

Decision-Making with Machine Prediction: Evidence from Predictive Maintenance in Trucking

Adam Harris (Cornell University) and Maggie Yellen

**FutureTech Workshop on the Economics of AI and Automation
December 6, 2024**

Prediction, decision-making, and artificial intelligence

- Professional human decision-makers face high-stakes choices.

Prediction, decision-making, and artificial intelligence

- Professional human decision-makers face high-stakes choices.
 - Many involve two steps: **prediction** and **judgment**.

Prediction, decision-making, and artificial intelligence

- Professional human decision-makers face high-stakes choices.
 - Many involve two steps: **prediction** and **judgment**.
- Increasingly, **artificial intelligence plays a role in decision-making**.

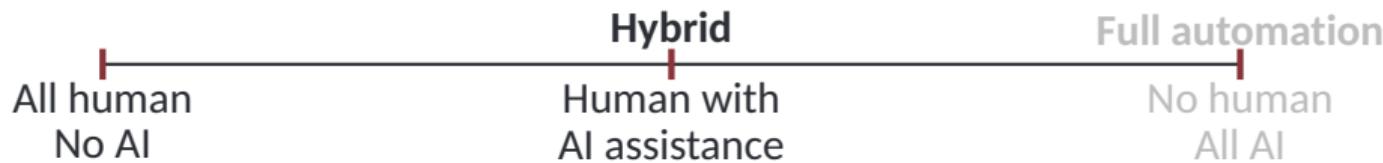
Prediction, decision-making, and artificial intelligence

- Professional human decision-makers face high-stakes choices.
 - Many involve two steps: **prediction** and **judgment**.
- Increasingly, **artificial intelligence plays a role in decision-making**.



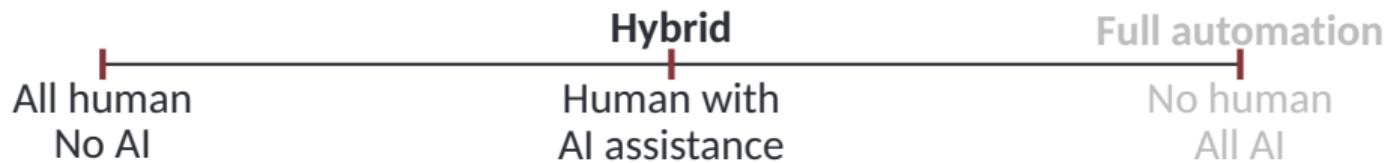
Prediction, decision-making, and artificial intelligence

- Professional human decision-makers face high-stakes choices.
 - Many involve two steps: **prediction** and **judgment**.
- Increasingly, **artificial intelligence plays a role in decision-making**.



Prediction, decision-making, and artificial intelligence

- Professional human decision-makers face high-stakes choices.
 - Many involve two steps: **prediction** and **judgment**.
- Increasingly, **artificial intelligence plays a role in decision-making**.



- Effect on decision **quality** is *a priori* unclear.

How do professional human decision-makers use predictive AI?

- **Quantitatively**, what is the effect of predictive AI on welfare?

How do professional human decision-makers use predictive AI?

- **Quantitatively**, what is the effect of predictive AI on welfare?
- **Setting**: Technicians making engine repair decisions for heavy-duty trucks.
 - Trade off repair cost and breakdown risk.
 - Relevant data is abundant.

How do professional human decision-makers use predictive AI?

- **Quantitatively**, what is the effect of predictive AI on welfare?
- **Setting**: Technicians making engine repair decisions for heavy-duty trucks.
 - Trade off repair cost and breakdown risk.
 - Relevant data is abundant.
 - We observe decisions from a fleet that introduced AI to predict breakdowns.

How do professional human decision-makers use predictive AI?

- **Quantitatively**, what is the effect of predictive AI on welfare?
- **Setting**: Technicians making engine repair decisions for heavy-duty trucks.
 - Trade off repair cost and breakdown risk.
 - Relevant data is abundant.
 - We observe decisions from a fleet that introduced AI to predict breakdowns.
- **Objective**: Use **observational data** on decisions to value AI assistance.

Using observational data to study AI & decision-making

Objective: Use **observational data** on decisions to value AI assistance.

Two challenges:

Using observational data to study AI & decision-making

Objective: Use **observational data** on decisions to value AI assistance.

Two challenges:

1. How to separately identify payoffs and beliefs?

Using observational data to study AI & decision-making

Objective: Use **observational data** on decisions to value AI assistance.

Two challenges:

1. How to separately identify payoffs and beliefs?
2. Dynamics + high-dimensional state?

Key findings

1. Predictive AI improves decision-making quality.

→ *Ceteris paribus*, AI reduces costs by **\$240-480 / truck / year** (85% of all feasible savings).

Key findings

1. Predictive AI improves decision-making quality.

→ *Ceteris paribus*, AI reduces costs by **\$240-480 / truck / year** (85% of all feasible savings).

2. How are these cost savings achieved?

— Without AI, technicians do costly, unnecessary repairs.

Key findings

1. Predictive AI improves decision-making quality.

→ *Ceteris paribus*, AI reduces costs by **\$240-480 / truck / year** (85% of all feasible savings).

2. How are these cost savings achieved?

- Without AI, technicians do costly, unnecessary repairs.
- Gains from AI come entirely from reduction in repair costs.

Related literature

1. AI + human decision-making:

- **Empirical:** Stevenson & Doleac (2021); Angelova et al. (2022); De-Arteaga et al. (2020); Albright (2023); Kleinberg et al (2018); Agarwal et al. (2023); Tschandl et al. (2020); Kim et al. (2020); Reverberi et al. (2022); and Jakubik et al. (2022).
- **Theoretical:** Agrawal, Gans, & Goldfarb (2019a-c, forthcoming); Athey, Gans, & Bryan (2020)

2. Estimation of dynamic discrete choice models.

- Rust (1987); Hotz & Miller (1993); Arcidiacono & Miller (2011).

3. Trucking.

- Harris & Nguyen (2021, 2023); Armitage et al. (2023); Yang (2022); Baker & Hubbard (2003, 2004); Hubbard (2000, 2001, 2003); Rose (1985, 1987).

Related literature

1. AI + human decision-making:

- **Empirical:** Stevenson & Doleac (2021); Angelova et al. (2022); De-Arteaga et al. (2020); Albright (2023); Kleinberg et al (2018); Agarwal et al. (2023); Tschandl et al. (2020); Kim et al. (2020); Reverberi et al. (2022); and Jakubik et al. (2022).
- **Theoretical:** Agrawal, Gans, & Goldfarb (2019a-c, forthcoming); Athey, Gans, & Bryan (2020)

2. Estimation of dynamic discrete choice models.

- Rust (1987); Hotz & Miller (1993); Arcidiacono & Miller (2011).

3. Trucking.

- Harris & Nguyen (2021, 2023); Armitage et al. (2023); Yang (2022); Baker & Hubbard (2003, 2004); Hubbard (2000, 2001, 2003); Rose (1985, 1987).

Related literature

1. AI + human decision-making:

- **Empirical:** Stevenson & Doleac (2021); Angelova et al. (2022); De-Arteaga et al. (2020); Albright (2023); Kleinberg et al (2018); Agarwal et al. (2023); Tschandl et al. (2020); Kim et al. (2020); Reverberi et al. (2022); and Jakubik et al. (2022).
 - **Theoretical:** Agrawal, Gans, & Goldfarb (2019a-c, forthcoming); Athey, Gans, & Bryan (2020)
- We study a new, **economically important** setting.

2. Estimation of dynamic discrete choice models.

- Rust (1987); Hotz & Miller (1993); Arcidiacono & Miller (2011).

3. Trucking.

- Harris & Nguyen (2021, 2023); Armitage et al. (2023); Yang (2022); Baker & Hubbard (2003, 2004); Hubbard (2000, 2001, 2003); Rose (1985, 1987).

Related literature

1. AI + human decision-making:

- **Empirical:** Stevenson & Doleac (2021); Angelova et al. (2022); De-Arteaga et al. (2020); Albright (2023); Kleinberg et al (2018); Agarwal et al. (2023); Tschandl et al. (2020); Kim et al. (2020); Reverberi et al. (2022); and Jakubik et al. (2022).
- **Theoretical:** Agrawal, Gans, & Goldfarb (2019a-c, forthcoming); Athey, Gans, & Bryan (2020)
- We study a new, **economically important** setting.
- We **quantify the economic value** added by AI assistance.

2. Estimation of dynamic discrete choice models.

- Rust (1987); Hotz & Miller (1993); Arcidiacono & Miller (2011).

3. Trucking.

- Harris & Nguyen (2021, 2023); Armitage et al. (2023); Yang (2022); Baker & Hubbard (2003, 2004); Hubbard (2000, 2001, 2003); Rose (1985, 1987).

Related literature

1. AI + human decision-making:

- **Empirical:** Stevenson & Doleac (2021); Angelova et al. (2022); De-Arteaga et al. (2020); Albright (2023); Kleinberg et al (2018); Agarwal et al. (2023); Tschandl et al. (2020); Kim et al. (2020); Reverberi et al. (2022); and Jakubik et al. (2022).
- **Theoretical:** Agrawal, Gans, & Goldfarb (2019a-c, forthcoming); Athey, Gans, & Bryan (2020)
- We study a new, **economically important** setting.
- We **quantify the economic value** added by AI assistance.

2. Estimation of dynamic discrete choice models.

- **Rust (1987); Hotz & Miller (1993); Arcidiacono & Miller (2011).**
- Estimating when **state is high-dimensional**.

3. Trucking.

- Harris & Nguyen (2021, 2023); Armitage et al. (2023); Yang (2022); Baker & Hubbard (2003, 2004); Hubbard (2000, 2001, 2003); Rose (1985, 1987).

Overview

1. Setting & data.
2. Descriptive evidence: Five facts.
3. Model and estimation.
4. Results
 - Unpacking estimates.
 - Evaluating counterfactual.
5. Discussion
 - Reflecting on AI's role
 - Looking ahead

Overview

1. Setting & data.
2. Descriptive evidence: Five facts.
3. Model and estimation.
4. Results
 - Unpacking estimates.
 - Evaluating counterfactual.
5. Discussion
 - Reflecting on AI's role
 - Looking ahead

The technician's decision problem

Question: Do you do an engine repair or send the truck out for its scheduled deliveries?

The technician's decision problem

Question: Do you do an engine repair or send the truck out for its scheduled deliveries?



The technician's decision problem

Question: Do you do an engine repair or send the truck out for its scheduled deliveries?



What do you see?

Sensor measurements

What do you see?

Faults

Sensor measurements

Acceleration forward or braking

-1.4...0.9G force



Acceleration side to side

-0.44...0.49G force



⋮

Engine intake manifold 1 temperature

68...138.2F



Engine load

0...100%



⋮

Odometer

537476.5...537867.9mi



Oil pressure

0...47.6psi



⋮

Vehicle programmed maximum road speed limit enabled (1 = enabled)

0...1



What do you see?

Sensor measurements

Oil pressure

0...47.6psi



2/01/23 22:05:39.427	25.5 psi
2/01/23 22:05:40.423	29 psi
2/01/23 22:05:43.423	27.3 psi
2/01/23 22:05:44.423	38.9 psi
2/01/23 22:05:46.423	38.3 psi
2/01/23 22:05:48.423	44.1 psi
2/01/23 22:05:49.423	40.6 psi
2/01/23 22:05:51.423	43.5 psi
2/01/23 22:05:52.423	39.5 psi
2/01/23 22:05:58.417	38.9 psi
2/01/23 22:06:00.417	34.8 psi
2/01/23 22:06:08.413	37.1 psi
2/01/23 22:06:14.423	34.2 psi
2/01/23 22:06:16.417	25.5 psi
2/01/23 22:06:19.423	26.1 psi
2/01/23 22:06:20.423	34.8 psi
2/01/23 22:06:23.417	36 psi

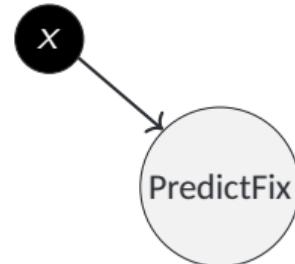
The AI tool

The AI tool

- In March 2020, fleet introduces **PredictFix** (AI tool), which generates *alerts*. [Details](#)

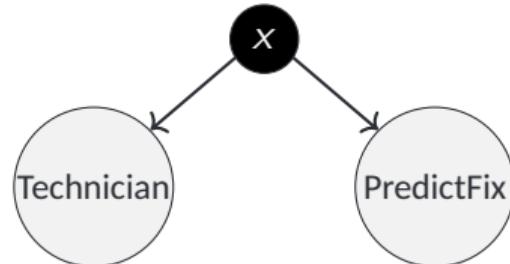
The AI tool

- In March 2020, fleet introduces **PredictFix** (AI tool), which generates *alerts*. [Details](#)
- Function of sensor data.



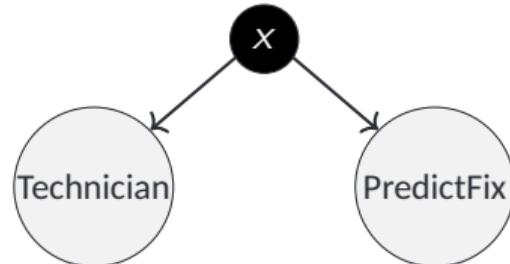
The AI tool

- In March 2020, fleet introduces **PredictFix** (AI tool), which generates *alerts*. [Details](#)
- Function of sensor data.



The AI tool

- In March 2020, fleet introduces **PredictFix** (AI tool), which generates *alerts*. [Details](#)
- **Function of sensor data** → no information that could not have been learned without PredictFix.



Why this setting?

Data

Why this setting?

Data

We observe:

- **Actions, outcomes:** Repairs, breakdowns at the work-order level.
- **AI output:** Alerts. [More](#)
- **State:** Everything that technicians see.
[Truck-generated data](#) [What techs see](#)

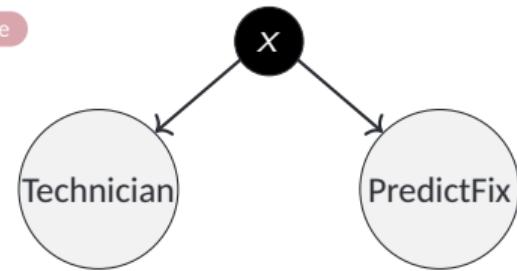
Why this setting?

Data

We observe:

- **Actions, outcomes:** Repairs, breakdowns at the work-order level.
- **AI output:** Alerts. [More](#)
- **State:** Everything that technicians see.

Truck-generated data What techs see

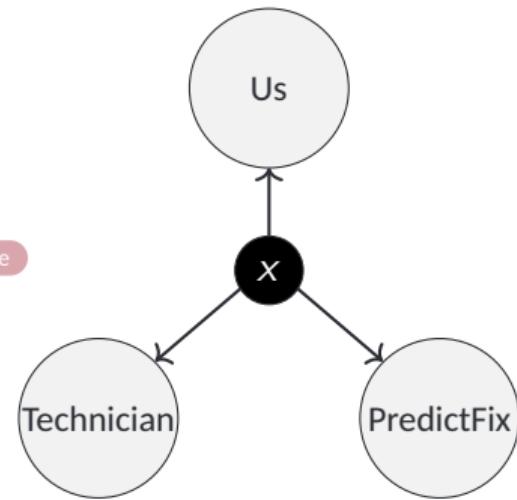


Why this setting?

Data

We observe:

- **Actions, outcomes:** Repairs, breakdowns at the work-order level.
- **AI output:** Alerts. [More](#)
- **State:** Everything that technicians see.
Truck-generated data What techs see

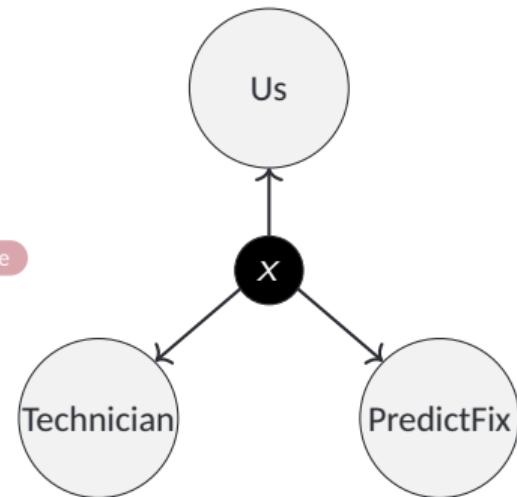


Why this setting?

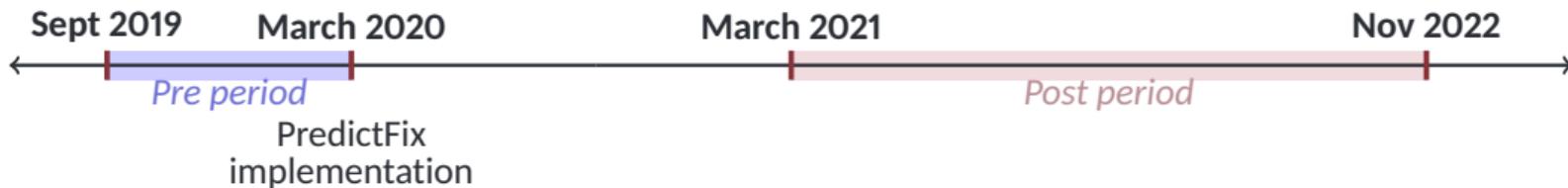
Data

We observe:

- **Actions, outcomes:** Repairs, breakdowns at the work-order level.
- **AI output:** Alerts. [More](#)
- **State:** Everything that technicians see.
Truck-generated data What techs see



Two disjoint periods:



Overview

1. Setting & data.
2. Descriptive evidence: Five facts.
3. Model and estimation.
4. Results
 - Unpacking estimates.
 - Evaluating counterfactual.
5. Discussion
 - Reflecting on AI's role
 - Looking ahead

Descriptive evidence: Five facts

Overview

Questions:

1.

2.

3.

4.

5.

Descriptive evidence: Five facts

Overview

Questions:

1. Are breakdowns predictable?
2. Is PredictFix a good predictor of breakdowns?
3. Do alerts change technician behavior?
- 4.
- 5.

Descriptive evidence: Five facts

Overview

Questions:

1. Are breakdowns predictable?
2. Is PredictFix a good predictor of breakdowns?
3. Do alerts change technician behavior?
- 4.
- 5.

}

Does PredictFix have the *potential* to improve decision-making quality?

Descriptive evidence: Five facts

Overview

Questions:

1. Are breakdowns predictable? **Yes.**

2. Is PredictFix a good predictor of breakdowns? **Yes.**

3. Do alerts change technician behavior? **Yes.**

4.

5.

}

Does PredictFix have the *potential* to improve decision-making quality? **Yes.**

Fact 1: Breakdown risk is predictable.

[More](#)[AUC Interpretation](#)

Fact 1: Breakdown risk is predictable.

[More](#)[AUC Interpretation](#)

Constructive argument:

1. Train ML model to predict breakdowns using sensor measurements.

2. Demonstrate its out-of-sample predictive quality using a **ROC curve**.
→ “Receiver Operating Characteristic.”

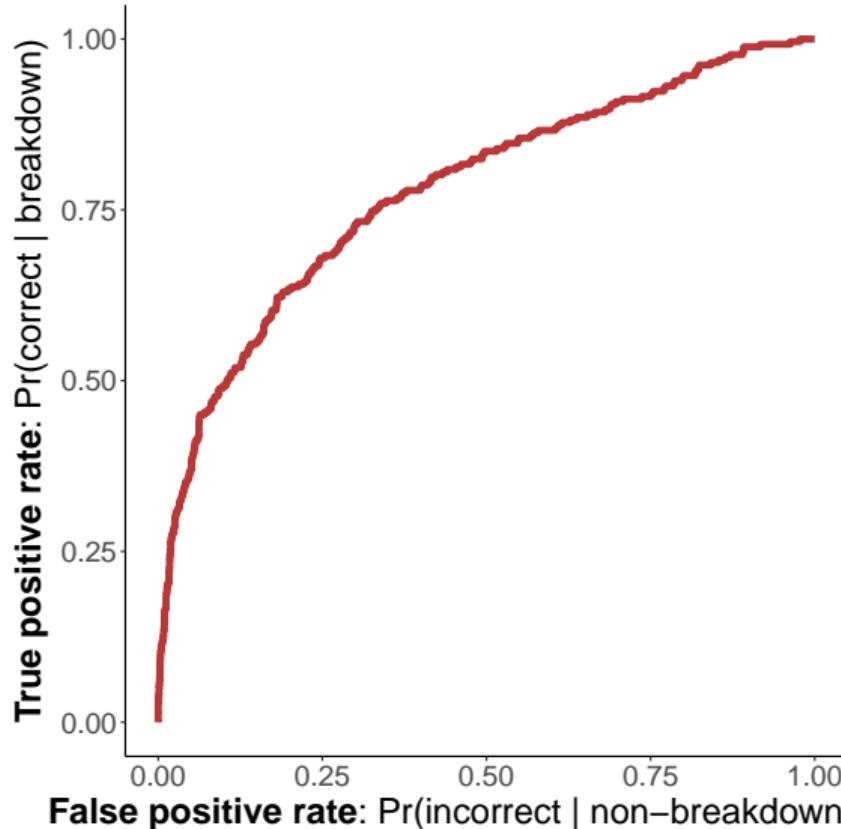
Fact 1: Breakdown risk is predictable.

[More](#)[AUC Interpretation](#)

Constructive argument:

1. Train ML model to predict breakdowns using sensor measurements.

2. Demonstrate its out-of-sample predictive quality using a **ROC curve**.
→ “Receiver Operating Characteristic.”



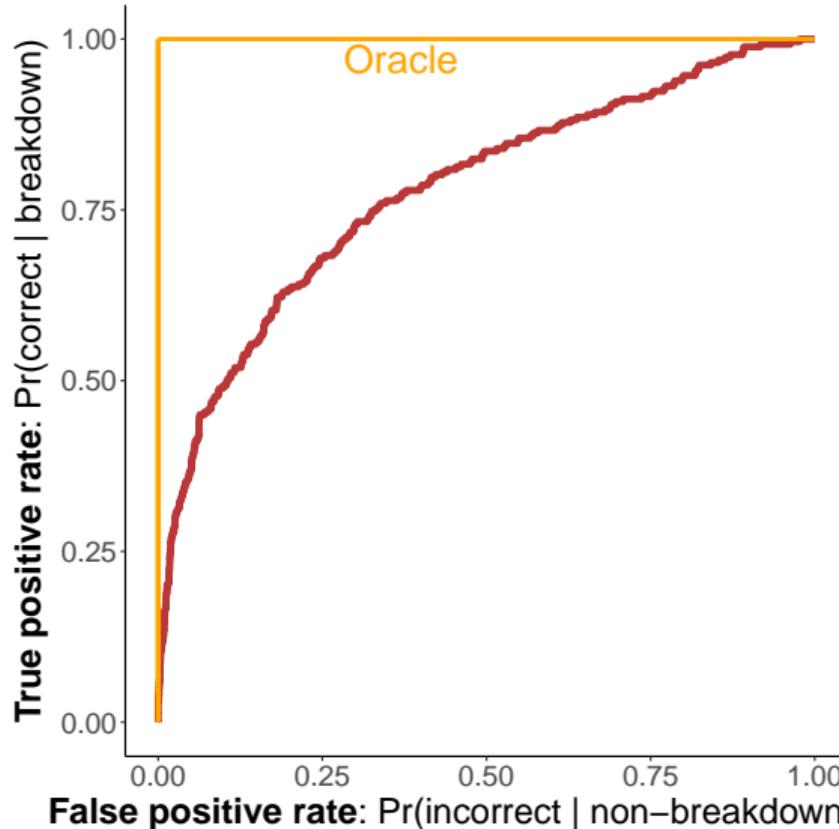
Fact 1: Breakdown risk is predictable.

