SELLING COOKIES

By

Dirk Bergemann and Alessandro Bonatti

October 2013
Revised September 2014

COWLES FOUNDATION DISCUSSION PAPER NO. 1920RR

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.econ.yale.edu/

Selling Cookies*

Dirk Bergemann† Alessandro Bonatti‡

August 14, 2014

Abstract

We propose a model of data provision and data pricing. A single data provider controls a large database that contains information about the match value between individual consumers and individual firms (advertisers). Advertisers seek to tailor their spending to the individual match value. The data provider prices queries about individual consumers' characteristics (cookies). We determine the equilibrium data acquisition and pricing policies. Advertisers choose positive and/or negative targeting policies. The optimal query price influences the composition of the targeted set. The price of data decreases with the reach of the database and increases with the fragmentation of data sales.

KEYWORDS: Data Providers, Data Pricing, Selling Information, Targeting, Online Advertising, Cookies, Media Markets.

JEL CLASSIFICATION: D44, D82, D83.

*The first authord acknowledges financial support through NSF Grants SES 0851200 and ICES 1215808. We would like to thank Bob Gibbons, Michael Grubb, Richard Holden, Duncan Simester, Andy Skrzypacz, David Soberman, K. Sudhir, Juuso Toikka, Mike Whinston, Jidong Zhou, as well as participants in various seminars and conferences for helpful discussions.

†Yale University, 30 Hillhouse Ave., New Haven, CT 06520, USA, dirk.bergemann@yale.edu

‡MIT Sloan School of Management, 100 Main Street, Cambridge MA 02142, USA, bonatti@mit.edu

The use of individual-level information is rapidly increasing in many economic and political environments, ranging from advertising (various forms of targeting) to electoral campaigns (identifying voters who are likely to switch or to turn out). In all these environments, the socially efficient match between individual and "treatment" may require the collection, analysis and diffusion of highly personalized data. A large number of important policy and regulatory questions are beginning to emerge around the use of personal information. To properly frame these questions, we must understand how markets for personalized information impact the creation of surplus, which is the main objective of this paper.

Much of the relevant data is collected and distributed by data brokers and data intermediaries ranging from established companies such as Acxiom and Bloomberg, to more recently established companies such as Bluekai and eXelate. Perhaps the most prevalent technology to enable the collection and resale of individual-level information is based on cookies and related means of recording browsing data. Cookies are small files placed by a website in a user's web browser that record information about the user's visit. Data providers use several partner websites to place cookies on user's computers and collect information. In particular, the first time any user visits a partner site (*e.g.*, a travel site), a cookie is sent to her browser, recording any action taken on the site during that browsing session (*e.g.*, searches for flights).[1] If the same user visits another partner website (*e.g.*, an online retailer), the information contained in her cookie is updated to reflect the most recent browsing history.

The data provider therefore maintains a detailed and up-to-date profile for each user, and compiles segments of consumer characteristics, based on each individual's browsing behavior. The demand for such highly detailed, consumer-level information is almost entirely driven by advertisers, who wish to tailor their spending and their campaigns to the characteristics of each consumer, patient, or voter.

The two distinguishing features of online markets for data are the following: (a) individual queries (as opposed to access to an entire database) are the actual products for sale,[2] and (b) linear pricing is predominantly used. In other words, advertisers specify which consumer segments and how many total users ("uniques") they wish to acquire, and pay a price proportional to the number of users.[3] These features are prominent in the market for cookies,

[1] This type of cookie is known as *third-party cookie*, because the domain installing it is different from the website actually visited by the user. Over half of the sites examined in a study by the Wall Street Journal installed 23 or more third-party cookies on visiting users' computers (*The Web's New Gold Mine: Your Secrets,* the Wall Street Journal, July 30, 2010).

[2] We formally define a database and a query in the context of our model in Subsection I.B.

[3] Information based on third-party cookies can be priced in two ways: *per stamp* (CPS), where buyers pay for the right to access information about an individual user, independent of the frequency of use of that data; and *per mille* (CPM), where the price of the information is proportional to the number of advertising impressions shown using that data. Most data providers give buyers a choice of the pricing criterion.

but are equally representative of many online and offline markets for personal information.

In all these markets, a general picture emerges where an advertiser acquires very detailed information about a segment of "targeted" consumers, and is rather uninformed about a larger "residual" set. This kind of information structure, together with the new advertising opportunities, poses a number of economic questions. How is the advertisers' willingness to pay for information determined? Which consumers should they target? How should a data provider price its third-party data? How does the structure of the market for data (*e.g.*, competition among sellers, data exclusivity) affect the equilibrium price of information? More specifically to online advertising markets, what are the implications of data sales for the revenues of large publishers of advertising space?

In this paper, we explore the role of data providers on the price and allocation of consumer-level information. We provide a framework that addresses general questions about the market for data and contributes to our understanding of recent practices in online advertising. We develop a simple model of data pricing that captures the key trade-offs involved in selling the information encoded in third-party cookies. However, our model also applies more broadly to markets for consumer-level information, and it is suited to analyze several offline channels as well.

The model considers heterogeneous consumers and firms. The (potential) surplus is given by a function that assigns a value to each realized match between a consumer and a firm (the match value function). The match values differ along a purely horizontal dimension, and may represent a market with differentiated products. In order to realize the potential match value, each firm must "invest" in contacting consumers. An immediate interpretation of the investment decision is advertising spending that generates contacts and eventually sales. We refer to the "advertising technology" as the rate at which investment into contacts generates actual sales, and to a "cookie" as the information required to tailor advertising spending to specific consumers.

We maintain the two distinguishing features of selling cookies (individual queries and per-user "bit" pricing) as the main assumptions. These assumptions can be stated more precisely as follows: (a) *Individual queries are for sale.* We allow advertisers to purchase information on individual consumers. This enables advertisers to segment users into a targeted group that receives personalized levels of advertising, and a residual group that receives a uniform level of advertising (possibly zero). More formally, this means the information structures available to an advertiser are given by specific partitions of the space of match values. (b) *Individual queries are priced separately.* We restrict the data provider to set a uniform unit price, so that the payment to the data provider is proportional to the number of users ("cookies") acquired.

There exist, of course, other ways to sell information, though linear pricing of cookies is a natural starting point. We address these variations in extensions of our baseline model. In particular, we explore alternative mechanisms for selling information, such as bundling and nonlinear pricing of data.

In Section II, we characterize the advertisers' demand for information for a given price of data. We establish that advertisers purchase information on two convex sets of consumers, specifically those with the highest and lowest match values. Advertisers do not buy information about every consumer. Instead, they estimate the match value within the residual group of consumers, and they exclude a convex set in order to minimize the prediction error. Under further conditions, the data-buying policy takes the form of a single cutoff match value. However, advertisers may buy information about all users above the cutoff value (*positive targeting*) or below the cutoff value (*negative targeting*). Each of these data-buying policies alleviates one potential source of advertising mismatch: wasteful spending on low-value matches, and insufficient intensity on high-value matches. The optimality of positive vs. negative targeting depends on the advertising technology and on the distribution of match values, *i.e.*, on properties of the complete information profit function alone.

The advertising technology and the distribution of match values have implications for the cross-price externalities between the markets for data and advertising. In particular, a consistent pattern emerges linking the advertisers' preferences for positive vs. negative targeting and the degree to which a publisher of advertising space benefits from the availability of consumer-level data.

In Section III, we turn to the data provider's pricing problem. We first examine the subtle relationship between the price of cookies and the cost of advertising. The cost of advertising reduces both the payoff advertisers can obtain through better information, and their payoff if uninformed. The overall effect on the demand for cookies and on the monopoly price is, in general, non-monotone. In a leading example, we establish that the price for cookies is single-peaked in the cost of advertising. This suggests which advertising market conditions may be more conducive for the data provider.

We then examine the role of market structure on the price of cookies. Surprisingly, concentrating data sales in the hands of a single data provider is not necessarily detrimental to social welfare. Formally, we consider a continuum of information providers, each one selling one signal exclusively. We find that prices are higher under data-sales fragmentation. The reason for this result is that exclusive sellers ignore the negative externality that raising the price of information about one consumer imposes on the demand for information about all other consumers. A similar mechanism characterizes the effects of an incomplete database, sold by a single firm. In that case, the willingness to pay for information increases with the size of the database, but the monopoly price may, in fact, decrease. This is contrast with the effect of a more accurate database.

In Section IV, we enrich the set of pricing mechanisms available to the data provider. In particular, in a binary-action model, we introduce nonlinear pricing of information structures. We show that the data provider can screen vertically heterogeneous advertisers by offering subsets of the database at a decreasing marginal price. The optimal nonlinear price determines exclusivity restrictions on a set of "marginal" cookies: in particular, second-best distortions imply that some cookies that would be profitable for many advertisers are bought by only by a small subset of high-value advertisers.

The issue of optimally pricing information in a monopoly and in a competitive market has been addressed in the finance literature, starting with seminal contributions by Admati and Pfleiderer (1986), Admati and Pfleiderer (1990) and Allen (1990), and more recently by García and Sangiorgi (2011). A different strand of the literature has examined the sale of information to competing parties. In particular, Sarvary and Parker (1997) model information-sharing among competing consulting companies; Xiang and Sarvary (2013) study the interaction among providers of information to competing clients; Iyer and Soberman (2000) analyze the sale of heterogeneous signals, corresponding to valuable product modifications, to firms competing in a differentiated-products duopoly; Taylor (2004) studies the sale of consumer lists that facilitate price discrimination based on purchase history; Calzolari and Pavan (2006) consider an agent who contracts sequentially with two principals, and allow the former to sell information to the latter about her relationship (contract offered, decision taken) with the agent. All of these earlier papers only allow for the complete sale of information. In other words, they focus on signals that revealed (noisy) information about all realizations of a payoff-relevant random variable. The main difference with our paper's approach is that we focus on "bit-pricing" of information, by allowing a seller to price each realization of a random variable separately.

The literature on the optimal choice of information structures is rather recent. Bergemann and Pesendorfer (2007) consider the design of optimal information structures within the context of an optimal auction. There, the principal controls the design of both the information and the allocation rule. More recently, Kamenica and Gentzkow (2011) consider the design of the information structure by the principal when the agent will take an independent action on the basis of the received information. In contrast to the persuasion literature, we endogenize the agent's information cost by explicitly analyzing the monopoly pricing of information rather than directly choosing an information structure.

In related contributions, Anton and Yao (2002), Hörner and Skrzypacz (2012), and Babaioff, Kleinberg and Paes Leme (2012) derive the optimal mechanism for selling information about a payoff-relevant state, in a principal-agent framework. Anton and Yao (2002) emphasize the role of partial disclosure; Hörner and Skrzypacz (2012) focus on the incentives to acquire information; and Babaioff, Kleinberg and Paes Leme (2012) allow both the seller and the buyer to observe private signals. Finally, Hoffmann, Inderst and Ottaviani (2014) consider targeted advertising as selective disclosure of product information to consumers with limited attention spans.

The role of specific information structures in auctions, and their implication for online advertising market design, are analyzed in recent work by Abraham et al. (2014), Celis et al. (2014), and Syrgkanis, Kempe and Tardos (2013). All three papers are motivated by asymmetries in bidders' ability to access additional information about the object for sale. Ghosh et al. (2012) study the revenue implications of cookie-matching from the point of view of an informed seller of advertising space, uncovering a trade-off between targeting and information leakage. In earlier work, Bergemann and Bonatti (2011), we analyzed the impact that changes in the information structures, in particular the targeting ability, have on the competition for advertising space.

I Model

A Consumers, Advertisers, and Matching

We consider a unit mass of uniformly distributed consumers (or "users"), $i \in [0,1]$, and advertisers (or "firms"), $j \in [0,1]$. Each consumer-advertiser pair (i,j) generates a (potential) match value for the advertiser j:

$$v : [0,1] \times [0,1] \to V, \tag{1}$$

with $v(i,j) \in V = [\underline{v}, \bar{v}] \subseteq \mathbb{R}_+$.

