


Quality Disclosure and Regulation: Scoring Design in Medicare Advantage

Benjamin Vatter

MIT

April 30th 2024

► Quality scores affect our everyday choices



FLORIDA DEPARTMENT OF
EDUCATION
fldoe.org

PORTAL HOME REPORT CARDS +

Compare Schools

BLANKNER K-8 Orange School District 2500 S Mills Ave, Orlando, FL 32806-4716 Traditional Public, <i>Combination School</i> 974 students	A	Compare <input type="checkbox"/>
CARVER MIDDLE Orange School District 4500 W Columbia St, Orlando, FL 32811-4052 Traditional Public, <i>Middle School</i> 832 students	C	Compare <input type="checkbox"/>
CATALINA ELEMENTARY Orange School District 2448 29th St, Orlando, FL 32805-6004 Traditional Public, <i>Elementary School</i> 733 students	C	Compare <input type="checkbox"/>
CENTRAL FLORIDA LEADERSHIP ACADEMY CHARTER Orange School District 427 N Primrose Dr, Orlando, FL 32803-5012	B	Compare <input type="checkbox"/>

Medicare.gov

Showing 1 - 15 of 62 hospitals

Sort by: Closest ▾

1. **Northshore University
Healthsystem - Evanston
Hospital** 

0.8 mi

ACUTE CARE HOSPITALS

2650 Ridge Ave
Evanston, IL 60201
(847) 432-8000

Overall star rating



Patient survey rating



[Compare](#)



2. **Presence Saint Francis Hospital**

2.4 mi



Overall star rating



► Quality scores affect our everyday choices



Earthjustice GIVE WITH CONFIDENCE ★★★★★	The Climate Reality Project GIVE WITH CONFIDENCE ★★★★★	Greenpeace Fund GIVE WITH CONFIDENCE ★★★★☆
Earth Day Network GIVE WITH CONFIDENCE ★★★★★	Natural Resources Defense Council GIVE WITH CONFIDENCE ★★★★★	Center for International Environmental Law (CIEL) GIVE WITH CONFIDENCE ★★★★★
350.org GIVE WITH CONFIDENCE ★★★★☆	As You Sow GIVE WITH CONFIDENCE ★★★★★	Climate Central GIVE WITH CONFIDENCE ★★★★☆
Sierra Club Foundation GIVE WITH CONFIDENCE ★★★★★	Union of Concerned Scientists GIVE WITH CONFIDENCE ★★★★★	Friends of the Earth GIVE WITH CONFIDENCE ★★★★☆



- ▶ **Quality scores affect our everyday choices**
 - > How to design them to maximize welfare?

- ▶ Two central mechanisms:
 - 1 Help consumers choose through added information (Dranove and Jin, 2010)
 - 2 Affect firms' incentives to invest in quality (Barahona et al., 2020)

- ▶ **Quality scores affect our everyday choices**
 - > How to design them to maximize welfare?
- ▶ Two central mechanisms:
 - 1 Help consumers choose through added information (Dranove and Jin, 2010)
 - 2 Affect firms' incentives to invest in quality (Barahona et al., 2020)
- ▶ Scores can be powerful policy tools, however
 - > No systematic guidance on how to design them
 - > Poor designs can backfire (gaming) (Feng Lu, 2012)

- Q: How to design welfare-maximizing scores for Medicare Advantage (MA)?**
- > Summarize medical and service quality of insurance plans using nine scores (stars)

Q: How to design welfare-maximizing scores for Medicare Advantage (MA)?

- > Summarize medical and service quality of insurance plans using nine scores (stars)

- ▶ Use yearly variation in scoring design between 2009 and 2015 to:
 - 1 Show that design affects demand and supply of health insurance

Q: How to design welfare-maximizing scores for Medicare Advantage (MA)?

- > Summarize medical and service quality of insurance plans using nine scores (stars)

- ▶ Use yearly variation in scoring design between 2009 and 2015 to:
 - 1 Show that design affects demand and supply of health insurance
 - 2 Estimate a model of demand, pricing, and quality investments
 - **Information asymmetries**: consumers' quality information is severely limited
 - **Inefficient quality provision**: too low on aggregate, distorted by private incentives (Spence, 1975)

Q: How to design welfare-maximizing scores for Medicare Advantage (MA)?

- > Summarize medical and service quality of insurance plans using nine scores (stars)

- ▶ Use yearly variation in scoring design between 2009 and 2015 to:
 - 1 Show that design affects demand and supply of health insurance
 - 2 Estimate a model of demand, pricing, and quality investments
 - **Information asymmetries**: consumers' quality information is severely limited
 - **Inefficient quality provision**: too low on aggregate, distorted by private incentives (Spence, 1975)

- ▶ Develop a general **empirical scoring design** methodology
 - > Combine computational methods with insights from information design (Kamenica and Gentzkow, 2011)
 - ⇒ Model + method deliver a welfare-improving design for MA

- ▶ **New design increases total welfare by 3.7 monthly premiums per consumer/year**
 - > Uses four scores: four stars with discrete increments
 - > One-star pools low and medium quality (↓ info) others partition high quality (↑ info)
 - > Consumers avoid one-star plans, firms respond by increasing investments (↑ quality)
 - > Reward more improvements in quality dimensions consumers' care about (↑ efficiency ↑ info)

- ▶ **New design increases total welfare by 3.7 monthly premiums per consumer/year**
 - > Uses four scores: four stars with discrete increments
 - > One-star pools low and medium quality (↓ info) others partition high quality (↑ info)
 - > Consumers avoid one-star plans, firms respond by increasing investments (↑ quality)
 - > Reward more improvements in quality dimensions consumers' care about (↑ efficiency ↑ info)
- ⇒ Consumers make more informed choices over higher quality products

► **New design increases total welfare by 3.7 monthly premiums per consumer/year**

- > Uses four scores: four stars with discrete increments
- > One-star pools low and medium quality (↓ info) others partition high quality (↑ info)
- > Consumers avoid one-star plans, firms respond by increasing investments (↑ quality)
- > Reward more improvements in quality dimensions consumers' care about (↑ efficiency ↑ info)

⇒ Consumers make more informed choices over higher quality products

► **Delivers broad lessons about scoring policies**

- > Scores are powerful mechanisms by which to regulate quality
- > Coarse, simple, scores can outperform full-information outcomes at small informational losses

1 Institutional Details and Data

- > Graphical representation of the scoring design problem

2 Model, Identification, and Estimates

- > Measurement of the frictions addressed by the scores

3 Scoring Design

- > Mechanisms by which optimal scores improve welfare

Three Facts About Medicare Advantage


- 1 National regulated private health insurance market
 - > All 65 million Medicare-eligible individuals can opt into MA, about half do
 - > Trade-off: greater access vs. better coverage
 - > Generous premium subsidies, risk-adjustments for insurers
- 2 Highly concentrated: 90% of average county enrollment controlled by 2 firms
 - > 4 firms account for 70% of national MA enrollment
- 3 Quality heterogeneity affects mortality, costs billions in subsidies (Abaluck et al., 2021)
 - > Challenging to assess if not for the quality scores