[More](#)[AUC Interpretation](#)

Constructive argument:

1. Train ML model to predict breakdowns using sensor measurements.

2. Demonstrate its out-of-sample predictive quality using a **ROC curve**.
→ “Receiver Operating Characteristic.”



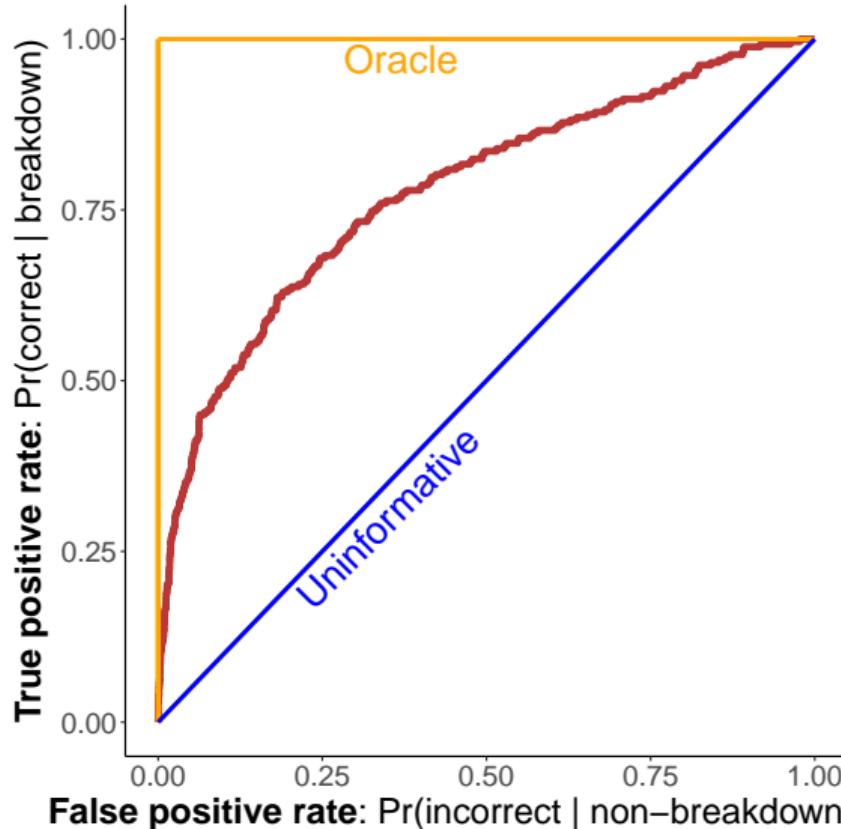
Fact 1: Breakdown risk is predictable.

[More](#)[AUC Interpretation](#)

Constructive argument:

1. Train ML model to predict breakdowns using sensor measurements.

2. Demonstrate its out-of-sample predictive quality using a **ROC curve**.
→ “Receiver Operating Characteristic.”

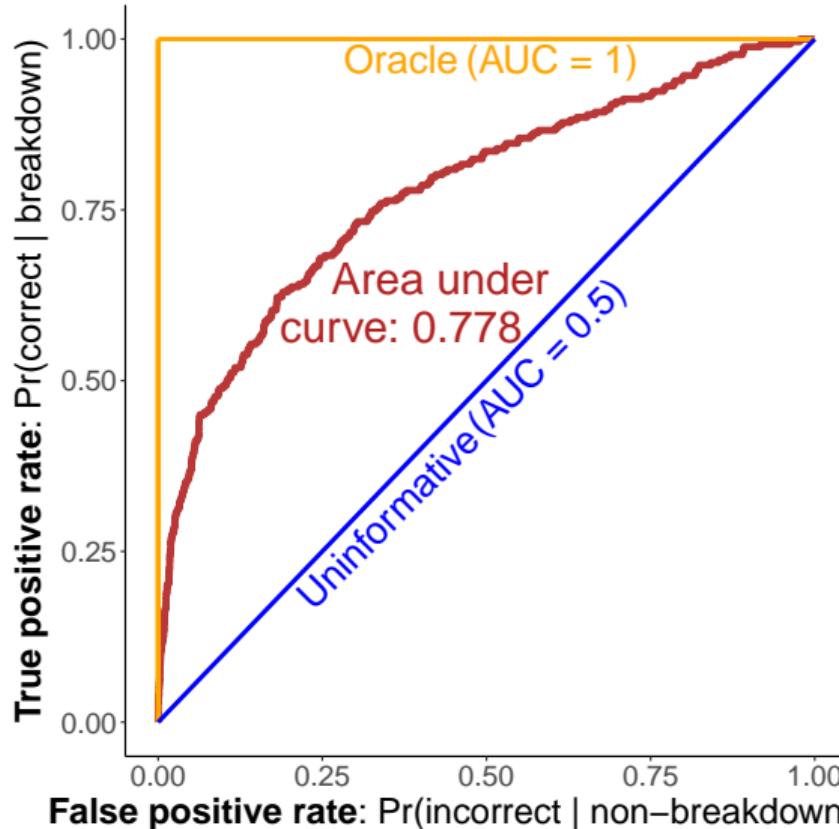


Fact 1: Breakdown risk is predictable.

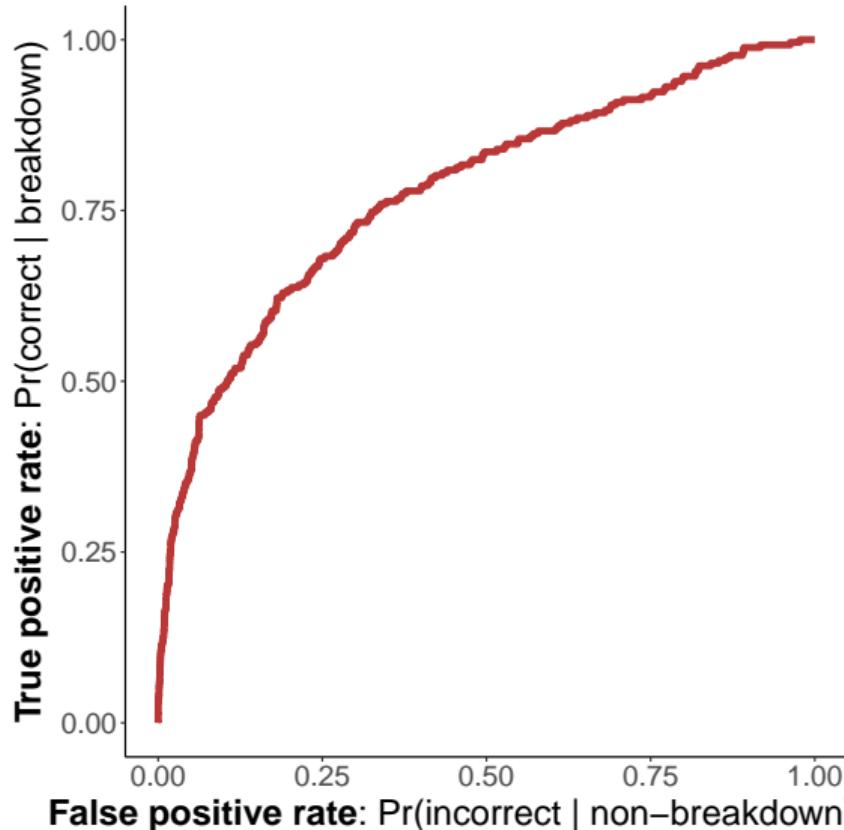
[More](#)[AUC Interpretation](#)

Constructive argument:

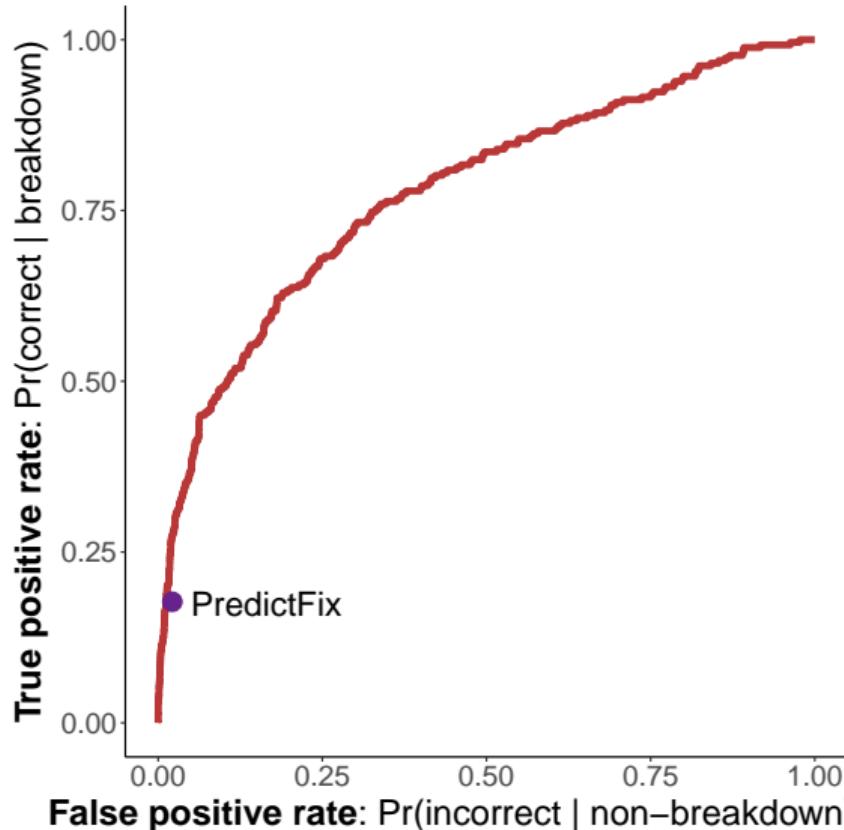
1. Train ML model to predict breakdowns using sensor measurements.
2. Demonstrate its out-of-sample predictive quality using a **ROC curve**.
→ “Receiver Operating Characteristic.”



Fact 2: PredictFix is a good predictor of breakdown risk.

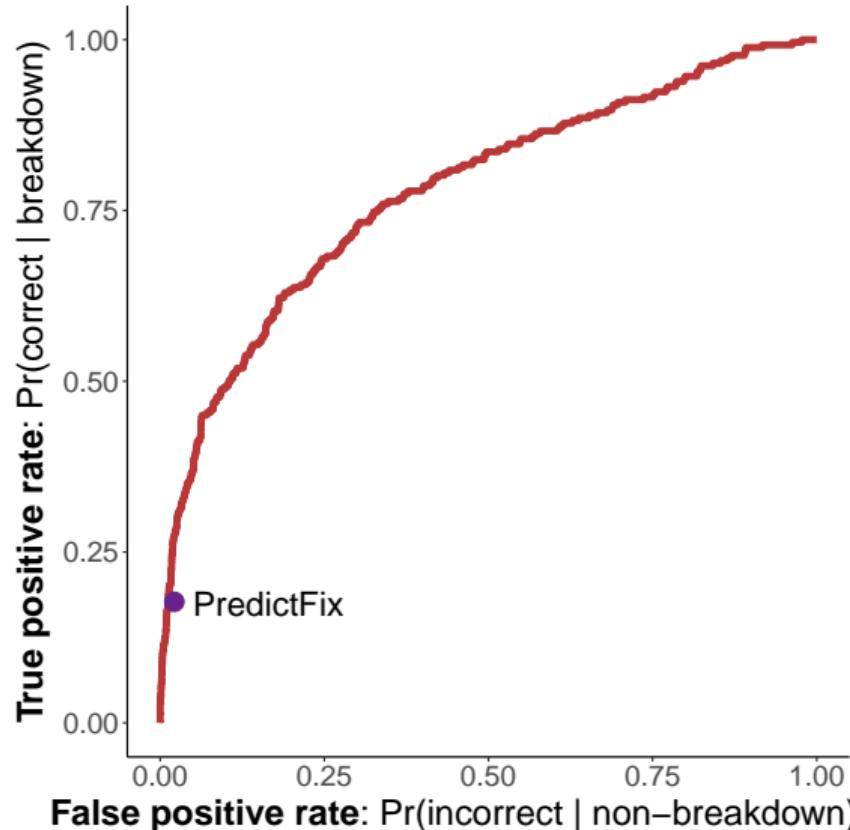


Fact 2: PredictFix is a good predictor of breakdown risk.



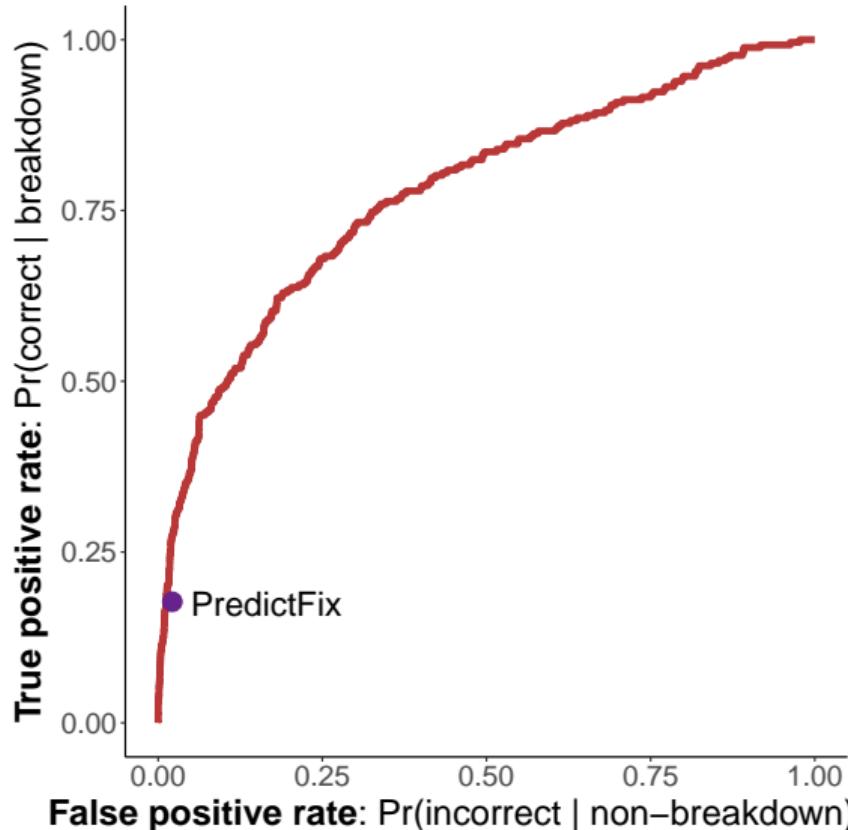
Fact 2: PredictFix is a good predictor of breakdown risk.

- Among best binary predictors that can be constructed.



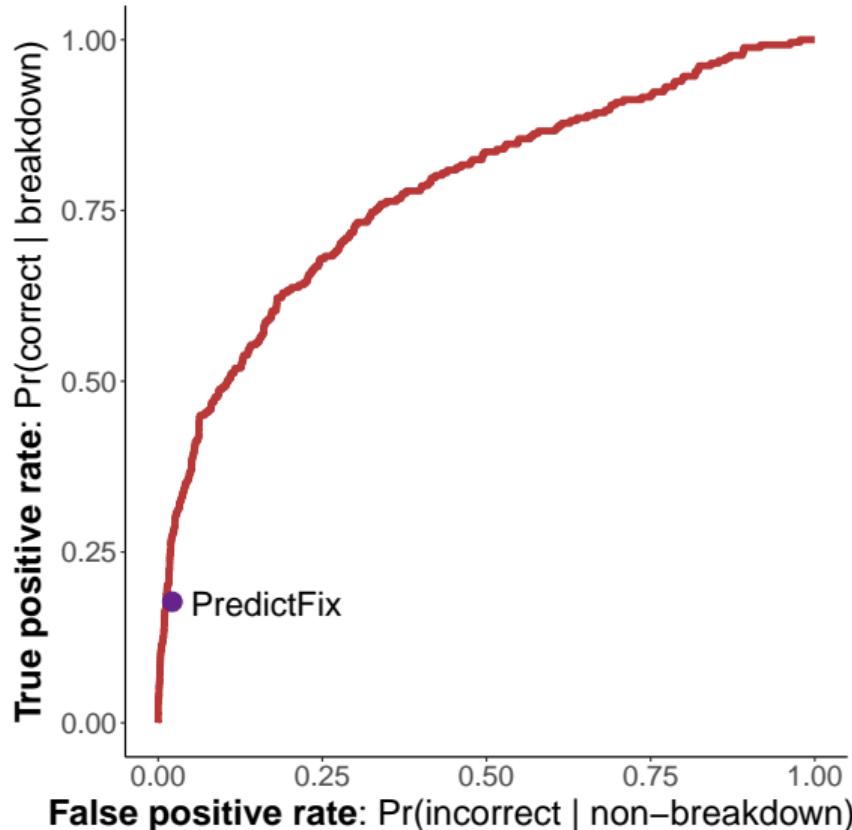
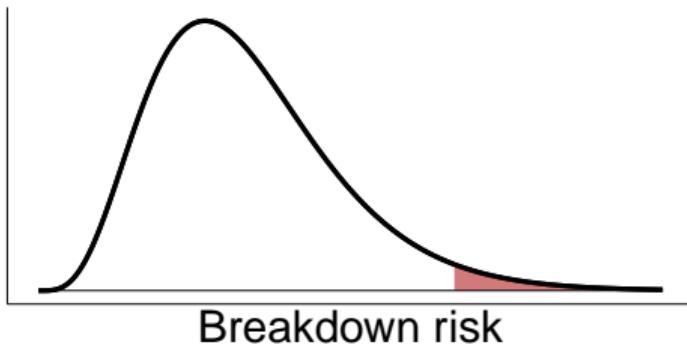
Fact 2: PredictFix is a good predictor of breakdown risk.

- Among best binary predictors that can be constructed.
- 2.5% of truck-weeks have an alert.



Fact 2: PredictFix is a good predictor of breakdown risk.

- Among best binary predictors that can be constructed.
- 2.5% of truck-weeks have an alert.



Fact 3: Technicians respond to PredictFix.

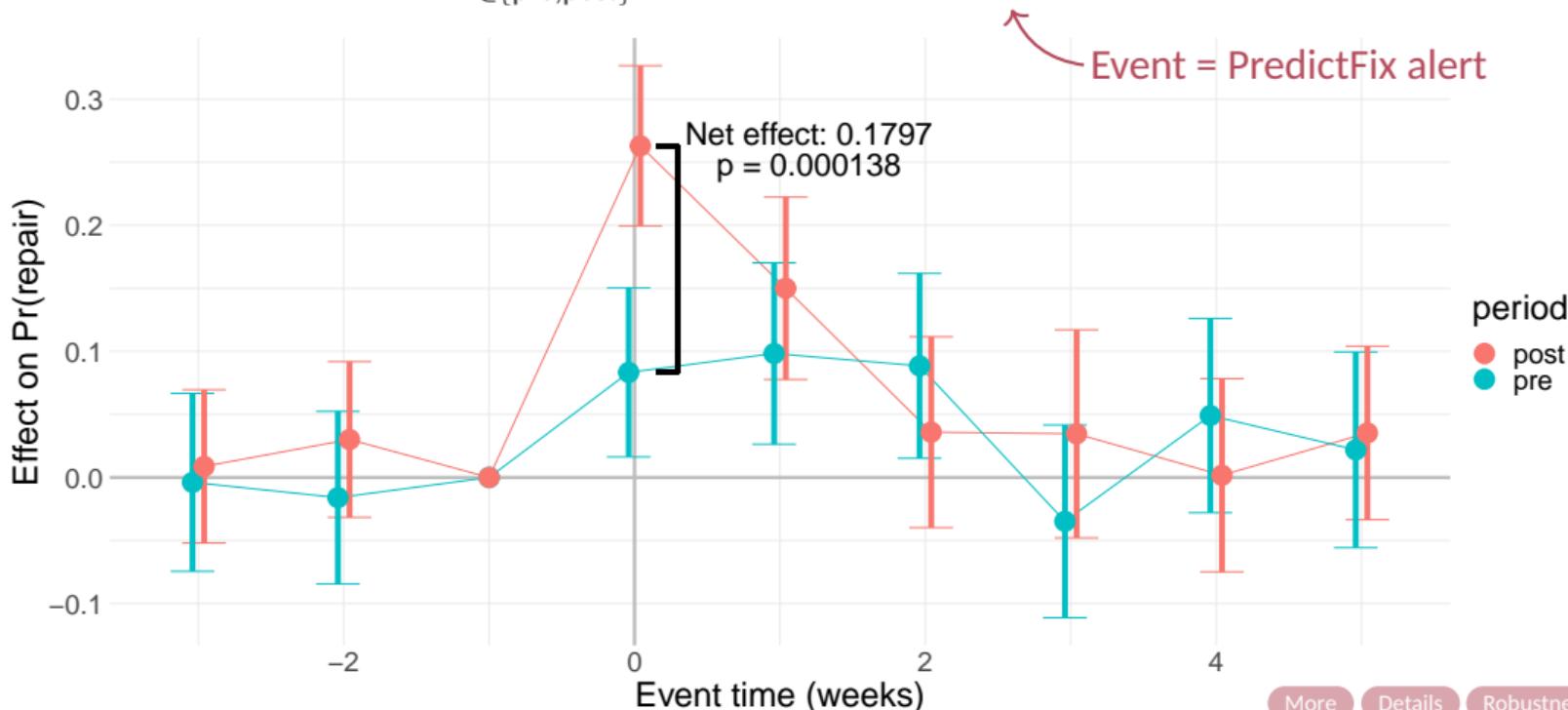
Fact 3: Technicians respond to PredictFix.

$$\text{Repair}_{i,t} = \alpha_0 + \sum_{k \in \{\text{pre,post}\}} \sum_{\tau=-3}^5 \beta_{\tau}^k \mathbb{1}\{t \in \mathcal{T}_k\} \widehat{\text{PredictFix}}_{i,t-\tau} + \alpha_i + \gamma_t + \epsilon_{i,t}$$

