PRIVACY PROTECTION

Work in Progress

Bruno Jullien
TSE

Yassine Lefouili
TSE

Michael Riordan
Columbia University

Questions

- How does the prospect of continued interaction affect website incentives to protect customer personal information?
- How do regulations restricting websites' collection and use of customer information affect economic welfare?

Motivation: Economics

- Privacy economics literature:
  - Quality attribute: O'Brien and Smith (*FTC WP*, 2014)
  - Asymmetric information: disclosure, reputation-building (Acquisti, Taylor & Wagman, *JEL*, 2015)
- Moral hazard:
  - Websites collect valuable personal consumer information.
  - Consumers imperfectly informed about website policies.
- Learning, signal-jamming (Judd and Riordan, *ReStud*,1994):
  - Consumers unsure about value of continued interaction.
  - Website actions affect "user's experience" and retention

Basic Model

- Unit mass of consumers visit a website in period 0 and may stay or not in period 1.
- Unit mass of third parties in each period, each with exactly one match.
  - A match is good (G) with probability λ, or bad (B) with probability $1-\lambda$;
  - Information allows to identify the match but not the type of match (needs inspection)
- The service is free, but the website:
  - obtains revenue a per user from advertising or merchandising;
  - can sell information to third parties at price v_0.

User's experience

- A consumer is vulnerable to a bad match with probability θ, which is unknown and can be high or low, $\theta = \theta_L$ or θ_H.
  - A good match gives a good experience G
  - A bad match gives a bad exeprience with probability θ and no experience $\emptyset$ otherwise
  - r_0 is consumers' prior belief that $\theta = \theta_L$
- Interpretation
  - targeted advertising;
  - spam, phishing,
  - deceptive ads, or malware.

Second period

- In the second period, the website sells the information for sure so that intrusion occurs
- Based on her experience, a consumer update her beliefs about low vulnerability to r_1
- A consumer returns to the website in period 1 with increasing probability $Q(r_1)$.
  - $Q(r) = \Pr(M_1(r) + \epsilon > 0)$
  - $M_1(r) = \lambda U_G + (1-\lambda)(r\theta_L + (1-r)\theta_H)U_B$
  - $U_1(r) = \delta E(\max(M_1(r) + \epsilon, 0))$

Why would the website protect consumers' personal data?

- Consumers experiencing a bad match are pessimistic about their vulnerability and less likely to visit the website in the future.
- This gives the website incentives to refrain from selling personal data to third parties.
  - Precaution: The website refuses to sell with probability X
- The website takes consumer beliefs ($r_{\emptyset}$, r_G, r_B) as given, but uses privacy policy to influence consumer's experience (probabilities of each event):

$$p_B = \theta(1-\lambda)(1-X) \qquad p_G = \lambda(1-X)$$

$$p_{\emptyset} = X + (1-\theta)(1-\lambda)(1-X)$$

Consumer learning

- Consumers take the website's privacy policy (*X*) as given and use Bayes Rule to form posterior beliefs:

$$r_G = r_0 \qquad r_B = \frac{\theta_L}{\overline{\theta}} r_0$$

$$r_{\varnothing} = \phi(X) \equiv \frac{X + (1-\theta_L)(1-\lambda)(1-X)}{X + (1-\overline{\theta})(1-\lambda)(1-X)} r_0$$

- No news is good news: $r_{\emptyset} > r_G > r_B$
- $\phi(X)$ is decreasing from $r_{\max}$ to r_0 .

Incentives to sell information

- Period 0 value of selling consumer data: v_0
- Period 1 value of retained consumer: $V_1=\delta(a+v_0)$
- Profit $(1-X)v_0+ E(Q(r_1))V_1$
- $X = 1$ is optimal if and only if :

$$\lambda[Q(r_\varnothing) - Q(r_G)] + (1-\lambda)\bar{\theta}[Q(r_\varnothing) - Q(r_B)] \geq v_0 / V_1$$

- The decision to sell information depends positively on consumer beliefs $r_\varnothing$

