Electronic Commerce Research and Applications 10 (2011) 17–25

Contents lists available at ScienceDirect

Electronic Commerce Research and Applications

journal homepage: www.elsevier.com/locate/ecra

Adverse selection in online "trust" certifications and search results

Benjamin Edelman

Harvard Business School, 1 Soldiers Field Rd., Boston, MA 02163, United States

ARTICLE INFO

Article history:
Received 19 October 2009
Received in revised form 19 May 2010
Accepted 23 June 2010
Available online 27 June 2010

Keywords:
Adverse selection
Certification
Reputation
Trust
Regulation

ABSTRACT

Widely-used online "trust" authorities issue certifications without substantial verification of recipients' actual trustworthiness. This lax approach gives rise to adverse selection: the sites that seek and obtain trust certifications are actually less trustworthy than others. Using an original dataset on web site safety, I demonstrate that sites certified by the best-known authority, TRUSTe, are more than twice as likely to be untrustworthy as uncertified sites. This difference remains statistically and economically significant when restricted to "complex" commercial sites. Meanwhile, search engines create an implied endorsement in their selection of ads for display, but I show that search engine advertisements tend to be less safe than the corresponding organic listings.

1. Introduction

When agents have hidden types, contract theory warns of bad results and potentially even market unraveling. Since Akerlof's "lemons" (1970), others have worried about similar problems in markets with hidden types – like bad drivers wanting more car insurance than good drivers (Chiappori and Salanie 2000), and healthy people disproportionately buying annuities (Finkelstein and Poterba 2004).

In general, it is difficult to empirically assess the significance of adverse selection problems. For example, used car markets are made more complicated by idiosyncratic details – unobservable car characteristics, local markets, and casual sellers. Some research manages to address these problems. For example, Chiappori and Salanie (2000) focus on novice drivers, who have less private information about their own type (since they have not yet started to drive), letting economists observe most relevant characteristics. But these special cases bring problems of their own. Researchers may be less interested in the absence of adverse selection among novice drivers' insurance purchases, and more interested in the adverse selection that might affect other drivers.

This paper applies an adverse selection model to a new market: web sites and their associated "trust"-type certifications. With a proprietary data source, I analyze characteristics generally unobservable both to consumers and to trust authorities. Unmasking sites' otherwise-hidden types provides an unusual opportunity to measure the magnitude of adverse selection occurring in this market.[1]

Beyond adverse selection, trust certifications are also of interest in their own right. These certifications have played an important role in the policy debate as to regulation of online privacy and safety, and typical Internet users see such certifications remarkably frequently. Yet adverse selection significantly taints trust certifications: my analysis indicates that low-quality sites disproportionately seek and receive certification, substantially reducing overall certification quality. In particular, in Section 6, I find that sites certified by the best-known authority, TRUSTe, are more than twice as likely to be untrustworthy as uncertified sites.

Distinct from trust certifications, search engines effectively endorse certain sites through their selection of ads to present with users' search results. Seeing prominent advertisements juxtaposed with sites presented as the "most relevant" for a given search, users may mistakenly believe the advertised sites deserve their trust. But in fact the true test for appearance in search advertisements is paying a fee, not satisfying substantive requirements. In Section 7, I show that search engine advertisements are systematically less safe than the corresponding organic results.

2. The basic web site safety problem

Consumers seeking online services face a serious difficulty in deciding what sites to use. Consumers could stick with "known-good" big names, but such a narrow focus would reduce match quality, denying users the rich diversity of Internet content. Exploring the broader Internet offers the potential for a better match, but with important risks: untrustworthy sites might send users spam (if users register or otherwise provide email addresses), infect users' computers with viruses or other harmful code (if users install the programs that sites offer), or simply fail to deliver the

E-mail address: bedelman@hbs.edu

[1] Despite similarity in name, trust certification authorities are entirely distinct from the certification authorities that offer SSL certificates, code-signing certificates, and encryption keys for secure communications.

1567-4223/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.elerap.2010.06.001

promised merchandise (if users make purchases). Ex ante, users have no easy way to know which sites to trust. A safe-looking site could turn out to be a wolf in sheep's clothing.

These online interactions reflect a two-sided market – with sites actively making decisions about how to present themselves. Good sites want to demonstrate their integrity. But as usual in adverse selection, bad sites pretend they are good.

Facing numerous untrustworthy or even malicious sites, some analysts call for government regulation. In principle, a government agency might examine web sites in search of spam, scams, and harmful programs. To some extent, the FTC and state attorneys general perform such investigations – though their efforts address only a small portion of bad actors. As a practical matter, government intervention seems inapt. For example, Tang et al. (2005) present a model of enforcement of online privacy breaches, finding mandatory government standards appropriate only for the most serious harms.

At the other extreme, users might be left entirely on their own. In complete *caveat emptor* ("buyer beware"), no regulator, computer maker, or IT department helps cure a user's problems. In some respects, *caveat emptor* is a reasonable description of the current state of affairs. (IT departments cannot protect users from getting ripped off, and even computer experts often feel powerless to stop spam.) But unaccountability carries substantial costs – leading users to take excessive precautions, and preventing the formation of otherwise-profitable relationships. Users would buy more products, join more sites, and download more programs were it not for their well-founded fears of fraud and abuse.