Advertiser j must take an action $q_{ij} \geq 0$ directed at consumer i to realize the potential match value $v(i,j)$. We refer to q as the *match intensity*. We abstract from the details of the revenue-generating process associated to matching with intensity q. The complete-information profits of a firm generating a match of intensity q with a consumer of value v are given by

$$\pi(v,q) \triangleq vq - c \cdot m(q). \tag{2}$$

The *matching cost function* $m : \mathbb{R}_+ \to \mathbb{R}_+$ is assumed to be increasing, continuously differentiable, and convex. In the context of advertising, q corresponds to the probability of generating consumer i's awareness about firm j's product. Awareness q is generated by buying an amount of advertising space $m(q)$, which we assume can be purchased at a unit price $c > 0$. If consumer i is made aware of the product, he generates a net present value to the firm equal to $v(i, j)$.

B Data Provider

Initially, the advertisers do not have information about the pair-specific match values $v(i, j)$ beyond the common prior distribution described below. By contrast, the monopolistic data provider has information relating each consumer to a set of characteristics represented by the index i, and each advertiser to a set of characteristics represented by the index j. The *database* of the data provider is simply the mapping (1) relating the characteristics (i, j) to a value of the match $v(i, j)$, essentially a large matrix with a continuum of rows (representing consumers) and columns (representing firms).

Advertisers can request information from the data provider about consumers with specific characteristics i. Now, from the perspective of advertiser j the only relevant aspect of the consumer's characteristic i is the match value $v(i, j)$. Thus, we refer to *cookie* v as the information necessary for advertiser j to identify all consumers with a realized match value $v = v(i, j)$. Similarly, we refer to *query* v as the request by advertiser j to identify all consumers in the database with characteristics i such that $v = v(i, j)$.[4]

Advertisers purchase information from the data provider in order to target their spending. For example, if advertiser j wishes to tailor his action q to all consumers with value v, then he queries for the identity of all consumers with characteristics i such that $v = v(i, j)$. More generally, each advertiser j can purchase information about any subset of consumers with match values $v \in A_j \subset V$. Thus, if advertiser j makes a query v (*i.e.*, purchases the cookie v), then the value $v = v(i, j)$ belongs to the set A_j, and the advertiser can target consumers with value v with a tailored level of match intensity. For this reason, we refer to the sets A_j and A_j^{C} as the *targeted set* and the *residual set* (or complementary set), respectively.

We assume that the data about individual consumer is sold at a constant linear price p per cookie.[5]

[4] A query v to the database thus requests the information contained in the cookie v, and in this sense we can use cookie and query as synonyms. To be precise, the cookie is the information technology that allows the database to record the characteristics of consumer i and the query retrieves the information from the database.

[5] This assumption reflects the pricing of data "per unique user" (also known as "cost per stamp"). It also matches the offline markets for data, where the price of mailing lists, or lists of credit scores is related to the number of user records.

C Distribution of Match Values

The (uniform) distribution over the consumer-firm pairs (i, j) generates a distribution of values through the match value function (1). For every measurable subset A of values in V, the resulting measure μ is given by:

$$\mu(A) \triangleq \int_{\{i,j\in[0,1]|v(i,j)\in A\}} \mathrm{d}i\mathrm{d}j.$$

Let the interval of values beginning with the lowest value be $A^v \triangleq [\underline{v}, v]$. The associated distribution function $F : V \to [0,1]$ is defined by

$$F(v) \triangleq \mu(A^v).$$

By extension, we define the *conditional* measure for every consumer i and every firm j by

$$\mu_i(A) \triangleq \int_{\{j\in[0,1]|v(i,j)\in A\}} \mathrm{d}j, \quad \text{and} \quad \mu_j(A) \triangleq \int_{\{i\in[0,1]|v(i,j)\in A\}} \mathrm{d}i,$$

and the associated conditional distribution functions $F_i(v)$ and $F_j(v)$. We assume that the resulting match values are identically distributed across consumer and across firms, *i.e.*, for all i, j, and v:

$$F_i(v) = F_j(v) = F(v).$$

Thus, $F(v)$ represents the common prior distribution for each firm and each consumer about the match values. Thus the price of the targeted set A_j is given by

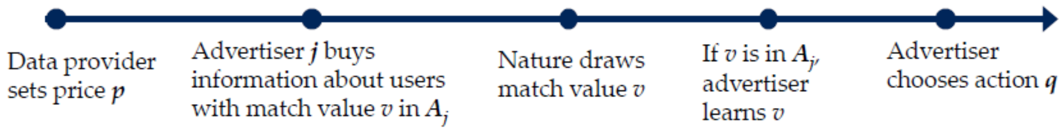
$$p(A_j) \triangleq p \cdot \mu(A_j). \tag{3}$$

Prominent examples of distributions that satisfy our symmetry assumption include: i.i.d. match values across consumer-firm pairs; and uniformly distributed firms and consumers around a unit-length circle, where match values are a function of the distance $|i - j|$. In other words, match values differ along a purely horizontal dimension. This assumption captures the idea that, even within an industry, the same consumer profile can represent a high match value to some firms and a low match value to others firms. This is clearly true for consumers that differ in their geographical location, but applies more broadly as well.[6]

[6] Consider the case of credit-score data: major credit card companies are interested in reaching consumers with high credit-worthiness; banks that advertise consumer credit lines would like to target individuals with average scores, who are cash-constraint, but unlikely to default; and subprime lenders such as used car dealers typically cater to individuals with low or non-existing credit scores, see Adams, Einav and Levin (2009) for a description and model of subprime lending.

Figure 1 summarizes the timing of our model.

FIGURE 1: TIMING



We note that the present model does not explicitly describe the consumer's problem and the resulting indirect utility. To the extent that information facilitates the creation of valuable matches between consumers and advertisers, as a first approximation, the indirect utility of the consumer may be thought of as co-monotone with the realized match value v. In fact, with the advertising application in mind, we may view q as scaling the consumer's willingness to pay directly, or as the amount of advertising effort exerted by the firm, which also enters the consumer's utility function. Thus, the profit function in (2) is consistent with the informative, as well as the persuasive and complementary views of advertising (see Bagwell, 2007).

A more elaborate analysis of the impact of information markets on consumer surplus and on the value of privacy would probably have to distinguish between information that facilitates the *creation of surplus*, which is focus of present paper, and information that impacts the *distribution of surplus*. For example, additional information could improve the pricing power of the firm and shift surplus from the consumer to the firm (as for example in Bergemann, Brooks, and Morris, 2013).

II Demand for Information

The value of information for each advertiser is determined by the incremental profits they could accrue by purchasing more cookies. Advertiser j is able to perfectly tailor his advertising spending to all consumers included in the targeted set A_j. In particular, we denote the complete information demand for advertising space $q^*(v)$ and profit level $\pi^*(v)$ by

$$\begin{aligned} q^*(v) &\triangleq \arg\max_{q\in\mathbb{R}_+}[\pi(v,q)], \\ \pi^*(v) &\triangleq \pi(v,q^*(v)). \end{aligned}$$

By contrast, for all consumers in the complement (or residual) set A_j^{C}, advertiser j must form an expectation over $v(\cdot,j)$, and choose a constant level of q for all such consumers.

Because the objective $\pi(v,q)$ is linear in v, the optimal level of advertising $q^*(A_j^{\mathrm{C}})$ is given by

$$q^*(A_j^{\mathrm{C}}) \triangleq \underset{q\in\mathbb{R}_+}{\arg\max}\mathbb{E}\left[\pi(v,q) \mid v \in A_j^{\mathrm{C}}\right] = q^*\left(\mathbb{E}\left[v \mid v \notin A_j\right]\right).$$

We can represent each advertiser's information acquisition problem as the choice of a measurable subset A of the set of match values V:

$$\max_{A\subseteq V}\left[\int_A \left(\pi\left(v, q^*\left(v\right)\right) - p\right)\mathrm{d}F\left(v\right) + \int_{A^{\mathrm{C}}} \pi(v, q^*(A^{\mathrm{C}}))\mathrm{d}F\left(v\right)\right], \tag{4}$$

where, by symmetry, we can drop the index j for the advertiser.

By including all consumers with match value v into the targeted set A, the advertiser can raise his gross profits from the uninformed choice to the informed choice of q, albeit at the unit cost p per consumer. In problem (4), the total price paid by the advertisers to the data provider is then proportional to the measure of the targeted set, or $p \cdot \mu(A)$. Next we characterize the properties of the optimal targeted set, as a function of the price of cookie p and of the cost of advertising c. We begin with a simple example.

A The Binary Action Environment

We start with linear matching costs and uniformly distributed match values; we then generalize the model to continuous actions and general distributions. Formally, let $F(v) = v$, with $v \in [0,1]$ and $c \cdot m(q) = c \cdot q$, with $q \in [0,1]$. The linear cost assumption is equivalent to considering a binary action environment, $q \in \{0,1\}$, as the optimal policy will only take those two values.

In this simplified version of the model, targeting is very coarse: under complete information, it is optimal to contact a consumer v (*i.e.*, to choose $q^*(v) = 1$) if and only if the match value v exceeds the unit cost of advertising c. Thus, the complete information profits are given by

$$\pi^*(v) \triangleq \max\{v - c, 0\}. \tag{5}$$

Likewise, the optimal action on the residual set is given by:

$$q^*(A^{\mathrm{C}}) = 1 \iff \mathbb{E}[v \mid v \in A^{\mathrm{C}}] \geq c.$$

As we show in Proposition 1, advertisers adopt one of two mutually exclusive strategies to segment the consumer population: (i) *positive targeting* consists of buying information on the highest-value consumers, contacting them and excluding everyone else; (ii) *negative targeting* consists of buying information on the lowest-value consumers, avoiding them and contacting everyone else. That is, advertisers choose a *constant* action $q \in \{0,1\}$ on the targeted set A and a *different constant* action on the residual set A^{C}. The actions differ across the targeted and the residual set as information about consumer v has a positive value only if it affects the advertiser's subsequent action.

The choice of the optimal targeting strategy and the size of the targeted set naturally depend on the cost of contact c and on the price of information p. We denote the optimal targeted set by $A(c,p)$. This set is defined by a threshold value v^* that either determines a lower interval $[\underline{v}, v^*]$ or an upper interval $[v^*, \bar{v}]$, depending on the optimality of either negative or positive targeting, respectively. The optimality of a threshold strategy follows from the monotonicity of the profit in v and the binary action environment.

We identify the size of the targeted set by consider the willingness to pay for the marginal cookie under each targeting strategy. If the advertiser adopts positive targeting, then he purchases information on all consumers up to the threshold v^* that leaves him with nonnegative net utility, or $v^* = c + p$. Conversely, if the advertiser adopts negative targeting, then at the marginal cookie, the gain from *avoiding* the contact, and thus saving $c - v$, is just offset by the price p of the cookie, and thus $v^* = c - p$. Under either targeting strategy, the advertiser trades off the magnitude of the error made on the residual set with the cost of acquiring additional information. Proposition 1 characterizes the optimal targeting strategy.