The MA Star Ratings


5 | 23


- Summarize medical and service quality in 1-to-5 stars, in half-star increments

There are 15 plans in **01757** that match your preferences. [View 10](#) [View 20](#) [View 50](#)

Compare Plans 

Sort Results by Lowest Estimated Annual Health and Drug Cost  **Sort** 

 **Tufts Medicare Preferred HMO Saver Rx (HMO) (H2256-028-0)**
Organization: Tufts Health Plan or Tufts Medicare Preferred

Estimated Annual Drug Costs:[?]	Monthly Premium: [?]	Deductibles [?] and Drug Copay[?] / Coinsurance:[?]	Health Benefits: [?]	Drug Coverage [?], Drug Restrictions[?] and Other Programs:	Estimated Annual Health and Drug Costs:[?]	Overall Star Rating:[?]	
Retail Pharmacy Status: Network Annual: \$29 Mail Order Annual: \$17	\$0.00 Drug: \$0.00 Health: \$0.00	Annual Drug Deductible: \$0 Health Plan Deductible: \$0 Drug Copay/ Coinsurance: \$5 - \$95, 33%	Doctor Choice: Plan Doctors Only Out of Pocket Spending Limit: \$6,700 In-Network 	All Your Drugs on Formulary: Yes Drug Restrictions: No No Additional Gap Coverage Lower Your Drug Costs	\$3,080	Rating definition for O ★★★★★ 4.5 out of 5 stars	Enroll

Scoring Design (simplified)

- 1 Measure plan's performance over five categories of quality
 - 1 Medical Outcomes
 - 2 Intermediate Medical Outcomes (chronic conditions)
 - 3 Access to Care
 - 4 Patient Experience
 - 5 Process Measures (preventive, diagnostic care)
- 2 Give a score of 1-5 to each plan and each category
- 3 Show consumers the rounded weighted average

Graphical Representation

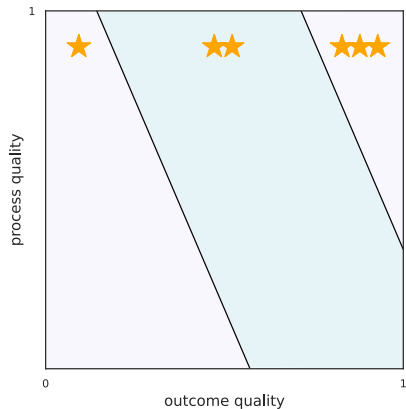
7 | 23

► **Design:** slope and location of hyper-planes

- > Slope = Weights, Location = Cutoffs
- > In two dimensions design is just lines →

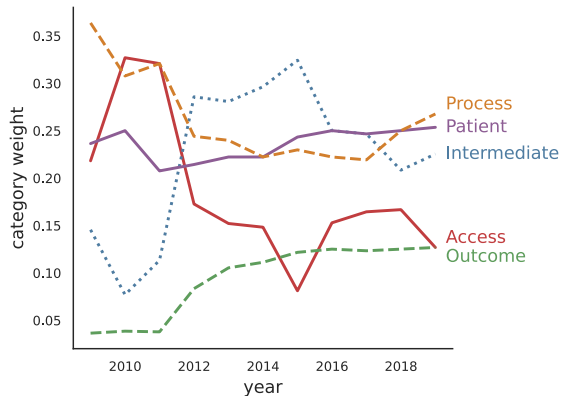
Q: Which lines to draw and how many?

► Scores reveal quality regions, not value

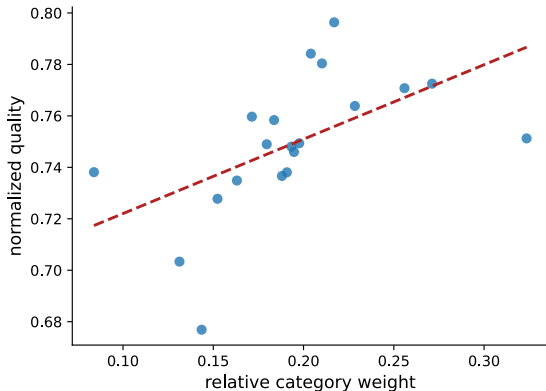


1 Scoring rules

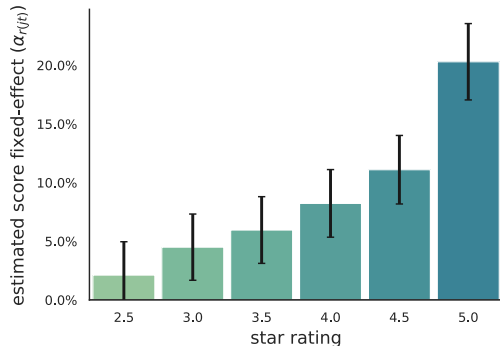
- > Hand collected from CMS
- > **Substantial variation in design**



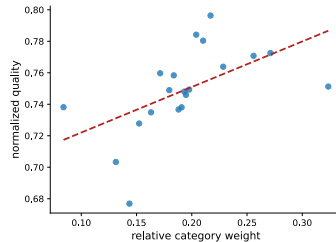
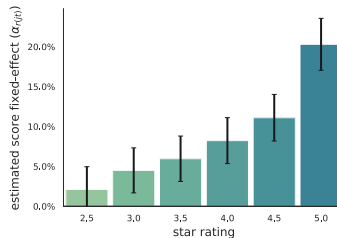
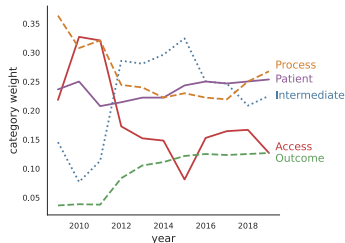
- 1 Scoring rules
- 2 Data on all plans
 - > Premiums, coverage, and benefits
 - > Total investment by contract (2015 only)
 - > **Quality: responds to design**



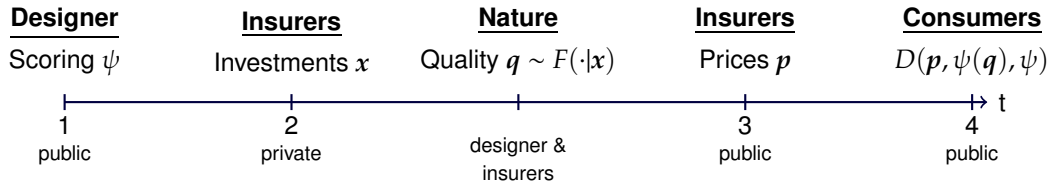
- 1 Scoring rules
- 2 Data on all plans
- 3 Enrollment data
 - > Individual-level representative panel
 - > 46,833 enrollment choices
 - > Linked claims
 - > **Consumers prefer higher-scoring plans**

