University of Antwerp
Antwerp Tax Academy

Survey experiment: perceived fairness of AI-use by tax administrations

Dr. Anouk Decuypere

Prof. dr. Anne Van de Vijver

October 2023

- Professor of Tax Law @ University of Antwerp, Belgium
- Chair of Antwerp Tax Academy, interfaculty institute of tax science, 7 faculties, 82 researchers
- Law, psychology & ethics

University of Antwerp
Antwerp Tax Academy

Overview

- Background

- Main variables: AI use -> fairness perception

- Other variables

- Transparency explanations (study 2)

- Study 1: results

- Study 2: results

- Conclusions

Background: Governmental level

- Increased AI-use by tax administrations (OECD, 2020, 2021)

- **GDPR** requirement for a human-in-the-loop, automation bias

- *How do we achieve AI-human synergy?*

Tax administrations using AI
Tax administrations developing AI models
No publicly documented use of AI
Not part of the EU

AItaxadmin

https://www.uantwerpen.be/en/research-groups/digitax/aitaxadmin/

University of Antwerp
Antwerp Tax Academy
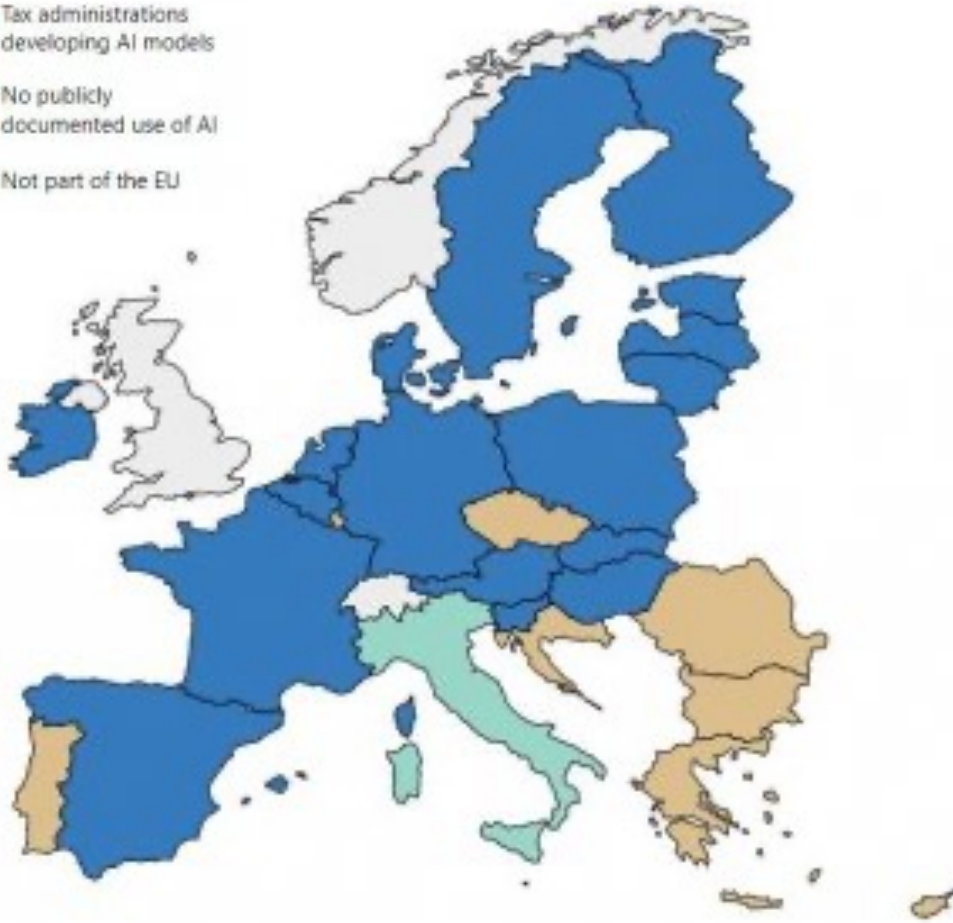
1. AI web-scraping: the SPF Finances makes use of an algorithm to automatically collect taxpayer data from e-commerce and e-sharing platforms, e.g. Amazon, Airbnb, eBay, 2emeMain, etc.