Event = PredictFix alert

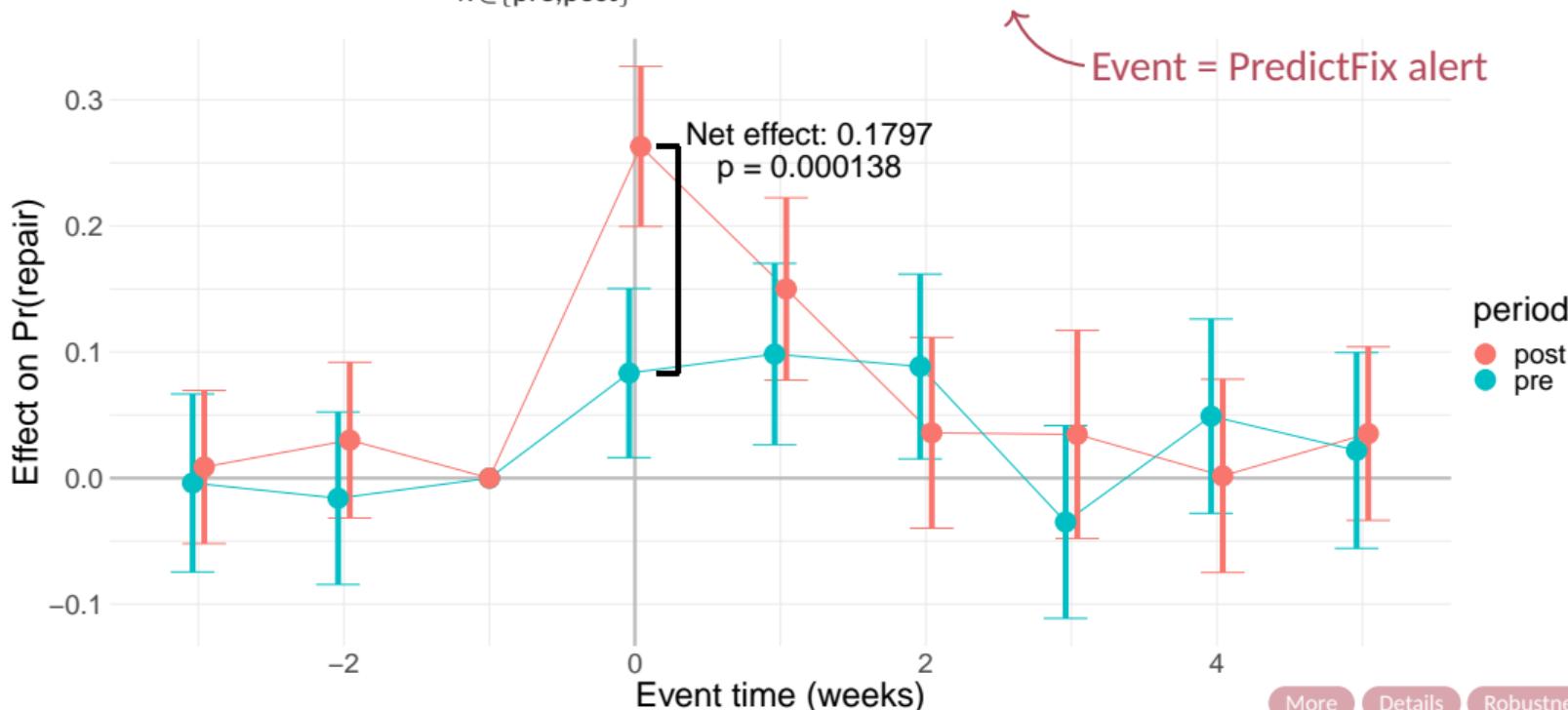
Fact 3: Technicians respond to PredictFix.

$$\text{Repair}_{i,t} = \alpha_0 + \sum_{k \in \{\text{pre,post}\}} \sum_{\tau=-3}^5 \beta_\tau^k \mathbb{1}\{t \in \mathcal{T}_k\} \widehat{\text{PredictFix}}_{i,t-\tau} + \alpha_i + \gamma_t + \epsilon_{i,t}$$



Fact 3: Technicians respond to PredictFix, but also ignore many alerts.

$$\text{Repair}_{i,t} = \alpha_0 + \sum_{k \in \{\text{pre,post}\}} \sum_{\tau=-3}^5 \beta_\tau^k \mathbb{1}\{t \in \mathcal{T}_k\} \widehat{\text{PredictFix}}_{i,t-\tau} + \alpha_i + \gamma_t + \epsilon_{i,t}$$



More

Details

Robustness

14/25

Descriptive evidence: Five facts

Overview

Questions:

1. Are breakdowns predictable? **Yes.**

2. Is PredictFix a good predictor of breakdowns? **Yes.**

3. Do alerts change technician behavior? **Yes.**

4.

5.

}

Does PredictFix have the *potential* to improve decision-making quality? **Yes.**

Descriptive evidence: Five facts

Overview

Questions:

1. Are breakdowns predictable? **Yes.**

2. Is PredictFix a good predictor of breakdowns? **Yes.**

3. Do alerts change technician behavior? **Yes.**

4.

5.

Does PredictFix have the *potential* to improve decision-making quality? **Yes.**

Does PredictFix *actually* improve decision-making quality? If so, what are the *quantitative* effects?

Descriptive evidence: Five facts

Overview

Questions:

1. Are breakdowns predictable? **Yes.**

2. Is PredictFix a good predictor of breakdowns? **Yes.**

3. Do alerts change technician behavior? **Yes.**

4. Benefits of PredictFix in aggregate data?

5.

Does PredictFix have the *potential* to improve decision-making quality? **Yes.**

Does PredictFix *actually* improve decision-making quality? If so, what are the *quantitative* effects?

Descriptive evidence: Five facts

Overview

Questions:

1. Are breakdowns predictable? **Yes.**

2. Is PredictFix a good predictor of breakdowns? **Yes.**

3. Do alerts change technician behavior? **Yes.**

4. Benefits of PredictFix in aggregate data? **No.**

5.

Does PredictFix have the *potential* to improve decision-making quality? **Yes.**

Does PredictFix *actually* improve decision-making quality? If so, what are the *quantitative* effects?

Descriptive evidence: Five facts

Overview

Questions:

1. Are breakdowns predictable? **Yes.**

2. Is PredictFix a good predictor of breakdowns? **Yes.**

3. Do alerts change technician behavior? **Yes.**

4. Benefits of PredictFix in aggregate data? **No.**

5. Cost conditions different in pre and post?

Does PredictFix have the *potential* to improve decision-making quality? **Yes.**

Does PredictFix *actually* improve decision-making quality? If so, what are the *quantitative* effects?

Descriptive evidence: Five facts

Overview

Questions:

1. Are breakdowns predictable? **Yes.**

2. Is PredictFix a good predictor of breakdowns? **Yes.**

3. Do alerts change technician behavior? **Yes.**

4. Benefits of PredictFix in aggregate data? **No.**

5. Cost conditions different in pre and post? **Yes.**

Does PredictFix have the *potential* to improve decision-making quality? **Yes.**

Does PredictFix *actually* improve decision-making quality? If so, what are the *quantitative* effects?

Where do we go from here?

Structural approach

Observed	
Pre	Post
Without PredictFix	With PredictFix
Pre costs	Post costs

Where do we go from here?

Structural approach

Observed	
Pre	Post
Without PredictFix	With PredictFix
Pre costs	Post costs

Next steps:

- 1. Identification:** Describe model of technician decision-making and conditions required for identification.

Where do we go from here?

Structural approach

Observed		Counterfactual
Pre	Post	Post'
Without PredictFix Pre costs	With PredictFix Post costs	With PredictFix Pre costs

Next steps:

- 1. Identification:** Describe model of technician decision-making and conditions required for identification.
- 2. Quantification:** Estimate model, evaluate counterfactual Post'.

Overview

1. Setting & data.
2. Descriptive evidence: Five facts.
3. Model and estimation.
4. Results
 - Unpacking estimates.
 - Evaluating counterfactual.
5. Discussion
 - Reflecting on AI's role
 - Looking ahead

Model & estimation overview

Model elements:

0. Technician's objective: Minimizing $\text{total cost} = \text{repair cost} + \text{breakdown cost}$.

Model & estimation overview

Model elements:

0. Technician's objective: Minimizing *total cost* = *repair cost* + *breakdown cost*.

← tradeoff →

Model & estimation overview

Model elements:

0. Technician's objective: Minimizing *total cost* = *repair cost* + *breakdown cost*.
 ↑—tradeoff—↑
1. Costs of repair, breakdown (pre and post).
2. Beliefs about breakdown risk (with and without PredictFix)

Model & estimation overview

Model elements:

- 0. Technician's objective: Minimizing *total cost* = *repair cost* + *breakdown cost*.

↑—tradeoff—↑

- 1. Costs of repair, breakdown (pre and post).

- 2. Beliefs about breakdown risk (with and without PredictFix): $\rho : \mathcal{X} \rightarrow [0, 1]$.

$\rho(x) \stackrel{?}{=} \pi(x)$ = *objective risk of breakdown*.

Model & estimation overview

Model elements:

0. Technician's objective: Minimizing $\text{total cost} = \text{repair cost} + \text{breakdown cost}$.

↑ tradeoff ↓

1. Costs of repair, breakdown (pre and post).

2. Beliefs about breakdown risk (with and without PredictFix): $\rho : \mathcal{X} \rightarrow [0, 1]$.

$\rho(x) \stackrel{?}{=} \pi(x)$ = objective risk of breakdown.

Objective: Estimate costs and beliefs from data.

Model & estimation overview

Model elements:

0. Technician's objective: Minimizing *total cost* = *repair cost* + *breakdown cost*.

↑ tradeoff ↑

1. Costs of repair, breakdown (pre and post).

2. Beliefs about breakdown risk (with and without PredictFix): $\rho : \mathcal{X} \rightarrow [0, 1]$.

$\rho(x) \stackrel{?}{=} \pi(x)$ = *objective risk of breakdown*.

Challenge 1: Separate identification. Details

Objective: Estimate costs and beliefs from data.

Model & estimation overview

Model elements:

0. Technician's objective: Minimizing *total cost* = *repair cost* + *breakdown cost*.

↑ tradeoff ↓

1. Costs of repair, breakdown (pre and post).

2. Beliefs about breakdown risk (with and without PredictFix): $\rho : \mathcal{X} \rightarrow [0, 1]$.

$\rho(x) \stackrel{?}{=} \pi(x)$ = objective risk of breakdown.

3. Dynamic considerations.

Challenge 1: Separate identification. Details

Objective: Estimate costs and beliefs from data.

Model & estimation overview

Model elements:

0. Technician's objective: Minimizing $\text{total cost} = \text{repair cost} + \text{breakdown cost}$.

↑ tradeoff ↑

1. Costs of repair, breakdown (pre and post).

2. Beliefs about breakdown risk (with and without PredictFix): $\rho : \mathcal{X} \rightarrow [0, 1]$.

$\rho(x) \stackrel{?}{=} \pi(x)$ = objective risk of breakdown.

3. Dynamic considerations. **Challenge 2: High-dimensional data.**

Details

Challenge 1: Separate identification.

Details

Objective: Estimate costs and beliefs from data.

Static model

Constant B

More on payoffs

Static model

Constant B

More on payoffs

		Potential breakdown	
		$s = 0$	$s = 1$
Repair action	$a = 0$	0	$-B$
	$a = 1$	$-c(v, x)$	$-c(v, x)$

State variables:

- v : cost-related variables.
- x : state of truck.

Static model

Constant B

More on payoffs

		Potential breakdown	
		$s = 0$	$s = 1$
Repair action	$a = 0$	0	$-B$
	$a = 1$	$-c(v, x)$	$-c(v, x)$

State variables:

- v : cost-related variables.
- x : state of truck.

- True probability of breakdown: $\pi(x) = \Pr(s = 1 | x)$.

Static model

Constant B

More on payoffs

		Potential breakdown	
		$s = 0$	$s = 1$
Repair action	$a = 0$	0	$-B$
	$a = 1$	$-c(v, x)$	$-c(v, x)$

- **True** probability of breakdown: $\pi(x) = \Pr(s = 1 | x)$.
- **Perceived** probability of breakdowns: $p(x) \stackrel{?}{=} \pi(x)$.
 - Not imposing specific model of how technicians form these beliefs.

State variables:

- v : cost-related variables.
- x : state of truck.

Static model

Constant B

More on payoffs

		Potential breakdown	
		$s = 0$	$s = 1$
Repair action	$a = 0$	0	$-B$
	$a = 1$	$-c(v, x)$	$-c(v, x)$

State variables:

- v : cost-related variables.
- x : state of truck.

- True probability of breakdown: $\pi(x) = \Pr(s = 1 | x)$.
- Perceived probability of breakdowns: $\rho(x) \stackrel{?}{=} \pi(x)$.
 - Not imposing specific model of how technicians form these beliefs.

Threshold: Repair if and only if

$$\rho(x) > \frac{c(v, x)}{B} \equiv \tau(v, x)$$

Static model

Constant B

More on payoffs

		Potential breakdown	
		$s = 0$	$s = 1$
Repair action	$a = 0$	0	$-B$
	$a = 1$	$-c(v, x)$	$-c(v, x)$

State variables:

- v : cost-related variables.
- x : state of truck.

- True probability of breakdown: $\pi(x) = \Pr(s = 1 | x)$.
- Perceived probability of breakdowns: $\rho(x) \stackrel{?}{=} \pi(x)$.
 - Not imposing specific model of how technicians form these beliefs.

Threshold: Repair if and only if

$$\rho(x) > \frac{c(v, x)}{B} \equiv \tau(v, x)$$

Key questions:

- Does PredictFix move ρ closer to π ?
- If so, to what extent do better decisions result?

► Results

18/25

Econometric challenge

Repair if and only if

$$\rho(x) > \frac{c(v, x)}{B} \equiv \tau(v, x)$$

Goal: Estimate preferences τ and beliefs ρ .

Fundamental challenge: How to separately identify **preferences and beliefs?**

Econometric challenge

Repair if and only if

$$\rho(x) > \frac{c(v, x)}{B} \equiv \tau(v, x)$$

Goal: Estimate preferences τ and beliefs ρ .

Fundamental challenge: How to separately identify **preferences and beliefs?**

Identification proof

- ① Restriction on private information.
- ② Exclusion restriction.
- ③ Restriction to identify level and scale.

Econometric challenge

Repair if and only if

$$\rho(x) > \frac{c(v, x)}{B} \equiv \tau(v, x)$$

Goal: Estimate preferences τ and beliefs ρ .

Fundamental challenge: How to separately identify **preferences and beliefs?**

Identification proof

- ① Restriction on private information. → No private information on state of truck. Details
- ② Exclusion restriction.
- ③ Restriction to identify level and scale.

Econometric challenge

Repair if and only if

$$\rho(x) > \frac{c(v, x)}{B} \equiv \tau(v, x) = g(w, x) + \epsilon \quad \text{where } v = (w, \epsilon), \epsilon \sim \text{Logistic}(\theta)$$

Goal: Estimate preferences τ and beliefs ρ .

Fundamental challenge: How to separately identify **preferences and beliefs?**

Identification proof

- ① Restriction on private information. → No private information on state of truck. Details
- ② Exclusion restriction.
- ③ Restriction to identify level and scale.

Econometric challenge

Repair if and only if

$$\rho(x) > \frac{c(v, x)}{B} \equiv \tau(v, x) = g(w, x) + \epsilon \quad \text{where } v = (w, \epsilon), \epsilon \sim \text{Logistic}(\theta)$$

$$\Rightarrow p(w, x) \equiv \Pr(a = 1 | w, x) = \Lambda(\theta[-g(w, x) + \rho(x)])$$

Goal: Estimate preferences τ and beliefs ρ .

Fundamental challenge: How to separately identify **preferences and beliefs?**

Identification proof

- ① Restriction on private information. → No private information on state of truck. [Details](#)
- ② Exclusion restriction.
- ③ Restriction to identify level and scale.

Econometric challenge

Repair if and only if

$$\rho(x) > \frac{c(v, x)}{B} \equiv \tau(v, x) = g(w, x) + \epsilon \quad \text{where } v = (w, \epsilon), \epsilon \sim \text{Logistic}(\theta)$$

$$\Rightarrow p(w, x) \equiv \Pr(a = 1 | w, x) = \Lambda(\theta[-g(w, x) + \rho(x)])$$

Goal: Estimate preferences τ and beliefs ρ .

Fundamental challenge: How to separately identify **preferences and beliefs?**

Identification proof

- ① Restriction on private information. → No private information on state of truck. [Details](#)
- ② Exclusion restriction.
- ③ Restriction to identify level and scale.

Econometric challenge

Repair if and only if

$$\rho(x) > \frac{c(v, x)}{B} \equiv \tau(v, x) = g(w, x) + \epsilon \quad \text{where } v = (w, \epsilon), \epsilon \sim \text{Logistic}(\theta)$$

$$\Rightarrow p(w, x) \equiv \Pr(a = 1 | w, x) = \Lambda(\theta[-g(w, x) + \rho(x)])$$

Goal: Estimate preferences τ and beliefs ρ .

Fundamental challenge: How to separately identify **preferences and beliefs?**

Identification proof

- ① Restriction on private information. → No private information on state of truck. [Details](#)
- ② Exclusion restriction. → $c(v, x) = c_{\text{tangible}}(v, x) + c_{\text{intangible}}(v)$.
- ③ Restriction to identify level and scale.

Econometric challenge

Repair if and only if

$$\rho(x) > \frac{c(v, x)}{B} \equiv \tau(v, x) = g(w, x) + \epsilon \quad \text{where } v = (w, \epsilon), \epsilon \sim \text{Logistic}(\theta)$$

$$\Rightarrow p(w, x) \equiv \Pr(a = 1 | w, x) = \Lambda(\theta[-g(w, x) + \rho(x)])$$

Goal: Estimate preferences τ and beliefs ρ .

Fundamental challenge: How to separately identify **preferences and beliefs?**

Identification proof

- ① Restriction on private information. → No private information on state of truck. [Details](#)
- ② Exclusion restriction. → $c(v, x) = c_{\text{tangible}}(v, x) + c_{\text{intangible}}(v)$.
- ③ Restriction to identify level and scale.

Econometric challenge

Repair if and only if

$$\rho(x) > \frac{c(v, x)}{B} \equiv \tau(v, x) = g(w, x) + \epsilon \quad \text{where } v = (w, \epsilon), \epsilon \sim \text{Logistic}(\theta)$$

$$\Rightarrow p(w, x) \equiv \Pr(a = 1 | w, x) = \Lambda(\theta[-g(w, x) + \rho(x)])$$

Goal: Estimate preferences τ and beliefs ρ .

Fundamental challenge: How to separately identify **preferences and beliefs?**

Identification proof

- ① Restriction on private information. → No private information on state of truck. Details
- ② Exclusion restriction. → $c(v, x) = c_{\text{tangible}}(v, x) + c_{\text{intangible}}(v)$.
- ③ Restriction to identify level and scale. → Two restrictions on ρ .

Separate identification of preferences and beliefs

Two restrictions on ρ

- Technician's perceived risk of breakdown risk:

1. is correct on average, i.e., $\mathbb{E}_x \rho(x) = \mathbb{E}_x \pi(x)$; and
2. attains the zero lower bound, i.e., $\min_x \rho(x) = 0$.

Separate identification of preferences and beliefs

Two restrictions on ρ

- Technician's perceived risk of breakdown risk:

1. is correct on average, i.e., $\mathbb{E}_x \rho(x) = \mathbb{E}_x \pi(x)$; and
 - ↳ Average over states.
2. attains the zero lower bound, i.e., $\min_x \rho(x) = 0$.

Separate identification of preferences and beliefs

Two restrictions on ρ

- Technician's perceived risk of breakdown risk:
 1. is correct on average, i.e., $\mathbb{E}_x \rho(x) = \mathbb{E}_x \pi(x)$; and
 - ↳ Average over states.
 2. attains the zero lower bound, i.e., $\min_x \rho(x) = 0$.
- Does not restrict how technicians **order** states by risk.

Dynamics: Toward estimation

Technician's problem is inherently dynamic. This week's action → future weeks' states.

Beliefs

Dynamics: Toward estimation

Technician's problem is inherently dynamic. This week's action → future weeks' states.

Beliefs

$$p(w_t, x_t) = \Lambda \left(\underbrace{\theta [-g(w_t, x_t) + \rho(x_t)]}_{\text{static payoffs}} + \underbrace{\delta (EV_1(w_t, x_t) - EV_0(w_t, x_t))}_{\text{ex-ante expected value functions}} \right)$$

Dynamics: Toward estimation

Technician's problem is inherently dynamic. This week's action → future weeks' states.

Beliefs

$$p(w_t, x_t) = \Lambda \left(\underbrace{\theta [-g(w_t, x_t) + \rho(x_t)}_{\text{static payoffs}} + \underbrace{\delta (EV_1(w_t, x_t) - EV_0(w_t, x_t))}_{\text{ex-ante expected value functions}} \right)$$

Question: How to bring this to the data?

More

Dynamics: Toward estimation

Technician's problem is inherently dynamic. This week's action → future weeks' states.

Beliefs

$$p(w_t, x_t) = \Lambda \left(\underbrace{\theta [-g(w_t, x_t) + \rho(x_t)]}_{\text{static payoffs}} + \underbrace{\delta (EV_1(w_t, x_t) - EV_0(w_t, x_t))}_{\text{ex-ante expected value functions}} \right)$$

Question: How to bring this to the data?

More

- Hotz and Miller (1993) + Arcidiacono and Miller (2011):

Details

Renewal

$$EV_0, EV_1 = \mathbb{E} [f(p(w_{t+1}, x_{t+1}))]$$

Dynamics: Toward estimation

Technician's problem is inherently dynamic. This week's action → future weeks' states.

Beliefs

$$p(w_t, x_t) = \Lambda \left(\underbrace{\theta [-g(w_t, x_t) + \rho(x_t)]}_{\text{static payoffs}} + \underbrace{\delta (EV_1(w_t, x_t) - EV_0(w_t, x_t))}_{\text{ex-ante expected value functions}} \right)$$

Question: How to bring this to the data?

More

- Hotz and Miller (1993) + Arcidiacono and Miller (2011):

Details

Renewal

$$EV_0, EV_1 = \mathbb{E} [f(p(w_{t+1}, x_{t+1}))]$$

- **Challenge:** State is high-dimensional.

Assumptions summary

► Results

21/25

Dynamics: Toward estimation

Technician's problem is inherently dynamic. This week's action → future weeks' states.

Beliefs

$$p(w_t, x_t) = \Lambda \left(\underbrace{\theta [-g(w_t, x_t) + \rho(x_t)]}_{\text{static payoffs}} + \underbrace{\delta (EV_1(w_t, x_t) - EV_0(w_t, x_t))}_{\text{ex-ante expected value functions}} \right)$$

Question: How to bring this to the data?

