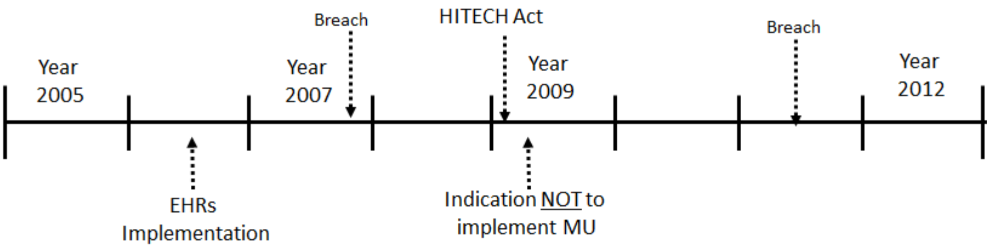
Online Appendix

1. Additional Information on Research Variables

i. Timing and Coding of EHRs/MU Initiative

Figure A1 illustrates the timing of events in our study as well as some particular coding examples.

Figure A1. Timing of Events and Coding Examples
(a) EHRs implemented in 2006 and No MU Initiative



Coding Example

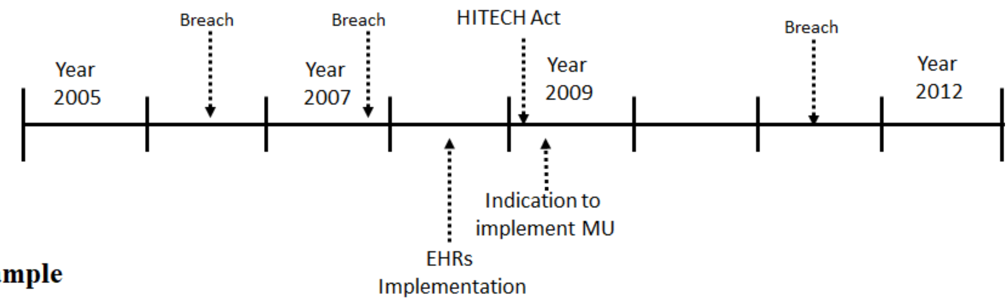
Cox Proportional Hazards Model (Main Analysis)

DV: Breach			1				1	
EHRs Adoption	0	0	1	1	1	1	1	1
Pursuit of MU	0	0	0	0	0	0	0	0

Panel Analysis (Probit Model)

DV: # of Breaches	0	0	1	0	0	0	1	0
EHRs Adoption	0	0	1	1	1	1	1	1
Pursuit of MU	0	0	0	0	0	0	0	0

(b) EHRs implemented in Year 2008 and MU Initiative



Coding Example

Cox Proportional Hazards Model (Main Analysis)

DV: Breach		1	1				1	
EHRs Adoption	0	0	0	0	1	1	1	1
Pursuit of MU	0	0	0	0	1	1	1	1

Panel Analysis (Probit Model)

DV: # of Breaches	0	1	1	0	0	0	1	0
EHRs Adoption	0	0	0	0	1	1	1	1
Pursuit of MU	0	0	0	0	1	1	1	1

ii. Number of Breaches and Classification Methods

Table A1 summarized the number of breach occurrences during the study period.

Table A1. Number of Breaches

Year	2005	2006	2007	2008	2009	2010	2011	2012
Overall	11	25	46	24	38	96	85	72
Accidental	8	19	45	15	24	69	55	50
Malicious	3	6	1	9	14	27	30	22

The breaches listed in the PRC, the main source of our breach data, shows eight types of breaches: unintended disclosure (DISC), hacking or malware (HACK), payment card fraud (CARD), insider (INSD), physical loss (PHYS), portable device (PORT), stationary device (STAT), and unknown or other (UNKN). Among these, HACK, CARD, and INSD relate to malicious breaches. The DISC type pertains to accidental breaches. However, other types may be either malicious or accidental. For these, the authors read all the details of the breach incidents to make sure that they were correctly classified. This coding was rather straightforward because a malicious breach contains such information as who gained unauthorized access to information and how the information was used. For example, an accidental breach may state, "a flash drive with the personal information of graduate medical residents and fellows was reported missing…" A malicious breach may mention something like, "Patient information was stolen from the [name of facility] and used to file fraudulent tax returns with the Internal Revenue Service.…" The breaches from the HHS website has similar classification such as theft, hacking/IT tools, unauthorized access/disclosure, etc. We coded them in a similar way.

iii. Validation of EHRs variable

To confirm that implementing EHRs in the AHA dataset refers to adopting more advanced EHRs components, we further analyzed our dataset, which can be found in Figure A2. For instance, the boxplot in blue shows the distribution of the number of EHRs components owned by the hospitals that implemented EHRs in the AHA dataset. That is, deploying EHRs in the AHA dataset corresponds to having a median of four major EHRs components out of five (computerized provider order entry, physician documentation, clinical data repository, clinical decision support systems, and order entry) in the HIMSS database and to be at a median of Level 2 out of 3 levels (the boxplot in red), and a median of Level 3 out of 4 levels (the boxplot in green), according to the EHRs level classification made by Furukawa et al. (2010) and Appari et al. (2012), respectively. Further, a majority of hospitals that implemented EHRs in the AHA survey as of 2010 also responded that they electronically exchange patient demographics (69%), clinical care records (72%), laboratory results (67%), medication history (72%), and radiology reports (67%) with hospitals in their health system. This confirms that EHRs implementation in the AHA dataset indicates a high degree of information exchange through integration, too.

