28,631	579,163	1,054,801	164,302	1,826,897
Inadvertent %	82%	47%	60%	72%	57%
Intentional %	16%	2%	4%	9%	4%

Variables Used in Models

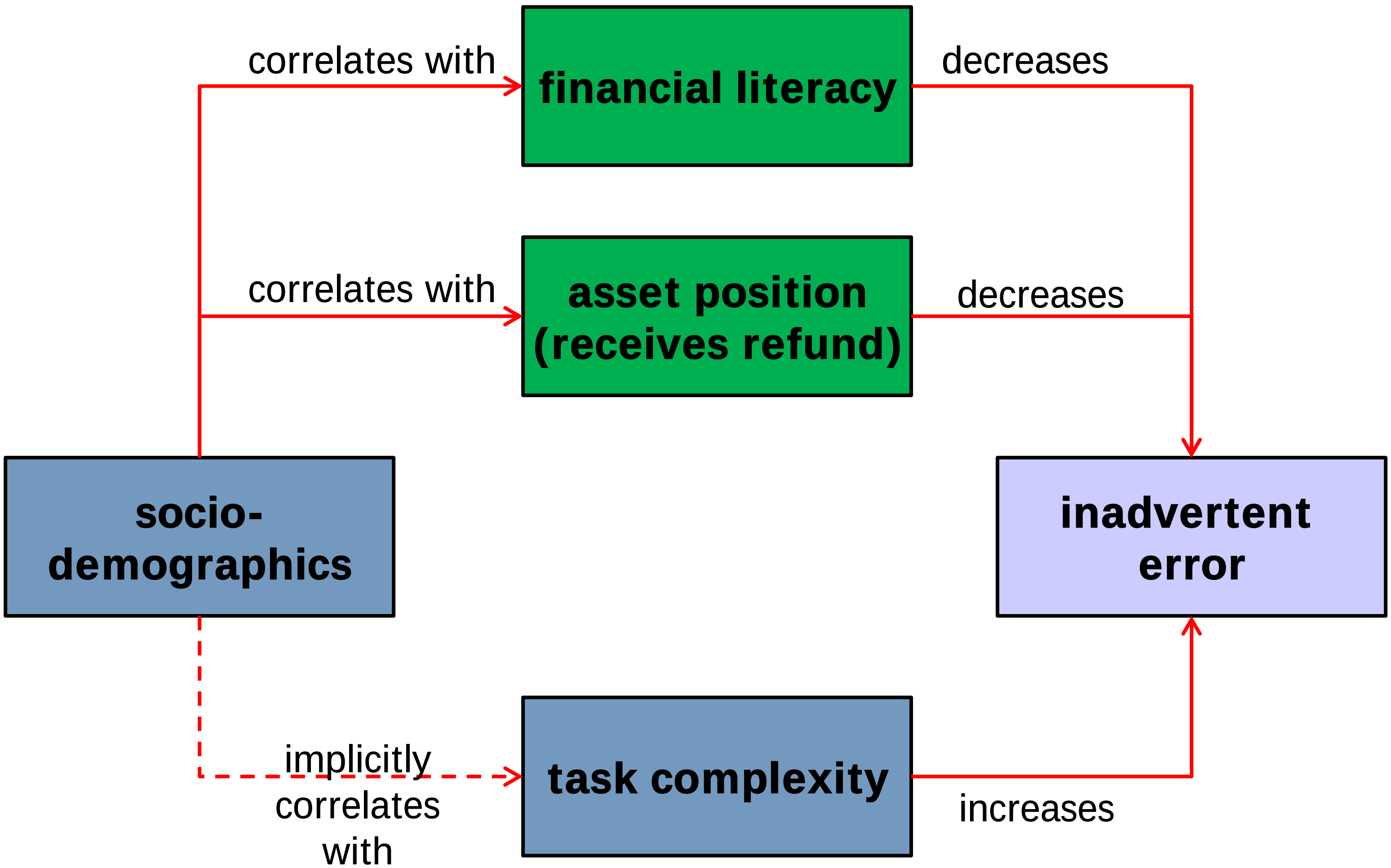
Variable	1st Principle Intentional	1st Principle Inadvertent	Statistical / Machine Learning
EITC	no	no	yes
Age	yes	yes	yes
Burden/Complexity	no	yes	yes
Late	no	no	yes
Filing Status	no	no	yes
Itemization	no	no	yes
Exemptions	no	no	yes
Preparer	no	no	yes
Error Amount	no	no	yes
Income	yes	yes	yes
Gender	yes	yes	no
Belief in obey law	yes	no	no
Education	yes	yes	no
Expect refund	no	yes	no

