

Identity Theft and Data Breaches

William Roberds¹ Stacey L. Schreft²

September 25, 2009

¹Federal Reserve Bank of Atlanta

²Mutual Fund Research Center, LLC

Identity theft: some facts

- 2006 FTC Survey (Synovate 2007)
 - 3.7% of U.S. households victimized
 - Estimated annual cost:
 - FTC: \$16 billion
 - Schreft (2007) adjusts up to \$64 billion
- Big question: *is this a market failure?*
 - (Anderson, Durbin, and Salinger JEP 2008)

[PRINT THIS](#)

Bernanke Victimized by Identity Fraud Ring

Exclusive: According to court documents, the Fed chairman and his wife were swindled in 2008 by a skilled team of crooks.

By **Michael Isikoff** | Newsweek Web Exclusive

Aug 25, 2009

If ever there were living proof that identity theft can strike the mighty and powerful as well as hapless consumers, look no further than the nation's chief banker: Ben Bernanke. The Federal Reserve Board chairman was one of hundreds of victims of an elaborate identity-fraud ring, headed by a convicted scam artist known as "Big Head," that stole more than \$2.1 million from unsuspecting consumers and at least 10 financial institutions around the country, according to recently filed court records reviewed by NEWSWEEK.

Last summer, just as he was dealing with the first rumblings of the financial crisis on Wall Street, Bernanke learned that a thief had swiped his wife's purse—including the couple's joint check book. Days later, someone started cashing checks on the Bernanke family bank account, the documents show. "It's fair to say he was not pleased," said one close associate of Bernanke, who asked not to be identified discussing what the Fed chairman considers a private matter.

ID theft a market failure?

- Popular wisdom: **YES**
 - “too much” data (PII) collected & stored
 - unauthorized access too easy
 - ID theft too common
- Legal literature: **YES**
 - Swire (2003): credit & payments industry has not delivered “efficient confidentiality” of PII
- Elected officials: **YES**, e.g.
 - U.S. 2003 FACT Act \implies 30+ pages of Federal regulations
 - State laws (breach notifications)

Industry view: ID theft **NOT** a market failure, because

- fraud losses **low** relative to usage of systems
 - e.g., \$3+ trillion card payments/year
- much fraud is **opportunistic** and **low-tech**
 - not caused by loss of commercial data
 - can be deterred by collecting & analyzing lots of data
- Implication
 - if ID theft a problem, solution is to collect **more** (e.g., biometric) data

- Theoretical examination of popular wisdom/ industry view using **economics of payments**
- EOP: study of mechanisms that allow people to trade when
 - 1 People want to consume at different times
 - 2 Limited enforcement of promises of future actions

Key tradeoff

Payment systems must deter 2 kinds of identity thieves

- *Unskilled frauds* (“opportunists”): **discouraged** by systems’ collection of PII, data security not important for deterrence
- *Skilled frauds* (“hackers”): possibly **enticed** by systems’ collection of PII, data security key for deterrence

Efficiency requires balance between data collection and data security

Model: basic features

- Infinite horizon, continuous time
- Large number of risk-neutral agents, congenitally split into multiple (2) **groups** G_A & G_B
 - All transactions occur within a group
- **Overlapping generations:** random subset of each group dies at dates $0, 1, 2, \dots$, is replaced by new agents

- Agents also partitioned according to **legitimacy** and **type**
 - **Legitimate agents:** can produce tradeable goods, no talent for fraud, measure $1 - F$
 - **Frauds:** cannot produce goods, but can impersonate others, measure F
- Agents distributed over **types** (virtual locations); many agents at each location
- (Legitimate) type $y \in [0, 1]$ agent can **produce** unit of nondurable good of type y at times $y, 1 + y, 2 + y, \dots$ at cost c
- At all other times $y' \in [0, 1], y' \neq y$, agent of type y wants to **consume** goods of (randomly selected person of) other types, generating flow utility $u > c$
 - *<consumption displacement>*

- Each agent has unique **identity**, time-invariant vector of personal data (not transactions history), effectively infinite dimension
- Subset of identity (**PII**) may be assembled, stored, secured at positive cost
- Agent's group, type, legitimacy & identity are **private information** subject to
 - **costly & imperfect verification** and/or
 - **revelation** through agent's behavior (not available instantaneously)
- *<informational frictions ⇒ imperfect enforcement of promises>*

Trade through payment networks

- No repeated interactions \Rightarrow unless agents' behavior can be tracked, **no agent would ever produce**
- Payment networks modeled as **clubs** for sharing information on agents' behavior
- One club for each group; **no info sharing across clubs**; club membership **voluntary**
- Information compiled by club:
 - (1) members' **production history** (has an agent produced goods for other members?)
 - (2) members' **PII** (so as to correlate individuals with histories, distinguish new members from old)
- Production information, PII available at discrete dates $0, 1, 2, \dots$