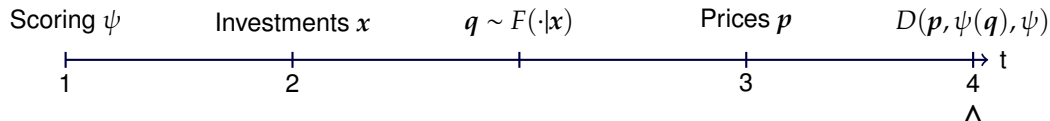


- ▶ Plentiful design variation reveals that scores:
 - 1 Shift demand across products
 - 2 Affect firms' quality investments
- ▶ To extrapolate to new designs, we must recover the social cost and value of quality
 - > **Costs**: from variation in scoring incentives to invest
 - > **Value**: from variation in WTP for scores



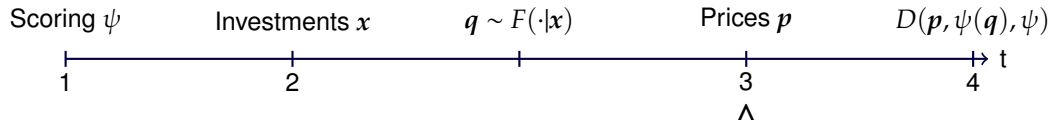
- 1 Institutional Details and Data
- 2 **Model, Identification, and Estimates**
- 3 Scoring Design





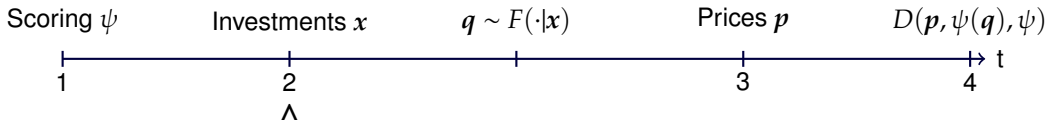
$$u_{ij} = \underbrace{\alpha_i p_j}_{\text{premium}} + \underbrace{\beta_i b_j}_{\text{coverage}} + \underbrace{\mathcal{E}_q[\gamma' q | \psi(q_j), \psi]}_{\text{quality}} + \underbrace{\lambda' z_{ij}}_{\text{Obs. attributes}} + \underbrace{\xi_j}_{\text{unobs. preferences}} + \underbrace{\varepsilon_{ij}}_{\sim \text{T1EV}}$$

- ▶ Choose among MA plans – or – Medicare + Part D (prescription drug coverage)
- ▶ Heterogeneity in WTP for quality (γ/α_i) \Rightarrow scoring granularity
- ▶ Subjective Bayesian non-parametric priors \Rightarrow scoring cutoffs and weights



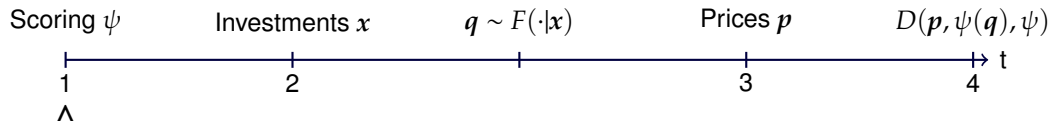
$$\pi_f(q, \psi) = \max_{\{p_j\}_{j \in J_f}} \sum_{j \in J_f} \underbrace{D_j(p, \psi(q))}_{\text{demand}} \underbrace{R_j(p_j)}_{\text{Mg. Revenue}} - \underbrace{C(q_j, z_j, \theta_j)}_{\text{Mg. Cost}}$$

- ▶ Multiproduct oligopolistic price competition with risk adjustment
- ▶ Quality affects insurance cost:
 - > Better hospitals increase claim prices ($\uparrow C$), preventive care reduces hospitalization ($\downarrow C$)



$$\max_{x_f \in \mathbb{R}^{|Q| \times |I_f|}} \underbrace{\int \mathbb{E}[\pi_f(q_f, q_{-f}, \psi)] dF(q_f | x_f)}_{\text{expected insurance profit}} - \underbrace{I_f(x_f)}_{\text{investment cost}}$$

- ▶ Choose investment for each product-category
- ▶ Rational expectations about rivals' investments based on market observables (Sweeting, 2009)
- ▶ Heterogenous convex investment costs \Rightarrow equilibrium quality effects



- No optimality imposed on designer's experimentation

- ▶ Supply model identified from profit optimality conditions

- ▶ Supply model identified from profit optimality conditions
- ▶ Revealed preferences identify consumers' WTP for scores
 - > **Cannot** tell if WTP comes from beliefs about quality or preferences
 - > Example: only readmission risk quality (scalar)
 - Consumers WTP \$100 for plan to have 4 instead of 3 stars, all else equal
 - $\Delta \mathcal{E}(q) = 1\%$ and $\gamma = \$100$ or $\Delta \mathcal{E}(q) = 5\%$ and $\gamma = \$20$?

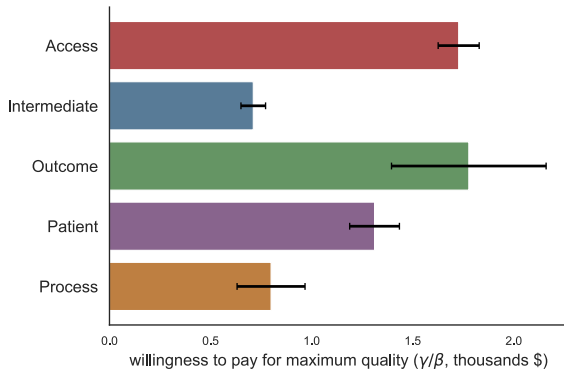
- ▶ Supply model identified from profit optimality conditions
- ▶ Revealed preferences identify consumers' WTP for scores
 - > **Cannot** tell if WTP comes from beliefs about quality or preferences
 - > Example: only readmission risk quality (scalar)
 - Consumers WTP \$100 for plan to have 4 instead of 3 stars, all else equal
 - $\Delta\mathcal{E}(q) = 1\%$ and $\gamma = \$100$ or $\Delta\mathcal{E}(q) = 5\%$ and $\gamma = \$20$?
- ▶ Intuition: if consumers understand design, posterior beliefs are bounded
 - > Bounds on beliefs + WTP \implies bounds on preferences
 - Consumers knows that $\psi(q) = 3 \iff q \in [0.8\%, 1\%)$ and $\psi(q) = 4 \iff q \in [0, 0.3\%)$
 - Therefore $\Delta\mathcal{E}(q) \in (0.5\%, 1\%) \implies \gamma \in (100, 200)$