Path Diagram of the First Principle Intentional Error Model



Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.149	0.248	0.304	**0.307**	0.368	0.481

First Principle Inadvertent Error Model



Data Used for Statistical / Machine Learning Model

- EOAD: 1,902,315 exams with matching IRTF data
- IRTF data : ~140M records available for matching with EOAD data for additional variables
- Complexity variable:
  - 0 – Simple: Form 1040, 1040A, or 1040EZ w/o schedules
  - 1 – Intermediate:
    - Form 1040A with schedules
    - 1040 with schedules A,B,D, Additional Child Tax Credit, Educational Credits, Child Care Credit, Credit for the Elderly or EITC
  - 2 – Complex: Form 1040 with schedules C,E or F or other schedules and all other specific Forms 1040, e.g. 1040PR, etc.

NOTE: We only have the line items considered in the exam to determine which schedules were used and so estimate complexity. Thus, we are probably underestimating complexity.

Inadvertent Error Predictions by Models Given Labeled Tax Returns

Income	Negative		Low		Middle		High	
	Confirmed Error	Potential Error	Confirmed Error	Potential Error	Confirmed Error	Potential Error	Confirmed Error	Potential Error
BNP	80.96%	16.51%	30.03%	16.51%	49.93%	23.51%	71.76%	28.24%
PL	80.98%	16.51%	29.80%	17.16%	49.42%	24.39%	70.29%	26.64%
BNP ∩ PL	80.51%	16.96%	28.24%	18.52%	47.02%	26.46%	70.29%	28.24%
BNP ∪PL	81.43%	16.06%	31.59%	15.14%	52.33%	21.44%	71.76%	26.64%
Confirmed Maximum	82.00%		47.00%		60.00%		72.00%	

∩ all models say case is a confirmed error
∪ at least one model says case is confirmed error
Confirmed maximum = Percentage cases marked as inadvertent by examiner

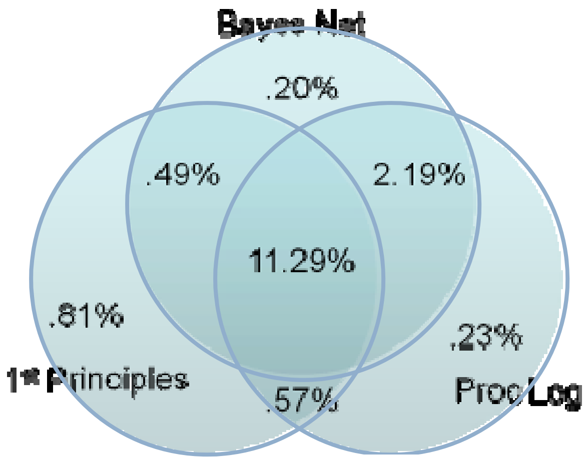
Intentional Error Predictions by Models Given Labeled Tax Returns

Income	Negative		Low		Middle		High	
	Confirmed Error	Potential Error	Confirmed Error	Potential Error	Confirmed Error	Potential Error	Confirmed Error	Potential Error
BNP	14.17%	56.02%	1.03%	4.36%	2.72%	13.14%	5.71%	24.92%
PL	14.28%	55.43%	1.03%	4.66%	2.68%	13.06%	6.09%	27.04%
FP	13.15%	59.41%	1.43%	42.20%	2.61%	37.80%	4.79%	33.43%
FP∩PL	11.86%	71.35%	0.94%	42.59%	2.02%	41.14%	3.73%	46.45%
FP ∪ PL	15.58%	43.49%	1.53%	4.27%	3.27%	9.72%	7.15%	14.01%
FP∩BNP	11.78%	71.98%	0.93%	42.52%	2.03%	41.40%	3.65%	44.52%
FP ∪ BNP	15.54%	43.45%	1.53%	4.04%	3.30%	9.54%	6.85%	13.82%
BNP∩PL	13.48%	62.01%	0.93%	5.33%	2.57%	14.28%	5.42%	29.93%
BNP ∪ PL	14.97%	49.45%	1.14%	3.69%	2.83%	11.93%	6.38%	22.03%
∩ all	11.29%	73.61%	0.85%	42.68%	1.95%	41.74%	3.51%	48.03%
Union All	15.77%	39.13%	1.55%	3.46%	3.34%	8.92%	7.29%	12.51%
Confirmed Maximum	16.00%		2.00%		4.00%		9.00%	