More

- Hotz and Miller (1993) + Arcidiacono and Miller (2011):

Details

Renewal

$$EV_0, EV_1 = \mathbb{E} [f(p(w_{t+1}, x_{t+1}))]$$

- Challenge: State is high-dimensional.

→ Our solution: $p(w_{t+1}, x_{t+1}) \mid w_t, x_t, a_t \sim \text{Beta}$ → closed-form expression for EV_0, EV_1 .

Formal statement

Parameterization

Details

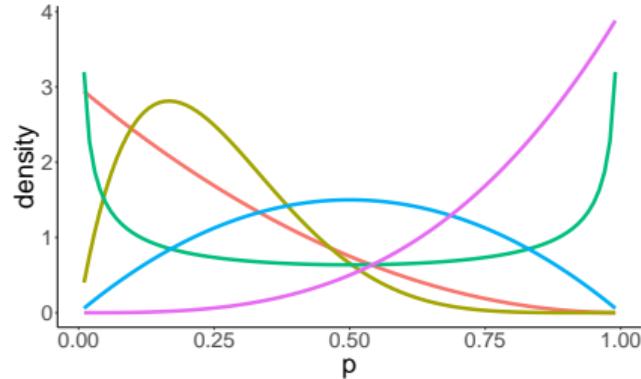
Assumptions summary

► Results

21/25

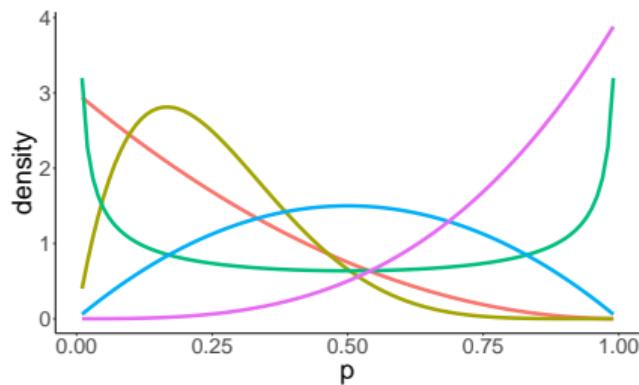
Beta distribution

(a) Examples: Beta for various parameters

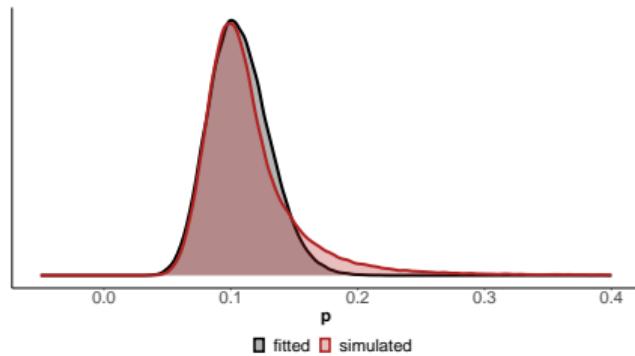


Beta distribution

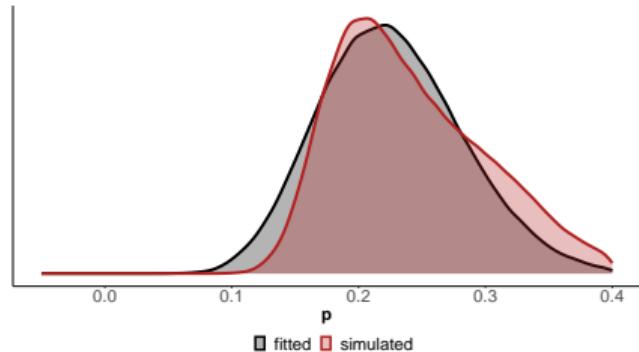
(a) Examples: Beta for various parameters



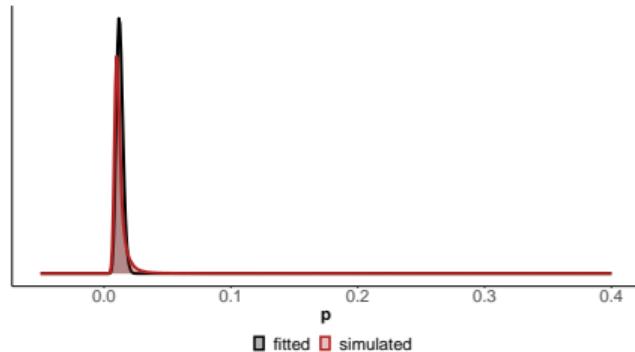
(b) Example 1



(c) Example 2



(d) Example 3



Model estimation

- Variable selection (x): ~ 2000 -dimensional \rightarrow 20-dimensional. [Details](#)
- Functional form for ρ : $\rho(x) = \Lambda(\lambda_0 + x'\lambda_1)$
- Functional form for g : $g(w, x) = \gamma_0 + w'\gamma_1 + \gamma_2\hat{g}_2(w, x)$. [Details](#)

Model estimation

- Variable selection (x): ~ 2000 -dimensional $\rightarrow 20$ -dimensional. [Details](#)
- Functional form for ρ : $\rho(x) = \Lambda(\lambda_0 + x'\lambda_1)$
- Functional form for g : $g(w, x) = \gamma_0 + w'\gamma_1 + \gamma_2\hat{g}_2(w, x)$. [Details](#)

Estimate $\beta = (\theta, \gamma, \lambda, \dots)$ **separately for pre and post using constrained maximum likelihood**:

Model estimation

- Variable selection (x): ~ 2000 -dimensional $\rightarrow 20$ -dimensional. [Details](#)
- Functional form for ρ : $\rho(x) = \Lambda(\lambda_0 + x'\lambda_1)$
- Functional form for g : $g(w, x) = \gamma_0 + w'\gamma_1 + \gamma_2\hat{g}_2(w, x)$. [Details](#)

Estimate $\beta = (\theta, \gamma, \lambda, \dots)$ **separately for pre and post using constrained maximum likelihood**:

$$\max_{\beta} \sum_i \log \mathcal{L}(a_i, w_i, x_i | \beta) \quad (\text{Likelihood of observed repair decisions})$$

where $\mathcal{L}(a_i, w_i, x_i | \beta) = \begin{cases} 1 - p(w_i, x_i | \beta) & \text{if } a_i = 0 \\ p(w_i, x_i | \beta) & \text{if } a_i = 1 \end{cases}$

Model estimation

- Variable selection (x): ~ 2000 -dimensional $\rightarrow 20$ -dimensional. [Details](#)
- Functional form for ρ : $\rho(x) = \Lambda(\lambda_0 + x'\lambda_1)$
- Functional form for g : $g(w, x) = \gamma_0 + w'\gamma_1 + \gamma_2\hat{g}_2(w, x)$. [Details](#)

Estimate $\beta = (\theta, \gamma, \lambda, \dots)$ **separately for pre and post using constrained maximum likelihood**:

$$\max_{\beta} \sum_i \log \mathcal{L}(a_i, w_i, x_i | \beta) \quad (\text{Likelihood of observed repair decisions})$$

subject to $\frac{1}{N} \sum_{i=1}^N \rho(x_i; \lambda) = \bar{\pi} \quad (\text{Correct on average})$

$$\min_x \rho(x; \lambda) \approx 0 \quad (\text{Zero lower bound})$$

where $\mathcal{L}(a_i, w_i, x_i | \beta) = \begin{cases} 1 - p(w_i, x_i | \beta) & \text{if } a_i = 0 \\ p(w_i, x_i | \beta) & \text{if } a_i = 1 \end{cases}$

Overview

1. Setting & data.
2. Descriptive evidence: Five facts.
3. Model and estimation.
4. Results
 - Unpacking estimates: beliefs and costs.
 - Evaluating counterfactual.
5. Discussion
 - Reflecting on AI's role
 - Looking ahead

The effect of PredictFix on technicians' beliefs

A key question: With PredictFix, do technicians exhibit a better understanding of breakdown risk?

The effect of PredictFix on technicians' beliefs

A key question: With PredictFix, do technicians exhibit a better understanding of breakdown risk?

→ Is \hat{p}_{post} a better predictor of breakdowns than \hat{p}_{pre} ?

The effect of PredictFix on technicians' beliefs

A key question: With PredictFix, do technicians exhibit a better understanding of breakdown risk?

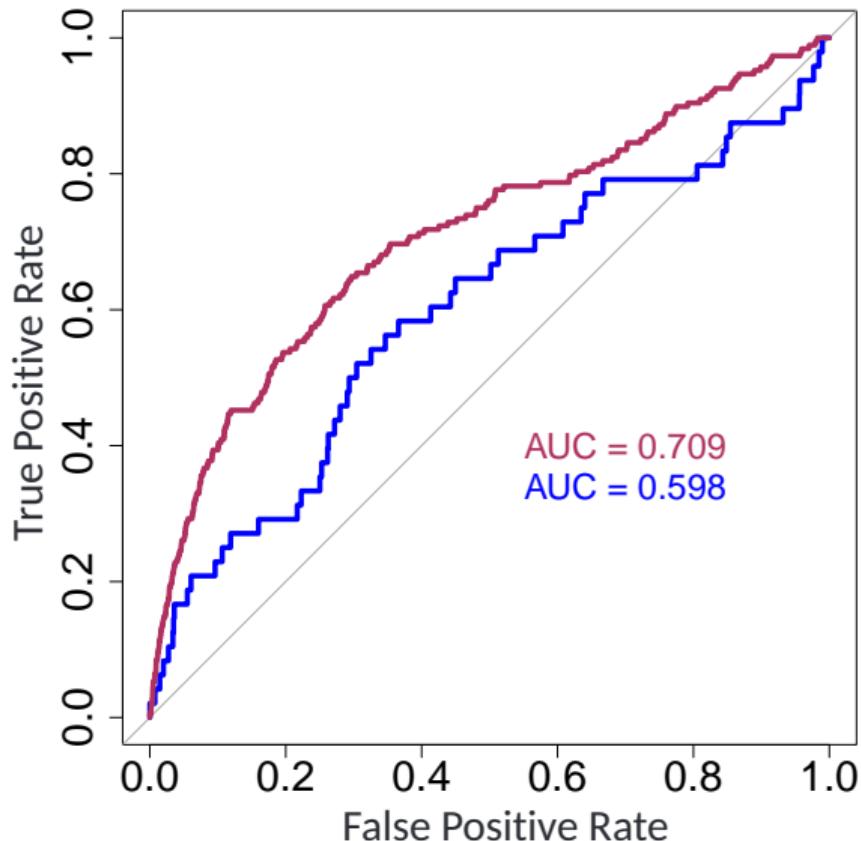
→ Is \hat{p}_{post} a better predictor of breakdowns than \hat{p}_{pre} ?

Tool: ROC curves.

Table: ROC curve specifications

Predictor	Outcome
\hat{p}_{pre}	Breakdowns
\hat{p}_{post}	Breakdowns

The effect of PredictFix on technicians' predictions (ρ)

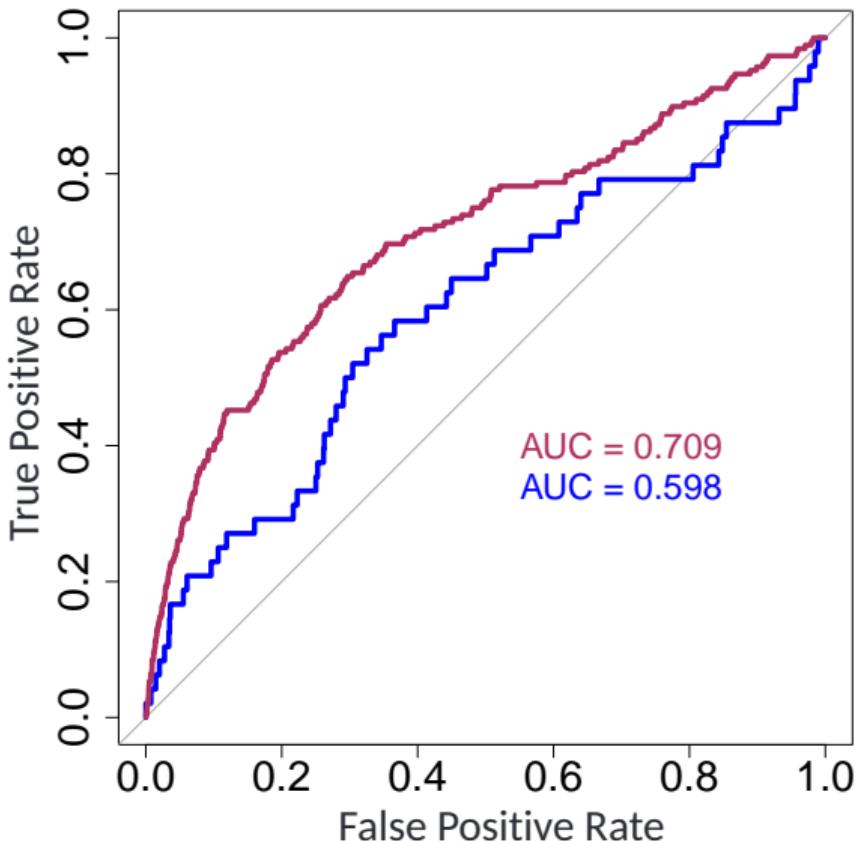


Predictor	
$\hat{\rho}_{\text{pre}}$	Beliefs w/o PredictFix
$\hat{\rho}_{\text{post}}$	Beliefs w/ PredictFix

Note:

- $\hat{\rho}_{\text{post}}$ strictly dominates $\hat{\rho}_{\text{pre}}$.

The effect of PredictFix on technicians' predictions (ρ)

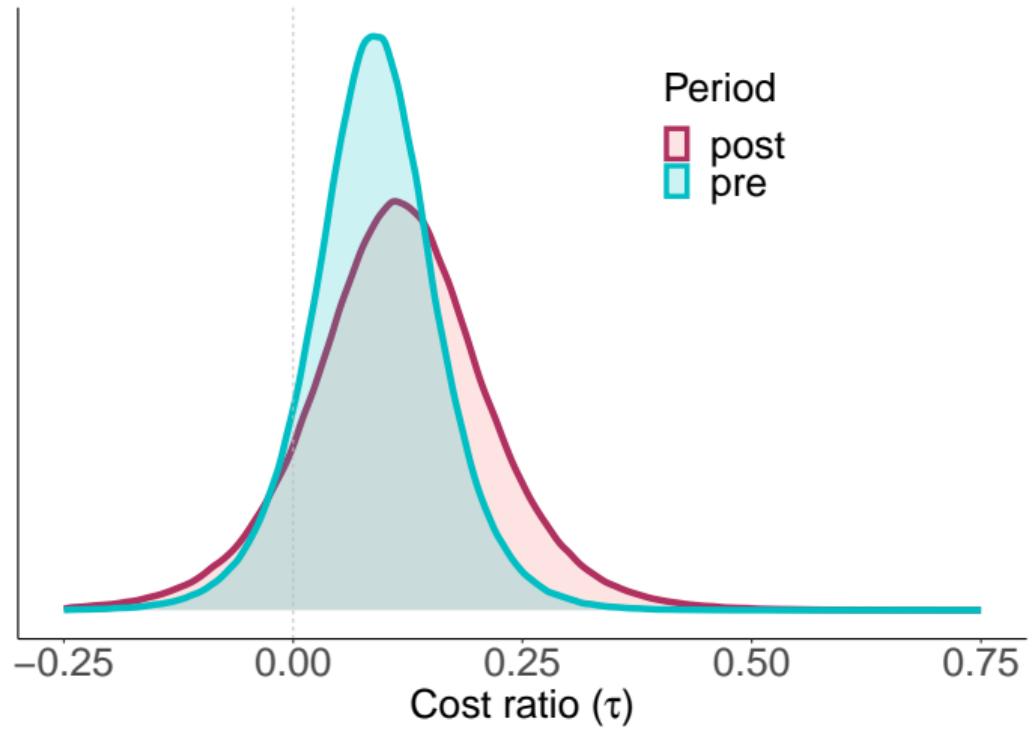


Predictor	
$\hat{\rho}_{\text{pre}}$	Beliefs w/o PredictFix
$\hat{\rho}_{\text{post}}$	Beliefs w/ PredictFix

Note:

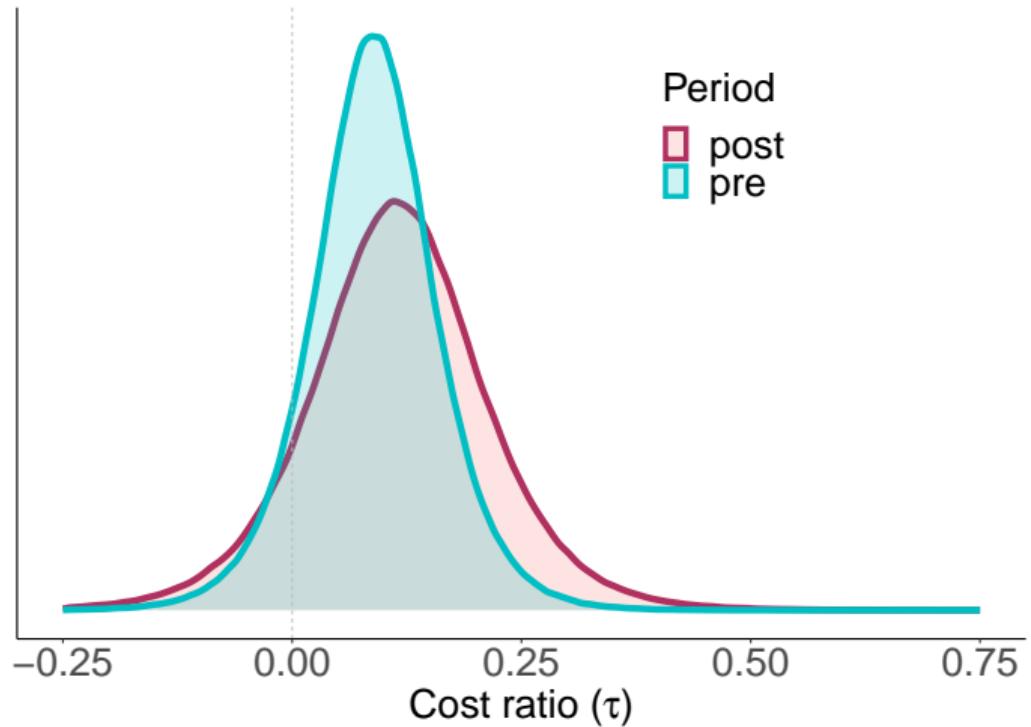
- $\hat{\rho}_{\text{post}}$ strictly dominates $\hat{\rho}_{\text{pre}}$.
- **Robust** to relaxation restrictions on beliefs: (1) correct on average and (2) zero lower bound.

Cost ratio (τ) = repair cost / breakdown cost



	Pre	Post
mean(τ)	0.0892	0.115
std. dev.(τ)	0.0718	0.102

Cost ratio (τ) = repair cost / breakdown cost



	Pre	Post
mean(τ)	0.0892	0.115
std. dev.(τ)	0.0718	0.102

- Supply chain disruptions.
- Fact 5: Tangible costs ↑.
- Repair completion time ↑.

Counterfactuals

Overview

Questions:

1. What is the value of PredictFix?
2. How is this value realized?

Counterfactuals

Overview

Questions:

1. What is the value of PredictFix?
2. How is this value realized?

Hold costs fixed, simulate decisions for:

- (i) $\rho = \rho_{\text{pre}}$ (without PredictFix)
- (ii) $\rho = \rho_{\text{post}}$ (with PredictFix)

Comparing (i) and (ii): How do *total costs* = *repair costs* + *breakdown costs* compare?

Counterfactuals

Overview

Questions:

1. What is the value of PredictFix?
2. How is this value realized?

Hold costs fixed, simulate decisions for:

- (i) $\rho = \rho_{\text{pre}}$ (without PredictFix)
- (ii) $\rho = \rho_{\text{post}}$ (with PredictFix)
- (iii) $\rho = \pi$ (optimal; technician knows true risk)

Comparing (i) and (ii): How do *total costs* = *repair costs* + *breakdown costs* compare?

Counterfactuals

Overview

Questions:

1. What is the value of PredictFix?
2. How is this value realized?

To evaluate counterfactuals, we need:

- Estimate of transition process. [Details](#)
- To get dollar interpretation, set value for B . Reasonable range: \$5,000-\$10,000.

Hold costs fixed, simulate decisions for:

- (i) $\rho = \rho_{\text{pre}}$ (without PredictFix)
- (ii) $\rho = \rho_{\text{post}}$ (with PredictFix)
- (iii) $\rho = \pi$ (optimal; technician knows true risk)

Comparing (i) and (ii): How do *total costs = repair costs + breakdown costs* compare?

Counterfactuals

Overview

Questions:

1. What is the value of PredictFix?
2. How is this value realized?

To evaluate counterfactuals, we need:

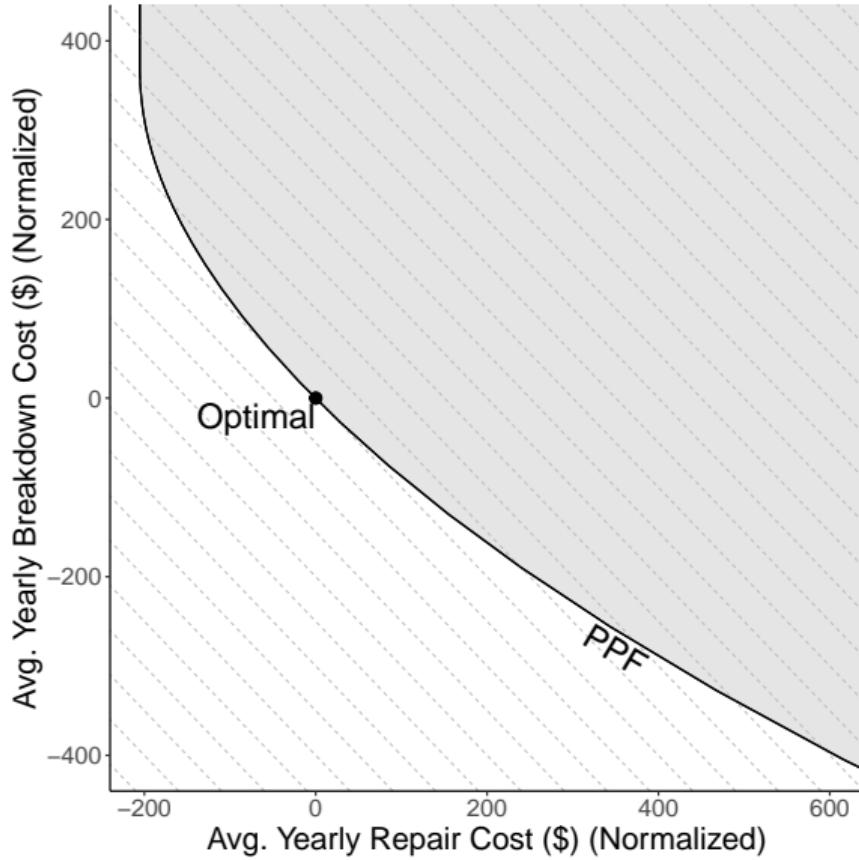
- Estimate of transition process. [Details](#)
- To get dollar interpretation, set value for B . Reasonable range: \$5,000-\$10,000.