Finally, there exists a middle approach between the extremes of government regulation and *caveat emptor*: a non-governmental rating organization. Such an organization would identify specific bad practices, then evaluate sites' behaviors. If evaluations were accurate and low-cost, such ratings might support an equilibrium where good firms receive positive evaluations, and where consumers use only sites with positive ratings. Tang et al. (2005) suggest that rating organizations are appropriate for a broad class of online interactions.

3. Trust authorities

Most prominent among non-governmental rating organizations are so-called "trust" certification authorities. These organizations set out specific criteria for membership, often focusing on privacy or on online safety more generally. The organizations reward their members by offering seals to be placed on recipients' web sites, typically on registration forms and checkout pages. To date, the best-known trust authorities are TRUSTe and BBBonline.

In principle, trust authorities might set and enforce substantive and procedural provisions sufficiently rigorous that certified members are highly likely to satisfy reasonable consumers' expectations of safety. But in practice, critics question the effectiveness of certain trust authorities. LaRose and Rifon (2002) offer a stinging critique: trust authorities have granted multiple certifications to firms under investigation by the FTC for privacy policy violations; trust authorities have declined to pursue complaints against major companies whose privacy breaches were found to be "inadvertent"; and in one case a trust authority even failed to abide by *its own* privacy policy. Ryan (2006) raises similar concerns: in a 2004 investigation after user complaints, TRUSTe gave Gratis Internet a clean bill of health. Yet subsequent New York Attorney General litigation uncovered Gratis' exceptionally far-reaching privacy policy violations – selling 7.2 million users' names, email addresses, street addresses, and phone numbers, despite a privacy policy exactly to the contrary.

As a threshold matter, trust authorities' substantive standards often seem to duplicate existing duties or practices. Consider the obligations in TRUSTe's Program Requirements. The first listed rule, requiring an email unsubscribe function, duplicates Sec.5.(a)(4)(A) of the federal CAN-SPAM Act. Similarly, credit card network rules exactly overlap with TRUSTe's requirement of SSL encryption (or similar technology) to protect sensitive credit card numbers. Boutin (2002) reports that TRUSTe initially lacked any substantive requirements whatsoever (requiring only the *presence* of a privacy policy). Low standards match the predictions of Lizzeri (1999), finding that, under general conditions, a certification intermediary prefers only to reveal whether quality exceeds some minimal standard.

Tellingly, strikingly few certificates have been revoked. For example, the TRUSTe Fact Sheet reports only two certifications revoked in TRUSTe's 10-year history. TRUSTe's small staff has little apparent ability to detect infractions. Instead, TRUSTe's posted procedures emphasize user complaints and sites' self-certifications. When violations have been uncovered, the proof has come from outside complaints, not from TRUSTe itself.

TRUSTe's "Watchdog Reports" also indicate a lack of focus on enforcement. TRUSTe's postings reveal that users continue to submit hundreds of complaints each month. But of the 3416 complaints received since January 2003, TRUSTe concluded that *not a single one* required any change to any member's operations, privacy statement, or privacy practices, nor did any complaint require any revocation or on-site audit. Other aspects of TRUSTe's watchdog system also indicate a lack of diligence.[2]

Finally, trust authorities are paid by the same companies they certify; in the language of Greenstadt and Smith (2005), trust authorities are "captured". With this revenue model, authorities have little short-run incentive to seek higher standards: any such pressure would discourage renewals and future applications – reducing revenues.

Even the creators of trust authorities report disappointment in their development. TRUSTe co-founder Esther Dyson called TRUSTe "a little too corporate", and said TRUSTe lacks the "moral courage" to criticize violations (Boutin 2002). Similarly, the Electronic Frontier Foundation, another TRUSTe co-founder, told the FTC that "it is time to move away from a strict self-regulation approach" (1999).

Table 1 reports selected untrustworthy sites certified by TRUSTe, along with a general statement of the sites' respective practices. As of January 2006, TRUSTe listed all these sites among its certified members.

Facing allegations of low substantive standards, lax enforcement, and ethical compromise, it is unclear what direct benefits site certifications offer to consumers. But at least some consumers seem to regard certification as a significant positive signal. For example, in recruiting web sites to get certified, TRUSTe offers an endorsement from certificate recipient Realty Tracker, which says TRUSTe "convey[ed] trust" and "built confidence" with site visitors, yielding "an increase in registrations". See TRUSTE's Realty Tracker Case Study.

Moreover, firms are well-equipped to evaluate claimed benefits to certification: firms can randomly include or omit a seal, thereby measuring whether a seal increases registrations and sales. Indeed, year after year, hundreds of firms seek and renew TRUSTe certification – suggesting that firms find certification valuable. Furthermore, in the related context of comparison shopping sites, Baye and Morgan (2003) empirically confirm the benefits of certification: merchants with seals can charge a price premium without losing customers.

Even well-known web sites tout their safety certifications. For example, the Microsoft's Online Privacy Policy index features the

[2] For example, TRUSTe failed to update its Watchdog Reports list between June 2004 and spring 2006, an omission corrected only after circulation of a draft of this article. Even in 2009, Watchdog Reports suffer broken links, missing reports, and contradictory document titles.