However, the EHRs variable in the AHA is not simply about how many EHRs components are owned by hospitals. Our EHRs variable from the AHA survey is not very highly correlated with owning more EHRs components in the HIMSS database although the correlation coefficient is positive (*Corr* = 0.39) (see Table A2). We believe that this is because the features of the EHRs defined in the AHA survey that require deriving clinical information from multiple sources and replacing paper-based records are not fully captured by mere ownership of EHRs components.

Similar in spirit to the EHRs variable, taking an MU initiative is positively correlated with implementation of EHRs components (*Corr* = 0.31) (See Table A2). Interestingly, the correlation between an MU initiative and actual accreditation for Stage 1 MU as of 2012 is only 0.44. In our dataset, only 30% of the hospitals with an MU initiative successfully attested for MU Stage 1 as of 2012. This shows that an MU initiative is both theoretically and empirically distinct from actual attestation of MU.

Figure A2. EHRs in the AHA Dataset and EHRs Components in the HIMSS Database

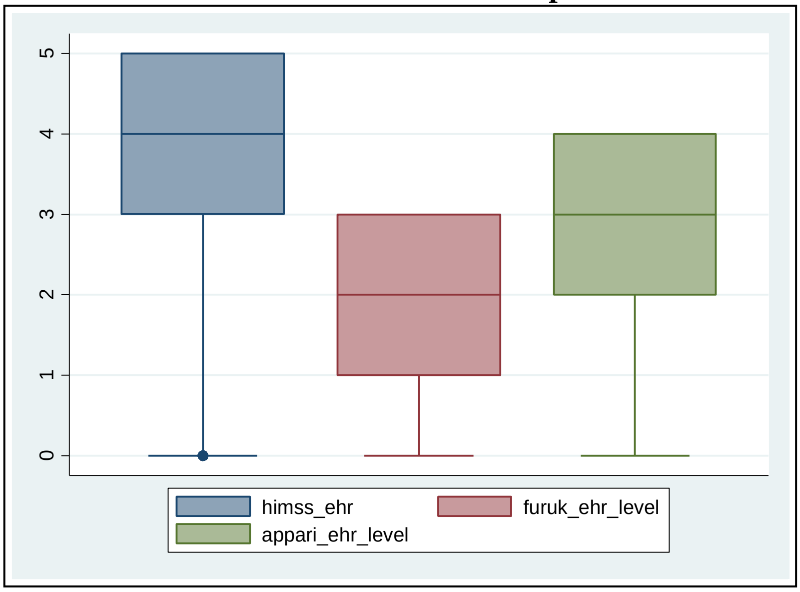


Table A2. Correlations between EHRs, MU Initiative, and Related Variables

	(1)	(2)	(3)	(4)	(5)
(1) EHRs in this study					
(2) MU Initiative in this study	0.33				
(3) EHRs components count	0.39	0.31			
(4) EHRs Level (Furukawa et al. 2010)	0.38	0.40	0.69		
(5) EHRs Level (Appari et al. 2012)	0.40	0.42	0.76	0.91	
(6) Stage 1 MU accreditation	0.15	0.44	0.25	0.31	0.31

Note: (1) EHRs variable based on the AHA survey, (2) MU initiative based on the AHA survey, (3) the number of major EHRs components based on the HIMSS database, (4) EHRs level defined based on (Furukawa et al. 2010), (5) EHRs level defined based on (Appari et al. 2012), (6) Actual accreditation for Stage 1 MU as of 2012 based on the HIMSS database. (4) and (5) were also defined based on the number of EHRs components owned by hospital in the HIMSS database.

2. Additional Information on the 2SRI Analysis

i. Relevance of Instruments and the First State Results

Table A3 shows the first stage results of the 2SRI analysis using the probit model. Table A3, columns 1 and 2 present the results for implementation of EHRs and a pursuit of MU, respectively. As shown, the coefficients for the instrumental variables are highly significant in the first stage. In column 1, the state-level adoption rate of EHRs and MU initiative were significant whereas the fiscal year ending month was insignificant. The Wald test also shows that the three instrumental variables were collectively significant ($\chi^2(3) = 427.06$; $p < 0.001$). In column 2, all the three instrumental variables are highly significant. Similarly, the Wald test also shows that the three instrumental variables were collectively significant ($\chi^2(3) = 1{,}059.7$; $p < 0.001$). In conclusion, we believe that our instrumental variables well satisfy the condition for instrument relevance.

