Competing under Information Heterogeneity: Evidence from Auto Insurance

Marco Cosconati Yi Xin Fan Yizhou Jin Wu

IVASS, Bank of Italy

Caltech

Caltech

University of Toronto

July 8, 2025

Motivation

- Firms increasingly differ in *information precision* (data access/analytics) and in *cost structures*.
- This creates information asymmetries *between* firms (beyond classic buyer–seller asymmetry).
- Policy interest: regulations that equalize or share consumer risk information (e.g., centralized "risk bureau").

Research Questions

- How does heterogeneous information across insurers shape pricing, sorting, and market power?
- What happens to prices, surplus, profits, and sorting if information is shared/standardized?
- Distributional effects: who gains (low vs. high risk)? Efficiency effects: matching and costs?

Contributions

- A tractable model of imperfect competition with firm-specific information precision and costs.
- New identification/estimation strategy using offered-price distributions and demand to recover signals.
- Evidence from Italian auto liability insurance with rich panel linking consumers across insurers.
- Counterfactuals: centralized risk bureau, full information, and privacy/high-variance restrictions.

Institutional Background: Italian Auto Liability (RCA)

- Mandatory, annual, exclusive contracts; insurers cannot reject consumers.
- Large market: $\approx$31M contracts in 2018; $\approx$50 national competitors.
- Key contract features widely standardized; little use of deductibles.

Data: IVASS IPER Microdata

- Nationally representative matched insurer–insuree panel with claims frequency/severity, premiums, coverage.
- Tracks policyholders across insurers and time $\Rightarrow$ measure risk using ex-post claims panel.
- Focus sample: new customers in Rome (2013–2021); top 10 firms + fringe group.

Sample & Summary Statistics

- $N \approx 124{,}428$ contracts; avg premium $\approx$ 478; within-year claim rate ≈ 0.08.
- Demographics/vehicle: 56% male; avg age 48; BM class ≈ 2; car age ≈ 8.3 years.

[Insert Table 1: Summary statistics (premiums, claims, characteristics).]

Stylized Facts: Price Variation & Sorting

- Large cross-firm variation in average premiums even at similar average risks/market shares.
- Firms with higher average claim costs attract riskier consumers $\Rightarrow$ sorting across firms.

[Insert Figure 1: Avg premium vs. avg claim payouts by firm; bubble size = share.]

Heterogeneity in Information Precision

- Measure how strongly each firm's premium responds to realized consumer risk (ex-post panel-based risk).
- Strong cross-firm differences in premium–risk slopes $\Rightarrow$ heterogeneous precision.

[Insert Figure 2: Coefficients of premium on risk by firm (95% CIs).]

Conceptual Framework (Overview)

- J insurers; standardized product; no outside option.
- Consumer true risk θ (expected cost/year) unobserved ex ante.
- Firm j observes a private signal $\hat{\theta}_j$ with precision that differs across firms.

Signal Structure

$$\hat{\theta}_j \sim \mathcal{N}(\theta,\ \sigma_j^2), \quad \text{independent across } j \mid \theta, \tag{1}$$

- Lower $\sigma_j^2 \Rightarrow$ higher information precision for firm j.
- Signals are used to form posterior beliefs about θ *conditional on selection*.

Risk Rating & Pricing

$$p_j(\hat{\theta}_j) \;=\; \alpha_j \;+\; \beta_j\, \mathbb{E}[\theta \mid \hat{\theta}_j,\, D = j], \tag{2}$$

- α_j: baseline markup; β_j: pass-through/sensitivity to risk rating.
- $\mathbb{E}[\theta \mid \hat{\theta}_j, D = j]$ embeds selection $\Rightarrow$ nonlinearity in $\hat{\theta}_j$.

Demand

- Consumers choose one insurer; utility depends on price and observable characteristics.
- No outside option (mandatory purchase) $\Rightarrow$ shares across J firms sum to 1.
- Preference parameters allowed to vary with observables and risk type.

Identification: Intuition

- Offered price is (strictly) increasing in the firm's private signal (auction-style monotonicity).
- Use observed *transaction* prices + demand model to invert to *offered-price* distributions by firm.
- How average prices move with risk identifies (α_j, β_j); residual dispersion $\Rightarrow \sigma_j^2$.

Estimation Steps

1. Estimate demand and map transaction prices/shares to offered-price CDFs (firm-specific, nonparametric).
2. Recover pricing coefficients (α_j, β_j) from price–risk relationships.
3. Use price dispersion to identify signal variance σ_j^2 (information precision).
4. Back out firm cost parameters from first-order conditions.