Proposition 1 (Targeting Strategy)
For all $c, p > 0$, the optimal targeted set $A(c,p)$ is the interval of values v given by:

$$A(c,p) = \begin{cases} [0, \max\{c-p, 0\}] & \text{if} \quad c < 1/2; \\ [\min\{c+p, 1\}, 1] & \text{if} \quad c \geq 1/2. \end{cases}$$

If the cost of advertising, *i.e.* the matching cost c, is particularly high, it is only profitable to bear the costs of generating awareness through advertising for very high-value customers, about which information is acquired from the data provider. Conversely, for low costs of advertising, all customers but the very low-value ones are profitable, about which information is purchased in order to exclude them from advertising.[7]

Proposition 1 establishes that the residual and the targeted set are both connected sets (intervals), and that advertisers do not buy information about every consumer. The binary environment illustrates some general features of optimal targeting and information policies. In particular, three implications of Proposition 1 extend to general settings: (a) the residual set is non-empty; (b) advertisers do not necessarily buy the cookies of high-value consumers;

[7] The value $c = 1/2$ of the threshold which determines the choice of targeting strategy happens to coincide with the threshold value that would determine whether advertisers contact all consumers, or none, under the prior information. This is a special feature of the uniform distribution.

and (c) the cost c of the advertising space guides their strategy. At the same time, the binary environment cannot easily capture several aspects of the model, including the following: the role of the distribution of match values (and of the relative size of the left and the right tail in particular); the role of precise tailoring and the need for more detailed information; the determinants of the advertisers' optimal targeting strategy; and the effect of the cost of advertising on the demand for information.

B The Continuous Action Environment

We now proceed to analyze the general version of our model, in which we consider a continuum of actions and a general distribution of match values. It is helpful to first describe the demand for advertising space when the value of the match v is known to the advertisers. Thus, we introduce the *complete information* decision and profits. We now allow for a general differentiable, increasing and strictly convex cost function $m(q)$ and assume that $m'(0) = 0$. This implies that the complete information demand for advertising is positive for all match values.

The complete information demand for advertising space, denoted by $q^*(v)$, is characterized by the first-order condition:

$$v = cm'(q^*(v)). \tag{6}$$

By contrast, if the advertiser has access to the distribution $F(v)$ only, the *prior-information* demand for advertising space $\bar{q}$ is given by

$$\bar{q} \triangleq q^*(A^{\mathrm{C}} = V) = q^*(\mathbb{E}[v]). \tag{7}$$

More generally, given a targeted set A, the optimal advertising level on the residual set A^{C} satisfies the following condition:

$$\mathbb{E}[v \mid v \in A^{\mathrm{C}}] = cm'(q^*(A^{\mathrm{C}})). \tag{8}$$

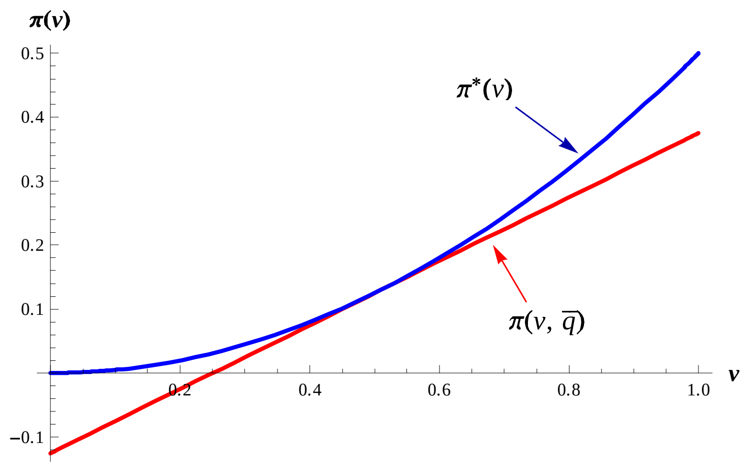
Thus, the continuous-action model has the two key features that advertisers: (a) differentiate spending levels within the targeted set, and (b) choose a uniform (strictly positive) advertising level for the residual set. Moreover, the optimal advertising level on the residual set $q^*(A^{\mathrm{C}})$ varies with the composition of the targeted set.[8]

It follows from (6) and from the strict convexity of $m(q)$ that the complete information demand $q^*(v)$ is strictly increasing. Since by the Envelope Theorem, $d\pi^*(v)/dv = q^*(v)$, the

[8] These advertising policies might arguably represent the choices of a large brand marketer who wishes to fine-tune spending on a group of consumers, while adopting "umbrella spending" on everyone else.

realized *complete information* profit $\pi^*(v)$ is strictly convex in v. In contrast, the realized profit under *prior information* is linear in v, and it is given by $\pi(v, \bar{q})$. Figure 2 describes the profit function under complete information $\pi^*(v)$ and prior information $\pi(v, \bar{q})$.

FIGURE 2: COMPLETE INFORMATION AND PRIOR INFORMATION PROFITS



As intuitive, under prior information, the firm chooses excessive (wasteful) advertising to low-value consumers and insufficient advertising to higher-value consumers. The firm therefore has a positive willingness to pay for information, *i.e.*, for cookies. The value of information for every match value v is visually described by the difference between the complete information and the prior information profit function:

$$\pi^*(v) - \pi(v, \bar{q}). \tag{9}$$

Figure 2 suggests that the value of information is highest for extreme match values.[9] Consequently, the next result establishes the optimality of a convex *residual set* of cookies. Each advertising firm purchases all cookies in a set:

$$A = [\underline{v}, v_1] \cup [v_2, \bar{v}].$$

The value of the lower and upper threshold are determined by c and p, thus $v_1 \triangleq v_1(c,p)$ and $v_2 \triangleq v_2(c,p)$, respectively. Proposition 2 confirms the intuition that the value of information is lowest for intermediate match values and highest for match values on the tails.

Proposition 2 (Convexity of Residual Set)
For all $c, p > 0$, the optimal residual set $A^C(c,p)$ is a non-empty interval $[v_1(c,p), v_2(c,p)]$.

[9] In this example, $cm(q) = q^2/2$, and $F(v) = v$, $v \in [0,1]$.

Proposition 2 allows us to rewrite the firm's problem (4) as the choice of two thresholds, v_1 and v_2, that define the targeted and residual sets, in terms of the gains relative to the complete information solution:

$$\begin{aligned} &\max_{v_1,v_2} \int_{v_1}^{v_2} \left[\pi\left(v, q^*\left([v_1, v_2]\right)\right) - \pi^*\left(v\right) + p\right] \mathrm{d}F\left(v\right), \\ &\text{s.t. } cm'\left(q^*\left([v_1, v_2]\right)\right) = \mathbb{E}\left[v \mid v \in [v_1, v_2]\right]. \end{aligned} \tag{10}$$

In program (10), as the bounds of the residual set are stretched (*e.g.*, as v_1 decreases), the advertiser earns a marginal benefit of p and incurs a marginal cost of $\pi^*\left(v_1\right) - \pi^*\left(v_1, q^*([v_1, v_2])\right)$. In addition, the advertiser adjusts the optimal action on the residual set to take the new inference problem into account. (Of course, this has no first-order effect on profits at the optimum.) The average match value $\mathbb{E}\left[v \mid v \in [v_1, v_2]\right]$ determines the demand for advertising space in the residual set $q^*\left([v_1, v_2]\right)$, which in turn affects the value of information.

Above, we described the value of information as the difference between the profit of an informed and an uninformed advertiser $\pi^*\left(v\right) - \pi\left(v, \bar{q}\right)$. This revenue comparison is conditional on the realization of the value v, and it is thus an ex-post comparison. For the complete determination of the optimal policy, the advertiser has to evaluate how large these gains from information are from an *ex-ante* point of view. The advertiser therefore has to weigh the likelihood of different realizations, represented by the distribution $F\left(v\right)$ of values, and the gains from responding to the information, represented by the convexity of the matching cost function $m\left(q\right)$. To understand the exact nature of these trade-offs, it is useful to begin with a "symmetric" environment for $F\left(v\right)$ and $m\left(q\right)$. In the context of negative vs. positive targeting, this corresponds to a symmetric distribution $F\left(v\right)$ around the mean $\mathbb{E}\left[v\right]$ and a quadratic matching cost function $m\left(q\right)$, such as in the example of Figure 2.

C Joint vs. Exclusive Targeting

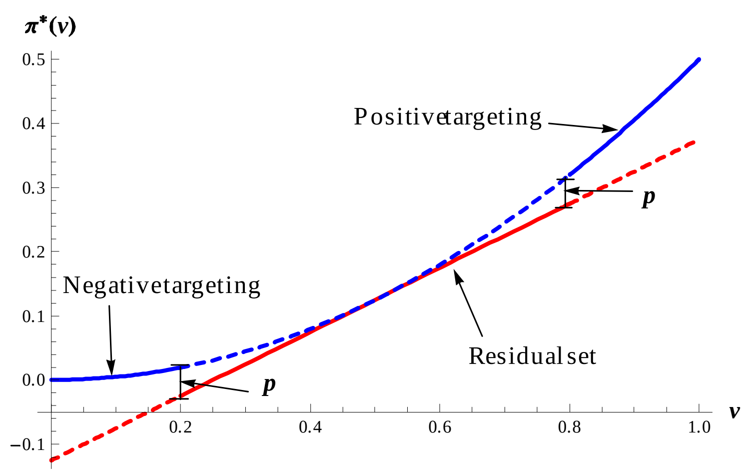
When matching costs are quadratic and match values are symmetrically distributed, advertisers always choose to target both low- and high-valuation consumers. In addition, under these symmetry conditions, the residual set (*i.e.*, the set of excluded valuations) is an interval centered on the prior mean $\mathbb{E}\left[v\right]$. With a quadratic matching cost function, the optimal complete information matching intensity is linear in v, or $q^*\left(v\right) = v/c$. Moreover, the *gains from information* relative to the optimal matching policy for the mean value $q^*\left(\mathbb{E}\left[v\right]\right)$ are identical for values equidistant from the mean, regardless of whether they are below or above the mean. The value of information arises from adjustments of the matching intensity relative to the mean, *i.e.*, increasing the matching intensity for values above the mean and decreasing the matching intensity for values below the mean. As the curvature of the cost function is constant in q when $m(q)$ is quadratic, this symmetry argument holds under any symmetric distribution $F(v)$. Proposition 3 verifies the above intuition.

Proposition 3 (Joint Targeting: Positive and Negative)
With symmetrically distributed match values and quadratic matching costs, the optimal residual set is given by:

$$A^C(c,p) = [\mathbb{E}[v] - 2\sqrt{cp},\ \mathbb{E}[v] + 2\sqrt{cp}].$$

The measure of the residual set is increasing in the product of the price of information p and the cost parameter c. Thus, an increase in either one depresses the number of cookies acquired, and shrinks the targeted set by expanding the residual set toward the tails of the distribution. Figure 3 illustrates the demand for cookies and the resulting profit levels in the quadratic environment. The bold segment represent the active policy for value v, the dashed line the inactive policy for value v.

Figure 3: Positive and Negative Targeting



The symmetry conditions introduced in Proposition 3 have important implications not only for the optimal location of the residual set, but also for its size. In particular, the expected match value in the residual set is equal to the prior mean $\mathbb{E}[v]$, regardless of the measure of the residual set A^C. Therefore, the quantity of signals purchased by the advertiser does not influence the uninformed action $\bar{q}$, and hence it does not affect the marginal value of information at any given v. This also implies that the willingness to pay for information about any consumer v is independent of the distribution of match values.

In turn, the interaction between the symmetric gains from information and the symmetry in the distribution suggest conditions under which either only positive or only negative targeting become optimal, as we establish in the next set of results.

While the residual set is always connected, as established by Proposition 2, the targeted set may be as well. In particular, the choice of a single (positive or negative) targeting policy depends on the value of information, and on its monotonicity properties over any interval. Proposition 4 establishes sufficient conditions under which firms demand cookies in a single interval, *i.e.*, they choose positive or negative targeting only.