2. Social Network Analysis (SNA): the SNA algorithm visually represents a network of individual taxpayers using graph theory. It represents a network of taxpayers as a combination of nodes for individuals or points of interests, and lines which quantitatively and/or qualitatively measure relations between the nodes.

3. Internal risk-management: The SPF Finances uses a suite of algorithms to predict the risk that taxpayers do not pay their taxes due, following a letter from the bailiff ('Pegasus'), or following a call from the outbound call center ('Iris'). These algorithms assist the SPF with their internal case management, and predict what course of action is most appropriate for the administration, based on historical taxpayer data, e.g. for taxpayers who are notoriously compliant/non-compliant these models prescribe a more coercive/cooperative course of action, and vice versa.

4. External risk-management (risk-scoring algorithms): The SPF Finances uses a suite of algorithms to predict specific risks of non-compliance of individual taxpayers, 'Hermes' predicts the risk of bankruptcy within a 12 months period for legal and self-employed persons, 'Delphi' predicts the solvency rate for natural, legal and self-employed persons. The SPF Finances also uses models to segment taxpayers into categories of risks to develop their annual audit plans, and select taxpayers with high-risks of non-compliance for further audits by human tax officials.

5. Nudging: the Belgian Tax Administration uses an algorithm to adapt the language of standard communication of taxpayers based on an analysis of individual taxpayer data.

AItaxadmin

https://www.uantwerpen.be/en/research-groups/digitax/aitaxadmin/

University of Antwerp
Antwerp Tax Academy

Background: Citizens' perceptions

- AI **aversion** (Burton et al., 2020)
  - Lower interpersonal fairness
  - Procedural fairness (Langer & Landers, 2021)
- AI **appreciation**
  - E.g., in lay people (Logg et al., 2019)

What are citizen fairness perceptions of current AI practices?

AI-use: vignette

Baseline information	Condition example study 1	Condition example study 2
One of the responsibilities of the Federal Tax Administration is to check the tax returns of citizens. These checks are performed annually by tax auditors. They base this on their accumulated expertise. In addition, tax auditors are supported by artificial intelligence (AI). The AI system relies on information from past fraudsters to make recommendations for audits of new returns.	[PROPORTION] Imagine the following situation: 100 people are chosen to check for fraud. The tax auditor may select [20] people, the AI system may select [80] people.	[TRANSPARANCY] The tax administration is not transparent about the specific data the AI takes into account [EXPLANATION] because they do not want tax payers to be able to evade fraud detection based on this information. [PROPORTION] Imagine the following situation: 100 people are chosen to check for fraud. The tax auditor may select [20] people, the AI system may select [80] people.

AI-use: vignette

- First study: random allocation to three different scenarios
  - *80/20 (group 1)*
  - *20/80 (group 2)*
  - *50/50 (group 3)*

- One vignette per person (carry-over effects)
- Positive formulation (negativity bias)
- Absolute numbers (perception bias; Kahneman, 2011)

Fairness

- Fairness
    - General fairness (Colquitt & Shaw, 2005 – adapted to tax context): e.g. "*The way this procedure works in tax administration is not fair.*"
    - Procedural fairness (Camps et al., 2023 – adapted to tax context): ethicality, accuracy, bias suppression, consistency: e.g. *"With this procedure, the selection of citizens is in line with general ethical guidelines"*
    - Seven-point likert scale (completely disagree – completely agree)

- Ideal proportion AI versus human decision making

Wat is volgens u de ideale verhouding tussen AI en belastingambtenaar?

0   10   20   30   40   50   60   70   80   90   100

% Selectie door AI

Fairness

- Studies on fairness perception of AI show mixed results
  - AI is perceived fairer in high-impact decisions such as health and justice (Starke et al., 2022)
  - Human decision-making for tasks requiring more human (contextualization) skills is perceived fairer (Nagtegaal, 2021)
  - Higher acceptance for general government services (Gesk & Leyer, 2022)
  - AI is seen as accurate (Araujo et al., 2020)

General hypothesis 1a: Higher AI-use by the tax administration is perceived as more fair in general.

General hypothesis 1b: Higher AI-use by the tax administration is perceived as more procedurally fair.

