Selling Cookies

Dirk Bergemann[y] Alessandro Bonatti[z]

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

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 e¢ cient match between individual and "treatment" may require the collection, analysis and di¤usion 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 …les 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 …rst 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 ‡ights).[1] If the same user visits another partner website (*e.g.*, an online retailer), the information contained in her cookie is updated to re‡ect the most recent browsing history.

The data provider therefore maintains a detailed and up-to-date pro…le 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 (i)6(1i(w)15c)9D(l)]TJr ar but are equally representative of many online and o- ine 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

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. Bergeformation 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 $q \geq 0$ 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* $m : \mathbb{R}_+ \to \mathbb{R}_+$

ing 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 …rm equal to $v(i;j)$.

B Data Provider

Initially, the advertisers do not have information about the pair-speci…c 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 …rms).

Advertisers can request information from the data provider about consumers with speci…c characteristics i. Now, from the perspective of advertiser j the only relevant aspect of the

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 is given by:

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

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

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

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

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

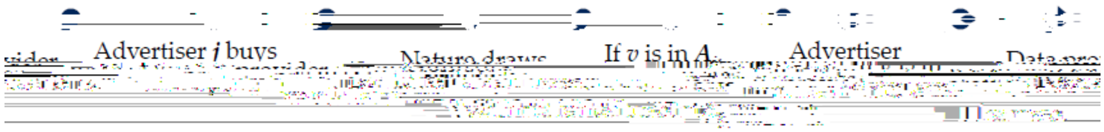
and the associated conditional distribution functions $F(v)$ and $F(v)$. We assume that the resulting match values are identically distributed across consumer and across firms, *i.e.*, for all i, j, and v:

$$F(v) = F(v) = F(v).$$

Thus, $F(v)$ represents the common prior distribution for each firm and each consumer about

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 …rst 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 e¤ort exerted by the …rm, which also enters the consumer's utility function. Thus, the pro…t 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 …rm and shift surplus from the consumer to the …rm (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 pro…ts 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 . In particular, we denote the complete information demand for advertising space $q\;(v)$ and pro…t level $\;(v)$ by

$$q\;(v) \triangleq \arg\max_{2R_+}[\;(v;q)];$$
$$(v) \triangleq \;(v;q\;(v)):$$

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

Because the objective (**v;q**) is linear in **v**, the optimal level of advertising **q** (**A**C is given by

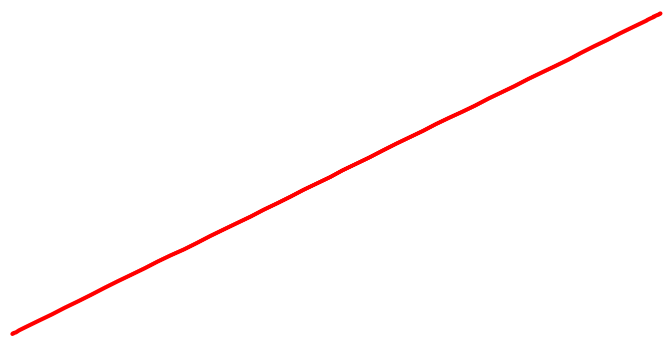
$$\boldsymbol{q}\ (\boldsymbol{A}^{C} \triangleq \arg\max_{2R} (\boldsymbol{v};\boldsymbol{q})\ j$$

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 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 e¤ect 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 …rst describe the demand for advertising space when the value of the match v is known to the advertisers. Thus, we introduce the realized *complete information* pro…t (***v***) is strictly convex in ***v***. In contrast, the realized pro…t under *prior information* is linear in ***v***, and it is given by (***v;q***). Figure 2 describes the pro…t function under complete information (***v***) and prior information (***v;q***).

Figure 2: Complete Information and Prior Information Profits



As intuitive, under prior information, the …rm chooses excessive (wasteful) advertising to low-value consumers and insu¢ cient advertising to higher-value consumers. The …rm therefore has a positive willingness to pay for information, *i.e.*, for cookies. The value of

d18d586(e)9(x)11(4680(h)120()6(49(t4)8(i)6(t)8(i)6Tf1)5((469205)1512(d)t)35)8(i312(d)2Tf9.,)]T86(e)9

Proposition 2 allows us to rewrite the …rm's problem (4) as the choice of two thresholds, $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$