- ▶ Supply model identified from profit optimality conditions
 - ▶ Revealed preferences identify consumers' WTP for scores
 - > **Cannot** tell if WTP comes from beliefs about quality or preferences
 - > Example: only readmission risk quality (scalar)
 - Consumers WTP \$100 for plan to have 4 instead of 3 stars, all else equal
 - $\Delta\mathcal{E}(q) = 1\%$ and $\gamma = \$100$ or $\Delta\mathcal{E}(q) = 5\%$ and $\gamma = \$20$?
 - ▶ Intuition: if consumers understand design, posterior beliefs are bounded
 - > Bounds on beliefs + WTP \implies bounds on preferences
 - Consumers knows that $\psi(q) = 3 \iff q \in [0.8\%, 1\%)$ and $\psi(q) = 4 \iff q \in [0, 0.3\%)$
 - Therefore $\Delta\mathcal{E}(q) \in (0.5\%, 1\%) \implies \gamma \in (100, 200)$
- \implies Variation in scoring design generates additional bounds and tightens identification

Key Estimates - Information Assymetry

12 | 23

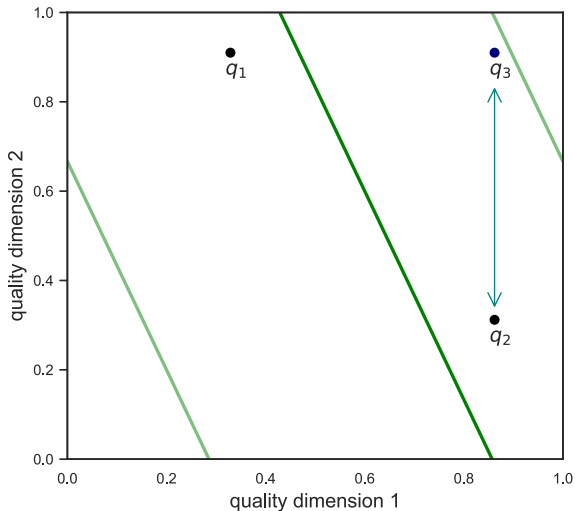
- ▶ Maximum Outcome quality \approx \$4,036 in OOP
- ▶ Incomplete info lowers surplus by \$199.3 (keeping supply fixed)
- ▶ Two sources of information asymmetry:



Key Estimates - Information Assymetry

12 | 23

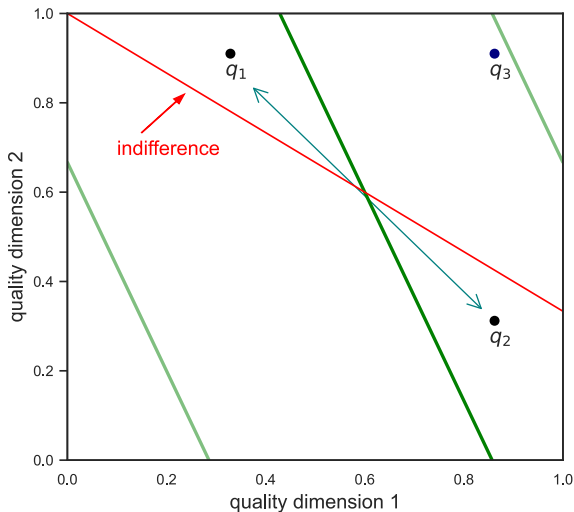
- ▶ Maximum Outcome quality \approx \$4,036 in OOP
- ▶ Incomplete info lowers surplus by \$199.3 (keeping supply fixed)
- ▶ Two sources of information asymmetry:
 - 1 Within-scores:
Best 4-star worth \$367.8 more than worst



Key Estimates - Information Assymetry

12 | 23

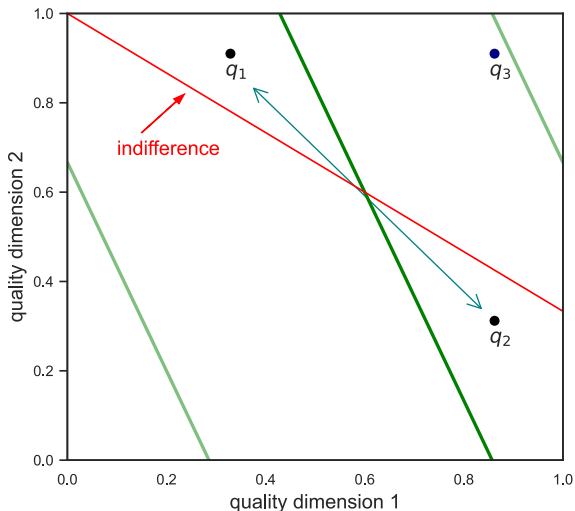
- ▶ Maximum Outcome quality \approx \$4,036 in OOP
- ▶ Incomplete info lowers surplus by \$199.3 (keeping supply fixed)
- ▶ Two sources of information asymmetry:
 - 1 Within-scores:
Best 4-star worth \$367.8 more than worst
 - 2 Across-scores:
22.7% of plans ranked opposite to preferences



Key Estimates - Information Assymetry

12 | 23

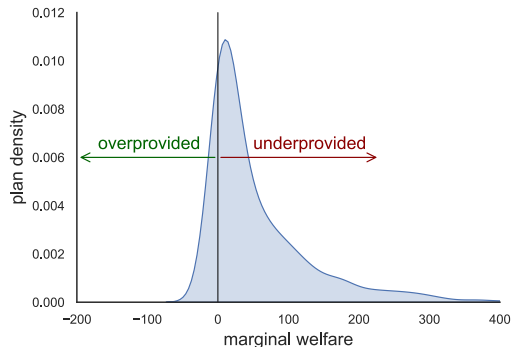
- ▶ Maximum Outcome quality \approx \$4,036 in OOP
- ▶ Incomplete info lowers surplus by \$199.3 (keeping supply fixed)
- ▶ Two sources of information asymmetry:
 - 1 Within-scores:
Best 4-star worth \$367.8 more than worst
 - 2 Across-scores:
22.7% of plans ranked opposite to preferences
 \Rightarrow 94.5% of losses come from across-score



Key Estimates - Quality provision

13 | 23

- ▶ Avg insurance markup of 10.5%
 - > For top insurers: avg marginal cost is \$758
 - > Curto et. al (2019): medical cost is \$680
- ▶ Median investment = 12% of insurance profits
- ▶ Quality is underprovided:
 - 1 On average, $dTW/dq \in [17.6, 84.9]$ million/contract
 - 2 Less so in more competitive markets (Spencian)
 - 3 Less so in categories with \uparrow weight (Design)



- 1 Institutional Details and Data
- 2 Model, Identification, and Estimates
- 3 Scoring Design

$$\max_{\psi \in \Psi} \mathbb{E}_q \left[\underbrace{CS(\psi, q)}_{\text{Consumer surplus}} + \underbrace{\sum_f V_f(\psi, q) - I(\mathbf{x}_f^*(\psi), \mu_f)}_{\text{Insurer profit}} | \mathbf{x}^*(\psi) \right]$$

- ▶ Subject to equilibrium behavior:
 - > Firms update investments, prices, beliefs about rivals
 - > Consumers update beliefs given design and realized scores
- ▶ Focus on deterministic, monotone, finite designs
 - > Includes MA, school letter grades, food labeling, ...