Table A3. First Stage Results: Probit Model

Variable	EHRs	MU
Intercept	9.460 *** (0.859)	-15.884 *** (0.945)
Number of Beds	4.555 *** (0.530)	5.991 *** (1.503)
State Breach Notification Law	0.014 (0.160)	1.005 *** (0.356)
Expense per FTE	0.002 ** (0.001)	0.000 (0.000)
Academic	-0.167 (0.332)	-0.405 (1.023)
PPO	-0.036 (0.415)	1.739 (1.123)
HMO	0.322 (0.495)	-0.155 (1.140)
ICU application	1.854 *** (0.149)	0.799 ** (0.311)
Encryption	-0.086 (0.195)	-0.299 (0.366)
Intrusion Detection System	0.300 * (0.164)	0.525 * (0.292)
Firewall	-0.055 (0.199)	0.240 (0.373)
Adoption of Cardiology IS	0.578 *** (0.154)	1.828 *** (0.363)
Adoption of ERP	0.702 *** (0.203)	-0.704 * (0.389)
Number of Computers	0.405 *** (0.094)	0.988 *** (0.175)
State Level Adoption Rate of EHRs	20.657 *** (1.004)	-28.004 *** (2.058)
State Level Pursuit of MU	-7.101 *** (1.127)	69.302 *** (2.525)
Fiscal Month	0.049 (0.030)	-0.205 *** (0.060)
Year Effects	YES	YES
Number of Hospitals	2,880	4,692
LL	-3950.56	-1646.71

Significant at 1 % ***, 5 % **, and 10% *.

ii. Exogeneity of Instruments

We believe that the three instrumental variables we chose are exogenous and satisfy the exclusion restriction as discussed in the main body of the paper. Unfortunately, it is unclear whether a test such as the Sargan-Hansen test of over-identifying restrictions can be applied to our context with the Cox proportional hazards model and the probit model in the first stage. Alternatively, we have tried the 2SLS model to run the Sargan-Hansen test only. That is, we used the annual number of breaches per hospital as the dependent variable with hospital and year fixed effects. In the first stage, the two treatment variables were regressed on the three instrumental variables. In this model, the Sargan's test statistic was 0.3199 (df = 1; p=0.989), and the null hypothesis that the instruments are valid could not be rejected. Although not completely accurate, this may show that our instrumental variables are exogenous and valid.

We also tried another test similar to the approach suggested by Murray (2006) in which a subset of instruments is used and compare the results across different subsets. Our instruments can be considered valid if we can obtain consistent results. Although the Murray (2006)'s suggestion was made for the case in which different sets of instruments could not be included in a single equation because of the data issue as in Levitt (2002), we believe that this could be a useful test. Our estimation results are summarized in Table A4. In the first case (Column (1)), the results remain largely unchanged compared to those in Column (2) of Table 4. In the second case (Column (2)), both EHRs and MU have become significant. The effect of EHRs adoption has become insignificant in the last case (Column (3)). However, the insignificance could be because of the increased standard errors by reducing the number of instruments (Murray 2006). Most importantly, the estimated hazards ratios are consistently greater than one, a similar case to Levitt (2002), and thus we conclude that our instruments can produce reasonably consistent results.

Table A4. Estimation Results

Variable	(1) State Level Adoption Rate of EHRs & Fiscal Year Ending Month as IVs	(2) State Level Adoption Rate of EHRs & Fiscal Year Ending Month as IVs	(3) State Level Pursuit of MU & Fiscal Year Ending Month as IVs
EHRs Adoption (AHA)	2.799 * (1.681)	3.381 ** (2.071)	2.367 (2.054)
Pursuit of Meaningful Use	1.262 (0.390)	2.124 *** (0.588)	1.196 (0.354)

Hazard ratios with p-values denoted by *** significant at 1 %, ** at 5 %, and * at 10%. The numbers in parentheses are robust (bootstrapped) standard errors. All other covariates used in Table 3 were included, but their estimated hazard ratios are not shown for brevity of presentation.

References

Appari, A., M.E. Johnson, D.L. Anthony. 2012. Meaningful Use of ElectronicHealth Record Systems and Process Quality of Care: Evidence from a Panel Data Analysis of U.S. Acute-CareHospitals. *Health Services Research* **48**(2 (Pt 1)) 354-375.

Furukawa, M.F., T.S. Raghu, B.B.M. Shao. 2010. Electronic Medical Records, Nurse Staffing, and Nurse-Sensitive Patient Outcomes: Evidence from California Hospitals, 1998 - 2007. *Health Services Research* **45**(4) 941-962.

Levitt, S.D. 2002. Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime: Reply. *American Economic Review* **92**(4) 1244-1250.

Murray, M.P. 2006. Avoiding Invalid Instruments and Coping with Weak Instruments. *Journal of Economic Perspectives* **20**(4) 111-132.