Hold costs fixed, simulate decisions for:

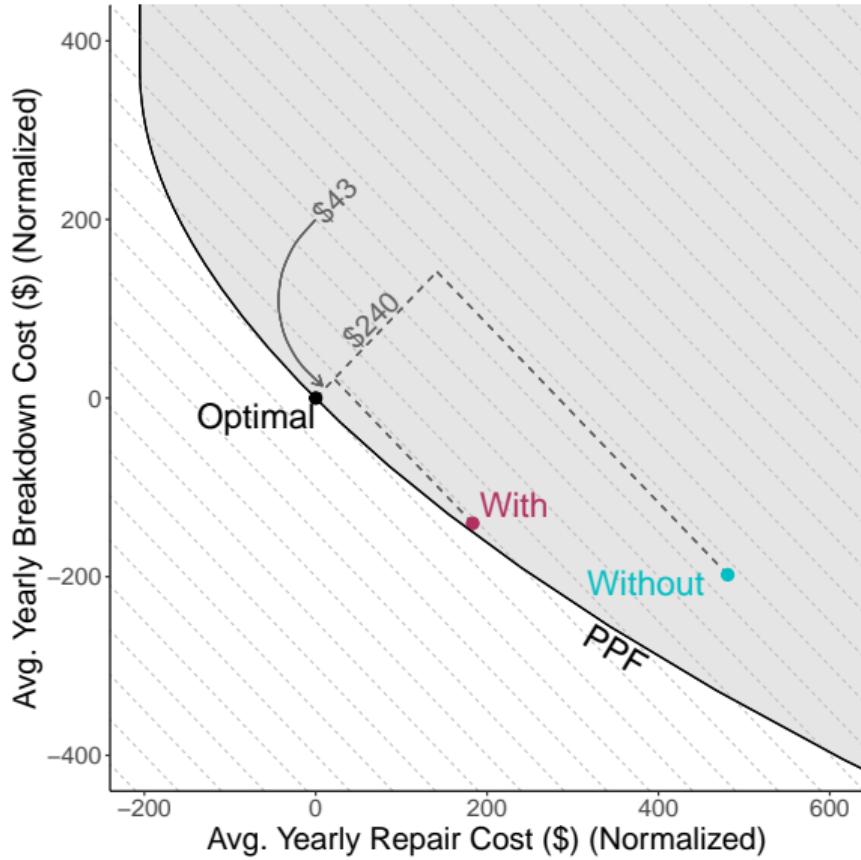
- (i) $\rho = \rho_{\text{pre}}$ (without PredictFix)
- (ii) $\rho = \rho_{\text{post}}$ (with PredictFix)
- (iii) $\rho = \pi$ (optimal; technician knows true risk)

Comparing (i) and (ii): How do *total costs = repair costs + breakdown costs* compare?

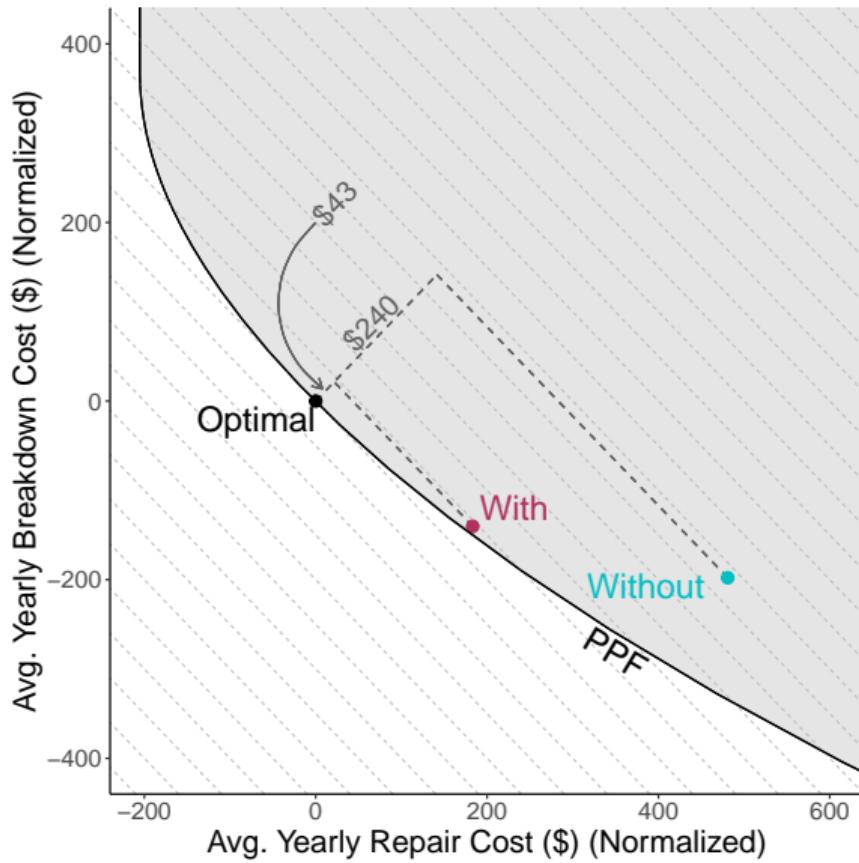
The value of PredictFix



The value of PredictFix



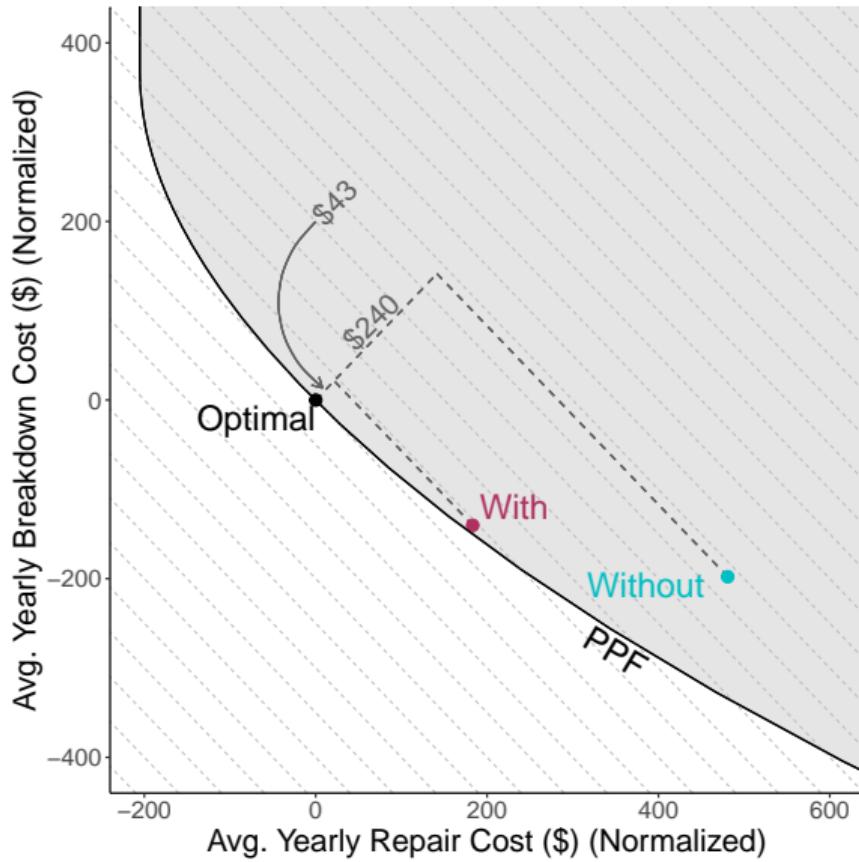
The value of PredictFix



Value of PredictFix:

- Total cost reduction: \$240.

The value of PredictFix



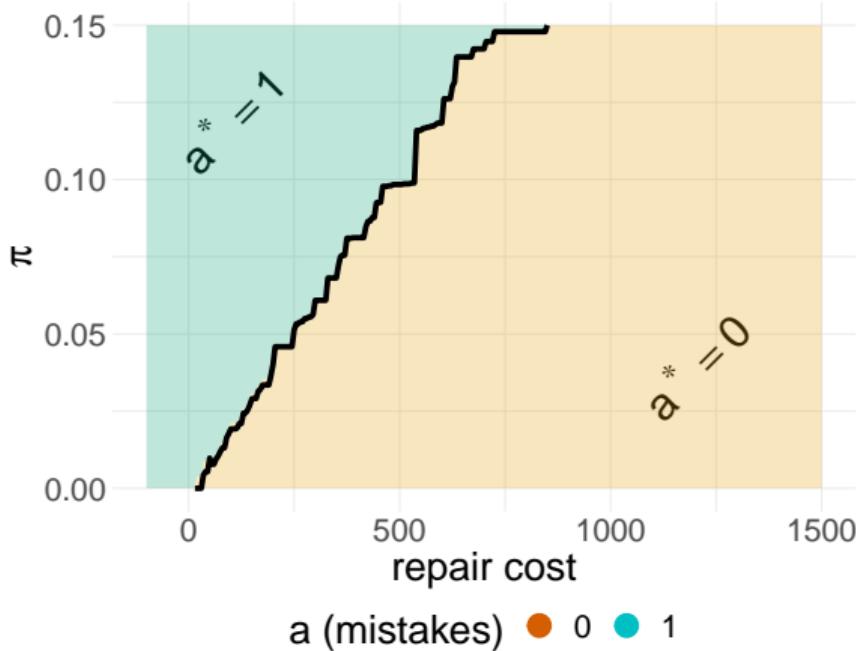
Value of PredictFix:

- Total cost reduction: **\$240**.
- Achieves $240 / (240 + 43) = 85\%$ of all feasible cost savings.

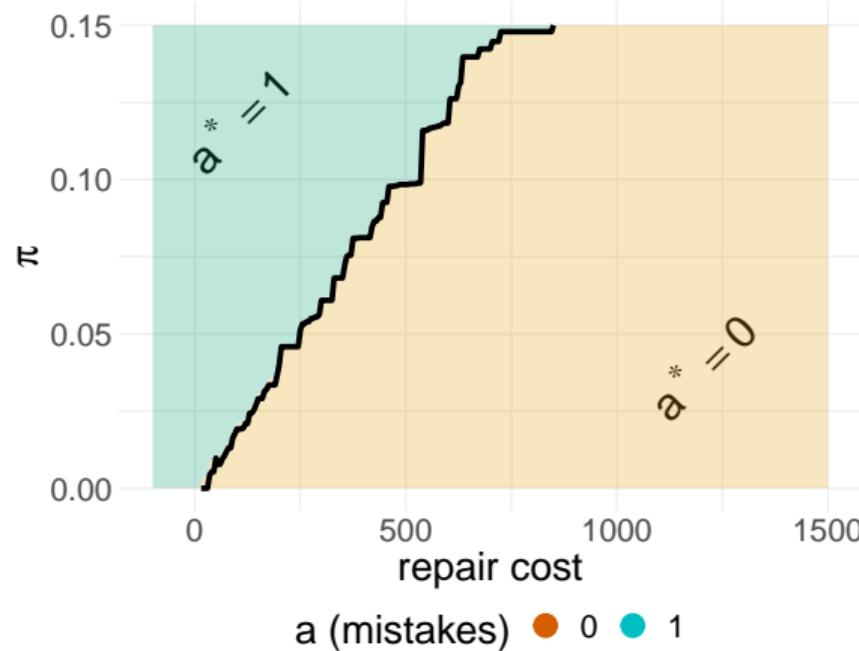
Unpacking the value of PredictFix

Mistakes with and without PredictFix

Without PredictFix



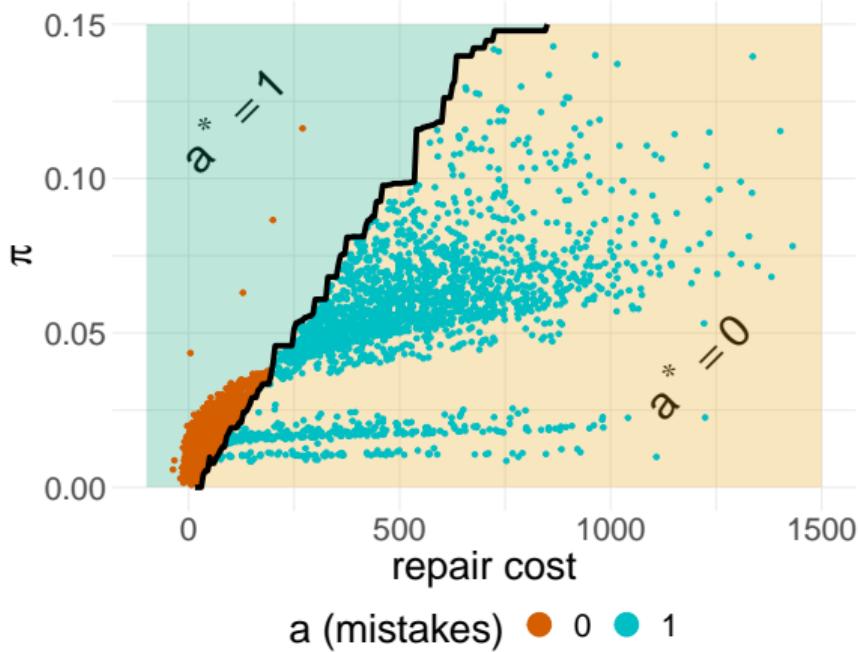
With PredictFix



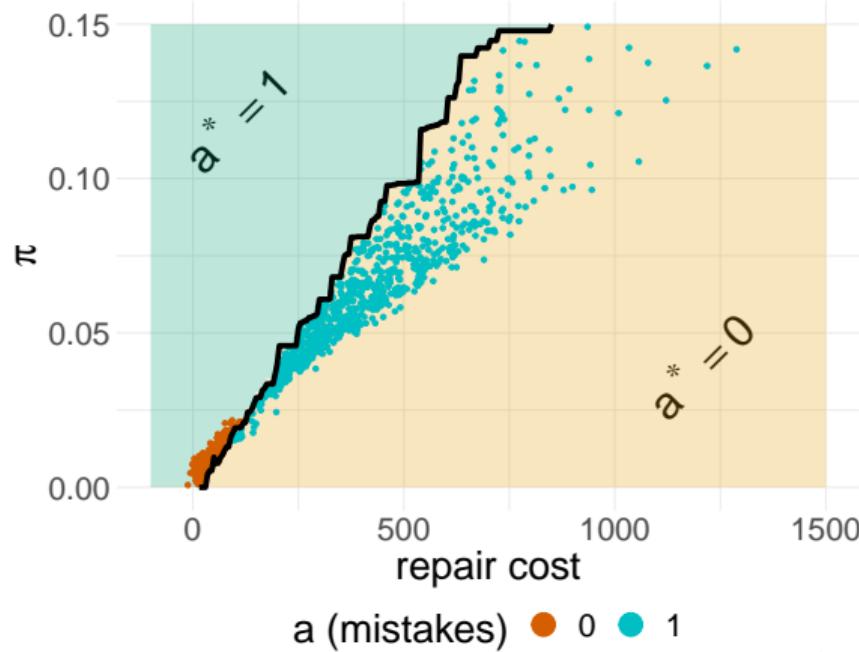
Unpacking the value of PredictFix

Mistakes with and without PredictFix

Without PredictFix



With PredictFix



Overview

1. Setting & data.
2. Descriptive evidence: Five facts.
3. Model and estimation.
4. Results
 - Unpacking estimates: beliefs and costs.
 - Evaluating counterfactual.
5. Discussion
 - Reflecting on AI's role
 - Looking ahead

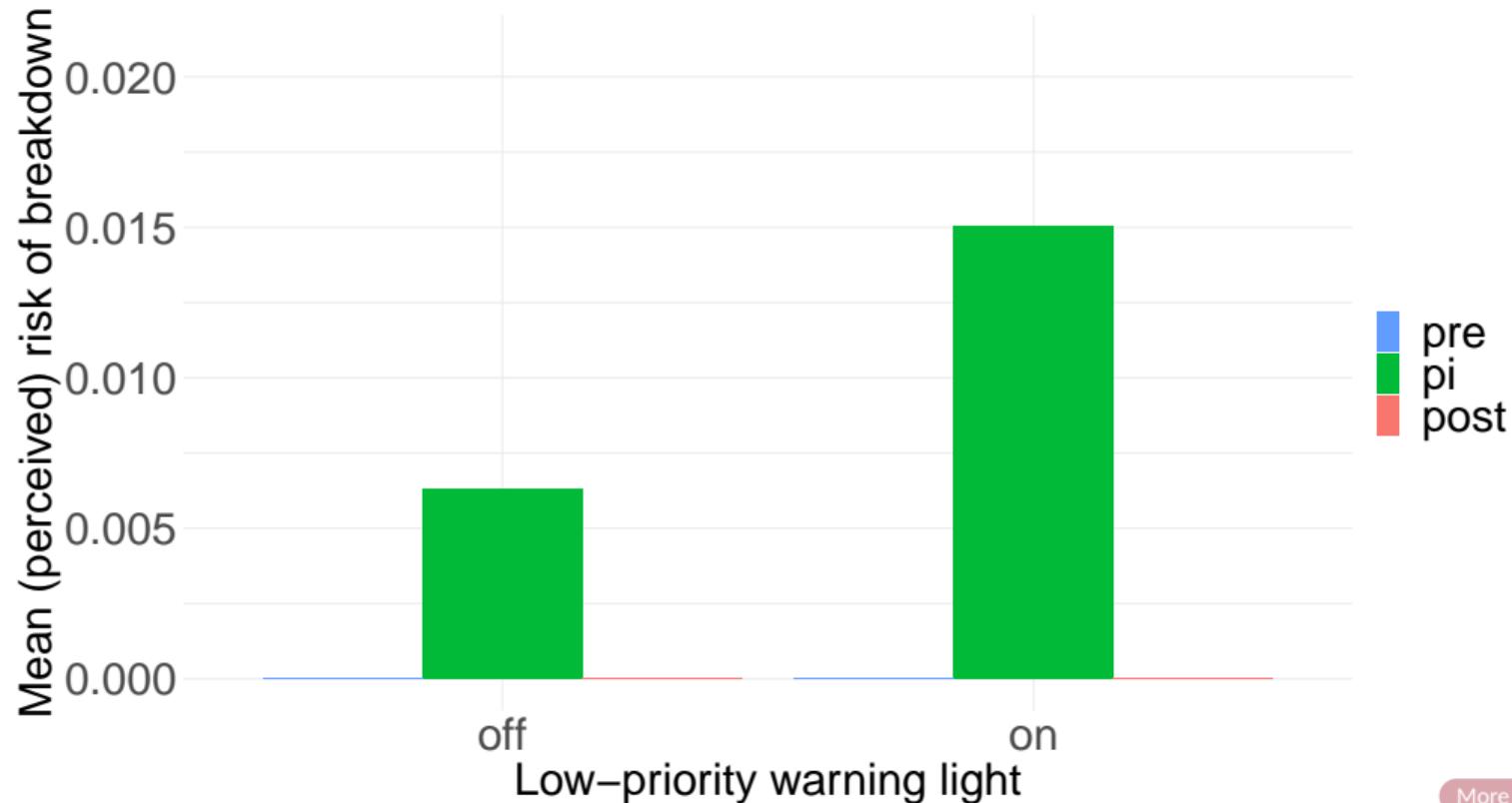
Where do technicians make mistakes?

An illustrative example: How technicians respond to dashboard warning lights



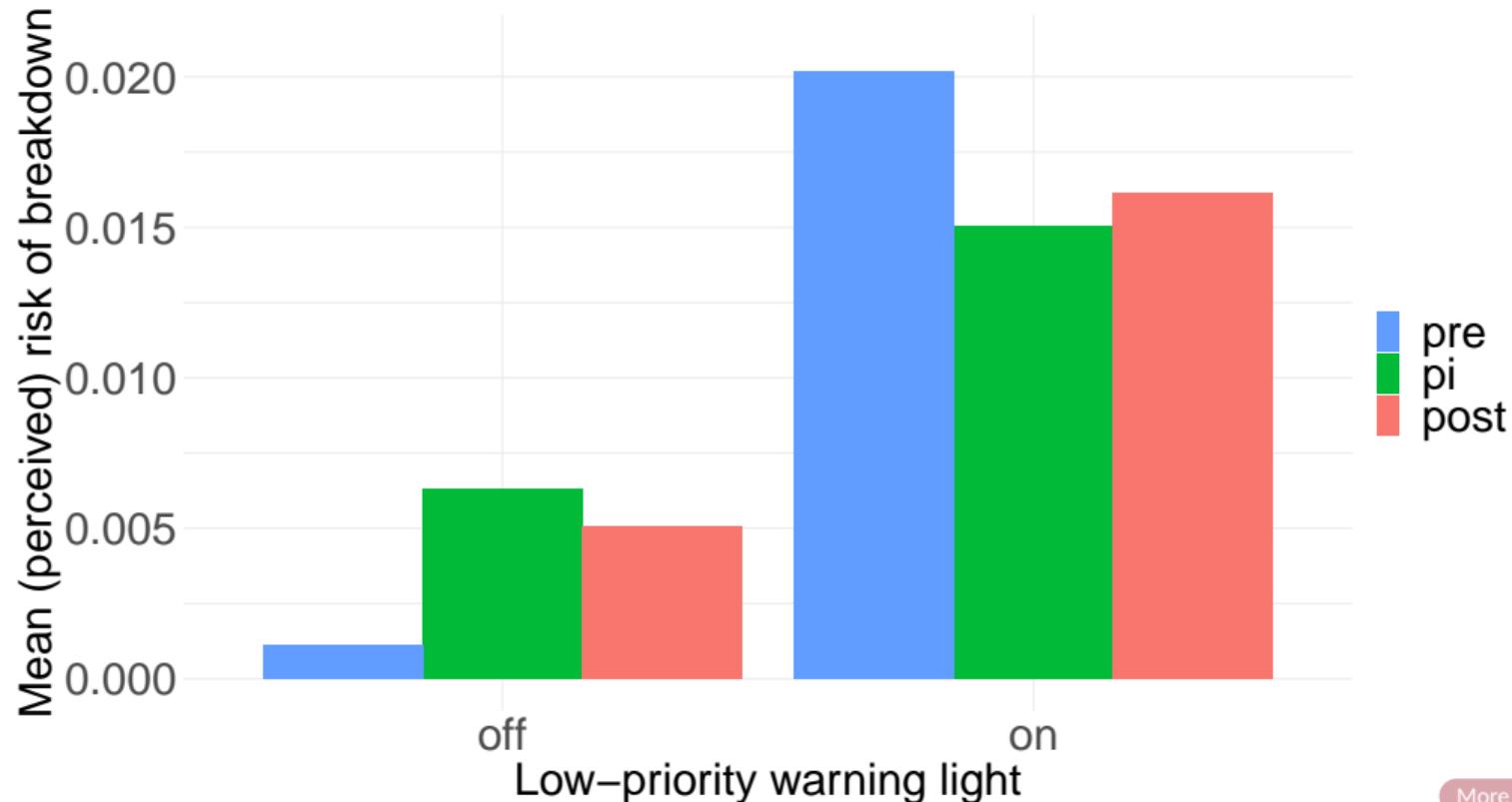
Where do technicians make mistakes?