Other variables: personality trait

- **Fairness propensity**: a tendency to view events, people and organizations as fair – measured with six-item scale (Colquitt et al., 2018)
    - Communal life, e.g., a strong sense of belonging
    - Basic personality traits

Hypothesis 2a: Fairness propensity moderates the impact of AI-use on procedural fairness, in such a way that higher fairness propensity yields higher procedural fairness scores as AI-use goes up

Other variables: trust

- **Trust**: the willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party (Mayer et al., 1995) – measured with different scales
  - Perceived competence
  - Benevolence
  - Integrity

Hypothesis 2b: Trust in the tax authorities moderates the impact of AI-use on procedural fairness, in such a way that higher trust leads to higher procedural fairness perceptions as AI-use goes up.

Other variables: demographics

- **Demographics**
  - Age and gender: mixed results
  - Higher computer programming knowledge diminished fairness perceptions (Lee & Baykal, 2017; Starke et al., 2022

Hypothesis 3: Professionals perceive AI procedures as less preferable in comparison to students.

Transparency (study 2)

Baseline information	Condition example study 1	Condition example study 2
One of the responsibilities of the Federal Tax Administration is to check the tax returns of citizens. These checks are performed annually by tax auditors. They base this on their accumulated expertise. In addition, tax auditors are supported by artificial intelligence (AI). The AI system relies on information from past fraudsters to make recommendations for audits of new returns.	[PROPORTION] Imagine the following situation: 100 people are chosen to check for fraud. The tax auditor may select [20] people, the AI system may select [80] people	**[TRANSPARANCY]** **The tax administration is not transparent about the specific data the AI takes into account** **[EXPLANATION]** **because they do not want tax payers to be able to evade fraud detection based on this information.** [PROPORTION] Imagine the following situation: 100 people are chosen to check for fraud. The tax auditor may select [20] people, the AI system may select [80] people.

Transparency (study 2)

- Legal transparency requirements (e.g. GDPR); Black box problem
- Informational fairness (Colquitt, 2001)
- Absence of transparency impacts governmental AI use (Schiff et al., 2022); Inconclusive results (Grimmelikhuijsen, 2023; Starke et al., 2022)

Hypothesis 4a: Transparency (yes/no) on the data used for AI is perceived as more (procedurally) fair.

Hypothesis 4b: Explanations for the reasoning (on the presence or absence of transparency) is perceived as more (procedurally) fair.

Hypothesis 4c: The combination of transparency and an explanation is perceived as more (procedurally) fair than only providing information on whether there is transparency.

2 studies

- Study 1: 3 experimental conditions (proportions)
- Study 2: same as study 1 + (3x2: proportions with transparency) + (3x2: proportions with explanation)

Vignette scenario factors	Vignette levels	Study 1	Study 2
Proportion of decision making by ADM (versus tax civil servants, i.e., "high human condition")	High human (20 vs. 80) Average (50 vs. 50) High ADM (80 vs. 20)	X	X
Transparency	Extra info: No transparency Transparency		X
Explanation	Extra info: No transparency + Explanation Transparency + Explanation		X

University of Antwerp
Antwerp Tax Academy

Study 1: results

- **Hypothesis 1b: supported**: statistically significant main effect of the vignette scenario on procedural fairness ($F_{(4,272)} = 2.24$, $p < .03$)

  - Bias suppression: high AI-use vignette ($M = 4.88$, $SD = 1.26$) versus high human vignette ($M = 3.87$, $SD = 1.47$)

- **Hypothesis 3: opposite effect**: ideal proportions of AI-use, average scores of professionals were significantly higher and in favor of more AI-use compared to students (72.13% vs. 54.54%, respectively)