Proposition 4 (Exclusive Targeting: Positive or Negative)

1. *If $m''(q)$ and $f(v)$ are decreasing, positive targeting is optimal:*

$$A(c,p) = [v_2(c,p), \bar{v}], \text{ and } v_2 > \underline{v}.$$

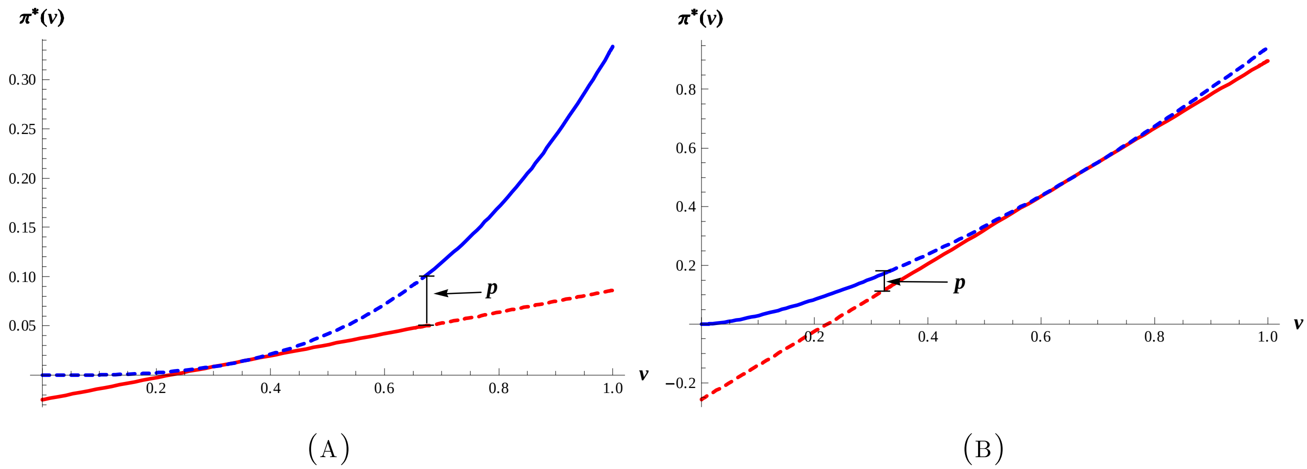
2. *If $m''(q)$ and $f(v)$ are increasing, negative targeting is optimal:*

$$A(c,p) = [\underline{v}, v_1(c,p)], \text{ and } v_1 < \bar{v}.$$

The sufficient conditions in Proposition 4 for exclusive targeting are perhaps best understood when viewed as departures from the symmetric conditions of Proposition 3. If, say, positive targeting is to dominate negative targeting, then the gains from information must be larger on the upside than on the downside of values. Recall that the gains from information given the realization v are equal to $\pi^*(v) - \pi(v, \bar{q})$. Thus, if the curvature of the matching cost function $m''(q)$ is decreasing, the gains from information for realizations v equidistant from the mean $\mathbb{E}[v]$ are larger above the mean than below. Now, this pairwise comparison and reasoning could be undone by the relative likelihood of these two events. Thus, for the sufficient conditions, we need to guarantee that the distribution of values supports this pairwise argument, and hence the corresponding monotonicity requirement on the density $f(v)$. Figure 4 shows the equilibrium profit levels under positive targeting (A) and negative targeting (B).[10]

[10] In both panels, $F(v) = v$, $v \in [0,1]$ and $m(q) = q^b/b$. In panel (A), $b = 3/2$, and in panel (B), $b = 3$.

Figure 4: Positive or Negative Targeting



The optimality of targeting consumers in a single interval can be traced back to the two sources of the value of information, *i.e.*, wasteful advertising for low types and insufficient advertising for valuable consumers. Proposition 4 relates the potential for mismatch risk to the properties of the match cost function. In particular, when the curvature of the matching cost function is increasing, it becomes very expensive to tailor advertising to high-value consumers. In other words, the risk of insufficient advertising is not very high, given the cost of advertising space. The firm then purchases cookies related to lower-valued consumers.[11]

When choosing a targeting strategy, the advertiser trades off the amount of learning over values in the residual set with the costs and benefits of acquiring information about values in the targeted set. The amount of learning is related to the range of the residual set $|v_2 - v_1|$, while the costs and benefits of information are related to the probability measure of the targeted set. Therefore, targeting a less likely subset of values requires a smaller expense (in terms of the cost of cookies) in order to generate a given amount of information. The distribution of match values then affects the optimality of positive vs. negative targeting: for example, under a matching cost function with constant curvature, decreasing density $f(v)$ leads to positive targeting, and vice-versa.

D Empirical Relevance

Both positive and negative targeting strategies are relevant for online advertising markets. In particular, negative targeting is explicitly allowed as a refinement option by most large

[11]Examples of matching cost functions with concave marginal costs include power functions, $m(q) = q^a$ with $a < 2$. Examples of convex marginal costs include those derived from the Butters (1977) exponential matching technology, *i.e.*, $m(q) = -a \ln(1-q)$, with $a > 0$, and power functions $m(q) = q^a$, with $a > 2$.

providers of advertising space, including Google, Yahoo!, and Facebook.[12] Clearly, an advertiser may adopt either or both strategies, and the choice of a strategy in any specific context will depend on the distribution of consumer values and on the cost of advertising. For instance, in the market for credit scores, a credit card company may want to acquire the profiles of consumers with the lowest scores, and make sure *not* to reach out to them; or it may select a small group of high credit-worthiness consumers, and reach out to them more aggressively.

Within the marketing and economics literature, Pancras and Sudhir (2007) document the use of both positive and negative targeting in the context of retail shopping. While the main focus of Pancras and Sudhir (2007) is on competition and information-sharing among catalogue merchants and manufacturers, they also examine the pricing of data by several intermediaries.[13]

More recent studies provide indirect evidence in favor of adopting negative targeting to exploit the consumers' purchase cycle. For example, in the context of sponsored-search advertising, Blake, Nosko and Tadelis (2013) document that eBay obtains a positive return on investment only for consumers who have not visited the eBay site in the last two months. A similar pattern for the profitability of different customers also appears in the case of (offline) direct-marketing companies documented by Anderson and Simester (2013). In both contexts, a cost-efficient strategy for retailers consists of acquiring information about consumers with recent purchases and appropriately reducing the amount of advertising directed at them. These consumers are both low-value (at this point in their purchase cycle) and low in number, relative to the overall population, which makes negative targeting especially profitable.[14]

Finally, as real-time bidding makes online data markets more integrated with the advertising exchanges, we can identify two contrasting forces in terms of our model. On the one hand, the combined sale of data and advertising favors positive targeting almost by construction. On the other hand, when the cost of the data is tied to the price paid for advertising, contacting high-value consumers becomes increasingly costly. If targeting through cookies results in a higher marginal cost of advertising, advertisers may specify lower bids for selected consumer segments (*i.e.*, adopt negative targeting) in order to reduce their total expenditure.

[12] For example, Facebook offers negative targeting based on third-party data as a "custom audience" selection criterion for advertisers. See https://developers.facebook.com/docs/ads-api/targeting and "Buy Signal: Facebook Widens Data Targeting," *The Wall Street Journal*, April 9, 2013.

[13] Our model so far abstracts from competition among advertisers. In the working paper Bergemann and Bonatti (2013), we introduce pecuniary externalities through a market-clearing price of advertising space. The characterization of the optimal targeting strategies of Proposition 4 is unchanged.

[14] While advertisers may be able to identify their own repeat shoppers, they need to purchase third-party information about their competitors' customers who are at a similar stage in their purchase cycle.

E Implications for Publishers

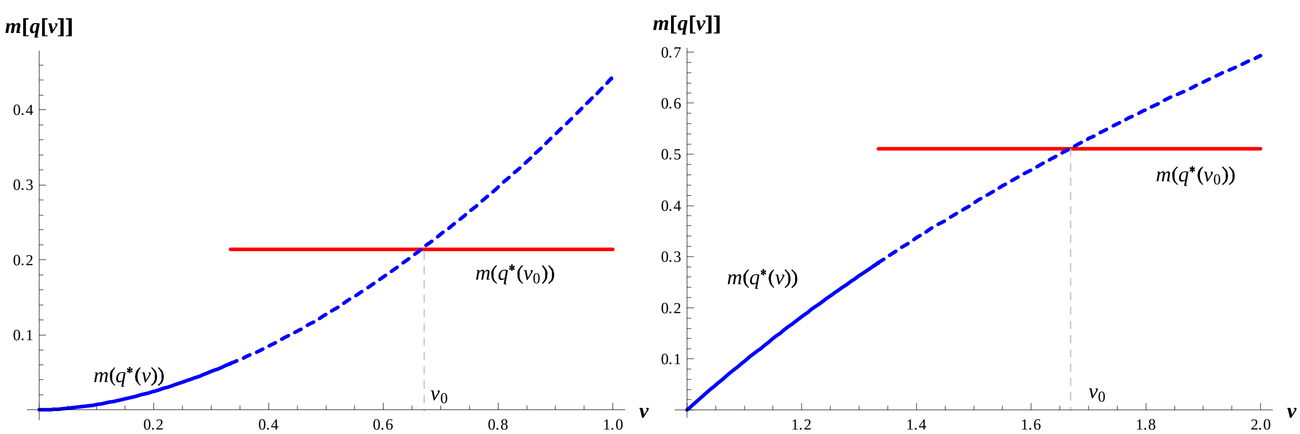
We conclude this section by examining the interaction between the markets for data and online advertising. In particular, we assess the effect of data sales on the demand for advertising space and the implications of vertical integration between publishers and data providers.

The effect of the price of data on the total demand for advertising space is unclear *a priori*. For instance, the demand for advertising space may increase or decrease in the amount of information available to advertisers, depending on whether the data is used for positive or negative targeting. To formalize this trade-off, consider the total demand for advertising space as a function of the targeted set $A(c, p)$. Because any advertiser who wishes to generate match intensity q with a consumer must purchase an amount of space equal to $m(q)$, the total demand for advertising is given by

$$M(A) \triangleq \int_A m(q^*(v))\,\mathrm{d}F(v) + \int_{A^{\mathrm{C}}} m(q^*(A^{\mathrm{C}}))\mathrm{d}F(v). \tag{11}$$

We are interested in the effect of the amount of data sold $\mu(A)$ on the total demand for advertising $M(A)$. Figure 5 considers the case of negative targeting, and compares the demand for advertising $m(q(v))$ for fixed targeted and residual sets, under two different matching cost functions.

Figure 5: Total Demand for Advertising



$m(q) = q^b/b, \quad b = 9/4$ $\qquad$ $m(q) = -\ln(1-q)$

As is intuitive, the total demand for advertising (*i.e.*, the area under the solid lines in Figure 5) is increasing in the measure of the targeted set A when the complete information demand for advertising $m(q^*(v))$ is convex in v. Our next result formalizes the interaction of the data and advertising markets by relating the sign of the cross-market externality to the properties of the matching cost function. In Proposition 5 (as well as in Propositions 7, 8, and 9), we assume that the distribution of match values and the matching cost function lead to exclusive targeting (positive or negative). Proposition 4 provides sufficient conditions.

Proposition 5 (Market Interaction)
Assume exclusive (positive or negative) targeting is optimal.

1. *If $m'(q)$ is log-concave, the demand for advertising $M(A(c,p))$ is decreasing in p.*

2. *If $m'(q)$ is log-convex, the demand for advertising $M(A(c,p))$ is increasing in p.*

The proof of Proposition 5 establishes that convexity of the complete-information demand for advertising is equivalent, in terms of the primitives of our model, to the log-concavity of the marginal cost of matching. Furthermore, the conditions in Proposition 5 are related to those for the optimality of exclusive targeting (Proposition 4). In particular, if positive targeting is optimal, the demand for advertising space is decreasing in p (but not vice-versa).