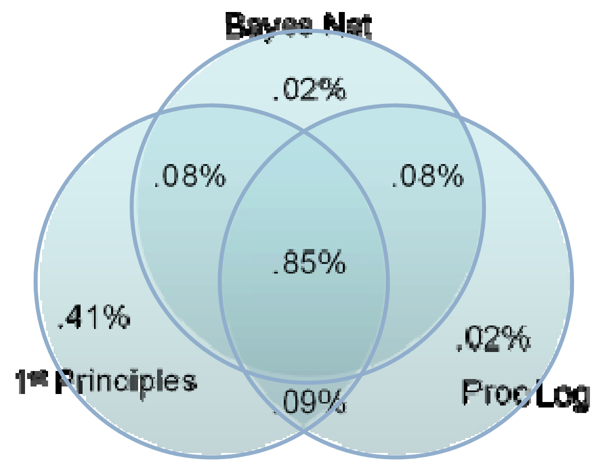
Confirmed maximum = Percentage cases marked as intentional by examiner

Result: Overlap of Intentional Error Models

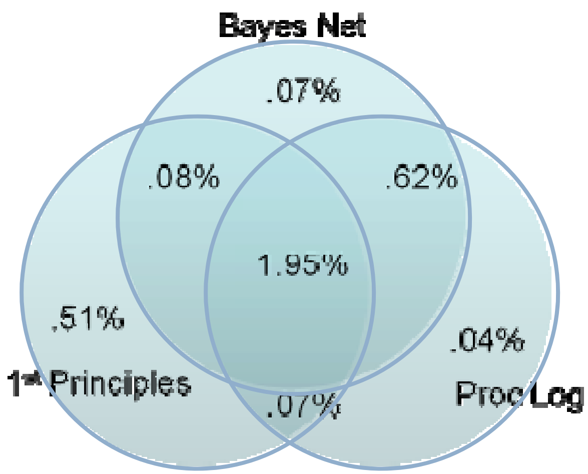
- Negative income