Clubs' operation

- At 0, 1, 2, ... clubs G_A & G_B open membership to all
- Agents wishing membership in club i must submit PII of dimension $d_i > 0$ (if not already on file)
- Each club member receives uncounterfeitable **credit card** entitling agent to goods produced by other (legitimate) club members
- In general, **clubs not viable** (not IR for legit agents) if legit agents must produce for all agents, including frauds
- \Rightarrow Clubs **exclude nonproducers** at discrete dates, when production info becomes available
- Can apply penalties to non producers (bill collection) but only if stored PII corresponds to "real identity"

- Club i collects, stores data d_i
- **One-time cost** K when member first joins, plus **proportional storage cost** kd_i per unit time
 - K, k include intangibles (“loss of privacy”)
 - if data not stored, initial verification cost must be incurred
- Club i applies **security level** (“skill threshold”) s_i at cost ℓs_i
- Hacking skills s have some distribution $\Phi(s)$ over population of frauds

How identity theft occurs in the model

Frauds can join clubs by impersonating a legitimate agent; frauds either **skilled** or **unskilled**

- Unskilled group i frauds ($s \leq s_j$) can join club i **w/o revealing true identity** at **effort cost** εd_i where $\varepsilon \geq 0$ has distribution Γ
- Skilled group i frauds ($s > s_j$) lower effort cost by stealing (breaching) data held by other club j (club i rejects duplicate identities) at **lower effort cost**

$$\varepsilon \max\{d_i - \eta d_j, 0\}$$

- $\eta \in (0, 1)$ measures **overlap** between 2 clubs' databases of members' PII; determines spillover effects
- *<note: successful ID theft always revealed after one period>*

Identity theft: costs per incidence of fraud

- 1 c : cost to legitimate members of club i of **providing goods to identity thieves** (e.g., FTC: median cost \$1,350/ stolen ID)
- 2 L : **additional cost** (time, inconvenience, intangible) to club i of resolving fraud (FTC: resolution time 10 hours/ stolen ID)
- 3 B : **cost** to club i when club j ID theft results from **breach** of club i 's data (Ponemon Institute 2006: $< \$100$ / record breached)

Model calculations assume $c + L > B$

Steady-state allocations and objectives

- **Allocation:** PII and security (d_i, s_i) for each club $i = G_A, G_B$
- **Objectives:** Each club i chooses (d_i, s_i) to maximize value of legitimate membership

$$\underbrace{(1 - F)(u - c)}_{\text{Transaction benefit}} - \underbrace{[(1 - \beta)K - kd_i - \ell s_i]}_{\text{data costs}} - \underbrace{[(\rho_i^U + \rho_i^S)(c + L) + \beta \rho_j^S B]}_{\text{costs of ID theft}}$$

(cf. Varian 2004, Grossklags, Christin, & Chuang 2008)

- **Symmetric Nash equilibrium** (d^*, s^*)
 - maximizes club i membership value when j also chooses (d^*, s^*)
- **(Constrained) efficient** (d_p, s_p)
 - maximizes steady-state value of legitimate club membership for both clubs
- Game plan: characterize potential “market failures” as deviations of Nash from efficient allocation

Nash equilibrium: sources of inefficiency

- Externalities present in both decision variables; each club
 - 1 Internalizes deterrence **benefits** of PII collection d but not **costs** to other club (facilitation of future skilled ID theft)
 - 2 Does not internalize full **benefits** of data security s (reduction in skilled ID theft to other club)

Nash equilibrium: manifestations of inefficiency

Main analytical results

With sufficiently high data overlap ($\eta \rightarrow 1$) and low data costs ($k, \ell \rightarrow 0$)

- 1 inefficient **overaccumulation of PII, inefficiently low levels of data security** applied
 - 2 Unskilled ID theft **inefficiently low** (because too much PII collected)
 - 3 Skilled ID theft **inefficiently high** (data undersecured)
 - 4 For (ℓ/k) bounded (e.g. persistent intangible privacy cost), **total ID theft inefficiently low** (unskilled effect dominates)
- I.e., each Nash club sacrifices members' privacy as defense against other club's insufficient security & data overcollection

Qualitative predictions of model

Variable	Eq. vs. efficient value	Popular wisdom?
Data length d	Higher	Yes
Data security s	Lower	Yes
Skilled ID theft rate	Higher	Yes
Unskilled ID theft	Lower	No
Total ID theft	Lower	No

Policy approaches

Summary of analytical / numerical results

- 1 Increase civil liabilities for a data breach (up to economic loss)
 - limited effectiveness; does not shut down substitution of data collection for security
- 2 Enforce higher security standards
 - can approximate efficient allocation but requires very high data security standard
- 3 Constrain PII collected
 - improves security incentives & welfare, but leads to high ID theft rate of both types

- 1 Paper develops a meaningful concept of “efficient confidentiality”
 - levels of PII and security that enable beneficial exchange at minimum cost
- 2 Characterizes potential market failures
 - can be consistent with stylized facts about ID theft
 - welfare losses may be “invisible” due to intangible costs, but real
- 3 Analyzes policy interventions
- 4 Provides generalizable framework