Finally, we can leverage the results of Proposition 5 to analyze the problem a company (*e.g.*, Google, Yahoo!, or Facebook) that acts as both data provider (by providing information that allows targeted advertising) and publisher (by allowing advertisers to contact consumers). In particular, under the sufficient conditions of Proposition 5, the publisher wants to allow either complete access or no access to the data (corresponding to $p^* \in \{0, \infty\}$). In other words, our analysis suggests which market conditions are conducive to the wide diffusion of user-level information among the advertisers, and conversely which conditions discourage sellers from offering precise targeting opportunities. In particular, when the demand for advertising space is *decreasing* in p, a publisher with access to data can benefit from the *indirect* sale of information, *i.e.* from bundling information and advertising space in order to drive up demand for the latter.[15]

III The Price of Data

In this section, we explore the determinants of the monopoly price of data. We begin with the cost of advertising c, before turning to the fragmentation of data sales, the size of the database, and the precision of the data provider's information. In the latter three cases, we highlight the role of the residual set in determining the willingness to pay for information, and of the ability of the monopolist to influence its composition.

[15] We could also endow the publisher with market power, *i.e.*, allow the publisher and the data provider to coordinate their actions, without qualitatively affecting this result.

An important implication of the demand analysis in Section II is that the advertisers' optimal targeting strategy is not influenced qualitatively by the price of data p. In particular, under the conditions of Propositions 1, 3 or 4, the price of data affects the size of the targeted set only. In other words, throughout this section, the monopolist takes the shape of the targeted set $A(c,p)$ as given, and chooses the revenue-maximizing price

$$p^* = \arg\max_p \left[p \cdot \mu\left(A\left(c,p\right)\right)\right].$$

A Data and Advertising: Complements or Substitutes?

From the point of view of an advertiser, the data provider and the publisher of advertising space are part of a value chain. It is therefore tempting to view the interaction of the data provider and publisher as a vertical chain (formed by strategic complements), and to associate with it the risk of double marginalization. This would suggest that an increase in the cost c of advertising would lead optimally to a partially offsetting decrease in the price of information $p^*(c)$. But at closer inspection, the relationship between the price of data and that of advertising is more subtle.

The purchase of data may allow the advertiser to concentrate the purchase of advertising space on a smaller but highly relevant segment. Thus, from the point of view of the advertiser, the data provides an option whose value might be increasing as the advertising space becomes more expensive.

Therefore, data purchases act as strategic complements to advertising purchases for high value realizations, but as strategic substitutes for low valuations, because after learning of a low-value consumer, the advertiser reduces his matching intensity. This subtlety in the interaction already appeared in the binary environment of Subsection II.A, to which we now return. The following results are an immediate consequence of Proposition 1.

Proposition 6 (Data and Advertising)

1. *For all $c \in [0,1]$, the monopoly price of a cookie is:*

$$p^*(c) = (1/2)\min\{c, 1-c\}.$$

2. *The equilibrium sales of cookies are given by the targeted set $A(c,p^*(c))$:*

$$A\left(c,p^*(c)\right) = \begin{cases} [0, c/2] & \text{if} \quad c < 1/2; \\ [(1+c)/2, 1] & \text{if} \quad c \geq 1/2. \end{cases}$$

3. *The equilibrium price, sales and profits of the data provider are single peaked in c.*

Recall the characterization of the advertiser's optimal targeting strategy in the binary-action setting (Proposition 1): positive targeting is adopted when the cost of advertising c is sufficiently high and negative targeting when the cost of advertising is low. Proposition 6 shows that both the price of the data and the profits of the data provider are non-monotone in c. Intuitively, the value of information is highest for intermediate levels of c. In the absence of information, advertisers choose either $q_0 = 0$ or $q_0 = 1$, depending on the cost of the advertising space c. In particular, for very low *and* very high values of c, the availability of data modifies there optimal action only on a limited set of consumers. Consequently, the willingness to pay for information is also limited.

The binary-action environment suggests which market conditions are more conducive to the profitability of a data provider. Perhaps contrary to a first intuition, niche markets with a high cost of advertising space and few profitable consumers are not necessarily the best environment. While the availability of data would have a large impact (demands for advertising would be nil without information), the data provider's profits are constrained by the low levels of surplus downstream. Instead, markets with relatively large fractions of both profitable and unprofitable consumers yield a higher value of information, which translates into higher prices for data and higher provider profits.

While general results on the comparative statics of the monopoly price are harder to obtain in the continuous-action environment, more intuition can be obtained from specific examples. For instance, if joint targeting is optimal, an immediate implication of Proposition 3 is that the monopoly price is *inversely proportional* to the cost of advertising space. If exclusive targeting is optimal, and match values are uniformly distributed, the log-concavity (log-convexity) of $m'(q)$ is a sufficient condition for the monopoly price to increase (decrease) with c.[16] Taken together, these partial results suggest that the non-monotonicity of the monopoly price in Proposition 6 is deeply tied to *changes* in the optimal targeting strategy induced by changes in the cost of advertising space.

In the following subsections, we take the cost of advertising space as given. We return to our continuous-action environment in order to illustrate the role of the composition of the targeted and residual sets.

B Data Sales Fragmentation

We have so far assumed a monopoly structure for the data industry. We now assess the consequences of competition among sellers, and of the structure of the data industry. In

[16] In the working paper Bergemann and Bonatti (2013), we show that the conditions of Proposition 5 also determine the effect of the cost c on the advertisers' marginal willingness to pay for any targeted set A. The comparative statics of the monopoly price require further assumptions on the distribution of values.

particular, we focus on the externality that each seller's price imposes on the other sellers through the composition of the advertisers' residual set. Our formulation follows closely the business model of the *data exchange*, where a data provider does not buy and resell information, but rather offers a platform for matching individual buyers and sellers, who set their own prices.[17]

Formally, we consider a continuum of data sellers, and we assume that each seller has exclusive information about one consumer segment i. Thus, each seller sets the price for one cookie only. We seek to characterize a symmetric equilibrium of the pricing game. In the following discussion, we assume that positive targeting is optimal (Proposition 4 provides sufficient conditions). Analogous results hold for the case of negative targeting, as stated in Proposition 7.

We begin by considering an advertiser's demand for information. Suppose all sellers but j charge price p_{-j}. Every advertiser then chooses the targeted set $A = [v_2, \bar{v}]$ where the threshold value $v_2(p_{-j})$ solves the condition

$$p_{-j} = \pi^*(v_2) - \pi(v_2, q^*([\underline{v}, v_2])).$$

Thus, the cookie sold by seller j will have a distribution of values across advertisers that depends on the other sellers' prices through their effect on the residual set. In particular, a symmetric price profile p_{-j} can be summarized by the threshold v_2 that it induces. Now consider an advertiser whose match value with the cookie of seller j is equal to v. This advertiser's willingness to pay is equal to the differential profit under the threshold strategy $v_2(p_{-j})$. Therefore, seller j faces the inverse demand function $p(v, v_2)$ given by

$$p(v, v_2) \triangleq \pi^*(v) - \pi(v, q^*([\underline{v}, v_2])). \tag{12}$$

Because match values with a given seller $v(\cdot, j)$ are identically distributed, we can reformulate the seller's problem as choosing a threshold v to maximize profits given the advertisers' threshold v_2. A symmetric equilibrium threshold then solves the following problem:

$$v_2 = \arg\max_{v} [p(v, v_2)(1 - F(v))].$$

The key difference with the monopoly problem lies in the residual advertising intensity $q^*([\underline{v}, v_2])$, which cannot be influenced by the price of any individual seller. More precisely, suppose the monopolist considers expanding the supply of cookies, hence lowering the

[17] We may also interpret the fragmentation of data sales as a market where individual users are able to sell their own data.

threshold v_2. By increasing supply, he would reduce the gap between complete and prior information profits for the marginal consumer v_2. Naturally then, the monopolist would have to lower the price. At the same time, the composition of the residual set will have changed. In fact, the average value on the residual set will have decreased, and thus the advertising level on the residual set will be lower. But this means that the value of information for the marginal consumer just below the targeted set has increased, and hence the marginal advertiser just below the threshold will have a higher value of information. Now, this effect provides an additional incentive to lower prices and expand supply for the monopolist. Competing sellers do not internalize the positive externality present across cookie sales. Higher prices under fragmented data sales are then due to the absence of a *composition effect*.

The fragmented data sales is illustrative of a more general result. Suppose we were to consider n symmetric data sellers, each holding information about a measure $1/n$ of consumers distributed identically according to $F(v)$. The n sellers set prices simultaneously. Consider now the trade-off facing a specific seller. She knows that, by lowering her price, all advertisers will purchase more from her, as well as from everyone else. This occurs because the action q on the residual set will decrease. However, as for the case of fragmented sales, the composition effect is attenuated in equilibrium by the fact that all other sellers are holding their prices fixed. Thus, in a symmetric equilibrium the price is increasing in n. As the number of sellers grows large, the equilibrium price approaches the price under fragmentation, where the action on the residual set is constant.

To summarize, we obtain the following comparative statics of the equilibrium price.

Proposition 7 (Equilibrium under Data Fragmentation)
Assume exclusive (positive or negative) targeting is optimal.

1. *The symmetric equilibrium price of cookies with a continuum of data sellers $\bar{p}$ is higher than the monopoly price p^*.*
2. *The symmetric equilibrium price with n independent and exclusive data sellers $p^*(n)$ is increasing in n, and approaches $\bar{p}$ as $n \to \infty$.*

Clearly, if many sellers would lead to a duplication in the datasets, sellers would only be able to capture the incremental value of their information, thus driving prices down. In particular, there will exist a critical level of duplication for which the monopoly and the oligopoly prices are equal.

C Reach of the Database

So far, we implicitly assumed that the monopolist's dataset covers all consumers, *i.e.,* that it has maximal *reach*. We now explore the implications of limited reach on the monopoly price of cookies, and on the equilibrium profits of the data provider and the advertisers.

We assume that the data provider owns information about a fraction $\beta < 1$ of all consumers. Advertisers know the distribution of match values of consumers present in the database, and of those outside of it. In real-world data markets, consumers in a database may have different characteristics from those outside of it, *i.e.*, the presence of a cookie on a given consumer is *per se* informative. For simplicity, we assume that the two distributions are identical, so that the measure of consumers in the dataset is given by $\beta F(v)$. We then have the following result.

Proposition 8 (Reach and Demand)
Assume exclusive (positive or negative) targeting is optimal. Then the advertisers' marginal willingness to pay $p(A, \beta)$ is increasing in β for all A.

Quite surprisingly, the demand function for information shifts out as more consumers are reached by the database. That is, an advertiser *marginal* willingness to pay for information increases with β. The reason behind this result can be traced back to the effects of a larger database on the optimal action in the residual set $q^*(A^{\mathrm{C}})$. When positive targeting is optimal (so that $A = [v_2, \bar{v}]$), the average type in the residual set A^{C} is given by

$$\mathbb{E}[v \mid v \in A^{\mathrm{C}}] = \beta \mathbb{E}[v \mid v \leq v_2] + (1 - \beta) \mathbb{E}[v]. \tag{13}$$

Because the average type is decreasing in β for all A, the quantity of advertising demanded on the residual set is decreasing in β. Thus, the willingness to pay for information on the *marginal* consumer v_2 increases. A similar argument applies to the case of negative targeting. Conversely, under the conditions of Proposition 3 (joint targeting), the quantity of advertising $\bar{q}$ demanded on the residual set is independent of β. Hence, the reach parameter β has no effect on the monopoly price.

Even under exclusive targeting, Proposition 8 does not imply that the monopoly price is increasing in the reach parameter β. On the contrary, as the reach of the database increases, the optimal monopoly price is pushed lower by two effects. First, the willingness to pay for any targeted set increases (Proposition 8), which makes raising price and restricting supply more costly. Second, the optimal action in the residual set is now more sensitive to the price of cookies. This is due to the composition effect: the average consumer outside the targeted set becomes less likely to have a high match value; as a consequence, the quantity of advertising demanded on the residual set decreases faster as the targeted set expands. Both these effects induce the monopolist to lower price and expand supply as the database becomes less limited.[18]

Two final remarks are in order. First, a reduction in price implies an increase in the *range* of data sold by the monopolist $[v_2, \bar{v}]$ as the reach β increases. Therefore, an increase in the reach β leads to higher data sales. Thus advertisers pay a lower price and access more information, which implies that their profits increase. This means that an increase in data availability can induce a Pareto improvement in the market for information.