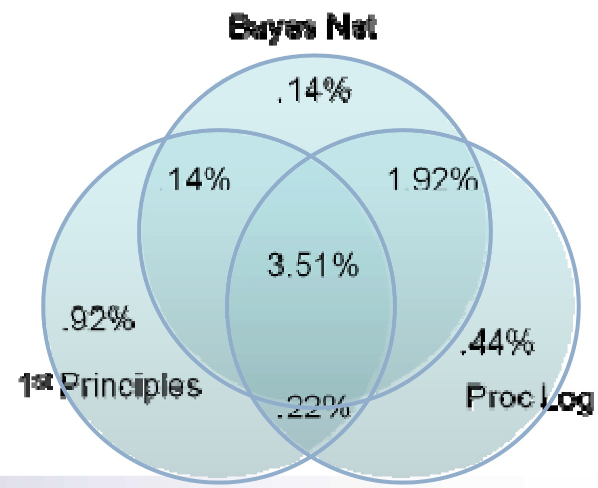
- Low income



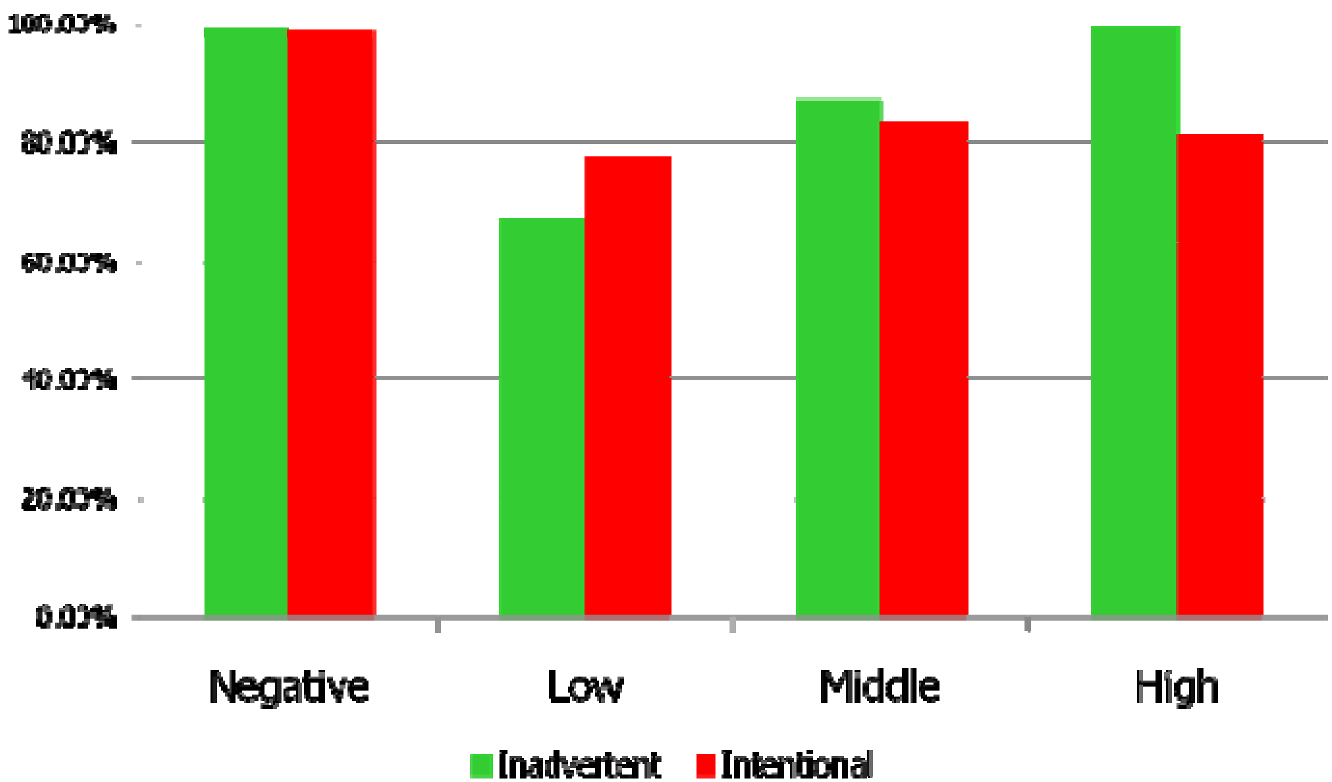
- Mid income



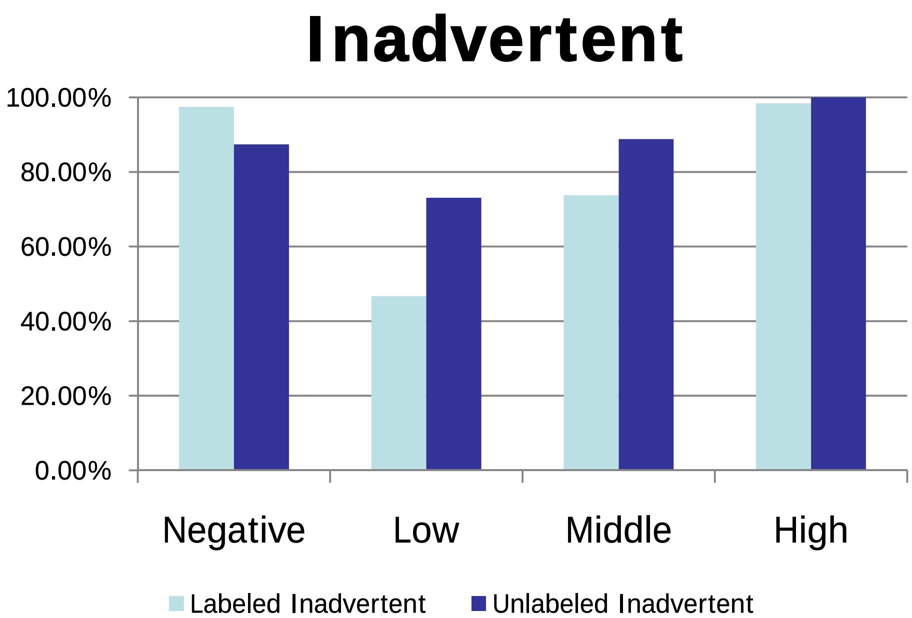
- High income

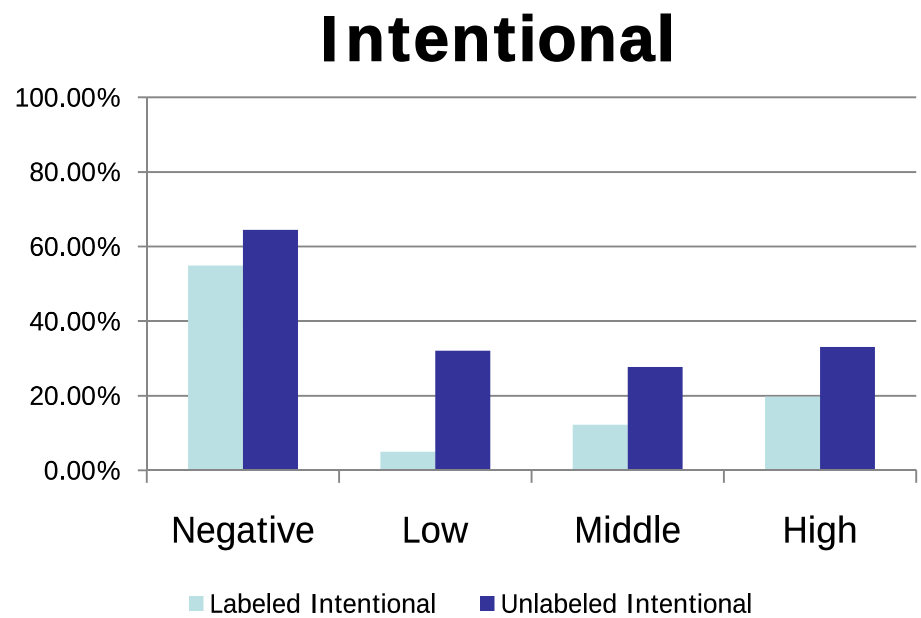


