Encryption and Data Loss

Amalia Miller and Catherine Tucker

University of Virginia and MIT Sloan

How does the adoption of encryption tools affect publicized data loss?

- Empirically, how does hospital adoption of encryption tools affected instances of publicized data loss?
- Direction of relationship is not clear
  - On the one hand, encryption may reduce instances of publicized data losses, as they are less likely to be reported on.
  - On the other hand, if encryption leads to carelessness in other dimensions, it may lead to more instances of publicized data loss.

Our results are surprising.

- Find no evidence that encryption tools prevent 'publicized' data losses
- Instead adoption is correlated with an increase in instances of publicized data loss due to loss of equipment or employee fraud
- Results hold when look at adoption motivated by external encryption exemptions in state data breach notification laws.

Even if data is encrypted its loss can still matter

- Encrypted data loss can still make the news (Rainbow Hospice)
- Encrypted data can still be breached. (Troy Beaumont Hospital in Detroit)
- Proposed Federal 'Data Breach Notification Act' (Senate Bill 139)
  - However, encryption without appropriate data management and user control policies may not be enough.

We have data on publicized data breaches by hospitals

- Data from 2005-2008 on publicized security breaches within the US.
- These data were collected by the 'Open Security Foundation' volunteers
- Robustness checks to ensure accuracy and representativeness

We have many explanatory variables in our regressions

- Hospital characteristic data from annual American Hospital Survey on [all] US hospitals.
- Encryption adoption data from 4 years of Healthcare Information and Management Systems Society (HIMSS) Analytics^{TM} Database (HADB) [most] US hospitals.
  - Very generalized survey question on encryption.
- Data on status of state data breach notification laws