Second, note that we have assumed in Proposition 8 that exclusive (positive or negative) targeting is optimal for all price levels. Informally, this means one source of advertising mismatch (wasteful spending or insufficient intensity) is particularly prominent. However, the price of information is not generally continuous or monotone in the reach parameter β. In particular, jumps may occur when the targeting policy induced by the monopolist switches from joint (both positive and negative) targeting for low reach values β to exclusive (positive *or* negative) targeting for high reach values β.

D Precision of the Database

We now assess the implications of the quality of the database, as measured by the data provider's ability to estimate the consumer's characteristics. Because advertisers are risk-neutral and the profit function (2) is linear in v, the choice of action q and willingness to pay for information depend only on the conditional expectation of the match value. It is then convenient to interpret v as a posterior mean. In particular, we assume the true match value is unknown to all, but the data provider has access to an informative signal. The data provider's signals induce posterior means v distributed according to $F(v)$. We then relate the precision of the data provider's information to the properties of the distribution of estimated match values.

We model the precision of the database through the spread of the distribution of match values. We then assume that the distribution of values is ordered by *increasing variance*.[19] Formally, this means that $\mathbb{E}_F[v]$ is constant in a spread parameter k, and that

$$F_k(v) \triangleq F\left(\left(v - \mathbb{E}[v]\right)/\sigma(k)\right), \text{ with } \sigma'(k) > 0. \tag{14}$$

We define the quantity of data demanded at price p as the measure of the optimal targeted set $A_k(p)$ under the distribution $F_k(v)$. We then derive comparative statics of the demand for information with respect to the precision of the database.

[18] In the working paper, Bergemann and Bonatti (2013), we identify sufficient conditions on the distribution of values under which the monopoly price of data is decreasing in its reach.

[19] This stochastic order is a specific instance of the rotation order, see Johnson and Myatt (2006). It obtains, for example, when the seller observes a truth-or-noise signal of the underlying value.

Proposition 9 (Demand for Information)

1. *Assume joint targeting is optimal. The marginal willingness to pay $p_k(A)$ is constant in k, and the quantity of data demanded is strictly increasing in k.*

2. *Assume exclusive targeting is optimal. The marginal willingness to pay $p_k(A)$ is increasing in k, and the quantity of data demanded is increasing in k if $\mathbb{E}[v] \notin A_k(p)$.*

Let us consider the cases of joint and exclusive targeting separately. Under joint targeting, the optimal action on the residual set $\bar{q}$ is unaffected by the spread of the distribution. However, as k increases, more probability mass is placed on the tails of the distribution, so that the total demand for information is higher. Under exclusive targeting, the optimal action on the residual set moves closer to $q^*(\bar{v})$ or $q^*(\underline{v})$ as the spread increases. The composition effect therefore raises the differential profits on any given v, hence the willingness to pay for the marginal cookie. This may lead to a higher or lower quantity of data sold, depending on the level of purchases. In particular, if the cookie $\mathbb{E}[v]$ is in the residual set, then a more spread out distribution implies higher sales.

We now turn attention to the monopoly price of cookies. We first analyze the case of joint targeting. In Proposition 10, we maintain the assumptions of Proposition 3, namely the *symmetry* of $F(v)$ and the *quadratic* profits.

Proposition 10 (Monopoly Price, Joint Targeting)
If joint targeting is optimal, the monopoly price of a cookie is increasing in k.

An important feature of the symmetric quadratic environment is that the distribution of types affects the monopolist's problem through the quantity of data demanded, and not through the marginal value of information. This allows for clean comparative statics of the monopoly price in Proposition 10. Notice, however, that the monopoly quantity responds differently from the price, and it may increase or decrease even if the price is increasing (see Johnson and Myatt (2006) for a thorough discussion).

The effects of information precision on the monopoly price under exclusive targeting are rather intricate. On the one hand, information precision affects advertisers' willingness to pay for the marginal cookie. On the other hand, the spread of the values distribution directly impacts the quantity of data sold. Under the joint restriction of power cost functions $m(q) = q^b$ and uniformly distributed types over $[(1-k)/2, (1+k)/2]$, it is possible to show that the spread k has a positive impact on the monopoly price and a negative impact on the quantity of data sold.[20]

[20] This result does not depend on whether positive or negative targeting is optimal (i.e. on whether the cost parameter $b > 2$). The details are available from the authors.

IV Beyond Linear Pricing

We have focused so far on a fairly specific set of information structures (cookies-based) and pricing mechanisms (linear prices). We now return to the monopoly environment, and we generalize our analysis of data sales to address two closely related questions: *(i)* What is the optimal mechanism for a monopolist to sell information? *(ii)* Are there conditions under which pricing of individual cookies can implement the optimal mechanism?

Up to now, we assumed that the advertisers are symmetric in the distribution of the match values. Moreover, the advertisers attached the same willingness to pay to a consumer with match value v. Thus, from an ex-ante point of view, the advertisers are all identical, and their common ex ante value of information is assumed to be known to all market participants, including the data provider. Therefore, it is as if the data provider has complete information about the preferences of the advertisers. In this setting, suppose the data provider could choose among *unrestricted* pricing mechanisms and information structures, *i.e.* mappings from consumer match values to signals for the advertisers. The data provider would then be able to extract the entire ex ante surplus from the advertisers, for example, by charging a bundle price for the entire database equal to the ex-ante value of information.

In this section, we allow for a private-information component in the advertisers' willingness to pay to match with a consumer with characteristics v. Thus, we consider advertisers who differ in their marginal willingness to pay, denoted by $\theta \in \Theta = [0,1]$. Extending the earlier expression (2), the net value of a match is now given by:

$$\pi(v,q,\theta) \triangleq \theta v q - c \cdot m(q).$$

The marginal willingness to pay θ is private information to each advertiser and is distributed in the population of advertisers according to a continuous distribution function $G(\theta)$ with density $g(\theta)$. For this section, we return to the binary decision environment of Subsection II.A, and restrict attention to binary decisions $q \in \{0,1\}$ of the advertiser (or alternatively linear matching cost $m(q) = q$). The net value of a match is then given by, extending the earlier expression (5):

$$\pi^*(v,\theta) \triangleq \max\{\theta v - c, 0\}.$$

Thus, for advertising to generate positive value, the realization of θ must exceed c as $v \in [0,1]$.

We now explore the data provider's ability to screen advertisers by offering different information structures (or "information policies"), and by pricing the amount of information in a nonlinear way. We begin our analysis with *noiseless* information structures, *i.e.*, deterministic mappings from the advertiser's payoff-relevant states $v \in V$ to a set of signals, and we characterize the optimal mechanism within this class.[21]

With binary actions, the socially efficient information policy can be induced by a threshold $v^*(\theta)$ that informs advertisers perfectly and without noise about the match value v if and only if v exceeds the threshold $v^*(\theta)$ given by:

$$v^*(\theta) = \frac{c}{\theta}. \tag{15}$$

In other words, the data provider can attain the efficient allocation of information through an information policy based on cookies. Under the efficient information policy, each advertiser receives information about every realization of v such that $v \geq v^*(\theta)$. Consequently, advertisers adopt positive targeting, *i.e.*, contact all consumers they receive information about, and ignore the residual users.[22] The expected gross value of the efficient information policy for an advertiser with willingness to pay θ is:

$$w^*(\theta) \triangleq \int_{\frac{c}{\theta}}^{1} (\theta v - c)\, \mathrm{d}F(v).$$

Now consider an arbitrary noiseless information policy with threshold x. The value of this information structure to an advertiser with willingness to pay θ is given by:

$$w(\theta, x) \triangleq \int_{x}^{1} (\theta v - c)\, \mathrm{d}F(v). \tag{16}$$

Note the submodularity property of $w(\theta, x)$, namely that $\partial^2 w(\theta, x)/\partial\theta\partial x = -v < 0$. Therefore, any implementable information policy leads to more data, and hence lower thresholds x, being assigned to advertisers with higher willingness to pay θ. Given the noiseless nature of the information policy, the above problem (16) is akin to a nonlinear pricing problem, where the quantity variable is the amount of information, or the number of cookies sold.

In the associated direct revelation mechanism, each advertiser communicates his willingness to pay, and in exchange is offered a set of cookies and a price for the bundle of cookies. The set of cookies is determined by the threshold $v^*(\theta)$ and hence the associated quantity of cookies is

$$Q(\theta) \triangleq 1 - F(v^*(\theta)),$$

and we denote the transfer payment in the direct mechanism by $T(\theta)$. As in the standard

[21] In the working paper Bergemann and Bonatti (2013), we establish that noiseless information structures remain optimal even when we consider arbitrary information structures. This result requires substantial additional language and notation, and is thus relegated to the working paper.

[22] In this binary action setting, providing information about the complement set and inducing negative targeting yields an identical outcome.

analysis of revenue-maximizing mechanisms, we impose a regularity condition such that the local incentive conditions generate the requisite monotone allocation, which in this context is simply the requirement that the "virtual utility"

$$\theta - \frac{1 - G(\theta)}{g(\theta)} \tag{17}$$

is increasing in θ. We maintain this restriction in Proposition 11, whose proof is in the working paper, Bergemann and Bonatti (2013).

Proposition 11 (Information Policy)
The optimal information policy is a noiseless information policy with threshold

$$v^*(\theta) = \frac{c}{\theta - \frac{1-G(\theta)}{g(\theta)}}. \tag{18}$$

Perhaps the surprising element in the determination of the information policy is that the distributional information about the match values (*i.e.*, $f(v)$ or $F(v)$) does not appear in the description of the optimal information policy. This results from the additivity of the utility of all types θ in the number of user contacts.[23]

The direct mechanism establishes some key properties of the information policy. In particular, $T(\theta)$ and $Q(\theta)$ are strictly increasing in θ, as shown in Proposition 12. A related, indirect mechanism speaks more directly to the problem of data selling and access to the database. Namely, the data provider could specify a nonlinear pricing scheme, or conversely a price for incremental access to the database. With $Q(\theta)$ strictly increasing in θ, we can define a nonlinear pricing scheme, which associates every quantity Q with the transfer of the corresponding type $Q^{-1}(\theta)$:

$$P(Q) \triangleq T\left(Q^{-1}(\theta)\right).$$

We define the price $p(Q)$ as the price for incremental access to the database, or the marginal price that we can readily interpret as the price of an additional cookie:

$$p(Q) \triangleq P'(Q).$$

Under slightly stronger regularity conditions than (17), Proposition 12 establishes that the incremental pricing $p(Q)$ implements the optimal information policy. In fact, the data provider offer access to additional cookies at a declining price that mirrors the logic of

[23] More specifically, the probability density at the threshold $f(v^*(\theta))$ increases the willingness to pay and the information rent at the same rate.

quantity discounts in Maskin and Riley (1984). The proof of this result can be found in the working paper.

Proposition 12 (Prices and Quantities)

1. *The number of cookies sold, $Q(\theta)$ and the transfer $T(\theta)$ are increasing in θ.*
2. *The incremental cookie price $p(Q)$ is decreasing in Q and decentralizes the direct optimal mechanism if $(1-G(\theta))/g(\theta)$ is decreasing.*

Thus, the data provider can decentralize the optimal direct mechanism by allowing advertisers to access a given portion of the database, with volume discounts for those who demand a larger amount of cookies. This establishes an equivalent implementation of the optimal mechanism, based on advertiser self-selection of a subset of cookies. We can then view the (constant) monopoly price p for cookies (which yields a total payment pQ) as a linear approximation of the optimal nonlinear tariff $T(Q)$ in this particular case.[24]

V Concluding Remarks

We analyzed the sale of individual-level information in a setting that captures the key economic features of the market for third-party data. Specifically, in our model, a monopolistic data provider determines the price to access informative signals about each consumer's preferences.