An illustrative example: How technicians respond to dashboard warning lights



Where do technicians make mistakes?

An illustrative example: How technicians respond to dashboard warning lights



Where do technicians make mistakes?

An illustrative example: How technicians respond to dashboard warning lights



Looking ahead

A few key questions

1. Heterogeneous effects of AI on human decision-makers?

Looking ahead

A few key questions

1. Heterogeneous effects of AI on human decision-makers?

→ Are predictive AI and human experience complements or substitutes?

Looking ahead

A few key questions

1. Heterogeneous effects of AI on human decision-makers?

→ Are predictive AI and human experience complements or substitutes?

Looking ahead

A few key questions

1. Heterogeneous effects of AI on human decision-makers?

→ Are predictive AI and human experience complements or substitutes?

2. Interaction between AI and incentives

Looking ahead

A few key questions

1. Heterogeneous effects of AI on human decision-makers?

→ Are predictive AI and human experience complements or substitutes?

2. Interaction between AI and incentives

3. Explaining different effects in different contexts

Looking ahead

A few key questions

1. Heterogeneous effects of AI on human decision-makers?

→ Are predictive AI and human experience complements or substitutes?

2. Interaction between AI and incentives

3. Explaining different effects in different contexts

→ The role of private information?

Conclusion

We study the role of AI in repair decisions made by human technicians.

- Use **observational data** to quantify economic value of AI assistance.
 - Separately identify preferences and beliefs.
 - Account for dynamics.
- With AI, expenditures reduced by **\$240-480/truck/year (85% of all feasible cost savings)**.

Conclusion

We study the role of AI in repair decisions made by human technicians.

- Use **observational data** to quantify economic value of AI assistance.
 - Separately identify preferences and beliefs.
 - Account for dynamics.
- With AI, expenditures reduced by **\$240-480/truck/year (85% of all feasible cost savings)**.

More broadly,

- AI prediction is stellar, but few settings where AI makes economic decisions alone.
- As long as humans remain in the loop, **understanding how they interact with AI is critical**.

Thank you!

adamharris@cornell.edu

www.adamharris.phd

The technician's problem

Objective: cost minimization. Costs of repairs/breakdowns include:

- Tangible costs: Labor, materials, towing, etc.
- Intangible costs:
 - Opportunity cost of truck not being on the road.
 - Capacity constraints (shadow costs).
 - Disruption costs of breakdowns:
 - » Damage to relationships with drivers.
 - » Damage to relationships with customers.

State of the truck

What data do technicians see?

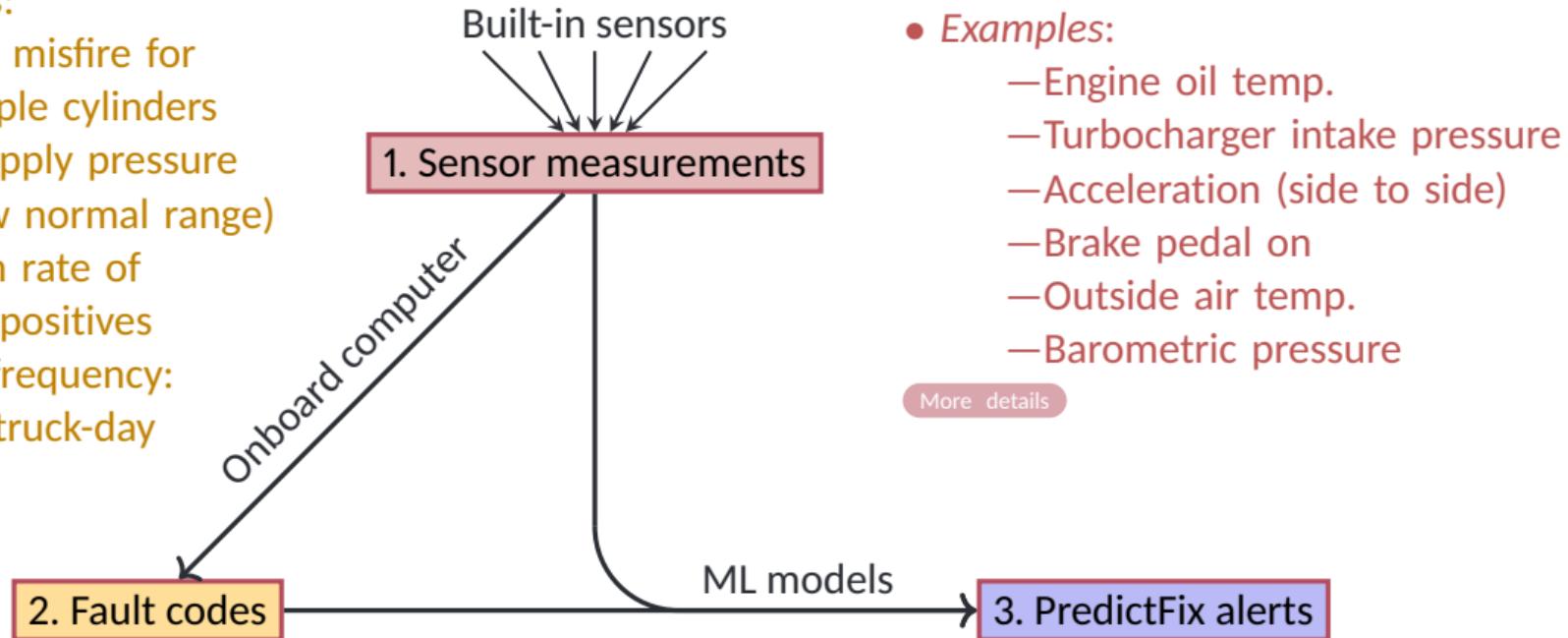
- Examples:

- Engine misfire for multiple cylinders
- Gas supply pressure (Below normal range)

- Very high rate of false positives

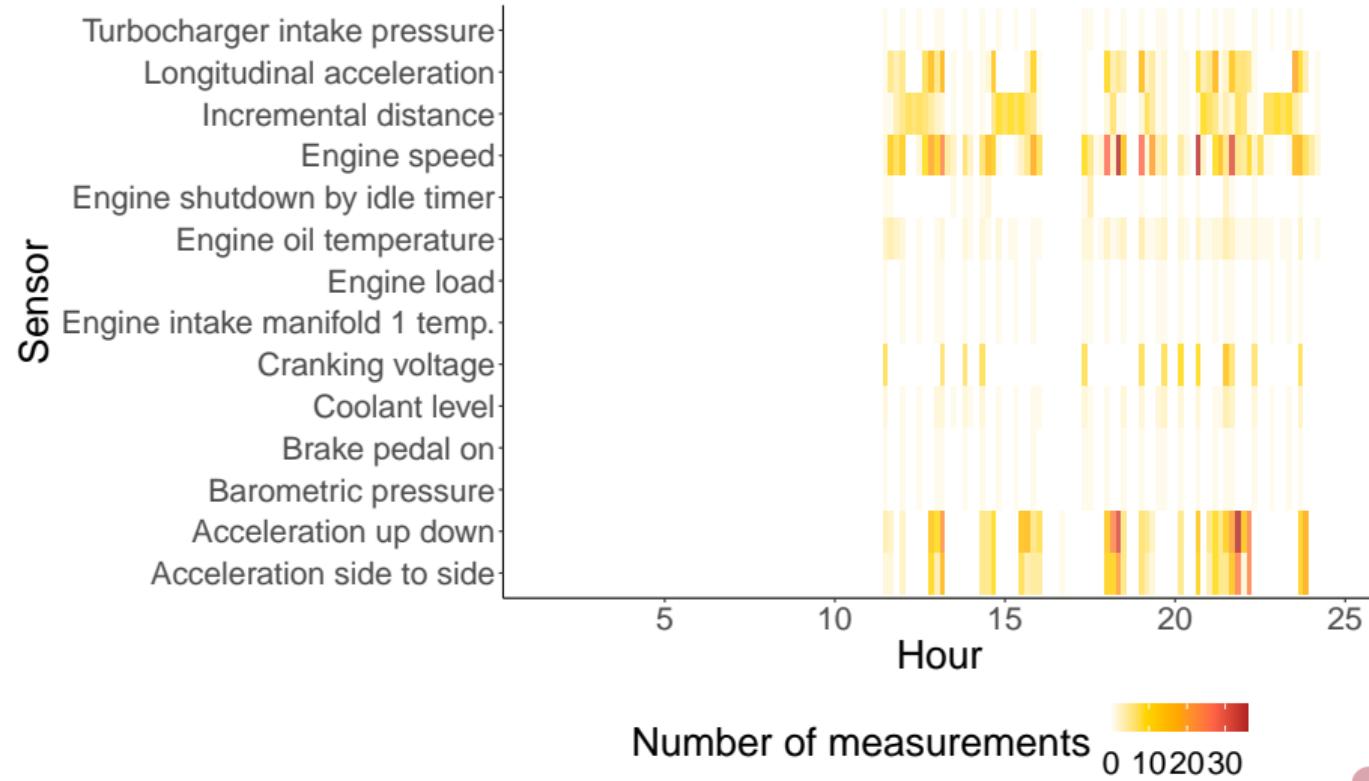
- Median frequency:
~ 10 per truck-day

[More details](#)



What do you see?

Sensor measurement frequency



State of the truck

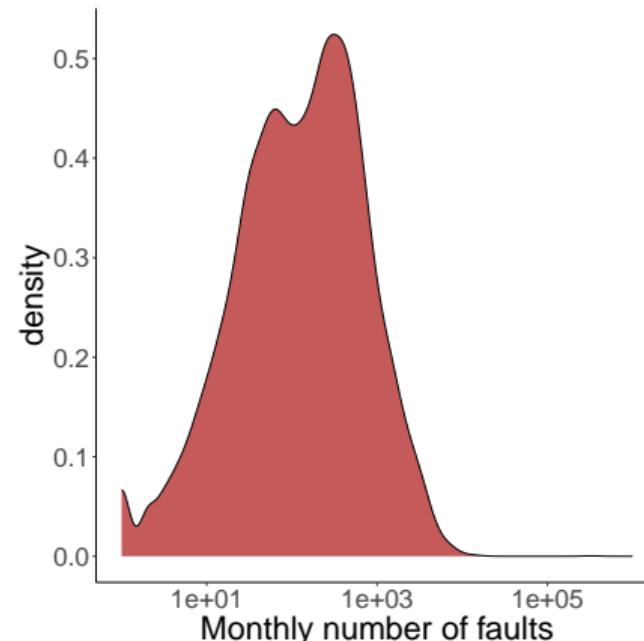
◀ Back

What data do technicians see?

Fault codes. Examples: “Engine misfire for multiple cylinders”,

“Gas supply pressure—Data valid but below normal operational range”.

Is **fault code → repair** the optimal policy? **No; very high rate of false positives.**



Combining PredictFix with other information from x

If PredictFix is a good predictor of breakdown risk, why might technicians still want to look at x ?

1. Combining x and PredictFix to form an optimal classifier.
 - PredictFix is not *on* the ROC curve, so it's not an *optimal* binary classifier.
 - Combine PredictFix alerts with x to form a binary classifier that is optimal.
2. Variation in cost threshold. Suppose PredictFix were an optimal binary classifier.

$$\text{PredictFix}_i \Leftrightarrow \pi(x_i) \geq \pi^*$$

$$a_i = 1 \Leftrightarrow \pi(x_i) > \frac{\text{Cost of repair}}{\text{Cost of breakdown}} \equiv \tau(v)$$

Table: Optimal decisions given optimal binary signal

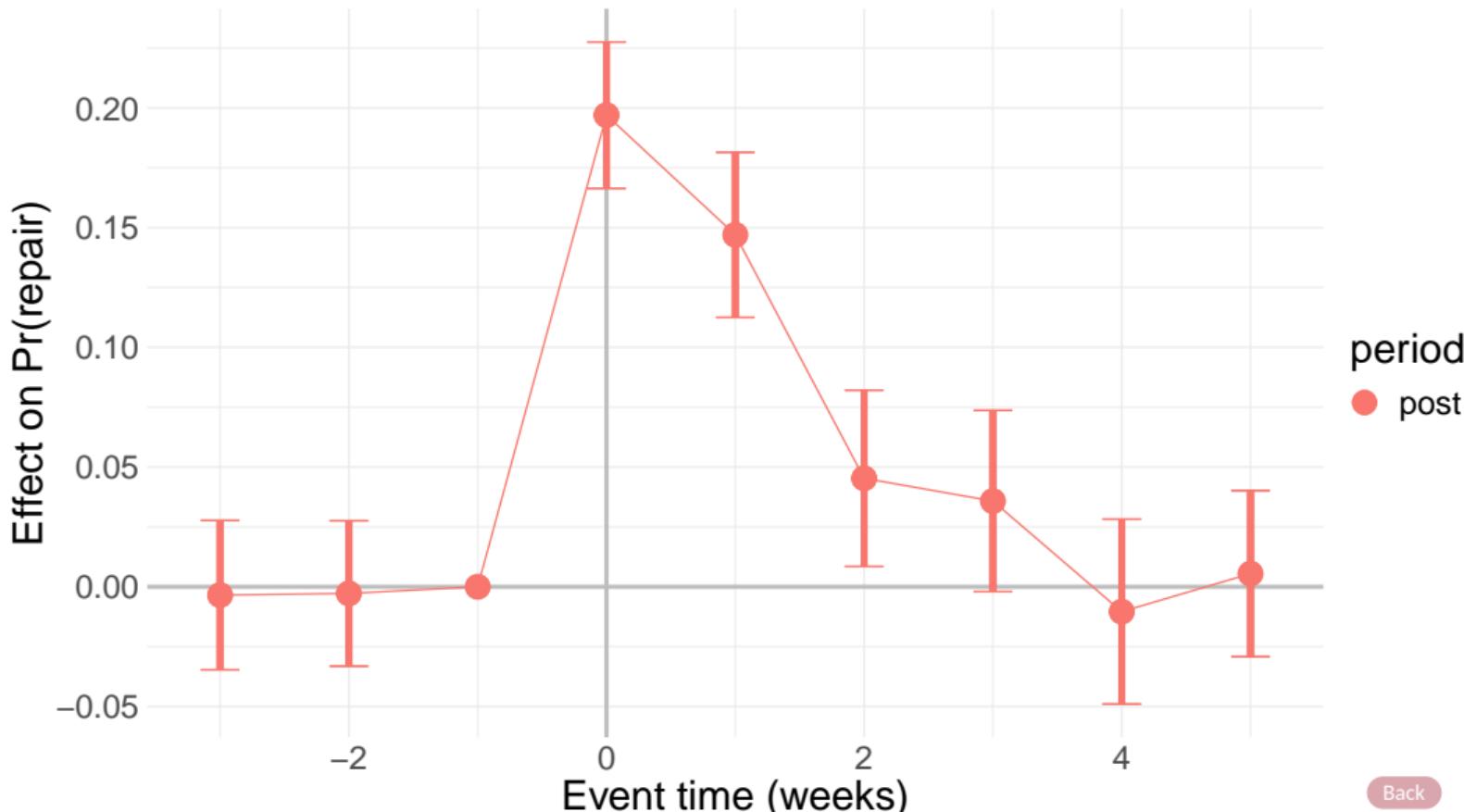
	$\tau(v) \leq \pi^*$	$\tau(v) > \pi^*$
PredictFix alert	Repair	?
No PredictFix alert	?	No repair

Fact 3: Technicians respond to PredictFix, but also ignore many alerts.

$$\text{Repair}_{i,t} = \alpha_0 + \sum_{k \in \{\text{pre,post}\}} \sum_{\tau=-3}^5 \beta_{\tau}^k \mathbb{1}\{t \in \mathcal{T}_k\} \widehat{\text{PredictFix}}_{i,t-\tau} + \alpha_i + \gamma_t + \epsilon_{i,t}$$

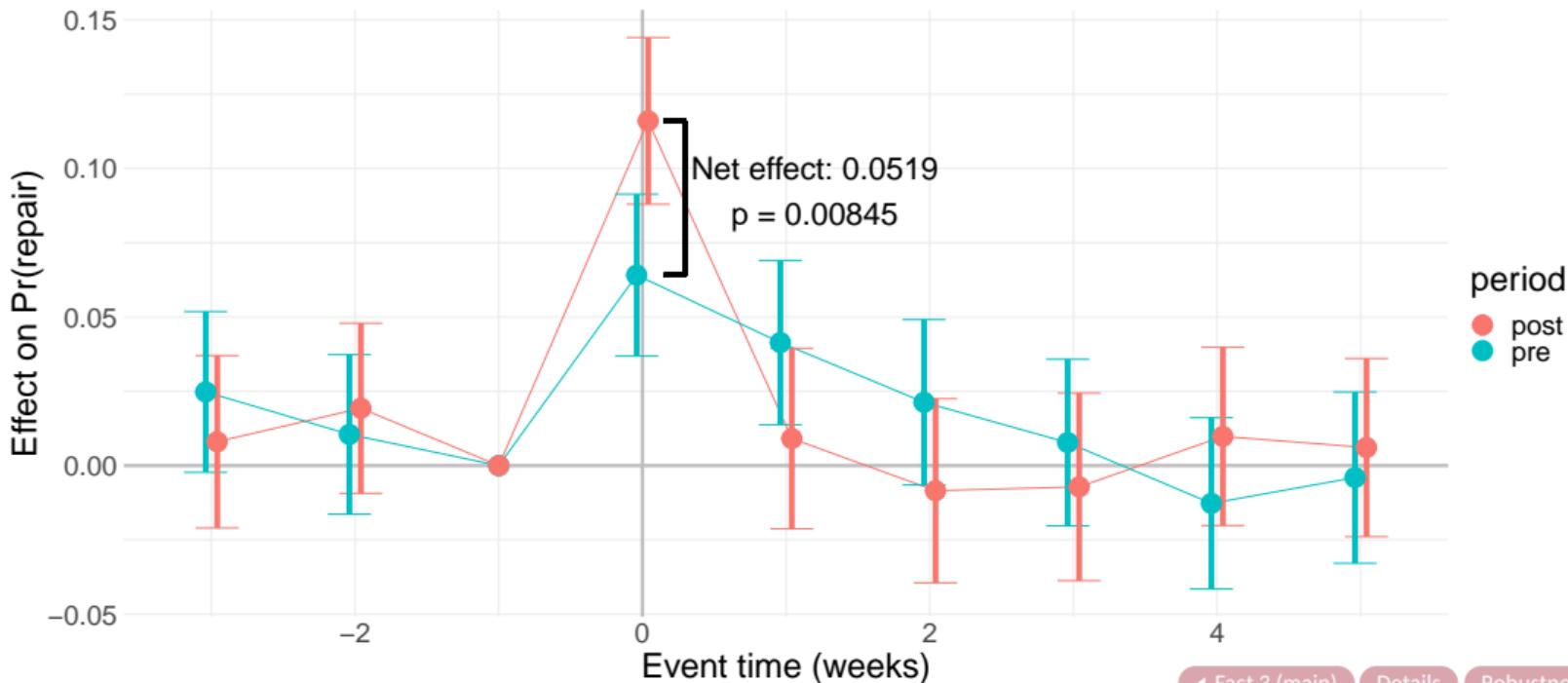
- What PredictFix alerts *would have happened* in pre period?
- Use ML to learn mapping: state → PredictFix alerts.
 - Very high degree of accuracy (AUC = 0.978). [Details](#) [Stability](#)

Robustness: Actual, rather than predicted, PredictFix alerts



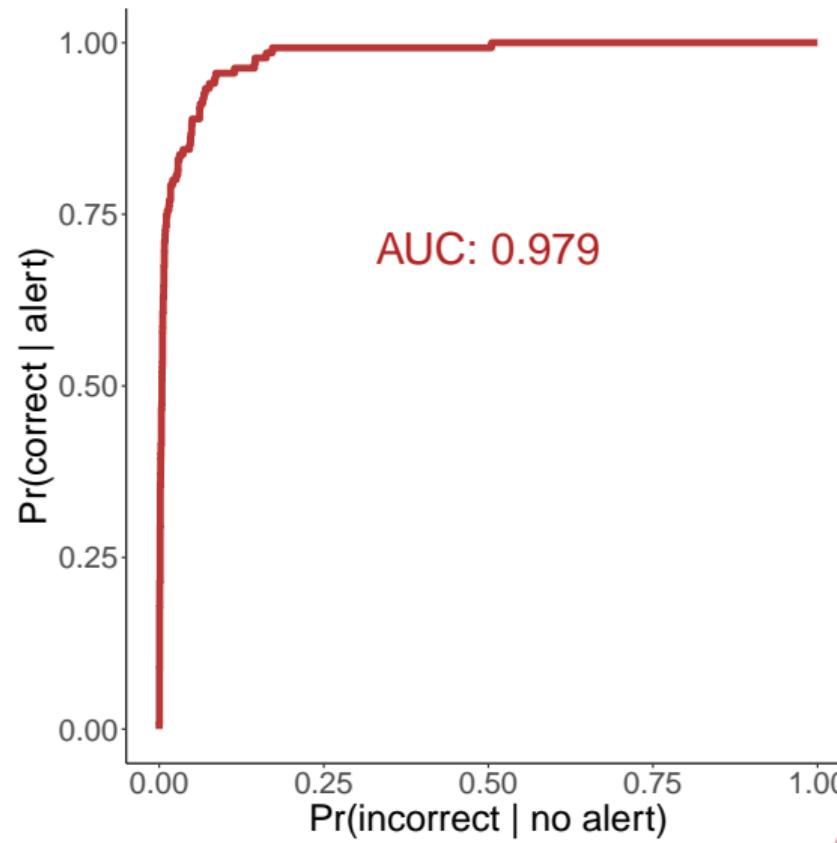
Fact 3 (Medium-Priority Alerts)

$$\text{Repair}_{i,t} = \alpha_0 + \sum_{k \in \{\text{pre,post}\}} \sum_{\tau=-3}^5 \beta_{\tau}^k \mathbb{1}\{t \in \mathcal{T}_k\} \widehat{\text{PredictFix}}_{i,t-\tau} + \alpha_i + \gamma_t + \epsilon_{i,t}$$