Figure: Growth in use of encryption software

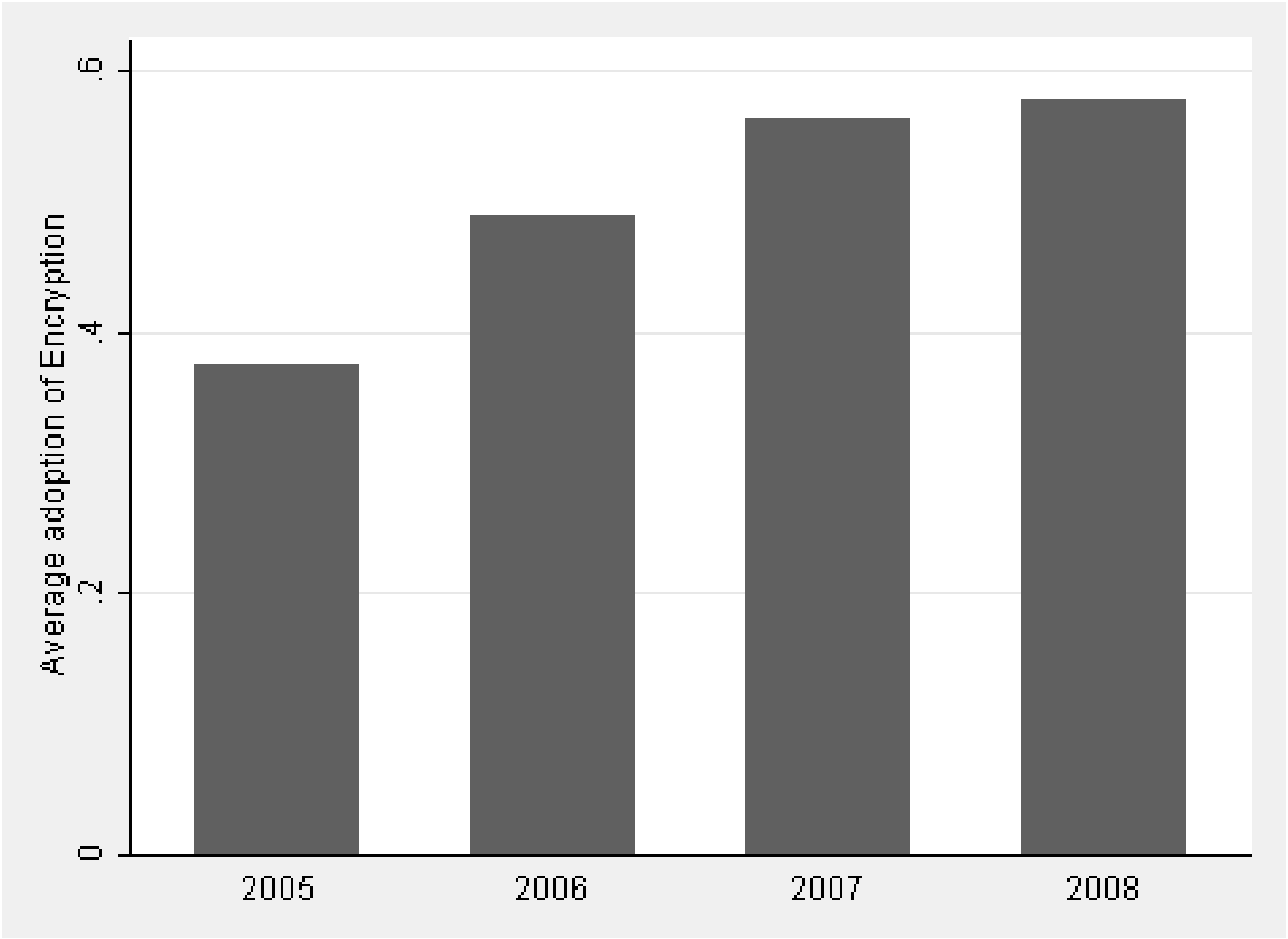

Table: Summary Statistics for Full Sample

	Mean	Std. Dev.	Min	Max
Any Data Breach	0.019	0.14	0	1
Data Breach: Lost Equipment	0.0066	0.081	0	1
Data Breach: Theft	0.0079	0.089	0	1
Data Breach: Fraud	0.0037	0.061	0	1
Encryption	0.50	0.50	0	1
Physician Documentation	0.24	0.43	0	1
Firewall	0.59	0.49	0	1
Clinical Data Repository	0.66	0.47	0	1
Data Warehouse Financial	0.22	0.42	0	1
Data Warehouse Clinical	0.17	0.38	0	1
EMPI (Enterprise Master Person Index)	0.30	0.46	0	1
State Data Breach Law	0.50	0.50	0	1
Encryption Exception	0.39	0.49	0	1
Payroll Expense per Patient ($000)	7.55	9.03	0.0027	589.1
Capital Expense per Patient ($000)	18.0	21.6	0.0068	1549.7
Admissions (000)	7.68	9.32	0.012	108.6
# Hospitals in System	21.7	40.9	0	170
Average Pay in County ($000)	34.3	10.0	13.5	102.2
Total Outpatient Visits (000)	128.8	187.1	0	3282.5
Full Time Employees (000)	0.95	1.31	0.011	17.8

17,300 observations for 4,325 hospitals over 4 years.

We use a binary choice model

$$\text{Prob}(DataBreach_{it} = 1 | Encryption_{it}, X_{it}) = \Phi(Encryption_{it}, X_{it}, \gamma) \quad (1)$$

- Robust to
  - Linear probability models with fixed effects

	(1) Data Breach	(2) Data Breach: Lost	(3) Data Breach: Theft	(4) Data Breach: Fraud
Encryption	0.403***	0.418**	0.0830	1.415***
Payroll Expense per Patient	-0.0484***	-0.0183*	-0.0897*	-0.0680**
Capital Expense per Patient	0.000399	0.000413	-0.00659	0.00855**
Admissions (000)	0.0141**	0.00484	0.0160*	0.0134
# Hospitals in System	-0.000545	-0.00132	-0.00343***	0.00497**
Average Pay in County	0.0150***	0.00908*	0.0141***	0.0263***
Total Outpatient Visits (000)	0.000490***	0.000410*	0.000537**	-0.000896
Full Time Employees	-0.0499	0.0272	-0.0254	0.0225
PPO	-0.566***	-0.348*	-0.886***	0.326
HMO	0.344***	0.0603	0.706***	-0.267
Physician Documentation	-0.111*	-0.176	-0.210**	0.113
Firewall	0.141	0.159	0.217	-0.635**
Clinical Data Repository	0.197**	0.206	0.122	0.438**
Data Warehouse Financial	0.180***	-0.256	-0.0128	0.922**
Data Warehouse Clinical	-0.0769	0.119	-0.373**	0.345*
EMPI	0.149**	0.470***	0.420**	-1.218***
State Data Breach Law	0.0346	0.286	-0.0888	3.426***
State Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Observations	16596	17300	17300	8888
Log-Likelihood	-1235.5	-480.3	-537.3	-201.6