- Not supported: hypotheses 1a (general fairness), 2a (fairness propensity) & 2b (trust)
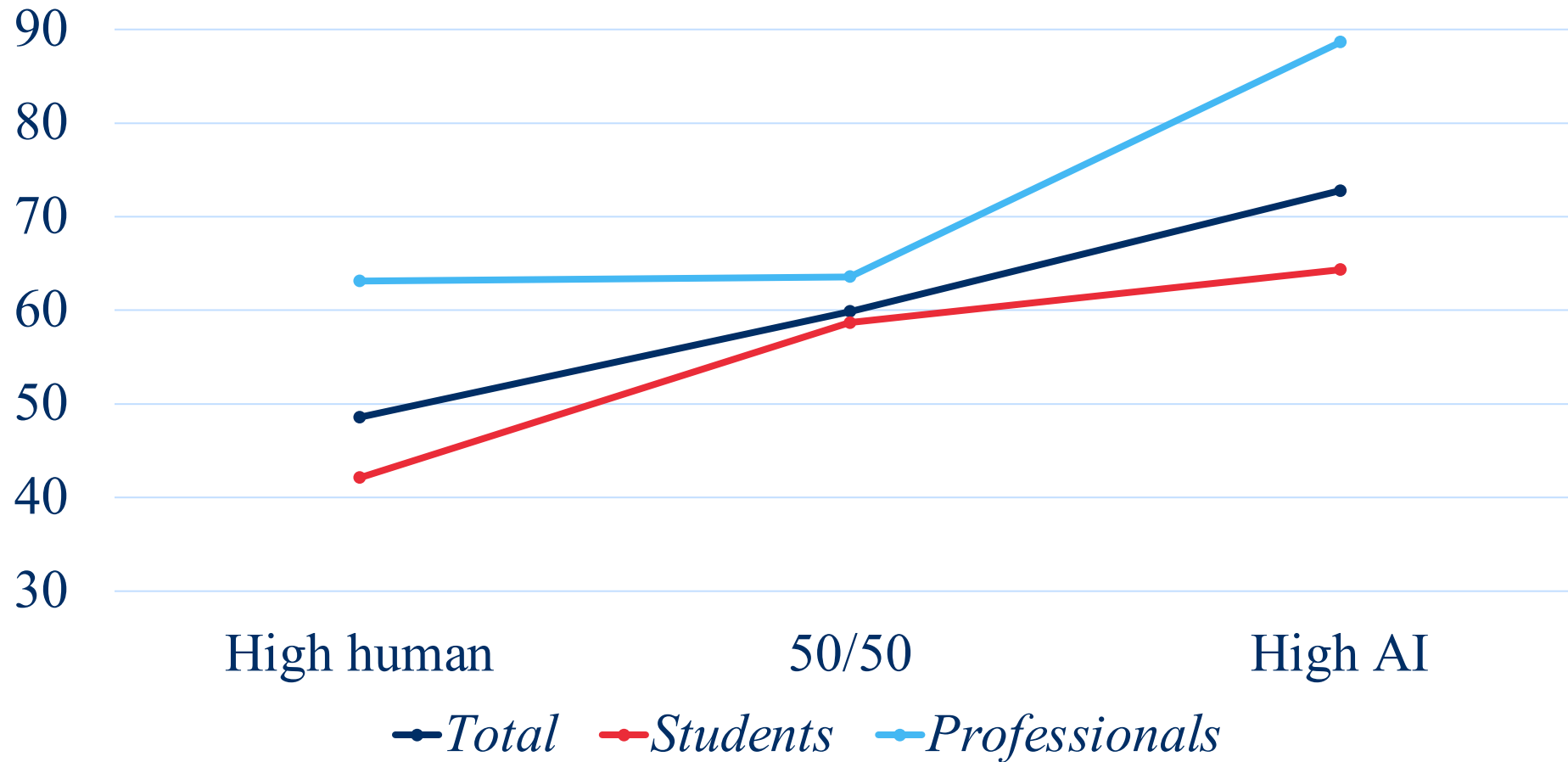
Study 1: results