1 Exploring the space (Ψ):

- > **Challenge:** no optimality conditions to guide us (ψ is discontinuous)

1 Exploring the space (Ψ):

- > **Challenge:** no optimality conditions to guide us (ψ is discontinuous)
- > **Solution:** divide into smaller, manageable problems
 - 1 ψ = polynomial aggregator \circ cutoffs
 - 2 Choose number of cutoffs, polynomial order of aggregator
 - 3 Problem is now finitely parameterized: solve and iterate

1 Exploring the space (Ψ):

- > **Challenge:** no optimality conditions to guide us (ψ is discontinuous)
- > **Solution:** divide into smaller, manageable problems
 - 1 ψ = polynomial aggregator \circ cutoffs
 - 2 Choose number of cutoffs, polynomial order of aggregator
 - 3 Problem is now finitely parameterized: solve and iterate

2 Evaluating the welfare value of ($TW(\psi)$):

- > **Challenge:** state-space for pricing subgame is huge: $[0, 1]^{|Q| \times |J|}$
 - ψ induces a distribution over state-space, requires costly integration for every guess

1 Exploring the space (Ψ):

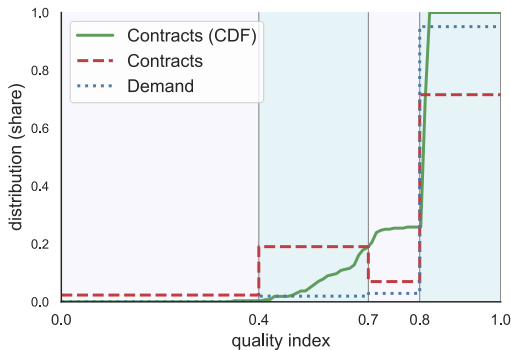
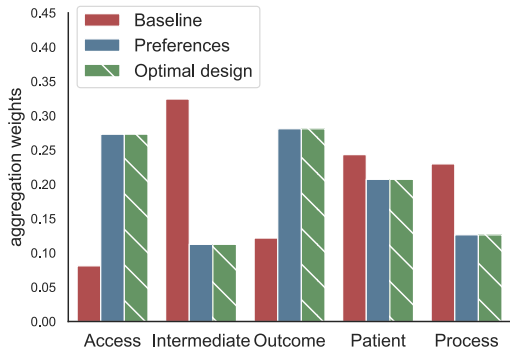
- > **Challenge:** no optimality conditions to guide us (ψ is discontinuous)
- > **Solution:** divide into smaller, manageable problems
 - 1 $\psi = \text{polynomial aggregator} \circ \text{cutoffs}$
 - 2 Choose number of cutoffs, polynomial order of aggregator
 - 3 Problem is now finitely parameterized: solve and iterate

2 Evaluating the welfare value of ($TW(\psi)$):

- > **Challenge:** state-space for pricing subgame is huge: $[0, 1]^{|Q| \times |J|}$
 - ψ induces a distribution over state-space, requires costly integration for every guess
 - > **Solution:** computation in Belief Space (Aumann and Maschler, 1995)
 - Drastically reduces dimensionality of state-space and integration costs
- ⇒ Solve large grid of independent equilibria, identify value of each score as a distribution over grid

Solution: Best Linear Design

16 | 23

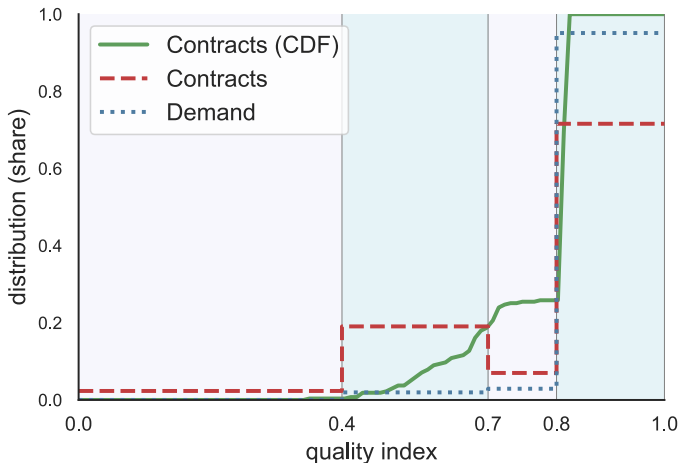


- 1 **Pooling at the bottom:** first score pools all low qualities
- 2 **Aggregator:** optimal weighting scheme aligned with preferences
- 3 **Limited granularity:** use only four scores; three partition higher quality

Decomposing the Design: Pooling at the Bottom

17 | 23

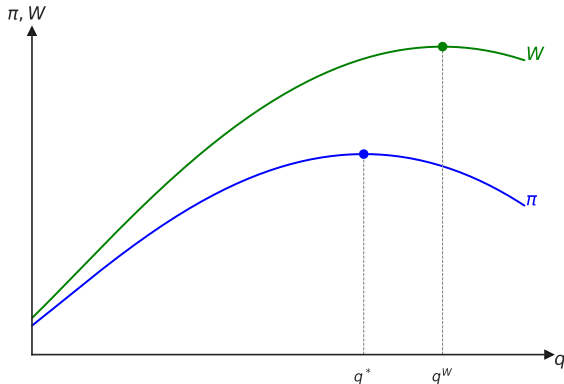
- Creates demand penalty for under-investment: \uparrow quality



Decomposing the Design: Pooling at the Bottom

17 | 23

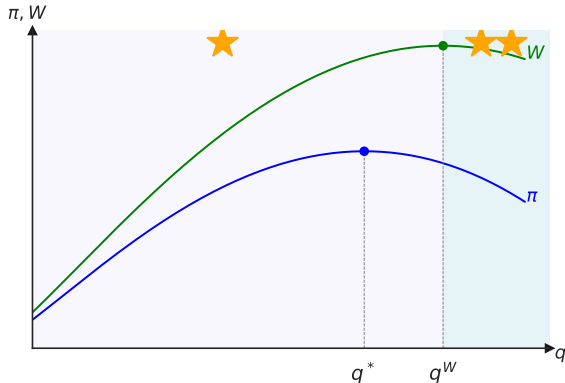
- ▶ Creates demand penalty for under-investment: \uparrow quality
- ▶ **Market power over quality** (Spence, 1975; Crawford et al., 2019) : firms under-invest even under full info