Equilibrium

- Consumer beliefs "best responds" to privacy policy

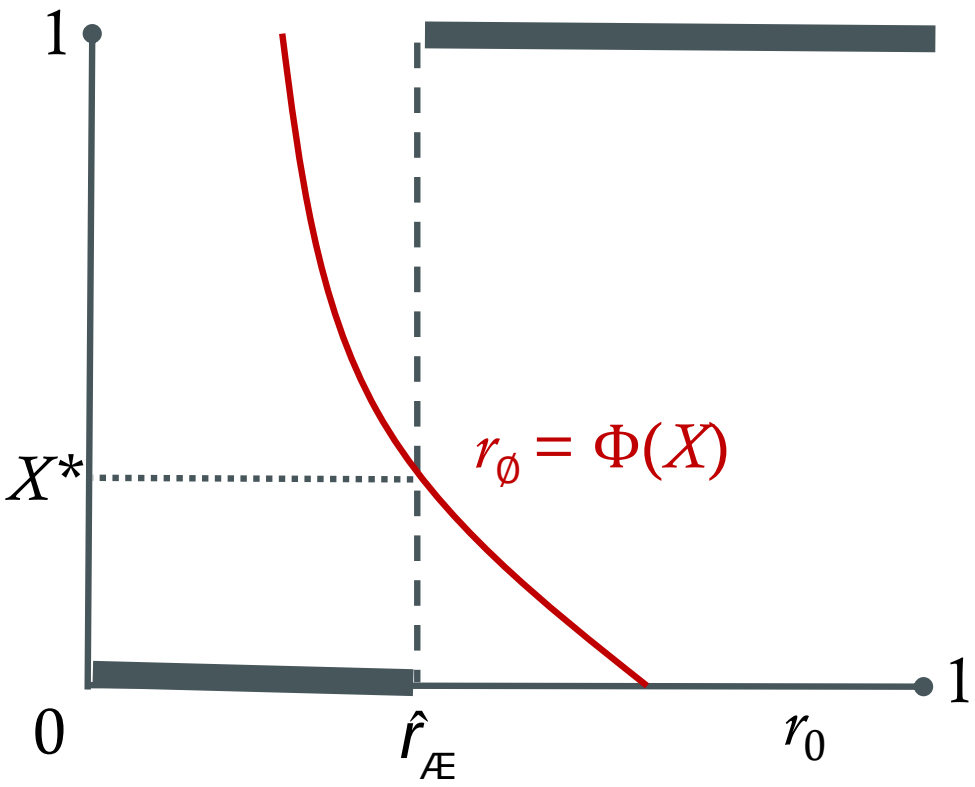
$$r_{\varnothing} = \phi(X^*)$$

- Website privacy policy "best responds" to consumer beliefs:
  - $X^* = 1$ if $r_{\emptyset}$ is above a critical value defined by

$$\lambda[Q(\hat{r}_{\varnothing}) - Q(r_G)] + (1-\lambda)\bar{\theta}[Q(\hat{r}_{\varnothing}) - Q(r_B)] = v_0 / V_1$$

- Lemma: The critical value is increasing in v_0.
- Proposition: Equilibrium precaution X^* is unique and it is nonincreasing in v_0 / V_1.

Random protection ($1 > X^* > 0$): medium ν_0



Competition on the market for information

- Suppose there are N websites with identical consumer demands (θ, ε are common to all websites)
- Consumers multi-home so there is no competition on the consumer side
- Websites compete to sell the consumer information to third parties
  - Assume simultaneous pricing of information by all websites
  - A price above v_0 means refusal to sell, x = *Prob(refusal to sell)*
  - *Total precaution is* $X=x^N$

First-period competition holding retention value constant $V_N = V_1$

- Assume the retention value is not affected by future competition so that $V_N = V_1$
- Then at most two symetric equilibria co-exist
  - **Coordination failure**: if $N > 1$, there always exists an equilibrium with zero price and no precaution.
  - **Coordinated equilibrium:** There also exists a symmetric equilibrium with total precaution X^* equal to the equilibrium value with only one website *(N=1).*
    - ✓ *But prices of information are lower when intrusion occurs*

Effect of competition on the market for information $V_N < V_1$

- Second period: all websites compete and the price of information fall to 0 $\rightarrow$ $V_N = a < V_1 = a + v_1$
- Because the retention value is lower with N websites, the total level of precaution is lower with multiple websites $\rightarrow$ $X = x^N < X^*$
- The total industry income from selling information is lower in both periods, due to price competition
- But if the alternative revenue a is large enough, total profit may increase

Policy/extensions

- Consumer privacy rights:
  - Transparency
  - Opt-out: refuse third-party sale of personal data
- Taxation
- Incentive to screen/inspect buyers of information

Consumer welfare

- In the Short Run (period 0), the consumers benefit from a good match G and are harmed by a bad match B.
  - Consumers are better off with <u>less</u> precaution if matches are beneficial on average:

$$/\,U_G + (1- /\,)\bar{q}U_B > 0$$

- Long Run utility $E(U_1(r_1))$ decreases with precaution because no-intrusion is less informative.