Our first set of results characterized the demand for such signals by advertisers who wish to tailor their spending to the match value with each consumer. We showed how properties of the complete information profit function determine the optimality of an information-purchasing strategy that achieves positive targeting, negative targeting, or both. We also explored the interaction between the markets for data and advertising, and we showed that a publisher of advertising space can, but need not, benefit from the availability of data to the advertisers.

Turning to monopoly pricing of cookies, we established that the ability to influence the composition of the advertisers' targeted and residual sets was the key driver of the optimal (linear) prices. As a consequence, both the reach of the monopolist's database and the concentration of data sales provide incentives to lower prices.

We then considered an environment in which advertisers differ in their willingness to pay, and we showed that cookies-based pricing can be part of an (approximate) optimal

[24]See Rogerson (2003) for bounds on the loss in profits from simpler mechanisms such as linear pricing.

mechanism for the sale of information. In particular, we showed that the data provider can decentralize the optimal mechanism by offering a nonlinear pricing schedule for cookies.

We, arguably, made progress towards understanding basic aspects of data pricing and data markets. We did so by making a number of simplifying assumptions. A more comprehensive view of data markets would require a richer environment. In the present model, the information supported the formation of valuable matches, and hence could be viewed as increasing the surplus of the consumer *and* the advertiser at the same time. But if information could also impact the division of surplus between them, then the value of information (and the corresponding value of privacy) would require a more subtle analysis.

In the present model neither the advertiser nor the publisher had access to any proprietary information about the consumers. In reality, advertisers and (more prominently) large publishers and advertising exchanges maintain databases of their own. Thus, the nature of the information sold and the power to set prices depend on the initial allocation of information across market participants. Moreover, online data transactions are inherently two-sided. Presently, we analyzed the price charged by the data provider to the advertisers. But there are cost of acquiring the data from individuals, publishers, or advertisers. Ultimately, the cost of acquiring information for the data provider should be related to the value of privacy, which may limit the availability of data or raise its price.

Appendix

Proof of Proposition 1. Suppose the advertisers' optimal action on the residual set is given by $q^*(A^{\mathrm{C}}) = 0$. The value of the marginal cookie is then given by $\max\{0, v - c\}$, which is increasing in v. We show that the value of information is strictly monotone in v. Notice that adding higher-v cookies to the targeted set does not change the optimal action on the residual set, because it lowers the expected value of a consumer $v \in A^{\mathrm{C}}$. Thus, if advertisers buy cookie v, they also buy all cookies $v' > v$. Conversely, if the optimal action on the residual set is given by $q^*(A^{\mathrm{C}}) = 1$, the value of the marginal cookie is $\max\{0, c - v\}$. By a similar argument, the value of information is strictly decreasing in v: if advertisers buy cookie v, they also buy all cookies $v' < v$.

Now consider the advertiser's profits under positive and negative targeting. In the former case, the advertisers' profits are given by

$$\pi_+(c,p) \triangleq \max_v \int_v^1 (x - c - p)\,\mathrm{d}F(x) = \int_{c+p}^1 (x - c - p)\,\mathrm{d}F(x).$$

In the latter case, profits are given by

$$\pi_-(c,p) \triangleq \max_v \left[\int_v^1 (x - c)\,\mathrm{d}F(x) - pF(v)\right] = \int_{c-p}^1 (x - c)\,\mathrm{d}F(x) - pF(c-p).$$

Now consider the difference

$$\pi_+(c,p) - \pi_-(c,p) = p\,(F(c-p) + F(c+p) - 1) - \int_{c-p}^{c+p} (v - c)\,\mathrm{d}F(v). \tag{19}$$

Under the uniform distribution, the second term in (19) is nil, while the first is equal to $p\,(2c-1)$, which establishes the result. ■

Proof of Proposition 2. Suppose towards a contradiction that the optimal residual set A^{C} is not an interval. Let $q_0 = q^*(A^{\mathrm{C}})$ denote the match intensity with all consumers in the residual set. By equation (8), we know q_0 is the optimal match intensity for the average type $v_A = \mathbb{E}[v \mid v \notin A]$. Suppose $v_A \in A$. Now consider two consumers with $v'' > v'$ and $q^*(v'') > q^*(v') > q_0$ such that the firm buys cookie v' but not v''. If A^{C} is not an interval, either such a pair exists, or there exists a pair with $v'' < v'$ and $q^*(v'') < q^*(v') < q_0$ such that the firm buys cookie v' but not v''. Consider the former case, and compute the change in profits obtained by swapping cookies, *i.e.*, purchasing (an equal number of) cookies v'' instead of cookies v'. Define the difference between complete and incomplete information profits as

$$\Delta(v, q_0) = v(q^*(v) - q_0) - c(m(q^*(v)) - m(q_0)),$$

and notice that $\Delta_v(v, q_0) = (q^*(v) - q_0)$. Therefore $q^*(v'') > q^*(v') > q_0$ implies $\Delta(v'', q_0) > \Delta(v', q_0)$. Because the advertiser gains $\Delta(v'', q_0)$ and loses $\Delta(v', q_0)$, it follows that the swap strictly improves profits. An identical argument applies to the case of $q^*(v'') < q^*(v') < q_0$. Finally, if $v_A \notin A$, then a profitable deviation consists of not purchasing v_A: advertisers avoid paying a positive price, and the optimal action on the residual set does not change. ■

Proof of Proposition 3. If costs are quadratic, so are the complete information profits. By symmetry of the distribution, $v_0 = \mathbb{E}[v \mid v \in [v_0 - \varepsilon, v_0 + \varepsilon]]$ for any $\varepsilon > 0$. The marginal value of information is then given by

$$p(v) = \pi^*(v) - (vq^*(v_0) - cm(q^*(v_0))) = (v_0 - v)^2 / 4c.$$

Solving for v_0 yields the optimal residual set as a function of p and c. ■

Proof of Proposition 4. Consider the necessary conditions for the optimal residual set A^{C} to be given by an interior interval $[v_1, v_2]$. Define the expected value on the residual set,

$$v_0 \triangleq \mathbb{E}[v \mid v \in A^{\mathrm{C}}].$$

For ease of notation, let $\pi(v) \triangleq \pi^*(v)$. It follows that $q_0 \triangleq q^*(A^{\mathrm{C}}) = q^*(v_0)$, and by the envelope theorem $q_0 = \pi'(v_0)$. The marginal value of information at v is then given by $\pi(v) - (\pi(v_0) + (v - v_0)\pi'(v_0))$, and its derivative with respect to v is given by $\pi'(v) - \pi'(v_0)$. Optimality of an interior residual set requires that the marginal value of information is equal to p at the two extremes *i.e.*,

$$\int_{v_1}^{v_2} (\pi'(v) - \pi'(v_0))\,\mathrm{d}v = 0.$$

Under concavity of $\pi'(v)$, however, we have

$$\int_{v_1}^{v_2} (\pi'(v) - \pi'(v_0))\,\mathrm{d}v \leq \int_{v_1}^{v_2} \pi''(v_0)(v - v_0)\,\mathrm{d}v,$$

which is non-positive if $f(v)$ is non-decreasing. This implies negative targeting. A similar last step implies positive targeting.

Finally, we relate the curvature of the profit function to that of the match cost function. The envelope theorem implies $\pi'(v) = q^*(v)$, and implicit differentiation of the first order condition yields

$$\pi''(v) = (cm''(q^*(v)))^{-1}.$$

Because $q^*(v)$ is strictly increasing, we conclude that $\pi'''(v) > 0$ if and only if $m'''(q) < 0$.■

Proof of Proposition 5. We first establish a property of the complete information demands for advertising. Differentiating $m(q^*(v))$ with respect to v, we obtain

$$\frac{\mathrm{d}m(q^*(v))}{\mathrm{d}v} = m'(q^*(v))\frac{\mathrm{d}q^*(v)}{\mathrm{d}v} = \frac{m'(q^*(v))}{cm''(q^*(v))}.$$

Therefore, the demand for advertising space is convex in v if $m''(q)/m'(q)$ is decreasing in q, *i.e.*, $m'(q)$ is log-concave. Conversely, $m(q^*(v))$ is concave in v if $m'(q)$ is log-convex.
(1.) We focus on the negative-targeting case $A = [\underline{v}, v_1]$, but all arguments immediately extend to the case of positive targeting. Now consider the publisher's revenues as a function of p. The total demand for advertising is given by

$$M(A) = \int_{\underline{v}}^{v_1} m(q^*(v))\,\mathrm{d}F(v) + (1 - F(v_1))\,m(q^*([v_1, \bar{v}])).$$

Letting $\hat{v} \triangleq \mathbb{E}[v \mid v \in [v_1, \bar{v}]]$, we have

$$\begin{aligned}\frac{\partial M}{\partial v_1} &= (m(q^*(v_1)) - m(q^*(\hat{v})))\,f(v_1) + (1 - F(v_1))\,m'(q^*(\hat{v}))\frac{\partial q^*(\hat{v})}{\partial \hat{v}}\frac{\partial \hat{v}}{\partial v_1} \\ &= f(v_1)(m(q^*(v_1)) - m(q^*(\hat{v}))) + f(v_1)\frac{m'(q^*(\hat{v}))}{cm''(q^*(\hat{v}))}(\hat{v} - v_1).\end{aligned}$$

This expression is positive if and only if $m''(q)/m'(q)$ is decreasing in q, *i.e.*, if $m(q^*(v))$ is convex. Because v_1 is decreasing in p, the publisher's revenue $c \cdot M$ is decreasing in p if $m'(q)$ is log-concave.
(2.) It is immediate to see that all results from part (1.) are reversed if $m'(q)$ is log-convex (so that $m''(q)/m'(q)$ is increasing in q and $m(q^*(v))$ is concave in v). ■

Proof of Proposition 6. (1.) We know from Proposition 1 that advertisers choose the following targeted set:

$$A(c,p) = \begin{cases} [0, \max\{c-p, 0\}] & \text{if} \quad c < 1/2; \\ [\min\{c+p, 1\}, 1] & \text{if} \quad c \geq 1/2. \end{cases} \tag{20}$$

Thus, under the uniform distribution, the monopoly price of cookies is given by

$$p^*(c) = \begin{cases} \arg\max_p [p(c-p)] & \text{if} \quad c < 1/2, \\ \arg\max_p [p(1-c-p)] & \text{if} \quad c \geq 1/2, \end{cases}$$

and therefore $p^*(c) = (1/2)\min\{c, 1-c\}$.
(2.) It follows from (20) that $A(c, p^*(c)) = [0, c/2]$ if $c < 1/2$ and $A(c, p^*(c)) = [(1-c)/2, 1]$ if $c \geq 1/2$.
(3.) The single-peakedness of prices $p^*(c)$, sales $\mu(A(c, p^*(c)))$, and hence profits, is immediate from parts (1.) and (2.). ■

Proof of Proposition 7. (1.) For the case of positive targeting, let

$$p(v, x) = \pi(v) - \pi(v, q^*([\underline{v}, x])).$$

A monopolist data provider chooses the marginal cookie v_2 to solve the following problem:

$$\max_v [p(v, v)(1 - F(v))].$$

The optimal v_2^* is then given by the solution v to the following first-order condition:

$$-p(v, v) f(v) + (1 - F(v))(\partial p(v, v)/\partial v + \partial p(v, v)/\partial x) = 0.$$

Conversely, in the symmetric equilibrium with a continuum of sellers, the equilibrium marginal cookie $\tilde{v}_2$ is given by the solution v to the following condition

$$-p(v, v) f(v) + (1 - F(v)) \partial p(v, v)/\partial v = 0.$$

However,

$$\frac{\partial p(v, v_2^*)}{\partial x} = -\frac{\partial \pi(v, q^*([\underline{v}, x]))}{\partial q} \frac{\partial q^*}{\partial v} \frac{\partial \mathbb{E}[v \mid v \leq x]}{\partial x} < 0,$$

because $q^*(v)$ is strictly increasing in v, and therefore $\partial \pi(v, q)/\partial q > 0$ for all $q < q^*(v)$. Therefore, the price under competition $\bar{p} \triangleq p(\tilde{v}_2, \tilde{v}_2)$ is higher than the monopoly price $p^* \triangleq p(v_2^*, v_2^*)$.