Decomposing the Design: Pooling at the Bottom

17 | 23

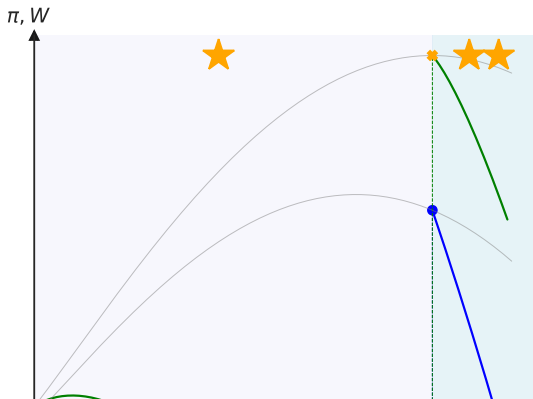
- ▶ Creates demand penalty for under-investment: \uparrow quality
- ▶ **Market power over quality** (Spence, 1975; Crawford et al., 2019) : firms under-invest even under full info



Decomposing the Design: Pooling at the Bottom

17 | 23

- ▶ Creates demand penalty for under-investment: \uparrow quality
- ▶ **Market power over quality** (Spence, 1975; Crawford et al., 2019) : firms under-invest even under full info
- ▶ Delegation equivalence (Zapechelnyuk, 2020) : certification $\iff q^w$ or 0

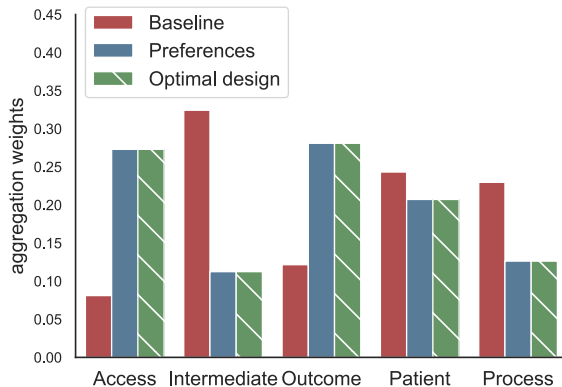


- ▶ Creates demand penalty for under-investment: \uparrow quality
- ▶ **Market power over quality** (Spence, 1975; Crawford et al., 2019) : firms under-invest even under full info
- ▶ Delegation equivalence (Zapechelnyuk, 2020) : certification $\iff q^w$ or 0
- ▶ **Accounts for 71.8% of welfare gain** (certification)
 - > 57% of contracts would receive <2 star in baseline, only 21% in equilibrium
 - > Serve only 1.9% of consumers
 - > **Quality is 4% higher in equilibrium, investment nearly triples**

Decomposing the Design: Aggregator

18 | 23

- ▶ New weights align with consumer preferences



Decomposing the Design: Aggregator

18 | 23

- ▶ New weights align with consumer preferences

Aggregation produces two problems:



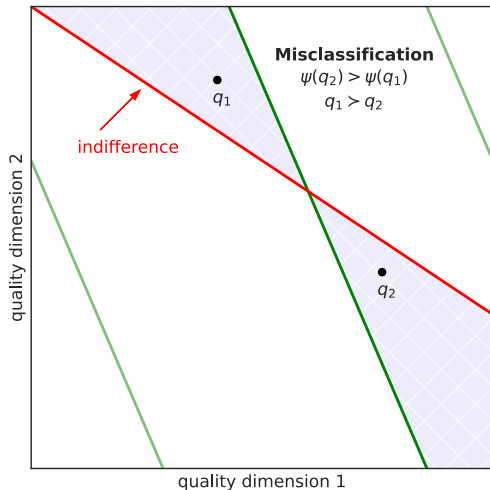
Decomposing the Design: Aggregator

18 | 23

- ▶ New weights align with consumer preferences

Aggregation produces two problems:

- 1 Across-scores information asymmetry:



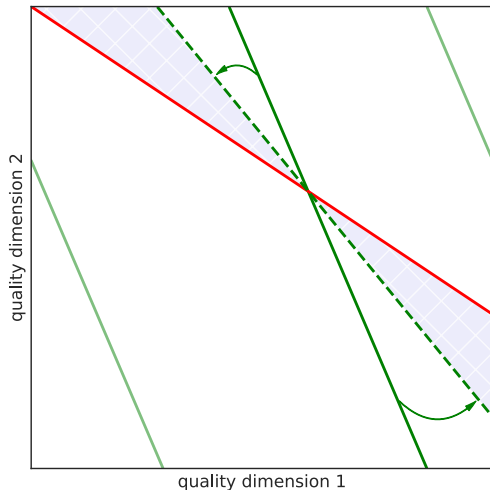
Decomposing the Design: Aggregator

18 | 23

- ▶ New weights align with consumer preferences

Aggregation produces two problems:

- 1 Across-scores information asymmetry:
 - > Eliminated by new weights



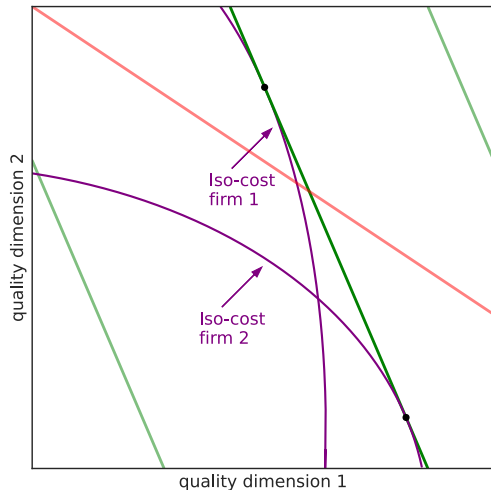
Decomposing the Design: Aggregator

18 | 23

- ▶ New weights align with consumer preferences

Aggregation produces two problems:

- 1 Across-scores information asymmetry:
 - > Eliminated by new weights
- 2 Multitasking moral hazard
(Holmstrom and Milgrom, 1991)
 - > Firms' allocations ignore preferences



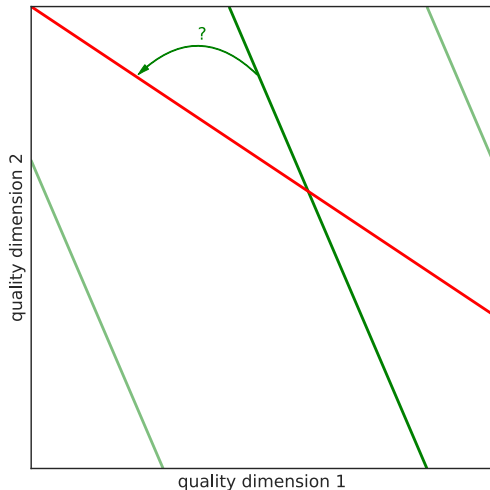
Decomposing the Design: Aggregator

18 | 23

- ▶ New weights align with consumer preferences

Aggregation produces two problems:

- 1 Across-scores information asymmetry:
 - > Eliminated by new weights
- 2 Multitasking moral hazard
(Holmstrom and Milgrom, 1991)
 - > Firms' allocations ignore preferences