I. Transparency / Commitment

- Suppose the law forces transparency of X
- The website is able observably to commit to a privacy policy.
- If a match is on average beneficial to consumers
  - The website sells with probability 1 in period 1
  - It chooses less precaution (X) in period 0 than the equilibrium level
  - Intuition: The website wants to increase $\Phi(X) \rightarrow$ Equilibrium signal jamming makes no-intrusion less informative about vulnerability, reducing second period profit
- Transparency benefits consumers and the website when matches are on average beneficial

Transparency / Commitment 2

- If a match is on average detrimental to consumers, then an additional effect arises :
  - The website may refrain from selling in period 1.
- If it sets $X_1>0$ then
  - The value of retention V_1 decreases
  - The retention rate $Q(r)$ increases
  - But the slope of $M_1(r)$ decreases
- Transparency leads to less precaution when matches are on average harmful, if Q is concave or weakly convex,
- The welfare effects are then ambiguous

II. Screening / inspection

- The website can screen the third party at random inspection cost z, drawn from $F(z)$
- The website inspects only if it intends to sell to G and refuse to sell to B.
- A privacy strategy is defined by a pair (X, Z):
  - $F(Z)$ is the probability of screening
  - X is the probability of selling information to an unscreened third party (the fraction of the population of consumers)

Signal jamming

- Probability that B obtains consumer data:

$$\pi_M(X,Z) = (1-\lambda)[1-F(Z)](1-X)\}$$

- Probability that neither G nor B obtains information:

$$\pi_N(X,Z) = (1-\lambda)\,F(Z) + [1-F(Z)]X$$

- Privacy policy to influence consumer experiences (B, G, ∅):

$$p_B = \theta\pi_M(X,Z) \qquad p_G = 1-\pi_M(X,Z)-\pi_N(X,Z)$$
$$p_\emptyset = \pi_N(X,Z) + (1-\theta)\pi_M(X,Z)$$

Consumer learning

- Consumers take the website's privacy policy (X, Z) as given and use Bayes Rule to form posterior beliefs:

$$r_{\varnothing} = \Phi(X,Z) \equiv \frac{(1-\theta_L)\pi_M(X,Z)+\pi_N(X,Z)}{(1-\bar{\theta})\pi_M(X,Z)+\pi_N(X,Z)} r_0$$

- No news is good news: $r_{\emptyset} > r_{\mathbf{G}} > r_{\mathrm{B}}$
- Φ decreases in X and in Z

Incentive to screen third parties

- Net benefit of selling information to G:

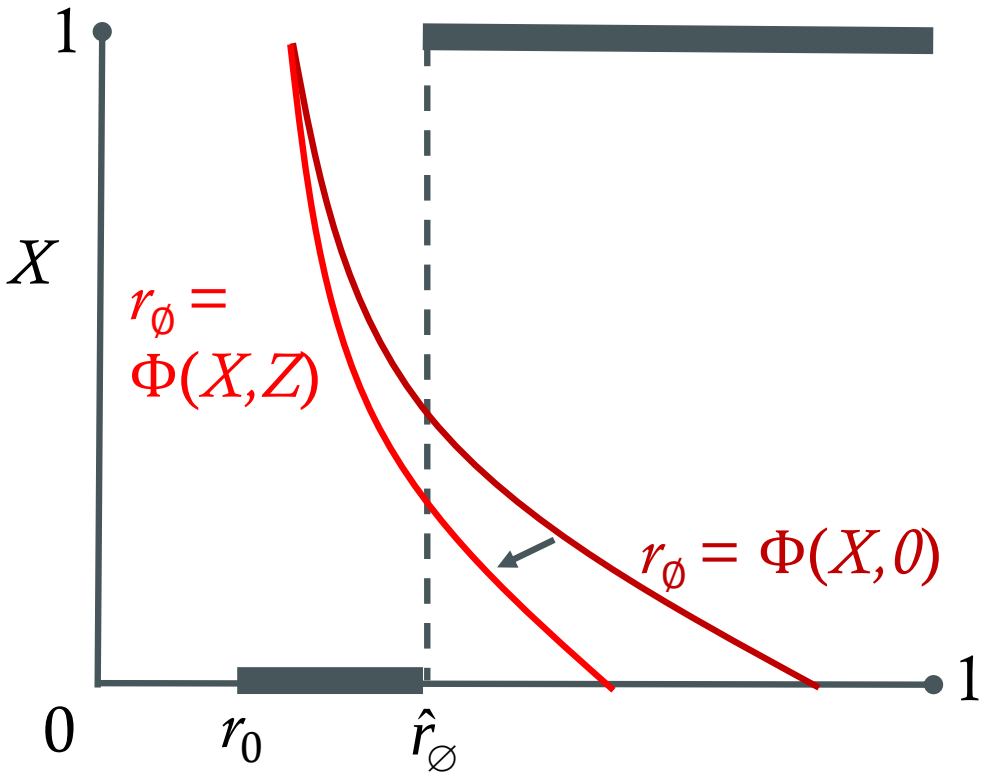
$$\Delta_G(r_\varnothing) = \lambda(v_0 - [Q(r_\varnothing) - Q(r_G)]V_1)$$

- Net benefit of denying information to B:

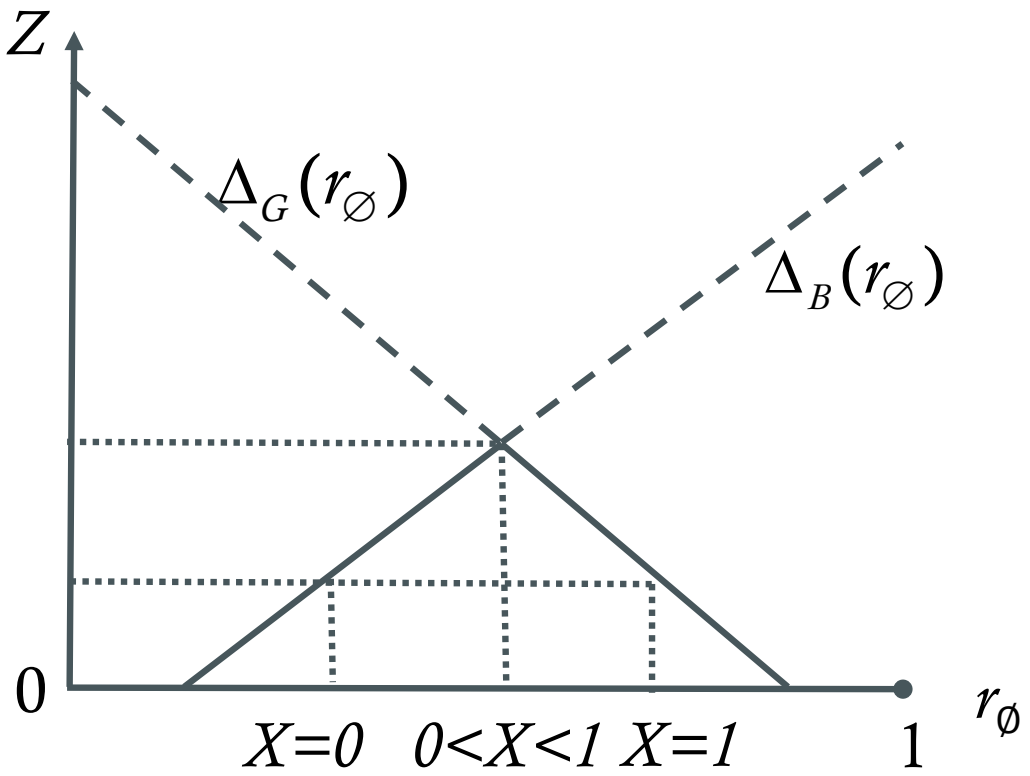
$$\Delta_B(r_\varnothing) = (1-\lambda)(\bar{\theta}[Q(r_\varnothing) - Q(r_B)]V_1 - v_0)$$