Data Construction of Risk (Two-Part Model)

- Panel regressions to estimate individual risk:
  - Frequency (accident counts with FE) and severity (conditional paid amount).
  - Multiply predicted frequency $\times$ predicted severity $\Rightarrow$ expected cost per year.
- Controls for contract features (coverage, restrictions, devices) mitigate moral-hazard confounds.

Results: Firm Heterogeneity

- Large differences across firms in information precision (σ_j^2), pricing sensitivity (β_j), and costs.
- Firms with less accurate risk-rating tend to have *more efficient* claim-processing costs.
- Baseline sorting: higher-risk consumers concentrate at firms with higher average claim payouts.

Results: Price Sensitivity & Markups

- Estimated β_j varies markedly: some firms' prices are much more responsive to risk.
- Baseline markups (α_j) differ, consistent with market power from information advantages.

[Insert plot: distribution of β_j and α_j across firms.]

Counterfactuals: Information Policies

- **Centralized Risk Bureau**: aggregate firms' signals (weighted by precision), share equally with all.
- **Full Information Benchmark**: firms observe true θ (eliminate information asymmetry).
- **Privacy/Restriction**: firms can only use basic information; set σ_j^2 to the worst observed.

Counterfactual Results: Prices & Welfare

- Average premiums fall by ∼21.6% (bureau) to ∼25.7% (full information).
- Consumer surplus rises by ∼15.7% (bureau), close to ∼16.9% (full information).
- Firm profits decline on average; losses largest for firms with advanced risk-rating tech.

[Insert bar chart: Δ premium, Δ CS, Δ profit under each scenario.]

Distributional Effects by Risk Type

- Bureau/full-info mainly benefit *low-risk* consumers via sharper risk-based pricing.
- Privacy/high-variance benefits *high-risk* consumers (harder to distinguish from low-risk).

[Insert plot: CS changes by risk decile under each scenario.]

Mechanism: Competition & Undercutting

- Equalizing information weakens incumbents' info-based market power.
- Common risk evaluation $\Rightarrow$ more effective undercutting $\Rightarrow$ stronger price competition.

Sorting & Efficiency

- With equal access to risk, firms more efficient at processing claims re-target higher-risk consumers.
- Sorting shifts from info advantages to cost specialization.
- Efficiency gains: avg cost $\downarrow$ by $\sim$3.7% (full info) and by $\sim$ 12 per contract (bureau).

[Insert figure: change in sorting patterns (risk $\times$ firm) vs. baseline.]

Robustness (Selected)

- Alternative risk measures and controls for contract features.
- Bootstrapped uncertainty accounting for generated regressors.
- Poisson checks: premiums predicting claim counts; similar cross-firm heterogeneity.

[Insert table/figure: robustness summaries.]

Policy Implications

- Centralized information can materially lower prices and raise consumer surplus.
- Distributional trade-offs: low-risk consumers gain more under information sharing; high-risk under privacy.
- Industry composition effects: advanced-screening firms lose profits; potential dynamic innovation effects.

Limitations

- Abstract from dynamic pricing/learning and multi-product cross-selling mechanisms.
- Treat signals as reduced-form precision differences (black box of algorithms/data).
- Focus on new customers (tenure $= 0$) to avoid dynamics $\Rightarrow$ external validity caveats.

Paths for Future Work

- Dynamic extensions with learning and switching costs.
- Endogenous investment in information precision and costs (innovation incentives under policy).
- Generalization to other selection markets (credit, health, annuities) under heterogeneous information.

Takeaways

- Information heterogeneity shapes pricing power, sorting, and efficiency.
- Centralized sharing can compress prices and reorient sorting toward cost efficiency.
- Welfare gains are sizable, with clear distributional patterns across risk types.

Appendix: Risk Construction Details

- Frequency model with individual fixed effects; severity model (log amounts).
- Predicted risk $= \widehat{\text{freq}} \times \widehat{\text{severity}}$; controls for contract features.

[Insert table/figure: frequency & severity regression summaries.]

Appendix: Identification Sketch

- Monotonicity of offers in signals $\Rightarrow$ order-preserving mapping to signal quantiles.
- Demand-implied mapping from transactions to offers recovers firm-specific offer CDFs.
- Price–risk slope pins down (α_j, β_j); residual dispersion identifies σ_j^2.

Q&A

Questions?