For the case of negative targeting, the monopolist maximizes $p(v, v) F(v)$, where

$$p(v, x) = \pi(v) - \pi(v, q^*([x, \bar{v}])).$$

The monopolist's first-order condition is then given by

$$p(v,v)\,f(v)+F(v)\left(\partial p(v,v)/\partial v+\partial p(v,v)/\partial x\right)=0.$$

In the symmetric equilibrium with a continuum of sellers, the equilibrium marginal cookie $\bar{v}_1$ solves the following condition

$$-p(v,v)\,f(v)+F(v)\,\partial p(v,v)/\partial v=0.$$

The solution $\tilde{v}_1$ is lower (the price is higher) than the monopolist's threshold v_1^* since

$$\frac{\partial p(v,\tilde{v}_1)}{\partial \tilde{v}_1}=-\frac{\partial \pi\left(v,q^*\left([\tilde{v}_1,\bar{v}]\right)\right)}{\partial q}\frac{\partial q^*}{\partial v}\frac{\partial \mathbb{E}\left[v\mid v\geq \tilde{v}_1\right]}{\partial \tilde{v}_1}>0,\text{ for all } v\leq \mathbb{E}\left[v\mid v\geq \tilde{v}_1\right].$$

(2.) We look for a symmetric equilibrium in the price-setting game with n data providers. Let $p_j=p_2$ for all $j\neq 1$ and characterize the advertisers' demand as a function of (p_1,p_2). If positive targeting is optimal, advertisers buy cookies $v\in[v_1,\bar{v}]$ from seller $j=1$ and $v\in[v_2,\bar{v}]$ from sellers $j\neq 1$. In particular, the thresholds v_j, $j=1,2$, satisfy

$$\pi(v_j)-\pi\left(v_j,q^*(\hat{v})\right)=p_j,$$

where

$$\hat{v}(p_1,p_2)=\frac{\mathbb{E}\left[v\mid v\leq v_1\right]+(n-1)\,\mathbb{E}\left[v\mid v\leq v_2\right]}{n}.$$

Note that $p_1>p_2$ implies $v_1>v_2$. Now rewrite the profit function of seller $j=1$ as

$$\Pi_1=\left(\pi(v_1)-\pi\left(v_1,q^*(\hat{v})\right)\right)\left(1-F(v_1)\right).$$

At a symmetric equilibrium where $v_j\equiv v$, the first-order condition of seller 1 is given by

$$\left(\pi(v_1)-\pi\left(v_1,q^*(\hat{v})\right)\right)\frac{f(v_1)}{1-F(v_1)}=\pi'(v_1)-\frac{\partial \pi\left(v_1,q^*(v)\right)}{\partial v_1}-\frac{\partial \hat{v}}{\partial v_1}\frac{\mathrm{d}q^*(\hat{v})}{\mathrm{d}\hat{v}}\frac{\partial \pi\left(v_1,q^*(v)\right)}{\partial q}.$$

Both $dq^*(\hat{v})/d\hat{v}$ and $\partial\pi\left(v_1,q^*(v)\right)/\partial q$ on the right-hand side are positive. Because

$$\frac{\partial \hat{v}}{\partial v_1}=\frac{1}{n}\frac{\partial \mathbb{E}\left[v\mid v\leq v_1\right]}{\partial v}$$

is decreasing in n, the symmetric equilibrium threshold $v^*(n)$ is increasing in n, and so is the price $p^*(n)$.

The analysis under negative targeting yields similar steps and is therefore omitted. ■

Proof of Proposition 8. Under positive targeting, the marginal willingness to pay $p(v,\beta)$ for a targeted set $A=[v,\bar{v}]$ is given by

$$p(v,\beta) \triangleq \pi^*(v) - \pi(v, q_0(v,\beta)),$$

where

$$q_0(v,\beta) \triangleq q^*(\beta \mathbb{E}_F[v' \mid v' \leq v] + (1-\beta)\mathbb{E}_F[v']).$$

The derivative of the inverse demand function with respect to the reach β is given by

$$\frac{\partial p(v,\beta)}{\partial \beta} = -(v - cm'(q_0(v,\beta)))\, q^{*\prime}(\cdot)\,(\mathbb{E}_F[v' \mid v' < v] - \mathbb{E}_F[v']). \tag{21}$$

The first two terms in (21) are positive: profits $\pi(v,q_0)$ are increasing in q because $q_0(v,\beta) < q^*(v)$; the complete information quantity $q^*(\cdot)$ is strictly increasing; and difference of the conditional and unconditional expected values is negative. Therefore, the marginal willingness to pay $p(v,\beta)$ is increasing in β. ■

Proof of Proposition 9. (1.) Under joint targeting, we know the optimal action on the residual set is given by $\bar{q}$ for all k. It follows that the willingness to pay for v is independent of the distribution. However, as k increases, both $F_k(v_1)$ and $1-F_k(v_2)$ increase, so the quantity of data demanded increases.
(2.) Consider the inverse demand for data in the case of negative targeting:

$$p(v_1) = v_1(q^*(v_1) - q^*([v_1,\bar{v}])) - c(m(q^*(v_1)) - m(q^*([v_1,\bar{v}]))).$$

As k increases, by second-order stochastic dominance, the conditional expectation $\mathbb{E}[v \mid v > v_1]$ increases as well. Therefore, $q^*([v_1,\bar{v}])$ increases, and because $q^*([v_1,\bar{v}]) > q^*(v_1)$, the willingness to pay $p(v_1)$ increases as well. Therefore, we know the threshold $v_1(p,k)$ is increasing in k. If in addition, $v_1 < \mathbb{E}[v]$, then $F_k(v_1)$ is increasing in k, and therefore $F_k(v_1(p,k))$ is increasing a fortiori. An identical argument applies to the case of positive targeting. ■

Proof of Proposition 10. Under the conditions of Proposition 3, the monopolist solves

$$\max_p \Pi(p,k) \triangleq pF_k(\mathbb{E}[v] - 2\sqrt{cp}) = pF(-2\sqrt{cp}/\sigma(k)). \tag{22}$$

The first-order condition for this problem can be written as

$$F(x) + xf(x)/2 = 0,$$

where $x := -2\sqrt{cp}/\sigma(k)$. This implies that the optimal p is proportional to $\sigma(k)^2$. ■

References

Abraham, Itai, Susan Athey, Moshe Babaioff, and Michael Grubb. 2014. "Peaches, Lemons, and Cookies: Designing Auction Markets with Dispersed Information." Microsoft Research, Stanford, and Boston College.

Adams, William, Liran Einav, and Jonathan Levin. 2009. "Liquidity Constraints and Imperfect Information in Subprime Lending." *American Economic Review*, 99(1): 49–84.

Admati, Anat R., and Paul Pfleiderer. 1986. "A monopolistic market for information." *Journal of Economic Theory*, 39(2): 400–438.

Admati, Anat R., and Paul Pfleiderer. 1990. "Direct and indirect sale of information." *Econometrica*, 58(4): 901–928.

Allen, Franklin. 1990. "The market for information and the origin of financial intermediation." *Journal of Financial Intermediation*, 1(1): 3–30.

Anderson, Eric T., and Duncan Simester. 2013. "Advertising in a Competitive Market: The Role of Product Standards, Customer Learning and Switching Costs." *Journal of Marketing Research*, 50(4): 489–504.

Anton, James J., and Dennis A. Yao. 2002. "The sale of ideas: Strategic disclosure, property rights, and contracting." *Review of Economic Studies*, 69(3): 513–531.

Babaioff, Moshe, Robert Kleinberg, and Renato Paes Leme. 2012. "Optimal Mechanisms for Selling Information." *EC '12*, 92–109.

Bagwell, Kyle. 2007. "The Economic Analysis of Advertising." *Handbook of industrial organization*, 3: 1701–1844.

Bergemann, Dirk, and Alessandro Bonatti. 2011. "Targeting in Advertising Markets: Implications for Offline Vs. Online Media." *RAND Journal of Economics*, 42(3): 417–443.

Bergemann, Dirk, and Alessandro Bonatti. 2013. "Selling Cookies." Cowles Foundation, Yale University CFDP 1920R.

Bergemann, Dirk, and Martin Pesendorfer. 2007. "Information Structures in Optimal Auctions." *Journal of Economic Theory*, 137(1): 580–609.

Bergemann, Dirk, Benjamin Brooks, and Stephen Morris. 2013. "The Limits of Price Discrimination." Cowles Foundation for Research in Economics, Yale University 1896R.

Blake, Thomas, Chris Nosko, and Steven Tadelis. 2013. "Consumer Heterogeneity and Paid Search Effectiveness: A Large Scale Field Experiment." eBay Research Labs, University of Chicago, and University of California Berkeley.

Butters, Gerard R. 1977. "Equilibrium Distributions of Sales and Advertising Prices." *Review of Economic Studies*, 44(3): 465–491.

Calzolari, Giacomo, and Alessandro Pavan. 2006. "On the optimality of privacy in sequential contracting." *Journal of Economic Theory*, 130(1): 168–204.

Celis, Elisa, Greg Lewis, Markus Mobius, and Hamid Nazerzadeh. 2014. "Buy-it-now or Take-a-chance: Price Discrimination through Randomized Auctions." *Management Science*, forthcoming.

García, Diego, and Francesco Sangiorgi. 2011. "Information Sales and Strategic Trading." *Review of Financial Studies*, 24(9): 3069–3104.

Ghosh, Arpita, Mohammad Mahdian, Preston McAfee, and Sergei Vassilvitskii. 2012. "To Match or Not to Match: Economics of Cookie Matching in Online Advertising." 741–753. ACM.

Hoffmann, Florian, Roman Inderst, and Marco Ottaviani. 2014. "Persuasion through Selective Disclosure: Implications for Marketing, Campaigning, and Privacy Regulation." J.W. Goethe University Frankfurt, UCL, and Bocconi.

Hörner, Johannes, and Andrzej Skrzypacz. 2012. "Selling Information." Cowles Foundation, Yale University CFDP 1743RR.

Iyer, Ganesh, and David Soberman. 2000. "Markets for product modification information." *Marketing Science*, 19(3): 203–225.

Johnson, Justin P., and David P. Myatt. 2006. "On the Simple Economics of Advertising, Marketing, and Product Design." *American Economic Review*, 96(3): 756–784.

Kamenica, Emir, and Matthew Gentzkow. 2011. "Bayesian Persuasion." *American Economic Review*, 101(6): 2590–2615.

Maskin, Eric, and John Riley. 1984. "Monopoly with Incomplete Information." *RAND Journal of Economics*, 15(2): 171–196.

Pancras, Joseph, and K. Sudhir. 2007. "Optimal marketing strategies for a customer data intermediary." *Journal of Marketing Research*, 44(4): 560–578.

Rogerson, William. 2003. "Simple menus of contracts in cost-based procurement and regulation." *American Economic Review*, 93(3): 919–926.

Sarvary, Miklos, and Philip M. Parker. 1997. "Marketing information: A competitive analysis." *Marketing Science*, 16(1): 24–38.

Syrgkanis, Vasilis, David Kempe, and Eva Tardos. 2013. "Information Asymmetries in Common-Value Auctions with Discrete Signals." Cornell University.

Taylor, Curtis R. 2004. "Consumer privacy and the market for customer information." *RAND Journal of Economics*, 35(4): 631–650.

Xiang, Yi, and Miklos Sarvary. 2013. "Buying and selling information under competition." *Quantitative Marketing and Economics*, 11(3): 321–351.