- Precaution: $X^* \in \arg\max_{0 \le X \le 1} \{X[\Delta_B(r_\varnothing) - \Delta_G(r_\varnothing)]\}$
- Screening: $Z^* = \max\{\min[\Delta_G(r_\varnothing), \Delta_B(r_\varnothing)], 0\}$

Inspection and precaution are substitutes



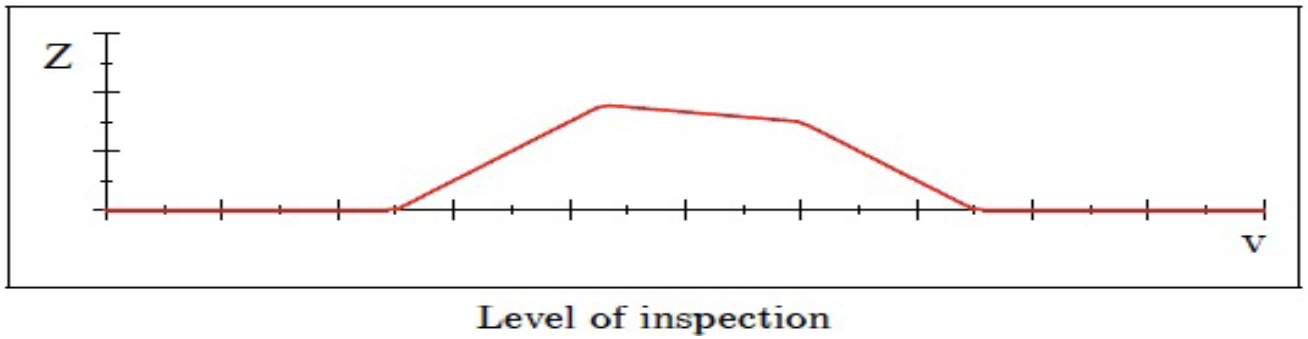
Choice of $X(r_\emptyset)$ and $Z(r_\emptyset)$



back

Screening: Summary

- Proposition: There exists a unique equilibrium (X^*, Z^*) such that
  - The posterior $r_\emptyset$ is nondecreasing in v_0.
  - X is nondecreasing in v_0 (if Q_Z is not too negative)
  - No screening if the average vulnerability is small
  - Z is nondecreasing in v_0 when small and nonincreasing in v_0 when large (random and weak protection).



Level of inspection

sensitivity

Screening: policy

- **Transparency/Commitment:** Suppose the website were able observably to commit to a privacy policy.
  - Then if matches are beneficial on average it would choose
    - ✓ less precaution (X) than the equilibrium level;
    - ✓ less screening *(Z)*
    - ✓ Consumer surplus may increase or decrease
- **Taxation** increases precaution and may increase inspection (when v_0 is large) or decrease it (when v_0 is small*)*