Model Predictions: Comparison of Confirmed Errors



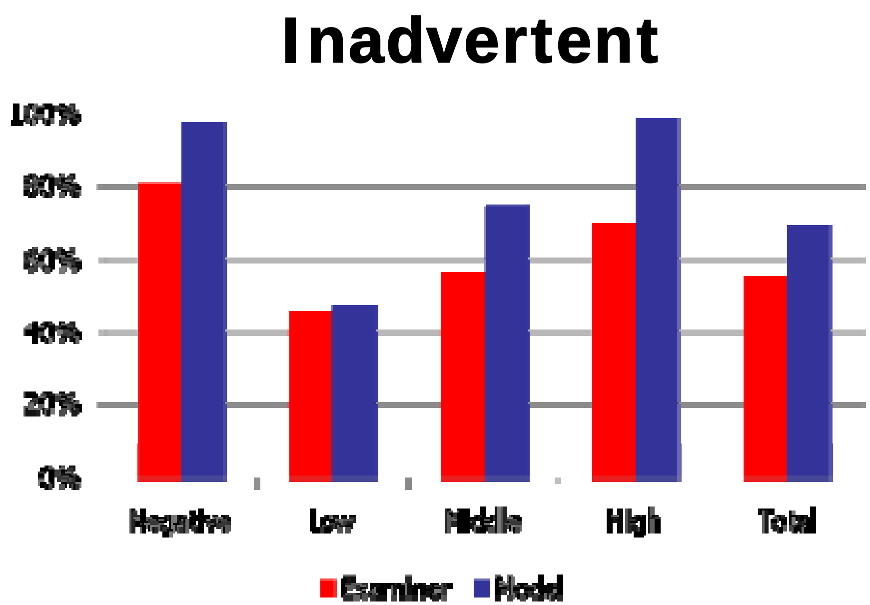
Model Predictions: Comparison of Percentage of Tax Returns Marked as Inadvertent or Intentional for Returns Labeled and Unlabeled by Examiner