- Bias suppression
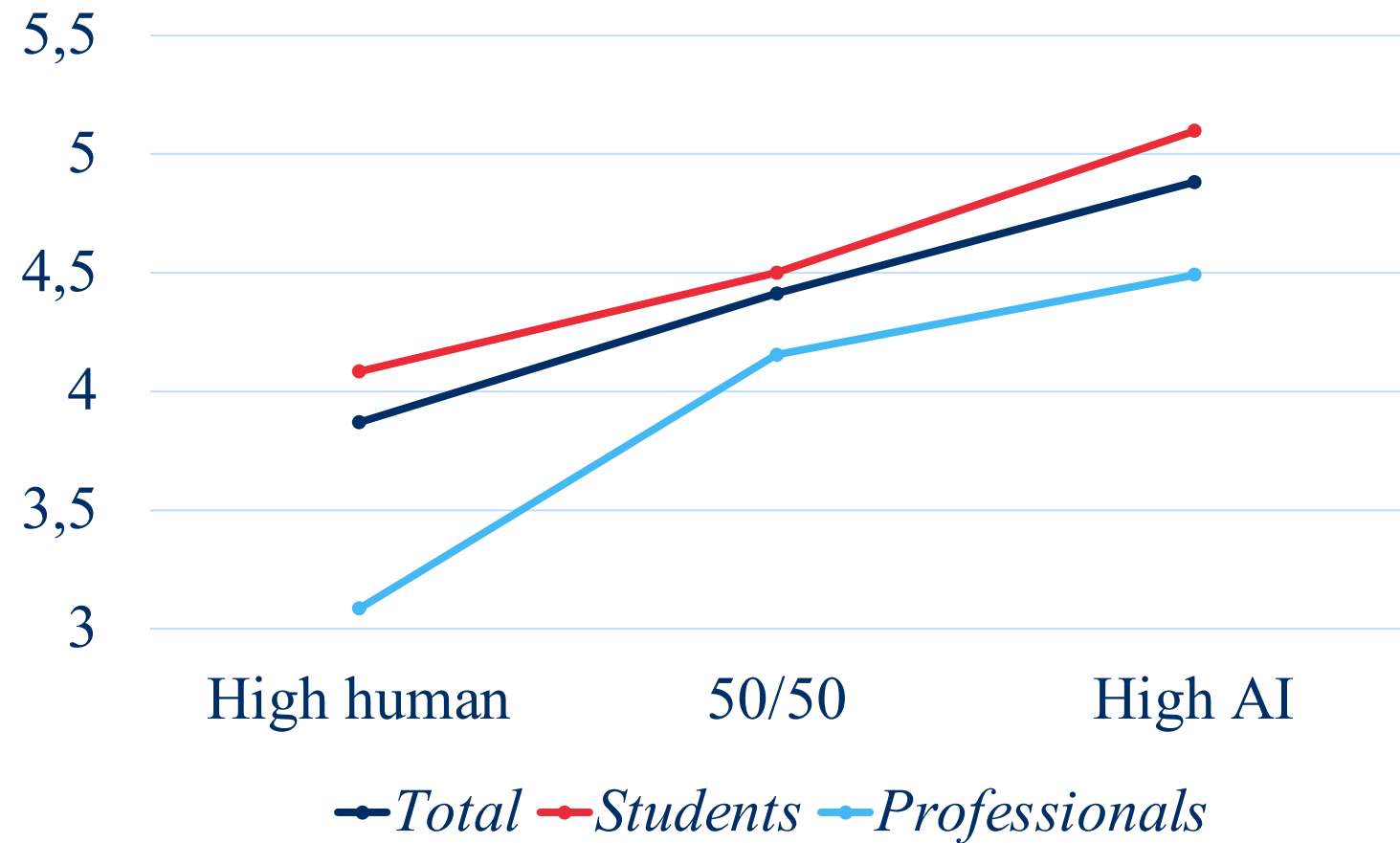
Study 1: results

Ideal % AI

Study 2: results

- Sample of 2366 respondents – average of 158 per vignette

- **Hypothesis 1b: supported**: statistically significant main effect of the vignette scenario on procedural fairness ($F(8,4772) = 12.18$, $p < .001$)

  - Bias suppression: low AI-use (M = 4.42 , SD = 1.23), 50/50 (M = 4.63, SD = 1.21), and high AI-use (M = 4.82, SD = 1.22)

  - Consistency: low AI-use scenario proportion (M = 4.36, SD = 1.22), 50/50 scenario (M = 4.47, SD = 1.21), and high AI-use (M = 4.64, SD = 1.23)