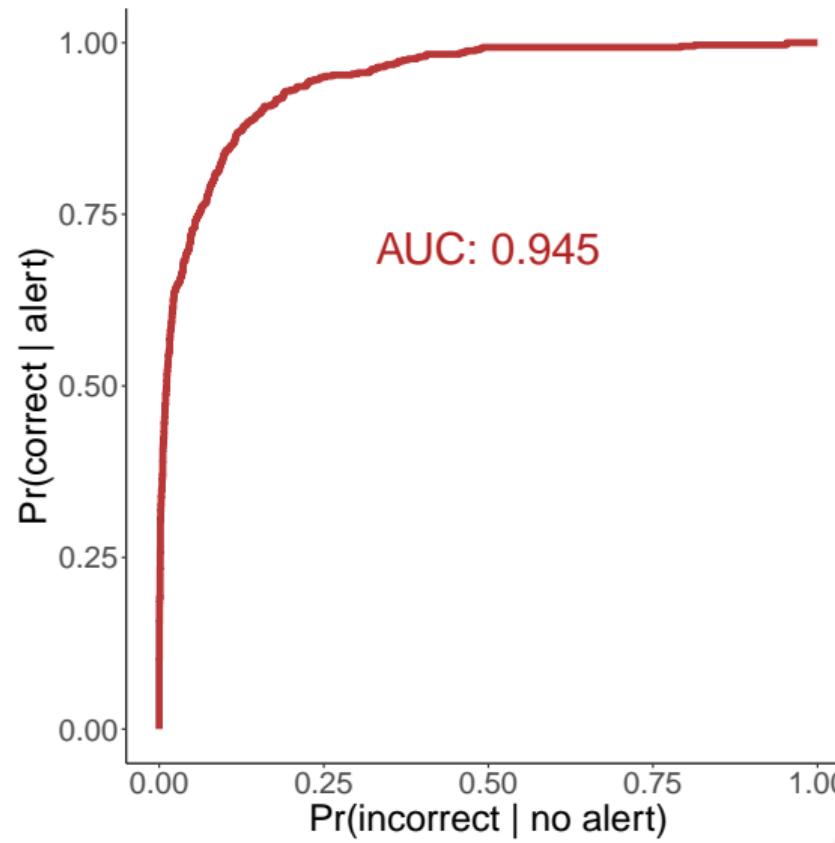
Predicting PredictFix

ROC curves

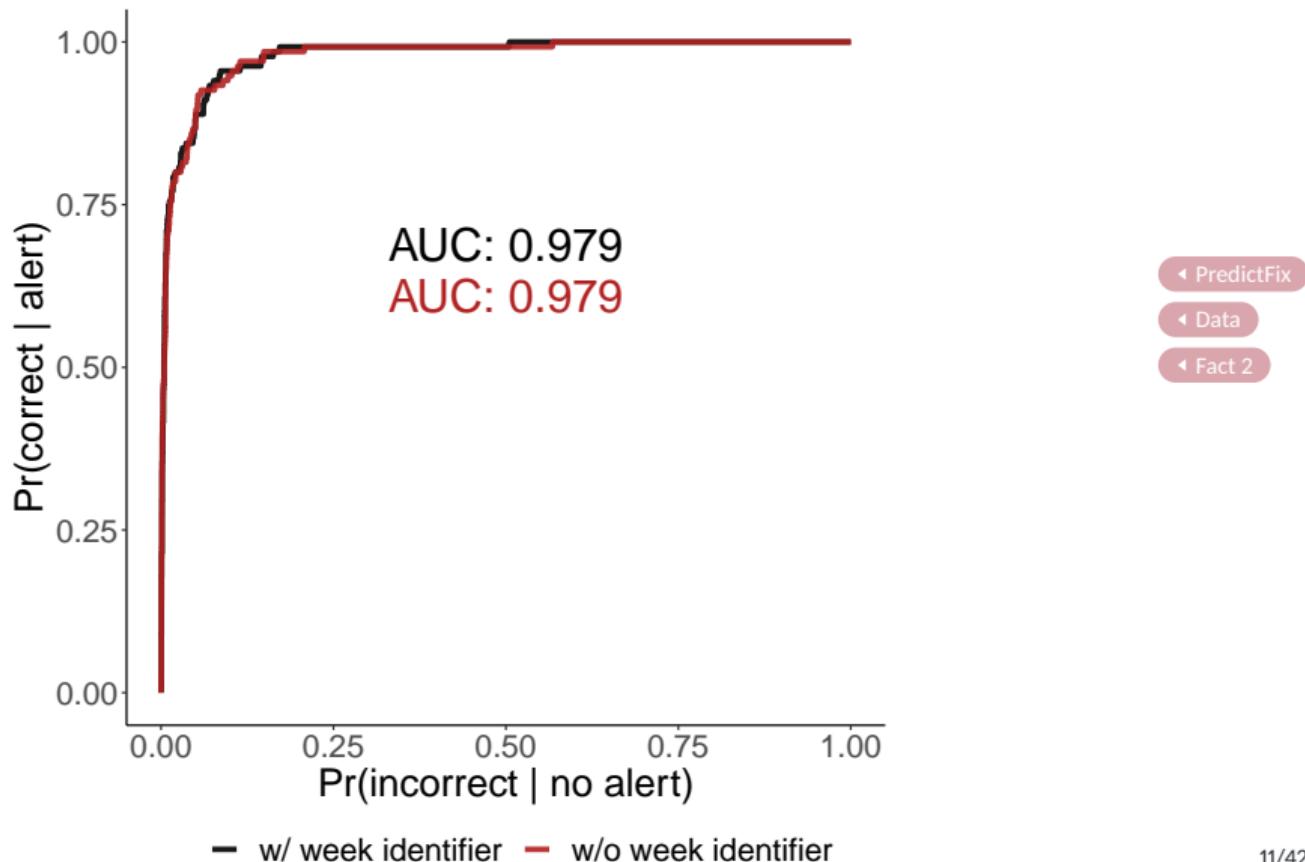


Predicting PredictFix

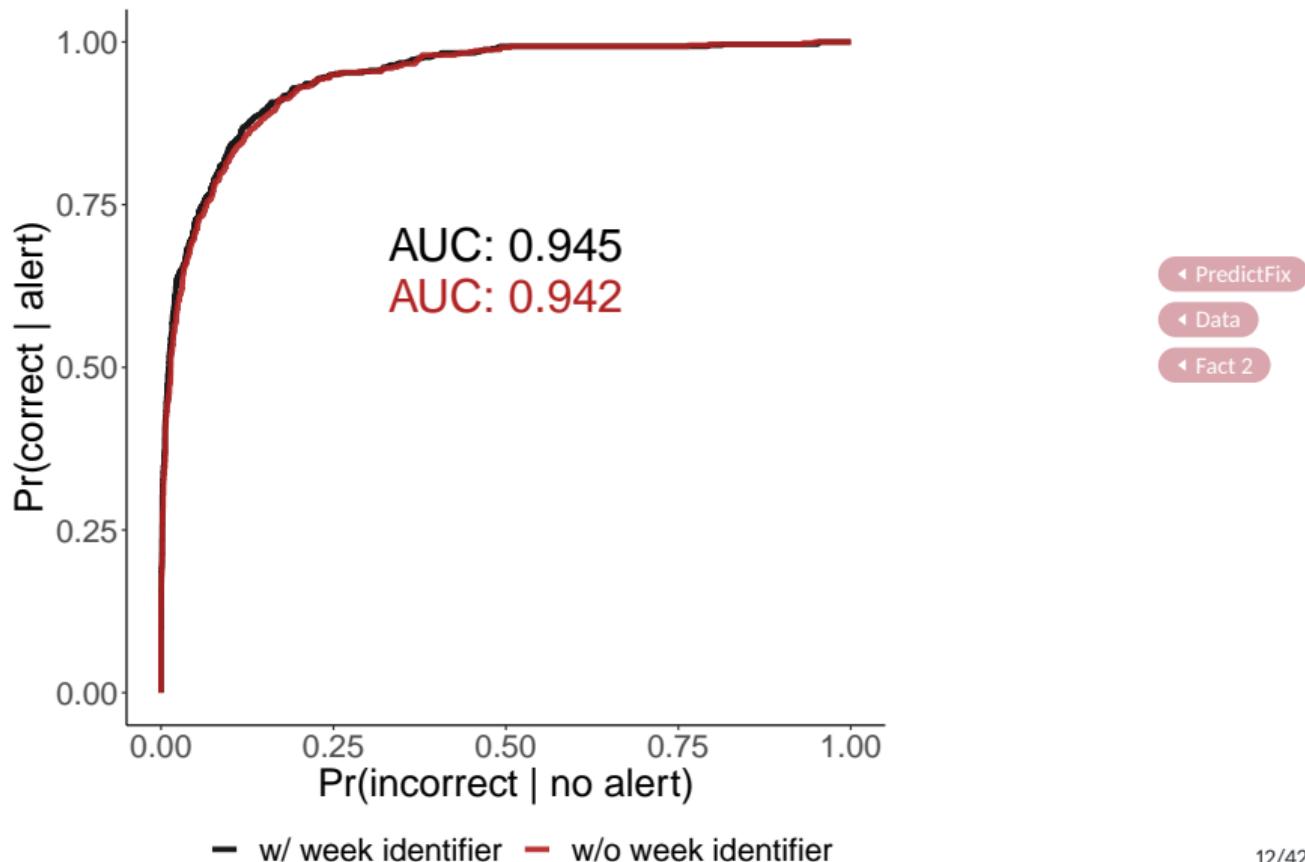
ROC curves



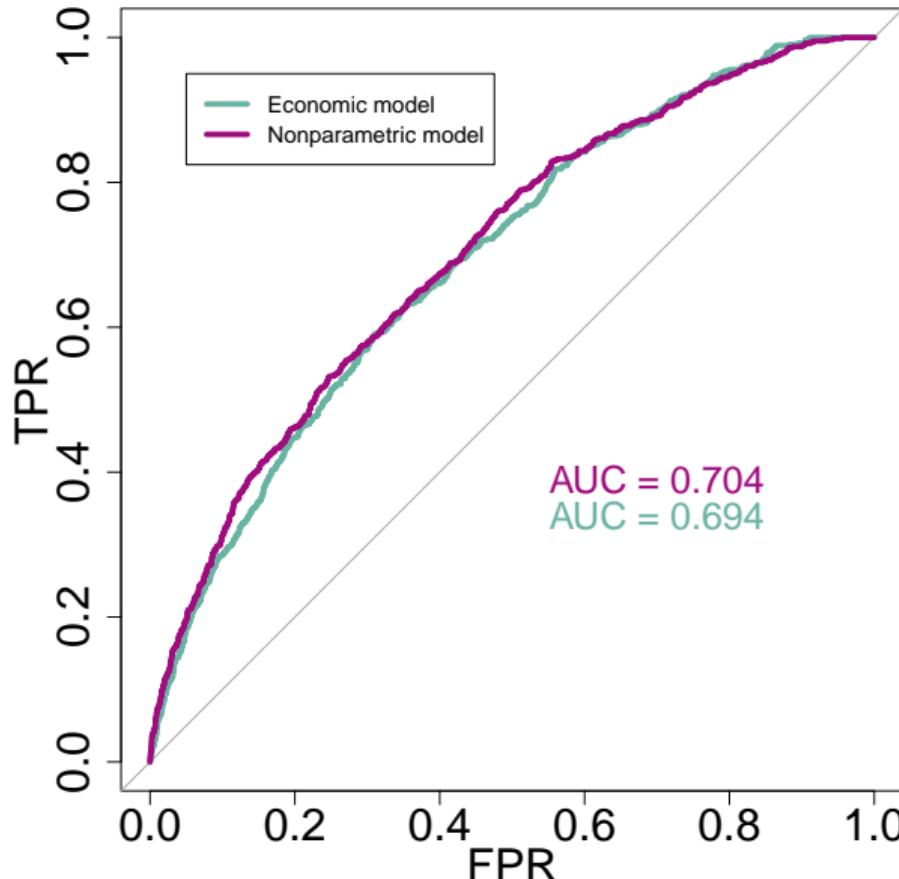
PredictFix is stable over the post period



PredictFix is stable over the post period



Model fit



Dynamics

Technician's problem is inherently dynamic. This week's action → future weeks' states.

Technician's beliefs about the future

- Potentially incorrect beliefs about future breakdown risk:

$$\rho(x_{t+1}) \stackrel{?}{=} \pi(x_{t+1})$$

- But technician knows the correct distribution of $t + 1$'s state given t 's state and action.

Toward estimation: The CCP approach

Back

Standard notation: The inclusive payoff from choosing a is

$$v_a(w, x) = u(a, w, x) + \delta EV_a(w, x)$$

where

$$EV_a(w, x) = \mathbb{E} \left[\max \left\{ v_0(w_{t+1}, x_{t+1}), v_1(w_{t+1}, x_{t+1}) + \epsilon_{t+1} \right\} \mid w_t = w, x_t = x, a_t = a \right]$$

Dynamic choice probability:

$$\begin{aligned} p(w, x) &= \Lambda(\theta [v_1(w, x) - v_0(w, x)]) \\ &= \Lambda(\theta [-g(w) + \rho(x) \\ &\quad + \delta (EV_1(w, x) - EV_0(w, x))]) \end{aligned}$$

where Λ is the Logistic function.

Question: How to bring this to the data?

1. Nested fixed-point approach: Rust (1987).
→ For each parameter set, solve for EV_0, EV_1 .
2. CCP approaches: Hotz and Miller (1993); Arcidiacono and Miller (2011).
→ $EV_0, EV_1 = f(\text{choice probabilities})$.

Toward estimation: The CCP approach

Back

Under this assumption

$$EV_1(w_t, x_t) - EV_0(w_t, x_t) = \Delta Eg(w_t, x_t) + \frac{1}{\theta} \Delta E \log p(w_t, x_t)$$

where

$$\Delta Eg(w_t, x_t) = \mathbb{E}[g(w_{t+1}) \mid w_t, x_t, a_t = 0] - \mathbb{E}[g(w_{t+1}) \mid w_t, x_t, a_t = 1]$$

$$\begin{aligned}\Delta E \log p(w_t, x_t) &= \mathbb{E}[\log(p(w_{t+1}, x_{t+1})) \mid w_t, x_t, a_t = 0] \\ &\quad - \mathbb{E}[\log(p(w_{t+1}, x_{t+1})) \mid w_t, x_t, a_t = 1]\end{aligned}$$

Choice probabilities:

$$p(w, x) = \Lambda(-\theta g(w) + \theta \rho(x) + \delta [\theta \Delta Eg(w, x) + \Delta E \log p(w, x)])$$

$$p(w, x) = \Lambda(-\theta g(w) + \theta p(x) + \delta [\theta \Delta E g(w, x) + \Delta E \log p(w, x)])$$

How to take this to the data?

- We need not consider $\Delta E g(w_t, x_t) = \mathbb{E}[g(w_{t+1}) \mid w_t, x_t, a_t = 0] - \mathbb{E}[g(w_{t+1}) \mid w_t, x_t, a_t = 1]$.
→ w evolves exogenously.

$$p(w, x) = \Lambda(-\theta g(w) + \theta p(x) + \delta [\theta \Delta E g(w, x) + \Delta E \log p(w, x)])$$

How to take this to the data?

- We need not consider $\Delta E g(w_t, x_t) = \mathbb{E}[g(w_{t+1}) \mid w_t, x_t, a_t = 0] - \mathbb{E}[g(w_{t+1}) \mid w_t, x_t, a_t = 1]$.
→ w evolves exogenously.
- $\Delta E \log p(w_t, x_t) = \mathbb{E}[\log(p(w_{t+1}, x_{t+1})) \mid w_t, x_t, a_t = 0] - \mathbb{E}[\log(p(w_{t+1}, x_{t+1})) \mid w_t, x_t, a_t = 1]$.
A typical approach:
 1. Nonparametrically estimate CCP function $p(w, x)$.
 2. Nonparametrically estimate transition process: Dist. of (w_{t+1}, x_{t+1}) conditional on (a_t, w_t, x_t) .
 3. Integrate $\log p(w, x)$ over the transition process.

$$p(w, x) = \Lambda(-\theta g(w) + \theta p(x) + \delta [\theta \Delta E g(w, x) + \Delta E \log p(w, x)])$$

How to take this to the data?

- We need not consider $\Delta E g(w_t, x_t) = \mathbb{E}[g(w_{t+1}) \mid w_t, x_t, a_t = 0] - \mathbb{E}[g(w_{t+1}) \mid w_t, x_t, a_t = 1]$.
→ w evolves exogenously.
- $\Delta E \log p(w_t, x_t) = \mathbb{E}[\log(p(w_{t+1}, x_{t+1})) \mid w_t, x_t, a_t = 0] - \mathbb{E}[\log(p(w_{t+1}, x_{t+1})) \mid w_t, x_t, a_t = 1]$.
A typical approach:
 1. Nonparametrically estimate CCP function $p(w, x)$.
 2. Nonparametrically estimate transition process: Dist. of (w_{t+1}, x_{t+1}) conditional on (a_t, w_t, x_t) .
 3. Integrate $\log p(w, x)$ over the transition process.
- When (w, x) is high-dimensional, 2. is clearly infeasible.

$$p(w, x) = \Lambda(-\theta g(w) + \theta p(x) + \delta [\theta \Delta E g(w, x) + \Delta E \log p(w, x)])$$

How to take this to the data?

- We need not consider $\Delta E g(w_t, x_t) = \mathbb{E}[g(w_{t+1}) | w_t, x_t, a_t = 0] - \mathbb{E}[g(w_{t+1}) | w_t, x_t, a_t = 1]$.
→ w evolves exogenously.
- $\Delta E \log p(w_t, x_t) = \mathbb{E}[\log(p(w_{t+1}, x_{t+1})) | w_t, x_t, a_t = 0] - \mathbb{E}[\log(p(w_{t+1}, x_{t+1})) | w_t, x_t, a_t = 1]$.
A typical approach:
 1. Nonparametrically estimate CCP function $p(w, x)$.
 2. Nonparametrically estimate transition process: Dist. of (w_{t+1}, x_{t+1}) conditional on (a_t, w_t, x_t) .
 3. Integrate $\log p(w, x)$ over the transition process.
- When (w, x) is high-dimensional, 2. is clearly infeasible. Need either
 - very strong functional form assumptions on transition; or
 - a new way to compute the integral $\Delta E \log p$ without estimating the transition process.

Taking stock of assumptions

Most substantive assumptions:

1. No private information.
2. Two known moments: $f_1(\vec{p}(X))$, $f_2(\vec{p}(X))$.
5. $p_{t+1}(w_{t+1}, x_{t+1}) \mid a_t, w_t, x_t \sim \text{Beta}(\mu(a_t, w_t, x_t), \nu)$

What we have *not* assumed:

- Technician's objective = the firm's objective.
 - Agency problems.
 - Technicians misunderstand firm's costs.
- Technician's risk preferences.
 - These don't affect ρ , they only affect the interpretation of τ .

(Starting in March 2020) **PredictFix alerts.**

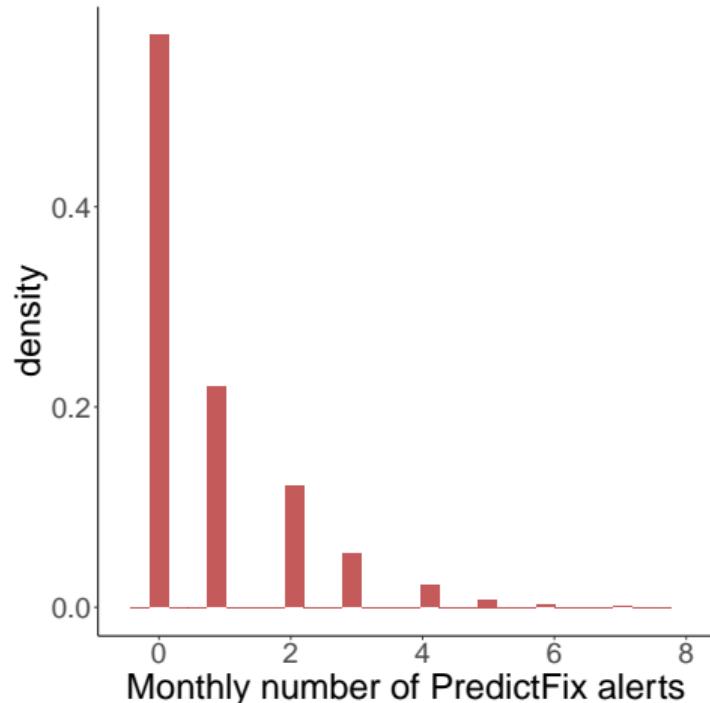
The PredictFix algorithm.

- PFC purchased PredictFix from a tech firm specializing in industrial ML prediction.
- Different models for different components.

Examples:

- “Cylinder issue” → High severity
- “Coolant leak” → Medium severity
- “Engine knocking” → Medium severity
- Each model has a PFC-assigned *severity level* (medium or high).
- Note: **Does not** provide new information.

Figure: Number of PredictFix alerts per truck-month



Fact 4: No change in aggregate outcomes.

Are positive effects of PredictFix evident in aggregate outcomes? **No.**

Table: Frequency of repairs and breakdowns

	Pre	Post
Engine Repairs	462 (10.5%)	1,778 (10.5%)
Engine Breakdowns	71 (1.43%)	297 (1.73%)

Fact 5: Repairs costs are higher, more variable in the post period.

- We observe (accounting estimates of) tangible repair costs.
- Comparison across periods:

	Pre	Post
Mean	\$621	\$721
Std. Dev	\$1118	\$1404

Constant breakdown cost

- Reflects thinking of PFC fleet management team.
- We need not know B .
- **Not** assuming $B_{\text{pre}} = B_{\text{post}}$.

Finite dependence

Assumption 3

For every component z of the state variable (w, x) , the transition process is such that z satisfies either of the following conditions:

- (i) z resets after a repair, i.e., the conditional distribution of $z_{t+1} | a_t = 1, w_t, x_t$ does not depend on (w_t, x_t) ; or
- (ii) z evolves independently and exogenously, i.e., the conditional distribution of $z_{t+1} | a_t, w_t, x_t$ only depends on z_t .

x satisfying (i):

- Sensor measurements
- Fault codes

x satisfying (ii):

- Weather-related variables
- Odometer readings

w all satisfy (ii):

- Facility's # of trucks.
- Facility's # of open work orders.
- Month FEs.

Restriction on private information

We observe: performance of truck components, truck use, and environmental conditions.

→ Leaves little scope for private information about breakdown risk.

Assumption 1 (No private information on state of truck)

Suppose ξ is observed by the technician but not by the econometrician. Then,

$$\Pr(\text{breakdown} | x, \xi) = \Pr(\text{breakdown} | x) = \pi(x)$$

where x is fully observed by the econometrician. Moreover, the technician's perceived risk of breakdown does not depend on ξ .

What this assumption buys us:

— Progress toward identification of ρ and τ .

— Identification of $\pi(x)$.

→ No selective labels problems.

→ Our ML predictor of breakdown risk has the interpretation of $\pi(x)$.

