The Design and Price of Information

Dirk Bergemann, Alessandro Bonatti, Alex Smolin

Discussed by Silvio Ravaioli

August 8, 2019

Overview of the Paper

- Model of data trading
- The fully informed monopolist data seller offers a menu of stochastic experiments
- Data buyers acquire additional information in order to improve own decision before taking an action
- Distribution of buyers: ex-ante identical, but with different interim beliefs (partially informed)

- Focus on the seller: properties of the revenue-maximizing menu of experiments
- Every experiment is a non-dispersed stochastic matrix
- Every menu contains a fully informative experiment
- Special case (binary states and actions): construction of optimal menu

Overview of the Paper

- Model of data trading
- The fully informed monopolist data seller offers a menu of stochastic experiments
- Data buyers acquire additional information in order to improve own decision before taking an action
- Distribution of buyers: ex-ante identical, but with different interim beliefs (partially informed)

- Focus on the seller: properties of the revenue-maximizing menu of experiments
- Every experiment is a non-dispersed stochastic matrix
- Every menu contains a fully informative experiment
- Special case (binary states and actions): construction of optimal menu

Motivation

- Unprecedented amount of information collected and traded: $20B to acquire/process consumer data in 2018
- What is the value of this data?

- Significant effort to link the design with real-world information products: Acxiom, Nielsen, Equifax, and Oracle [marketing and risk-mitigation purpose]
- Typically data buyers choose what to acquire from a standard database, such as potential consumers with prespecified characteristics (rows), or desired consumer attributes (columns)
- Rephrase the design of a statistical experiment in terms of hypothesis testing. Equifax (data seller) can provide characteristics (*red flags*) to improve an action (grant/deny a loan), with a trade-off between type I and type II errors

Literature and Contribution

- Novel with respect to the literature because of the setup and the insights it provides
  - Information is useful as long as it guides the actions
  - Horizontal and vertical differentiation of data buyers
  - Explanation for "no garbling" observed in real world settings

- **Information trading**

  Admati and Pfleiderer 1986: homogeneous data buyers, imperfectly informed, trading an asset with a common value
  Babaioff, Kleinberg, and Paes Leme 2012: similar setup but different timing; the payoff function depends on two state variables (private information for buyers and sellers), the seller offers a contract based on the observed variable

- **Discriminatory information disclosure**

  Lizzeri 1999: the seller decides whether to disclose match value
  Eso and Szentes 2007: the seller can make the payment contingent on statistical experiment and buyer's action

Literature and Contribution

- Novel with respect to the literature because of the setup and the insights it provides
  - Information is useful as long as it guides the actions
  - Horizontal and vertical differentiation of data buyers
  - Explanation for "no garbling" observed in real world settings

- **Information trading**
  Admati and Pfleiderer 1986: homogeneous data buyers, imperfectly informed, trading an asset with a common value
  Babaioff, Kleinberg, and Paes Leme 2012: similar setup but different timing; the payoff function depends on two state variables (private information for buyers and sellers), the seller offers a contract based on the observed variable

- **Discriminatory information disclosure**
  Lizzeri 1999: the seller decides whether to disclose match value
  Eso and Szentes 2007: the seller can make the payment contingent on statistical experiment and buyer's action

Literature and Contribution

- Bergemann and Bonatti 2015 "Selling cookies" model
- Focus on the data buyer (advertiser) optimal behavior
- Data buyers are identical in the prior (no private information), and they differ in the pair-specific match value $v(i, j)$, that is the profit of firm j being matched with consumer i
- Information structures are restricted to "queries": reveal or not the state realization
- The paper characterizes the optimal queries given the distribution of match values and the cost of advertising
- The price of information is determined by a competitive market

- Starting from Bergemann-Bonatti-Smolin's setup, we can consider further dimensions of heterogeneity across data buyers: action space, actions' costs, match value, preference for timing of information, ...