Panel data from 2005-2008 for 4,325 hospitals in the US. Probit specification.
Robust standard errors clustered at the state level. * $p < 0.10$, ** $p < 0.05$,*** $p < 0.01$ Dependent variable in Columns (1) is an indicator variable for whether there was any data breach at the hospital. Dependent variable in Columns (2)-(4) are indicator variables for whether there was a data breach due to equipment loss, theft or fraud.

We use data on changes in state laws to address concerns about endogeneity

- Concern that hospitals may have adopted encryption because they knew they were at increasing risk.
- Focus on exogenous adoption that can be explained by states adopting a state breach notification law with an exception for encryption, relative to state laws that made no exception.
  - Already include data breach laws in main specification.
  - Only exclude whether there is an exception for encryption in the law.

Table: The effect of breach notification encryption exceptions on encryption software adoption

	Encryption adoption before law	Encryption adoption after law	Difference
States with no encryption exception	0.50	0.54	-0.038
States with encryption exception	0.38	0.52	-0.13

Table: Biprobit Specification

	(1) Data Breach	(2) Data Breach: Lost	(3) Data Breach: Theft	(4) Data Breach: Fraud
Loss of Data				
Encryption	0.384**	0.406*	0.0767	1.406***
Payroll Expense per Patient	-0.0458**	-0.0184**	-0.0897*	-0.0681**
Admissions (000)	0.0137**	0.00494	0.0161*	0.0135
Average Pay in County	0.0162***	0.00910*	0.0141***	0.0263***
Total Outpatient Visits (000)	0.000493**	0.000411**	0.000538**	-0.000895
PPO	-0.572**	-0.348*	-0.886***	0.326
HMO	0.357**	0.0614	0.707***	-0.267
Clinical Data Repository	0.200	0.207	0.122	0.437**
Data Warehouse Financial	0.176	-0.256	-0.0127	0.922**
Data Warehouse Clinical	-0.0816	0.119	-0.374**	0.345*
State Data Breach Law	0.0688	0.286	-0.0887	2.349***
Encryption Software Adoption				
Encryption Exception	0.374***	0.212***	0.198***	0.227***
Payroll Expense per Patient	-0.0187**	-0.0201***	-0.0201***	-0.0214***
Admissions (000)	0.0266***	0.0270***	0.0258***	0.0241***
Average Pay in County	0.00440*	0.00468**	0.00440**	0.00439*
Total Outpatient Visits (000)	0.000279	0.000300**	0.000288	0.000324**
PPO	-0.0173	-0.0339	-0.0229	-0.0373
HMO	0.213***	0.230***	0.228***	0.243***
State Data Breach Law	-0.321***	-0.129***	-0.185***	-0.151***
State Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Other Hospital Controls	Yes	Yes	Yes	Yes
Observations	17300	17300	17300	17300
Log-Likelihood	-12022.8	-11300.4	-11356.7	-11051.4

We do many checks

- Traditional IV with fixed effects.
- Robust to excluding data that came from state mandated reporting.
- No pre-trend in states with encryption exemptions.
- No positive effect on encryption adoption if hospitals are excluded from law.

Encryption by itself may not be a solution.

- Surprisingly, encryption does not seem empirically to reduce the likelihood of a hospital experiencing a publicized data loss.
- This matters because
  - Bad publicity is bad publicity.
  - Still risks associated with loss of encrypted data.
  - Firm and Government Policy has focused on encryption, but may need also to encompass policies on data management (especially on portable devices) and user access controls.