◀ Fact 1

◀ Fact 1 (ROC)

◀ Econometric Challenge

24/42

AUC

Probability interpretation

- Let \mathcal{I}_0 and \mathcal{I}_1 represent the set of non-breakdown and breakdown observations, respectively.
- Randomly draw i_0 from \mathcal{I}_0 and i_1 from \mathcal{I}_1 .
- Then $\text{AUC} = \Pr(\hat{\pi}(x_{i_1}) > \hat{\pi}(x_{i_0}))$.

1. Identifying $\theta\alpha$:

$$\frac{d}{dw_2} \log \frac{p(x, w)}{1 - p(x, w)} = -\theta\alpha \frac{d}{dw_2} g_2(w_1, w_2, x)$$

2. Partially identifying g_1 : WLOG, $g_1(w_1) = \gamma_0 + \tilde{g}_1(w_1)$ where $\tilde{g}(w_1^0) = 0$ for some w_1^0 . Then

$$\frac{d}{dw_1} \log \frac{p(x, w)}{1 - p(x, w)} = -\theta \nabla \tilde{g}_1(w_1)$$

3. Partially identifying ρ :

$$\log \frac{p(w, x)}{1 - p(w, x)} + \theta \tilde{g}_1(w_1) + \theta\alpha g_2(w_1, w_2, x) = -\theta\gamma_0 + \theta\rho(x) \equiv \tilde{\rho}(x)$$

Recall: $\tilde{\rho}(x) \equiv -\theta\gamma_0 + \theta\rho(x)$ is identified.

4. **Point identification of g_1, ρ :** Note that the mean and minimum of $\tilde{\rho}(x)$ can be written as

$$\begin{aligned}\mathbb{E}_x \tilde{\rho}(x) &= -\theta\gamma_0 + \theta\mathbb{E}_x \rho(x) \\ \min_x \tilde{\rho}(x) &= -\theta\gamma_0 + \theta \min_x \rho(x)\end{aligned}$$

Writing this in matrix form,

$$\underbrace{\begin{bmatrix} -1 & \mathbb{E}_x \rho(x) \\ -1 & \min_x \rho(x) \end{bmatrix}}_A \begin{pmatrix} \theta\gamma_0 \\ \theta \end{pmatrix} = \begin{pmatrix} \mathbb{E}_x \tilde{\rho}(x) \\ \min_x \tilde{\rho}(x) \end{pmatrix}$$

DDC estimation with high-dimensional state

Assumption 5

The transition process for state variables (w, x) and the technician's conditional choice probability function $p(\cdot, \cdot)$ are such that

$$p(w_{t+1}, x_{t+1}) \mid a_t, w_t, x_t \sim \text{Beta}(\mu(a_t, w_t, x_t), \nu)$$

where $\mu : \{0, 1\} \times \mathcal{W} \times \mathcal{X} \rightarrow [0, 1]$ and $\nu \in \mathbb{R}^+$.

Implies that

$$\Delta E \log p(w_t, x_t) = \psi(\mu(0, w_t, x_t)\nu) - \psi(\mu(1, w_t, x_t)\nu)$$

where ψ is the digamma function, and

$$\mu(a_t, w_t, x_t) = \Pr(a_{t+1} = 1 \mid a_t, w_t, x_t)$$

Minimally restrictive:

- Beta is a flexible distribution.
- μ is function of (a_t, w_t, x_t) .
- ν to be estimated.

→ Easy to estimate (offline using GBDT).

- We use the “mean-precision” parameterization of the Beta distribution.
- First parameter can be arbitrary function of state and action.

$$p(w_{t+1}, x_{t+1}) \sim \text{Beta}(\mu(a_t, w_t, x_t), \nu)$$

Dynamics: Toward estimation

$$p(w_t, x_t) = \Lambda(\theta [-g(w_t) + \rho(x_t) + \delta (EV_1(w_t, x_t) - EV_0(w_t, x_t))])$$

Under the finite dependence assumption,

$$EV_1(w_t, x_t) - EV_0(w_t, x_t) = \frac{1}{\theta} \left(\mathbb{E} [\log p(w_{t+1}, x_{t+1}) \mid a_t = 0, w_t, x_t] - \mathbb{E} [\log p(w_{t+1}, x_{t+1}) \mid a_t = 1, w_t, x_t] \right)$$

Challenge: State (w, x) is high-dimensional.

$$\mathbb{E} [\log p(w_{t+1}, x_{t+1}) \mid a_t, w_t, x_t] = \int \log p(w', x') dF(x', w' \mid a_t, w_t, x_t)$$

Trick: If $p(w_{t+1}, x_{t+1}) \sim \text{Beta}(\mu(a_t, w_t, x_t), \nu)$, then

$$\mathbb{E} [\log p(w_{t+1}, x_{t+1}) \mid a_t, w_t, x_t] = \psi(\mu(a_t, w_t, x_t)\nu) - \psi(\nu)$$

Technical Condition

There exists some state (w, x) such that

$$\text{sign} \left(\frac{d^2}{dw^j dx^k} f(w, x) \right) \neq \text{sign} \left(\frac{d}{dw^j} f(w, x) \frac{d}{dx^k} f(w, x) \right)$$

where

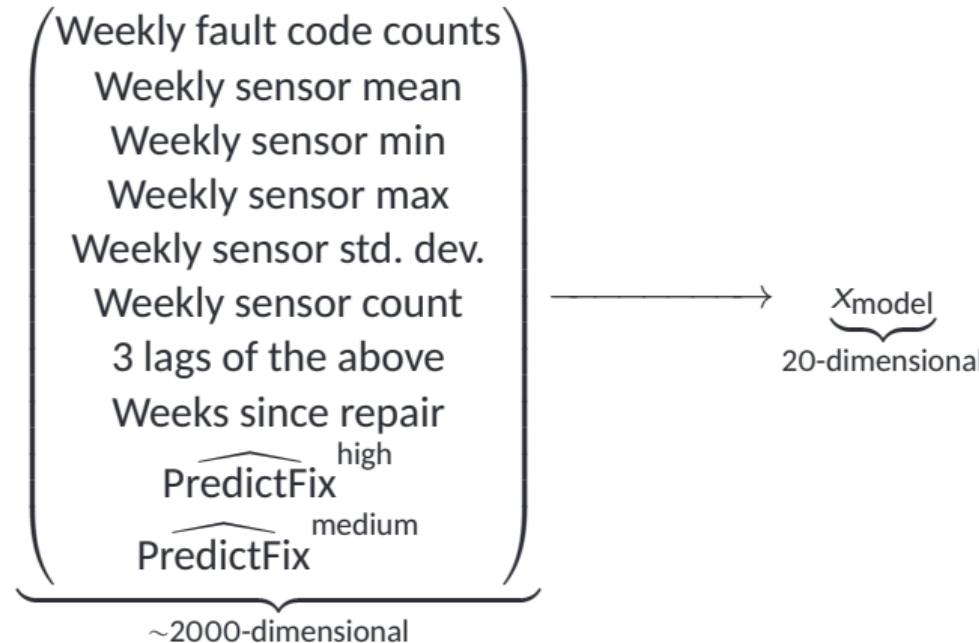
$$f(w_t, x_t) = \mathbb{E} [\log p(w_{t+1}, x_{t+1}) | w_t, x_t, a_t = 0],$$

and x^k is a resetting variable and w^j is an exogenously and independently-evolving variable.

Note: Need only hold for one state $(w, x) \in \mathcal{W} \times \mathcal{X}$.

Pre-estimation step

Variable selection



Method: Train GBDT to predict a_{it} conditional on (w_{it}, x_{it}) .

→ Take the $20 \times$ variables with the highest “gain.”

Back

Preferences and agency issues

If the technician has **risk-neutral preferences**,

$c(v) \rightarrow$ repair cost to technician

$B \rightarrow$ breakdown cost to technician

If there are **no agency issues**,

$c(v) \rightarrow$ cost of repair to firm

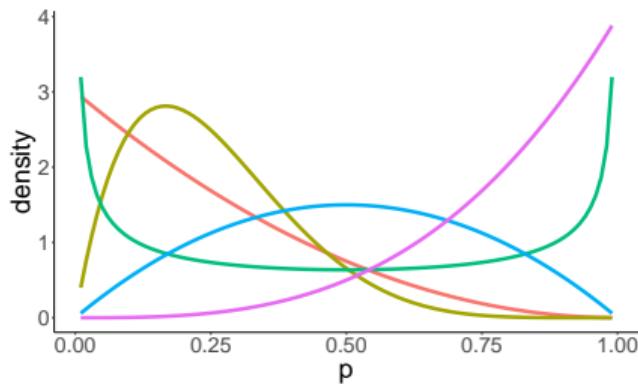
$B \rightarrow$ cost of breakdown to firm

Yet identification/estimation **do not** require these assumptions.

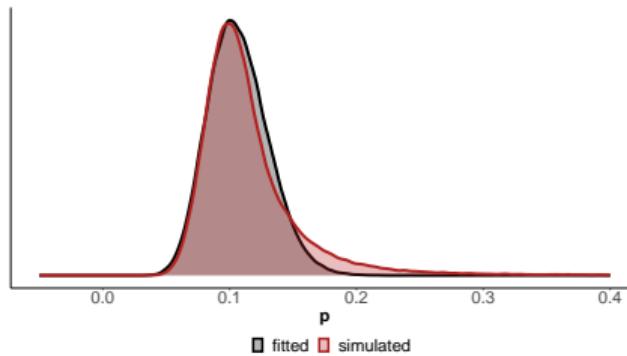
Beta distribution

[Back](#)

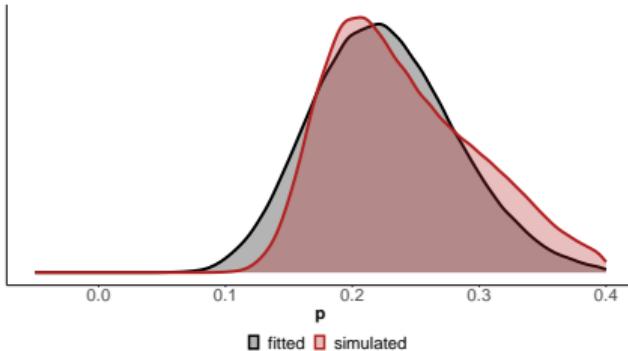
(a) Examples: Beta for various parameters



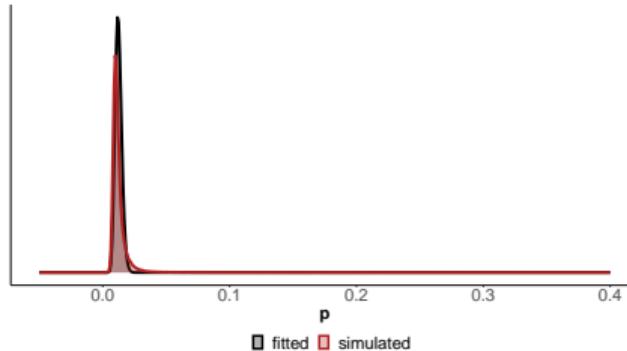
(b) Example: $a_{it} = 0, g_{it} = 0.1, \pi_{it} = 0.014$



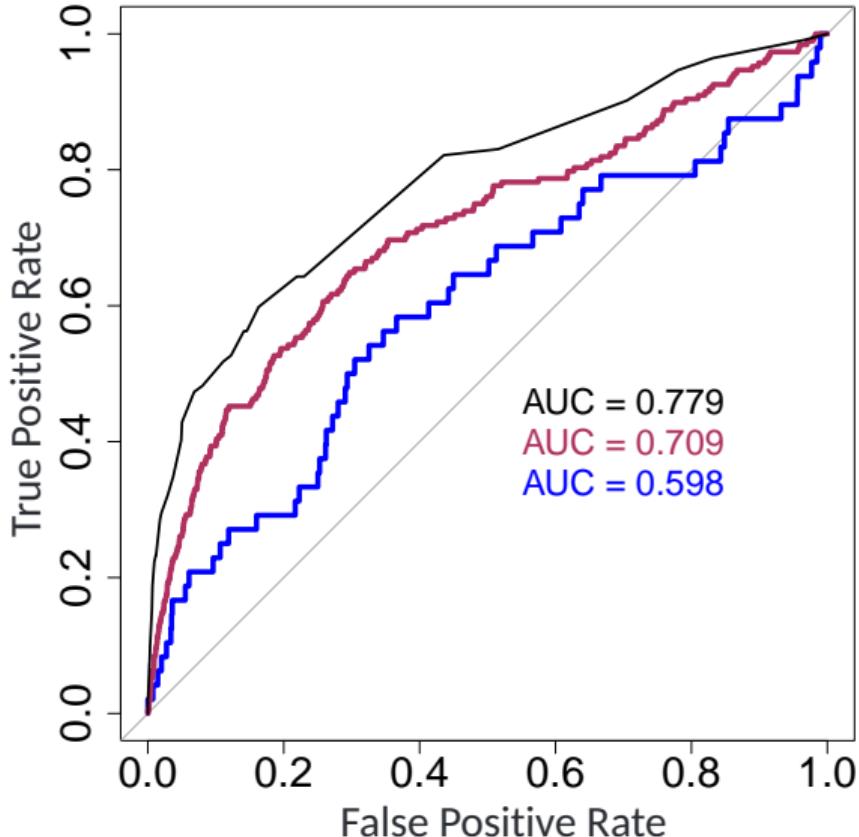
(c) Example: $a_{it} = 0, g_{it} = 0.04, \pi_{it} = 0.05$



(d) Example: $a_{it} = 0, g_{it} = 0.24, \pi_{it} = 0.004$



The effect of PredictFix on technicians' predictions (ρ)



Predictor

$\hat{\rho}_{\text{pre}}$

Beliefs w/o PredictFix

$\hat{\rho}_{\text{post}}$

Beliefs w/ PredictFix

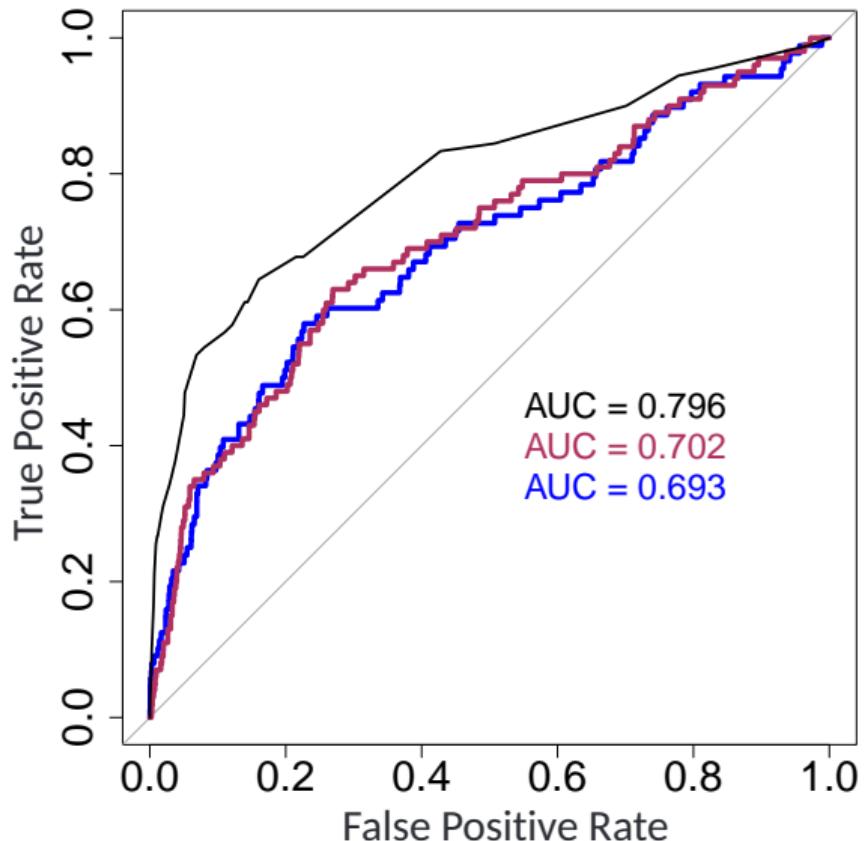
$\hat{\pi}_{\text{ML}}$

Benchmark

Note:

- $\hat{\rho}_{\text{post}}$ strictly dominates $\hat{\rho}_{\text{pre}}$.
- **Robust** to relaxations of “correct on average” and “zero lower bound” assumptions.

Differences in ρ within the post period



Predictor	Beliefs
$\hat{\rho}_{\text{post}} \text{ (first half)}$	
$\hat{\rho}_{\text{post}} \text{ (second half)}$	Benchmark

[Back to ROC curves](#)

What do we need to evaluate these counterfactuals?

- Need to know the transition process.
- Express it in terms of lower-dimensional objects:

$$g_{it+1}, \rho_{it+1}, \pi_{it+1} \mid g_{it}, \rho_{it}, \pi_{it}$$

[Details](#)

rather than

$$w_{it+1}, x_{it+1} \mid w_{it}, x_{it}$$

- Flexibly estimated (using Gaussian Mixture Regression.)
- To get dollar interpretation, set value for B . A reasonable range: \$5,000-\$10,000.

What do we need to evaluate these counterfactuals?

- Need to know the transition process.
- Express it in terms of lower-dimensional objects:

$$g_{it+1}, \rho_{it+1}, \pi_{it+1} \mid g_{it}, \rho_{it}, \pi_{it}$$
Details

rather than

$$w_{it+1}, x_{it+1} \mid w_{it}, x_{it}$$

- Flexibly estimated (using Gaussian Mixture Regression.)
- To get dollar interpretation, set value for B . A reasonable range: \$5,000-\$10,000.

What do we need to evaluate these counterfactuals?

- Need to know the transition process.
- Express it in terms of lower-dimensional objects:

$$g_{it+1}, \rho_{it+1}, \pi_{it+1} \mid g_{it}, \rho_{it}, \pi_{it}$$

[Details](#)

rather than

$$w_{it+1}, x_{it+1} \mid w_{it}, x_{it}$$

- Flexibly estimated (using Gaussian Mixture Regression.)
- To get dollar interpretation, set value for B . A reasonable range: \$5,000-\$10,000.

What do we need to evaluate these counterfactuals?

- Need to know the transition process.
- Express it in terms of lower-dimensional objects:

$$g_{it+1}, \rho_{it+1}, \pi_{it+1} \mid g_{it}, \rho_{it}, \pi_{it}$$

[Details](#)

rather than

$$w_{it+1}, x_{it+1} \mid w_{it}, x_{it}$$

- Flexibly estimated (using Gaussian Mixture Regression.)
- To get dollar interpretation, set value for B . A reasonable range: \$5,000-\$10,000.

What do we need to evaluate these counterfactuals?

- Need to know the transition process.
- Express it in terms of lower-dimensional objects:

$$g_{it+1}, \rho_{it+1}, \pi_{it+1} \mid g_{it}, \rho_{it}, \pi_{it}$$

[Details](#)

rather than

$$w_{it+1}, x_{it+1} \mid w_{it}, x_{it}$$

- Flexibly estimated (using Gaussian Mixture Regression.)
- To get dollar interpretation, set value for B . A reasonable range: \$5,000-\$10,000.

Gaussian Mixture Regression of y on x (Sung (2004))

Suppose that the vector (x, y) is distributed as a mixture of M Gaussians, i.e., its pdf is

$$f_{X,Y}(x, y) = \sum_{m=1}^M \kappa_m \phi((x, y)' ; \mu_m, \Sigma_m)$$

where $\{\kappa_m\}$ are weights with $\sum_{m=1}^M \kappa_m = 1$. We estimate $\{\kappa_m, \mu_m, \Sigma_m\}$ using the EM algorithm.

Then, the distribution of y conditional on an observed value x is

$$f_{Y|X}(y|x) = \sum_{m=1}^M \omega_m(x) \phi(x; \mu_{mX}, \Sigma_{mX})$$

where the mixing weights $\{\omega_m\}$ are derived using Bayes' Rule:

$$\omega_m(x) = \frac{\kappa_m \phi(x; \mu_{mX}, \Sigma_{mX})}{\sum_{m'=1}^M \kappa_{m'} \phi(x; \mu_{m'X}, \Sigma_{m'X})}$$

Counterfactual details

Joint transition process for (π_{it}, ρ_{it}) :

- Let $\iota_{it} = \Lambda^{-1}(\pi_{it})$.
- $\iota_{it+1} | \iota_{it} \sim F^{\iota}(\cdot; \iota_{it}, a_{it})$, where we estimate F^{ι} using Gaussian mixture regression.
- Relationship between ι and ρ : For each $j \in \{\text{pre, post}\}$,

$$\pi_{it} = \Lambda\left(\phi_0^j + \phi_1^j \Lambda^{-1}(\rho_{jit}) + \zeta_{it}\right) \quad \text{or, equivalently,} \quad \rho_{jit} = \Lambda\left(\frac{\iota_{it} - \phi_0^j - \zeta_{it}}{\phi_1^j}\right)$$

where $\zeta_{it} \sim N(0, \sigma_j)$.

- Find $(\phi_0^j, \phi_1^j, \sigma_j)$ such that simulation matches actual AUC and logistic regression results.

The effect of PredictFix on technicians' predictions

A key question: With PredictFix, do technicians exhibit better ability to predict breakdowns?

Tool: ROC curves.

Table: ROC curve specifications

Predictor	Outcome	Sample restriction
$\hat{\rho}_{\text{pre}}$	Breakdowns	$a_{it} = 0$
$\hat{\rho}_{\text{post}}$	Breakdowns	$a_{it} = 0$
$\hat{\pi}$	Breakdowns	Test sample and $a_{it} = 0$

[Back: Table](#)

[Back: ROC Curves](#)

The effect of PredictFix on technicians' predictions (ρ)

What about miscalibration?

Logistic regression:

$$\Pr(\text{Breakdown}_{it} | \rho_{jit}) = \Lambda(\phi_0 + \phi_1 \Lambda^{-1}(\rho_{jit}))$$

The effect of PredictFix on technicians' predictions (ρ)

What about miscalibration?

Logistic regression:

$$\Pr(\text{Breakdown}_{it} | \rho_{jit}) = \Lambda(\phi_0 + \phi_1 \Lambda^{-1}(\rho_{jit}))$$

	$\rho = \pi$	ρ_{pre}	ρ_{post}
$\Lambda^{-1}(\rho)$	1		
Constant	0		

The effect of PredictFix on technicians' predictions (ρ)

What about miscalibration?

Logistic regression:

$$\Pr(\text{Breakdown}_{it} | \rho_{jit}) = \Lambda(\phi_0 + \phi_1 \Lambda^{-1}(\rho_{jit}))$$

	$\rho = \pi$	ρ_{pre}	ρ_{post}
$\Lambda^{-1}(\rho)$	1	0.233*** (0.0435)	0.814*** (0.0764)
Constant	0	-3.198*** (0.218)	-0.817* (0.318)
N		19091	19091

Why do technicians overestimate risk?