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:*

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

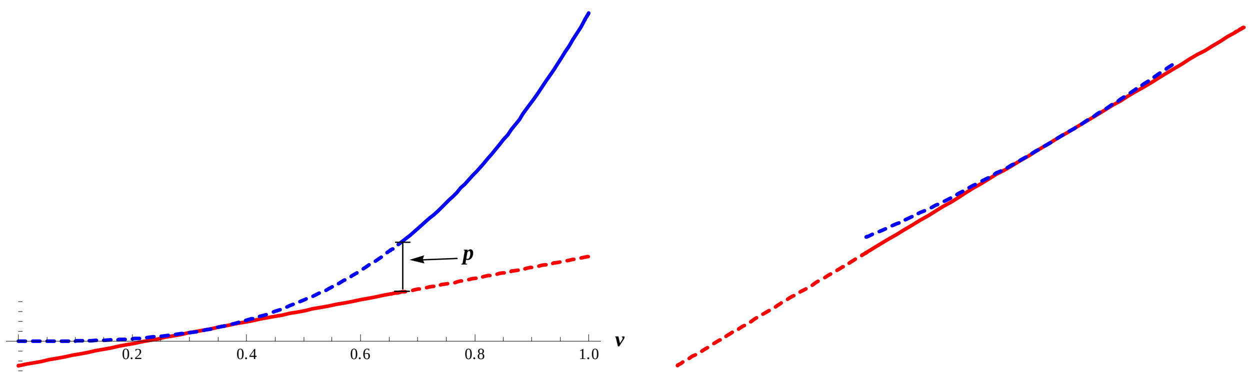
2. *If $m''(q)$ and $f(v)$ are increasing, negative targeting is optimal:*

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

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 $(v) \quad (v;q)$. 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



(a) (b)

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, $m(q) = q^{\alpha}$ with $\alpha < 2$. Examples of convex marginal costs include those derived from the Butters (1977) exponential matching technology, *i.e.*, $m(q) = -\alpha \ln(1-q)$; with $\alpha > 0$, and power functions $m(q) = q^{\alpha}$, with $\alpha > 2$.

providers of advertising space, including Google, Yahoo!, and Facebook.[12]

E Implications for Publishers

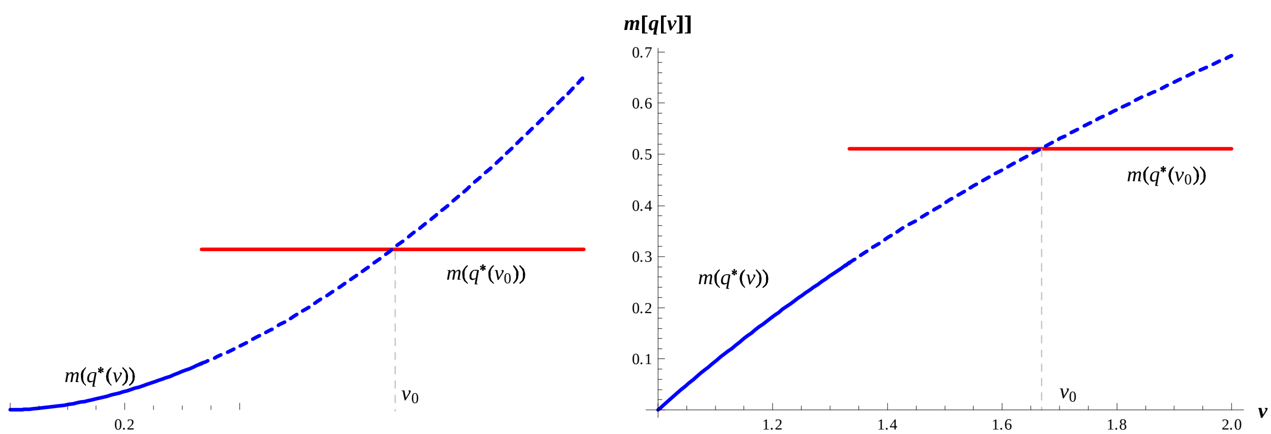
We conclude this section by examining the interaction between the markets for data and online advertising. In particular, we assess the effect of data sales on the demand for advertising space and the implications of vertical integration between publishers and data providers.