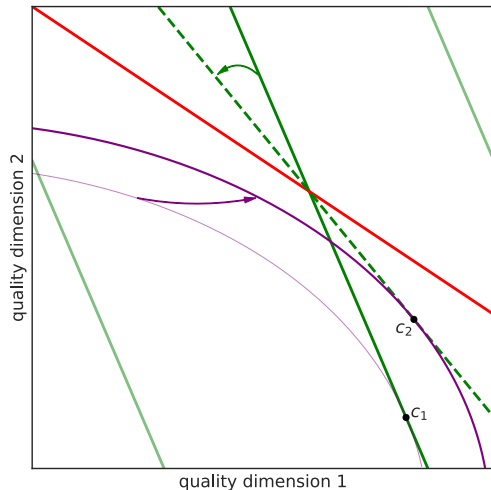
Decomposing the Design: Aggregator

18 | 23

- ▶ New weights align with consumer preferences

Aggregation produces two problems:

- 1 Across-scores information asymmetry:
 - > Eliminated by new weights
- 2 Multitasking moral hazard
(Holmstrom and Milgrom, 1991)
 - > Firms' allocations ignore preferences



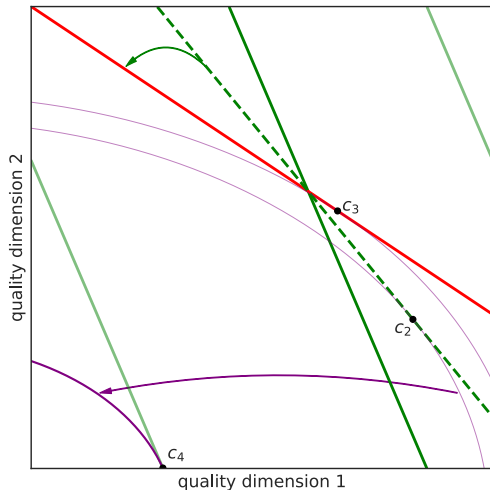
Decomposing the Design: Aggregator

18 | 23

- ▶ New weights align with consumer preferences

Aggregation produces two problems:

- 1 Across-scores information asymmetry:
 - > Eliminated by new weights
- 2 Multitasking moral hazard
(Holmstrom and Milgrom, 1991)
 - > Firms' allocations ignore preferences
- 3 Firm cost heterogeneity crucial for solution
 - > Otherwise, alignment leads to quality losses

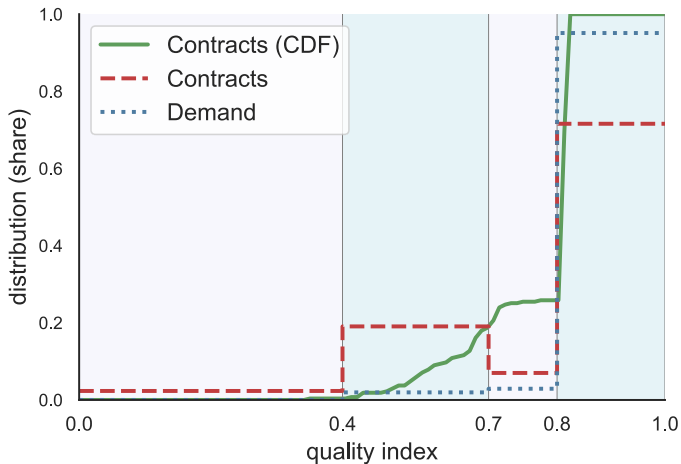


- ▶ New weights align with consumer preferences
 - ▶ **Pooling at the bottom + optimal aggregator account for 98.2% of welfare gains**
 - > Pooling increases overall investment
 - > Optimal aggregation improves informativeness and allocative efficiency of investments
- ⇒ High welfare value from optimal certification

Decomposing the Design: Granularity

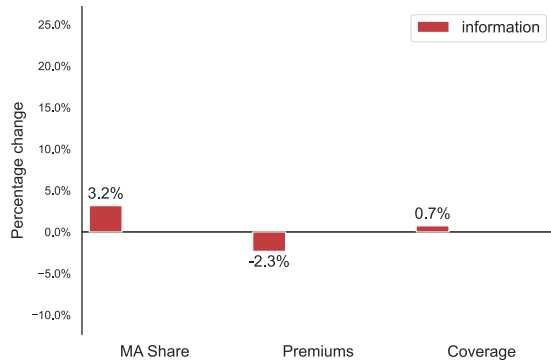
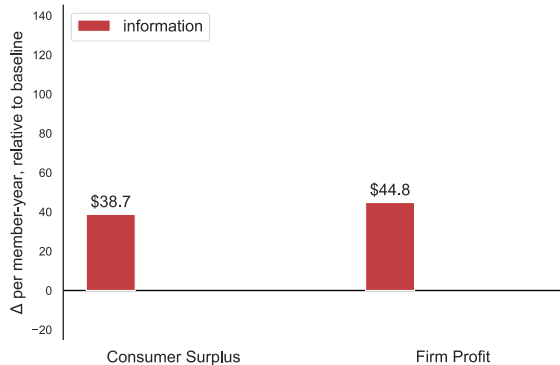
19 | 23

- Why only three scores at the top?



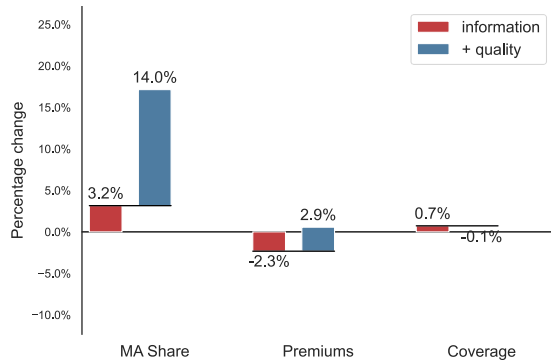
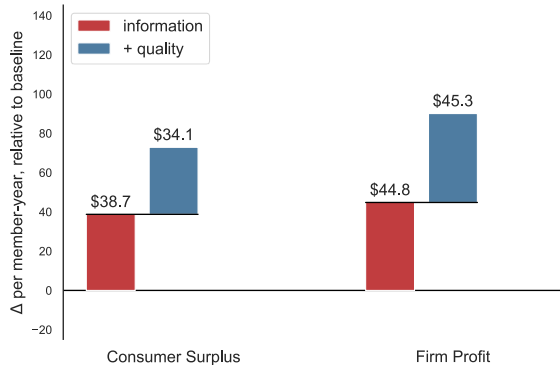
- ▶ Why only three scores at the top?
- ▶ Trade-off: efficiency vs. product variety
 - > More scores allow more investment actions for firms (delegation equivalence)
 - > More actions allow for more heterogeneity: lower quality at lower prices
 - > But also more deviations away from efficient production and towards profit maximization