Conclusions

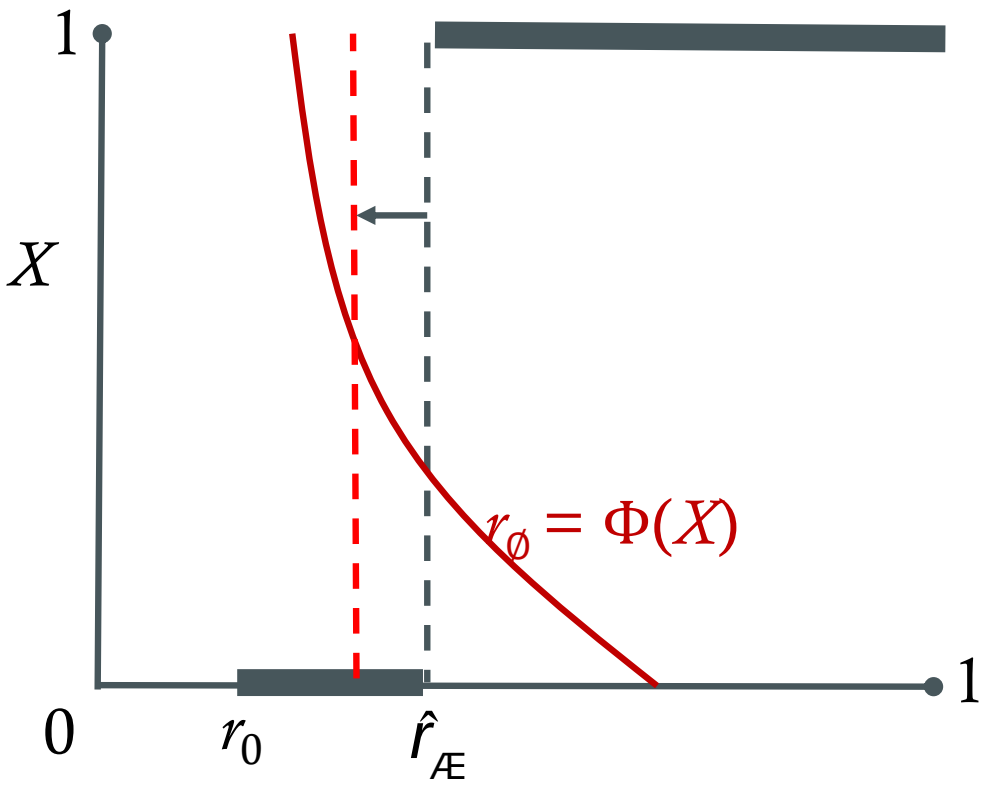
- Market forces provide positive but imperfect incentives for privacy protection.
  - Equilibrium incentives can lead to excessive precaution and deficient screening when the short-run value of selling information to third parties is sufficiently small, and if consumers on average benefit from matching with third parties.
- Transparency is welfare enhancing if consumers on average benefit from matching with third parties but could be welfare reducing otherwise
- Inspection (screening G and B) and precaution are substitutes.
- Taxation of information market raises equilibrium precaution
- Allowing consumers to opt-out of sharing their personal data leads to more precaution and has ambiguous welfare effects
- TO BE DONE: No history based discrimination (OLG model)

SUPPLEMENTARY MATERIAL

III. Taxation

- Suppose a tax t is imposed on the sale of information
- The ratio v_0/V_1 decreases to $\frac{(1-t)v_0}{a+(1-t)v_0}$
- Imposing a tax on the market for information raises the level of precaution (provided that there are other sources of revenue).
- The tax reduces the website's profit
- It raises consumer surplus if matches are detrimental on average and consumers are impatient
- It reduces consumer surplus if matches are beneficial on average

Taxation



IV. Opt out regime

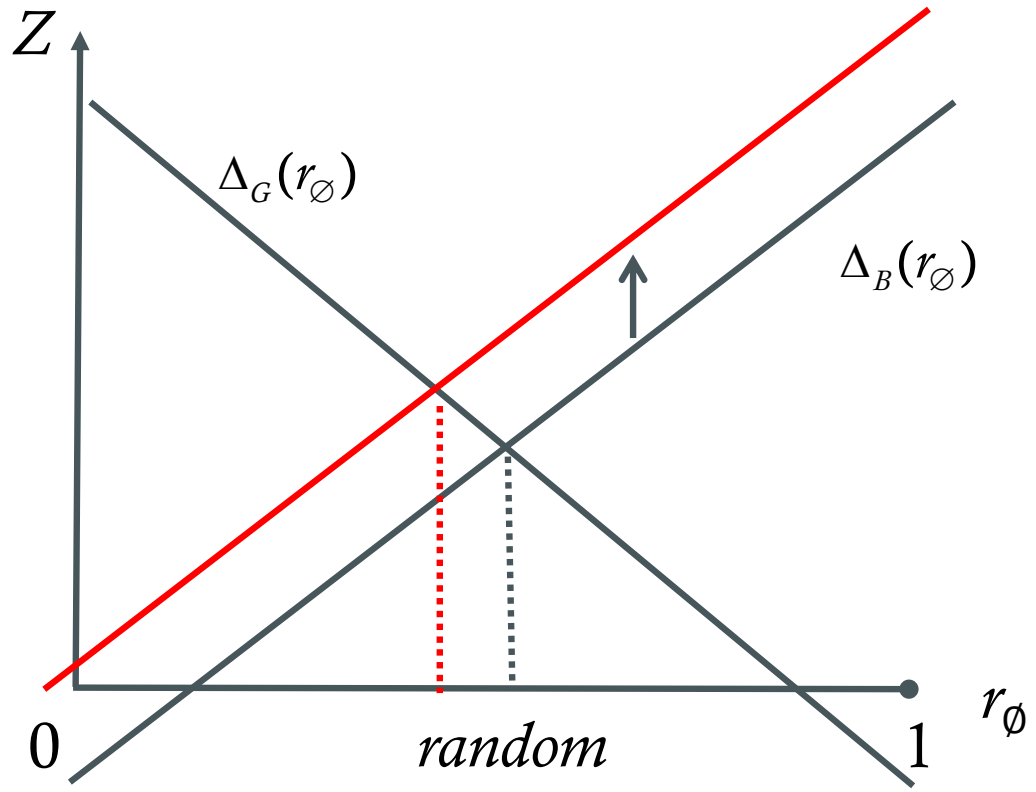
- Suppose consumers can "opt out", i.e. require that personal information not be sold in Period 1.
  - The consumer opts out if sufficiently pessimistic $r_1 < \check{r}$.
  - Retention rate is $Q(\check{r})$ and increases because $Q(\check{r}) > Q(r_1)$, while the value of retention is reduced to $V < V_1 = 1$.