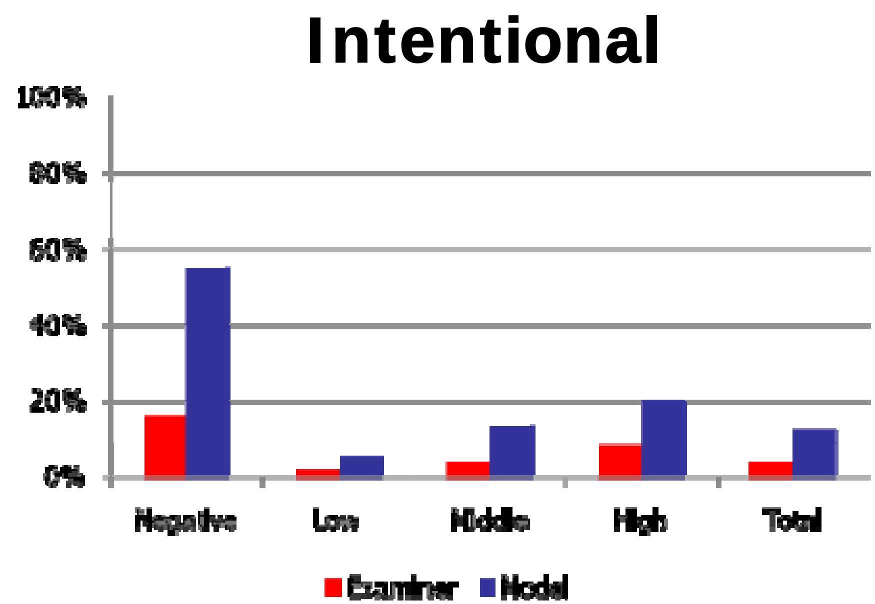




Models suggest that many of the unlabeled exams could have been labeled, particularly in the Low income category

Comparison of Tax Returns Identified as Containing an Inadvertent or Intentional Error by Examiner and Model





- Models suggest that almost all of the Negative and High income tax returns contain an inadvertent error.
- Models identify more of the exams as having intentional errors.

Model Identified Profiles for Tax Returns with Confirmed Errors as Compared with all other Returns

Inadvertent

	Age	Use Paid Preparer	Burden	Income	EITC	Itemized	Late	Filing Status	Exemp-tions	Error Amount
Negative	Slightly Older	Yes	High	NA	Mixed	Mixed	Mixed	Joint	--	--
Low	Mixed	No	High	Mixed	Yes	Yes	Mixed	Mixed	--	--
Middle	Older	No	High	Higher	Yes	Yes	Extension	Married-J	--	--
High	Slightly Older	Yes	High	NA	NA	Mixed	Mixed	Mixed	--	--

Intentional

	Age	Use Paid Preparer	Burden	Income	EITC	Itemized	Late	Filing Status	Exemp-tions	Error Amount
Negative	--	--	--	--	--	Yes	Extension	Married-J	Mixed	High
Low	--	--	--	--	--	No	Extension & No File	Single & Married-J	<2	Very High and Low
Middle	--	--	--	--	--	Yes	Extension	Single & Married-J	Mixed	Very High and Low
High	--	--	--	--	--	Yes	Mixed	Married-J	Mixed	Very High and Low

Study Limitations

- Intentional and Inadvertent errors defined by examiners.
  - Bias: human error
  - Mitigation: error jittering, no appreciable change in results
- Expectation of “intentionality” impacted type of exam; e.g., field or campus.
  - Bias: unknown factors
  - Mitigation: all exams were considered collectively with controls for types of exams considered
- Data only included tax returns thought to be in error.
  - Bias: selection on the dependent variable – error
  - Mitigation: future work should take the proposed models and test against a random sample of all tax returns

Conclusions

- Behavioral patterns are different by income levels
- Many unlabeled exams have the potential to be labeled
- Returns that are likely to have inadvertent errors are different from those likely to have intentional errors
  - 1st principle models: belief in obeying laws decreases intentional errors; whereas, complexity suggests inadvertent errors
  - Machine learning models: filing late, exemptions, and larger errors suggest intentional; whereas, age, paid preparers usage, EITC, and complexity suggest inadvertent
- Ensemble techniques should improve model accuracy
  - Gains should be larger for intentional errors
- Pattern of errors on line items may be diagnostic

Next Steps

- Test: Use profiles to extract new tax returns and check for errors, also test against random sample of tax returns
- Refine: models with findings
- Extend:
  - Develop models for core line items
  - Develop ensemble model using line item assessment and overall return assessment
- Forecast: Encapsulate combined model into Construct simulator for enabling forecasts by city