- ▶ Why only three scores at the top?
- ▶ Trade-off: efficiency vs. product variety
 - > More scores allow more investment actions for firms (delegation equivalence)
 - > More actions allow for more heterogeneity: lower quality at lower prices
 - > But also more deviations away from efficient production and towards profit maximization
- ▶ Granularity governed by:
 - 1 Value: consumers' heterogeneity in WTP for quality
 - 2 Cost: ability to generate separating choices for firms



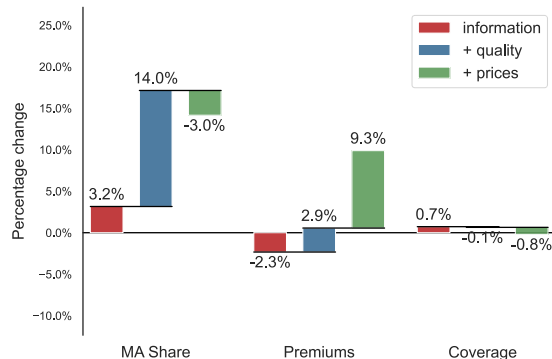
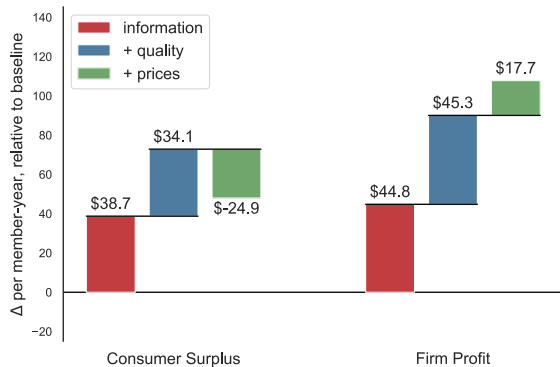
► Holding prices and quality change information:

- > Products are easier to choose, fewer mistakes
- > MA expansion: Consumers select quality that offsets systematic preferences



► Holding quality, change information, and prices:

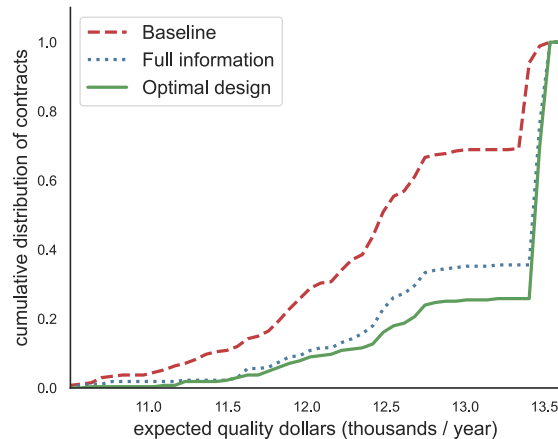
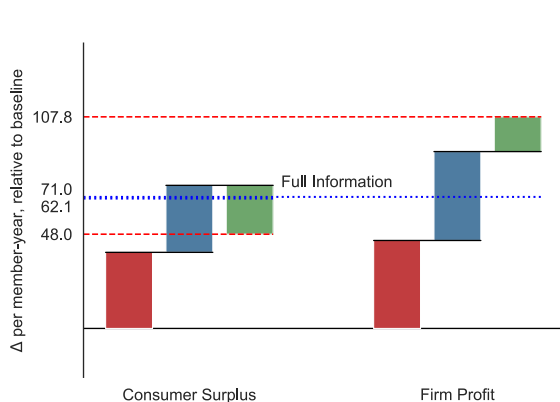
- > New information reveals vertical differentiation across products
- > Firms exert market power over prices, capturing surplus



► Full equilibrium changes:

- > Total welfare increases by \$155.7 per beneficiary/year, firms' benefit from coordination effect
- > Compensating variation of: quality = \$90.14 > \$70.45 = information

⇒ Quality regulation is key driver of welfare gains



- ▶ Full information allows exercise of market power over quality, reduces welfare
- ▶ New scores dominate only because of equilibrium quality effects

Why is CMS's design systematically different than the optimal?

- 1 Strong preferences for quality chronic care (Intermediate) and lower-cost hospitals (Outcome)
 - > Paternalism or dynamic considerations for future subsidized care
 - > Nudging the market with scores is enormously costly:
 - ⇒ Outperformed by a subsidy that generated 8 cents of investments per dollar spent

Why is CMS's design systematically different than the optimal?

- 1 Strong preferences for quality chronic care (Intermediate) and lower-cost hospitals (Outcome)
 - > Paternalism or dynamic considerations for future subsidized care
 - > Nudging the market with scores is enormously costly:
 - ⇒ Outperformed by a subsidy that generated 8 cents of investments per dollar spent
- 2 CMS might be risk averse to misrepresenting consumers' preferences
 - > CMS might also believe that consumers are naive (ignorant of policy changes)
 - > Medicare plays a delicate political and social role, objective might be $\max_{\psi \in \Psi} \min_{\gamma \in \Gamma} TW(\psi, \gamma)$
 - ⇒ CMS's design outperforms best (linear) monotone partitional design
 - > Assumptions of the setting are rejected by the data, yet presents credible rationale for status quo

- 1 New methodology delivers aggregators that offset multitasking moral hazard
 - > “Gaming” has been documented extensively in nursing homes, energy, schooling (Feng Lu, 2012; Clay et al., 2021; Neal and Schanzenbach, 2010)

- 1 New methodology delivers aggregators that offset multitasking moral hazard
 - > “Gaming” has been documented extensively in nursing homes, energy, schooling (Feng Lu, 2012; Clay et al., 2021; Neal and Schanzenbach, 2010)

- 2 Scores should be designed with quality goals in mind, not only informativeness
 - > Quality promoting initiatives exist alongside scores in healthcare, schooling, electric appliances,...
 - > Properly designed scores can enhance these efforts; poorly designed ones, counteract

- 1 New methodology delivers aggregators that offset multitasking moral hazard
 - > “Gaming” has been documented extensively in nursing homes, energy, schooling (Feng Lu, 2012; Clay et al., 2021; Neal and Schanzenbach, 2010)
- 2 Scores should be designed with quality goals in mind, not only informativeness
 - > Quality promoting initiatives exist alongside scores in healthcare, schooling, electric appliances,...
 - > Properly designed scores can enhance these efforts; poorly designed ones, counteract
- 3 Coarse, simple scores can improve welfare at small informational cost
 - > Longstanding concern about ability of consumers to process complex quality data
 - > Inherent value for simplicity in quality disclosure policies

- ▶ **Scores are powerful quality regulation policies:**
 - > Adapting MA's design to equilibrium effects increases welfare by \$8.8 billion
- ▶ Suggests potential for redesigning scores using theory and empirical work
 - > Challenges policy focus on granularity, (ex-ante) informativeness, cognitive bias considerations
 - ⇒ A simple, well-designed sticker can outperform full information outcomes
- ▶ **Empirical Scoring Design** methodology for disclosure policies
 - > Data-driven solution for an extensive policy problem

Thank You!

`bvatter@mit.edu`