- If $r_B < \check{r} < r_0$, then consumers:
  - opt-in in Period 0;
  - opt out in Period 1 after a bad match. The equilibrium characterization is similar to to the baseline case, except $Q(r_B)V_1$ replaced with $Q(\check{r})V$.

Opt out

- Proposition: Assume $r_B < \check{r} < r_0$ and $Q(r_B)V_1 > Q(\check{r})V$. Then permitting consumers to opt out increases precaution
- The website is worse off.
- The effect on consumers welfare is ambiguous.
  - Direct benefit of opting out when pessimistic about vulnerability.
  - Welfare loss from less informative signal about vulnerability due to greater precaution.

Opt out, 0<X<1


Z
$\Delta_G(r_\varnothing)$
$\Delta_B(r_\varnothing)$
$r_\varnothing$
0
random
1

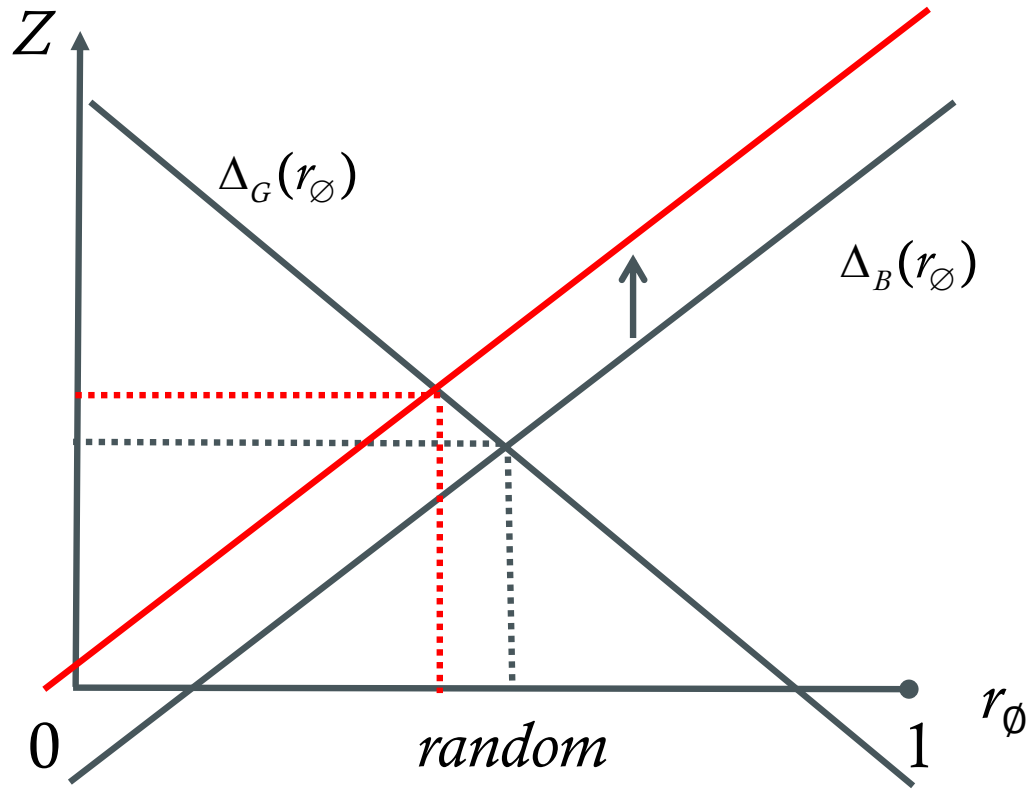
Opt-out in other cases

- Assume consumers opt-in in period 0 and $Q(r_B)V_1 > Q(\check{r})V$.