- **Hypothesis 2b: supported**: trust dampens down the increase in bias suppression scores as the use of AI-use increased in the vignettes
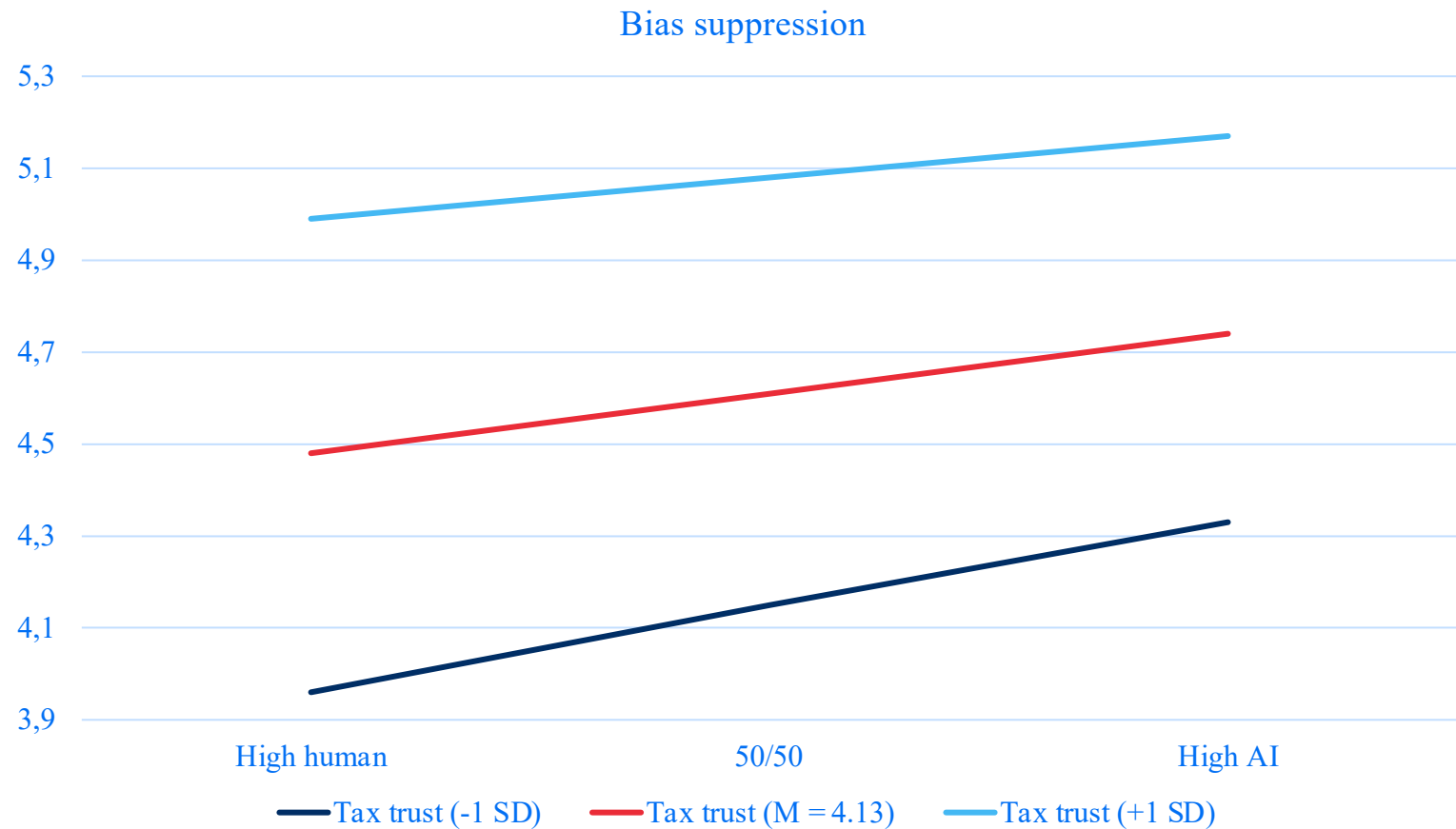
University of Antwerp
Antwerp Tax Academy

Study 2: results

- Bias suppression & consistency

Study 2: Replication

Study 2: results

- Interaction with trust

Bias suppression

Tax trust (-1 SD) — Tax trust (M = 4.13) — Tax trust (+1 SD)

Study 2: results

- Ideal proportion of AI-based decision making: 50/50 split is preferred – but a lot of disagreement in sample

- Not supported: hypotheses 1a (general fairness), 2a (fairness propensity) & 4 (transparency)

University of Antwerp
I Antwerp Tax Academy

Conclusions

- Positive perception towards AI, mostly carried by bias suppression

- Trust in the tax administration

- Tax professionals, automation bias

Questions?