Literature and Contribution

- Bergemann and Bonatti 2015 "Selling cookies" model
- Focus on the data buyer (advertiser) optimal behavior
- Data buyers are identical in the prior (no private information), and they differ in the pair-specific match value $v(i, j)$, that is the profit of firm j being matched with consumer i
- Information structures are restricted to "queries": reveal or not the state realization
- The paper characterizes the optimal queries given the distribution of match values and the cost of advertising
- The price of information is determined by a competitive market

- Starting from Bergemann-Bonatti-Smolin's setup, we can consider further dimensions of heterogeneity across data buyers: action space, actions' costs, match value, preference for timing of information, ...

1. Contracts are contingent on Experiments only

- In the model, the seller commits to a disclosure policy [standard assumption in the Bayesian persuasion literature]
- Motivation: when the buyer acquires the database, she does not know its content and it may be profitable or not. Similarly, companies commission a record without knowing if its content will be helpful

- The commitment assumption (contracts are contingent on the experiment, not on action/utility) is strong, and often difficult to motivate
- We have other examples [e.g. Eso and Szentes 2007] in which the seller provides a "bundle" of service+information. For example, Google ads are paid per click, and the seller provides a service (show the ad to a group of users) and uses own information (select targets based on available data)

1. Contracts are contingent on Experiments only

- In the model, the seller commits to a disclosure policy [standard assumption in the Bayesian persuasion literature]
- Motivation: when the buyer acquires the database, she does not know its content and it may be profitable or not. Similarly, companies commission a record without knowing if its content will be helpful

- The commitment assumption (contracts are contingent on the experiment, not on action/utility) is strong, and often difficult to motivate
- We have other examples [e.g. Eso and Szentes 2007] in which the seller provides a "bundle" of service+information. For example, Google ads are paid per click, and the seller provides a service (show the ad to a group of users) and uses own information (select targets based on available data)

2. Common Knowledge and Inference from Tariff

- It is standard in mechanism design to offer a menu of options as the one presented in the paper
- But in this framework the "product" that is sold is *information*, and we can be concerned about the menu providing some information itself

- If the seller is informed about buyers' type distribution, but buyers are not, can they make some inference about other buyers' types, and therefore about the value of data?
- Suppose that the buyer updates own beliefs after observing the menu: what does she learn about the buyers' type θ [and ω] distribution? what are her "new" interim beliefs? Do the constraints change based on these interim beliefs?

2. Common Knowledge and Inference from Tariff

- It is standard in mechanism design to offer a menu of options as the one presented in the paper
- But in this framework the "product" that is sold is *information*, and we can be concerned about the menu providing some information itself

- If the seller is informed about buyers' type distribution, but buyers are not, can they make some inference about other buyers' types, and therefore about the value of data?
- Suppose that the buyer updates own beliefs after observing the menu: what does she learn about the buyers' type θ [and ω] distribution? what are her "new" interim beliefs? Do the constraints change based on these interim beliefs?

3. Cost for Collecting Information

- In the model, the seller already has a database, and the marginal cost for sharing information is zero
- What if we add a third population of agents (users) who decide whether to provide own data?
- Conclusions: a richer model should include the cost of acquiring the information (duplicate, distribute, and potentially degrade the available information)
  - Users may have heterogeneous WTA for providing own data
  - WTA may be a combination of own cost (e.g. fill a survey) and benefit (e.g. be part of a newsletter)
  - Based on the seller-user IC, only some users will provide own data, and they will self-select in the database
  - The seller-user IC constraint and database composition should also change based on buyer distribution

3. Cost for Collecting Information

- In the model, the seller already has a database, and the marginal cost for sharing information is zero
- What if we add a third population of agents (users) who decide whether to provide own data?
- Conclusions: a richer model should include the cost of acquiring the information (duplicate, distribute, and potentially degrade the available information)
  - Users may have heterogeneous WTA for providing own data
  - WTA may be a combination of own cost (e.g. fill a survey) and benefit (e.g. be part of a newsletter)
  - Based on the seller-user IC, only some users will provide own data, and they will self-select in the database
  - The seller-user IC constraint and database composition should also change based on buyer distribution

The Design and Price of Information

Dirk Bergemann, Alessandro Bonatti, Alex Smolin

Discussed by Silvio Ravaioli

August 8, 2019