- If $r_0 < \check{r} < \Phi(1)$, then opt out provides greater incentives for precaution;
- If $\Phi(1) < \check{r} < \Phi(0)$, then opt out leads to random precaution if v_0 is below a critical value, and no privacy otherwise;
- If $\check{r} > \Phi(0)$, then opt out eliminates any incentives for protecting privacy.

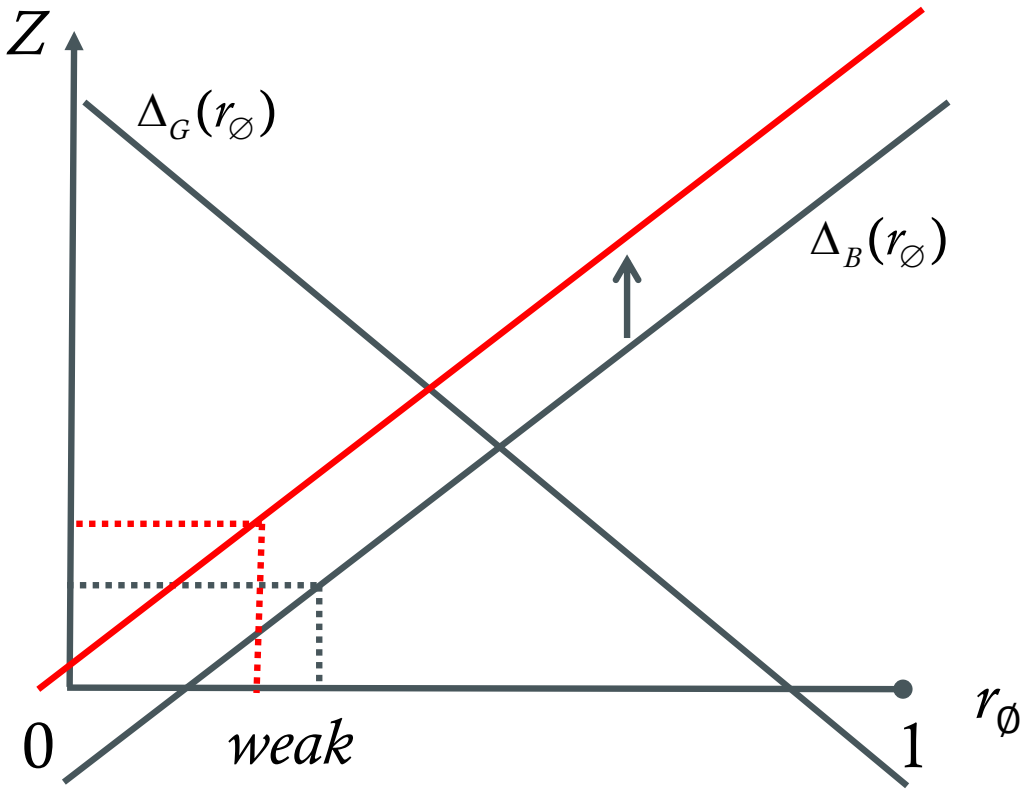
Opt-out with screening

- Suppose $r_B < \check{r} < r_0$ and $Q(r_B)V_1 > Q(\check{r})V$.
- Opt-out (weakly) increases screening because it increases the value of preventing a bad match.
- If $1 > X > 0$, then opt-out has an ambiguous effect on precaution, because more screening mitigate the increase in precaution (higher Z reduces X).

Opt out and screening, 0<X<1



Opt out and screening, X=0



Opt out with screening: precaution