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

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

We are interested in the effect of the amount of data sold (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



$m(q) = q/b; \quad b = 9/4$ $\qquad$ $m(q) = \ln(1 \quad q)$

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 $p \in \{0, 1\}$). 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

$$p \;\; = \arg\max \, [p \quad (A(c;p))] :$$

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

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 su¢ ciently high and negative targeting when the cost of advertising is low. Proposition 6 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 o¤ers 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 su¢ cient 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 . Every advertiser then chooses the targeted set $A = [v_2; \bar{v}]$ where the threshold value $v_2(p)$ solves the condition

$$p = (v_2) \quad (v_2; q([\underline{v}; v_2])):$$

Thus, the cookie sold by seller j will have a distribution of values across advertisers that depends on the other sellers' prices through their e¤ect on the residual set. In particular, a symmetric price pro…le p 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 di¤erential pro…t under the threshold strategy $v_2(p)$. Therefore, seller j faces the inverse demand function $p(v; v_2)$ given by

$$p(v; v_2) \triangleq (v) \quad (v; q([\underline{v}; v_2])): \tag{12}$$

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

$$v_2 = \arg\max\,[p(v; v_2)(1 \quad F(v))]\,.$$

The key di¤erence with the monopoly problem lies in the residual advertising intensity $q([\underline{v}; v_2]);$ which cannot be in‡uenced 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.

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 < 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 $F(v)$. We then have the following result.

Proposition 8 (Reach and Demand)

becomes less limited.[18]

Two …nal remarks are in order. First, a reduction in price implies an increase in the *range* of data sold by the monopolist $[\boldsymbol{v}_2; \boldsymbol{v}]$ as the reach increases. Therefore, an increase in the reach leads to higher data sales. Thus advertisers pay a lower price and access more information, which implies that their pro…ts 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)

Proposition 9 (Demand for Information)

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 we characterize the optimal mechanism within this class.[21]

With binary actions, the socially e¢ cient information policy can be induced by a threshold ***v***

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,* $Q(\ )$ *and the transfer* $T(\ )$ *are increasing in* .
2. *The incremental cookie price* $p(Q)$ *is decreasing in* Q *and decentralizes the direct optimal mechanism if* $(1 \quad G(\ ))=g(\ )$ *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 o¤ering 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 $q(A^C) = 0$. 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 $v \in A^C$. 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 $q(A^C) = 1$, 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 $v' < v$.

pro…ts as

$$(v; q_0) = v(q \ (v) \quad q_0) \quad c(m(q \ (v)) \quad m(q_0));$$

and notice that $(v; q_0) = (q \ (v) \quad q_0):$ Therefore $q \ (v'') > q \ (v') > q_0$ implies $(v''; q_0) > (v'; q_0)$. Because the advertiser gains $(v''; q_0)$ and loses $(v'; q_0)$, it follows that the swap strictly improves pro…ts. An identical argument applies to the case of $q \ (v'') < q \ (v') < q_0$. Finally, if $v \ \ 2 \ A$, then a pro…table deviation consists of not purchasing v : 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 pro…ts. By symmetry of the distribution, $v_0 = \mathbb{E}[v \text{ j } v$

condition yields

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

Because $q^*(v)$ is strictly increasing, we conclude that $\psi'''(v) > 0$ 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 $c\,M$ 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:

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

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

$$p(v; \ ) \triangleq \quad (v) \quad (v; q_0(v; \ ));$$

where

$$q_0(v; \ ) \triangleq q \ (\ \mathbb{E} \ [v' \mid v' \quad v] + (1 \quad) \mathbb{E} \ [v']):$$

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

$$\frac{\partial p(v; \ )}{\partial} = \quad (v \quad cm'(q_0(v; \ )))\, q'(\ )(\mathbb{E} \ [v' \mid v' < v] \quad \mathbb{E} \ [v']): \qquad (21)$$

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

Proof of Proposition 9. (1.) Under joint targeting, we know the optimal action on the residual set is given by q for all k. It follows that the willingness to pay for v is independent of the distribution. However, as k increases, both $F(v_1)$ and $1 \quad F(v_2)$ increase, so the quantity of data demanded increases.

(2.) Consider the inverse demand for data in the case of negative targeting:

$$p(v_1) = v_1(q(v_1) \quad q([v_1; v])) \quad c(m(q(v_1)) \quad m(q([v_1; v]))):$$

As k increases, by second-order stochastic dominance, the conditional expectation $\mathbb{E}[v \mid v > v_1]$ increases as well. Therefore, $q([v_1; v])$ increases, and because $q([v_1; v]) > q(v_1)$, the willingness to pay $p(v_1)$ increases as well. Therefore, we know the threshold $v_1($

Rogerson, William. 2003. "Simple menus of contracts in cost-based procurement and regulation." *American Economic Review*