Data Breaches: Measurement Efforts and Issues

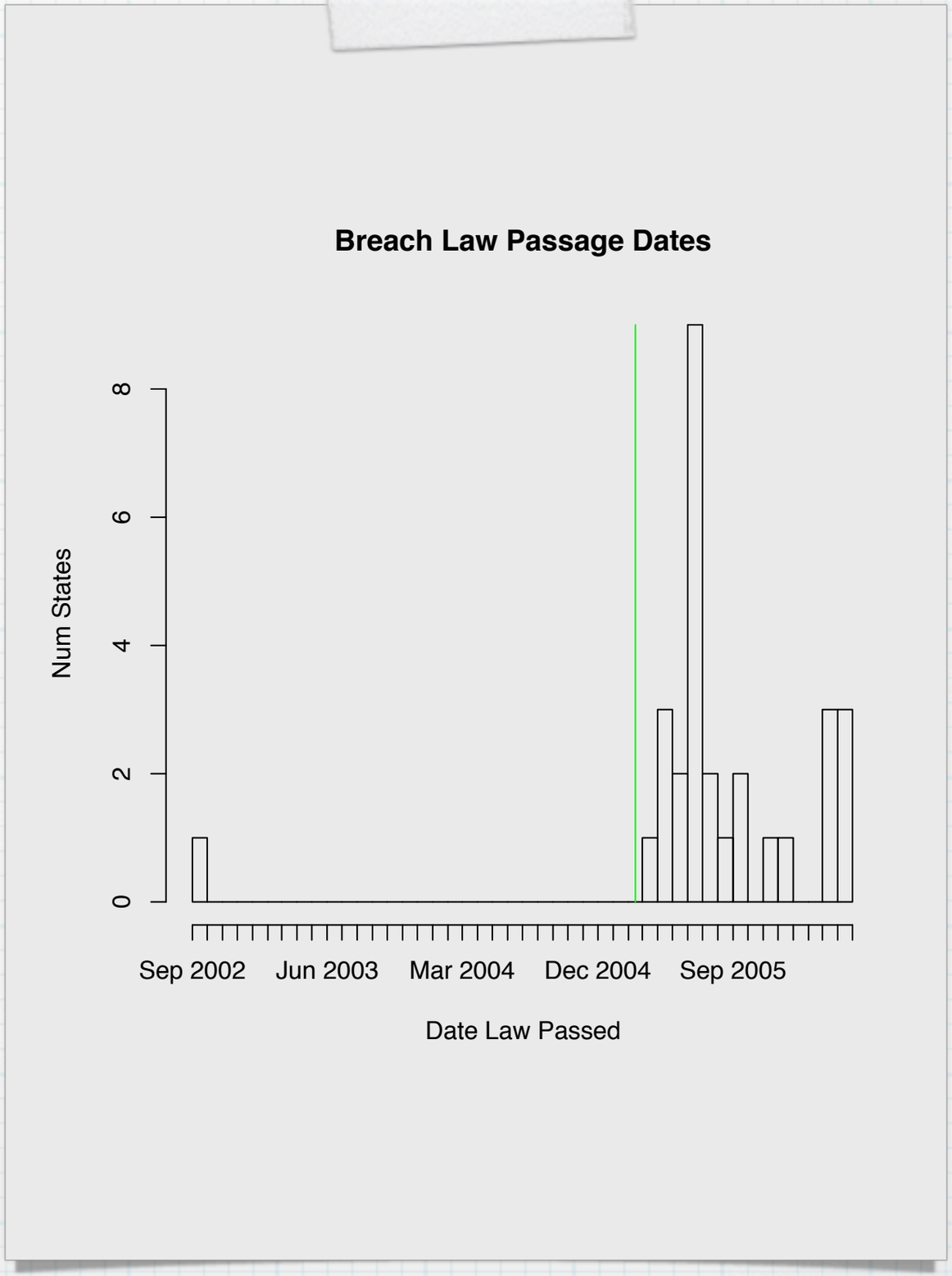
Chris Walsh
chris@cwalsh.org

ChoicePoint as Impetus

Breach focused attention, spurred legislative action

But, what can we actually measure, and how?

How big is the problem, and how costly the solution(s)?



Data Breaches and Identity Theft

* Relationship clearly(?) exists
* How much of either is there?
* Are on-line breaches a significant source of ID theft?

Today's ID Theft Measures

* Illustrative/Anecdotal: "Let's call him Joe"
* Retrospective Surveys
  * Estimate P("Identity Theft") for population, subgroups thereof
  * Summary statistics on losses
  * Whodunit?
* Industry fraud figures, FTC complaint volume

Today's Data Breach Measures

* Lists: Do raw data a "metric" make?
* Aggregated: x breaches, at y locations, affecting z people
* Econometric: Statistical estimates of impact on breached organization/firm.
* Survey: Samples of convenience, "illustrative" results.

Lists, Aggregates

* Dataloss
* Emergent Chaos
* Privacyrights.org

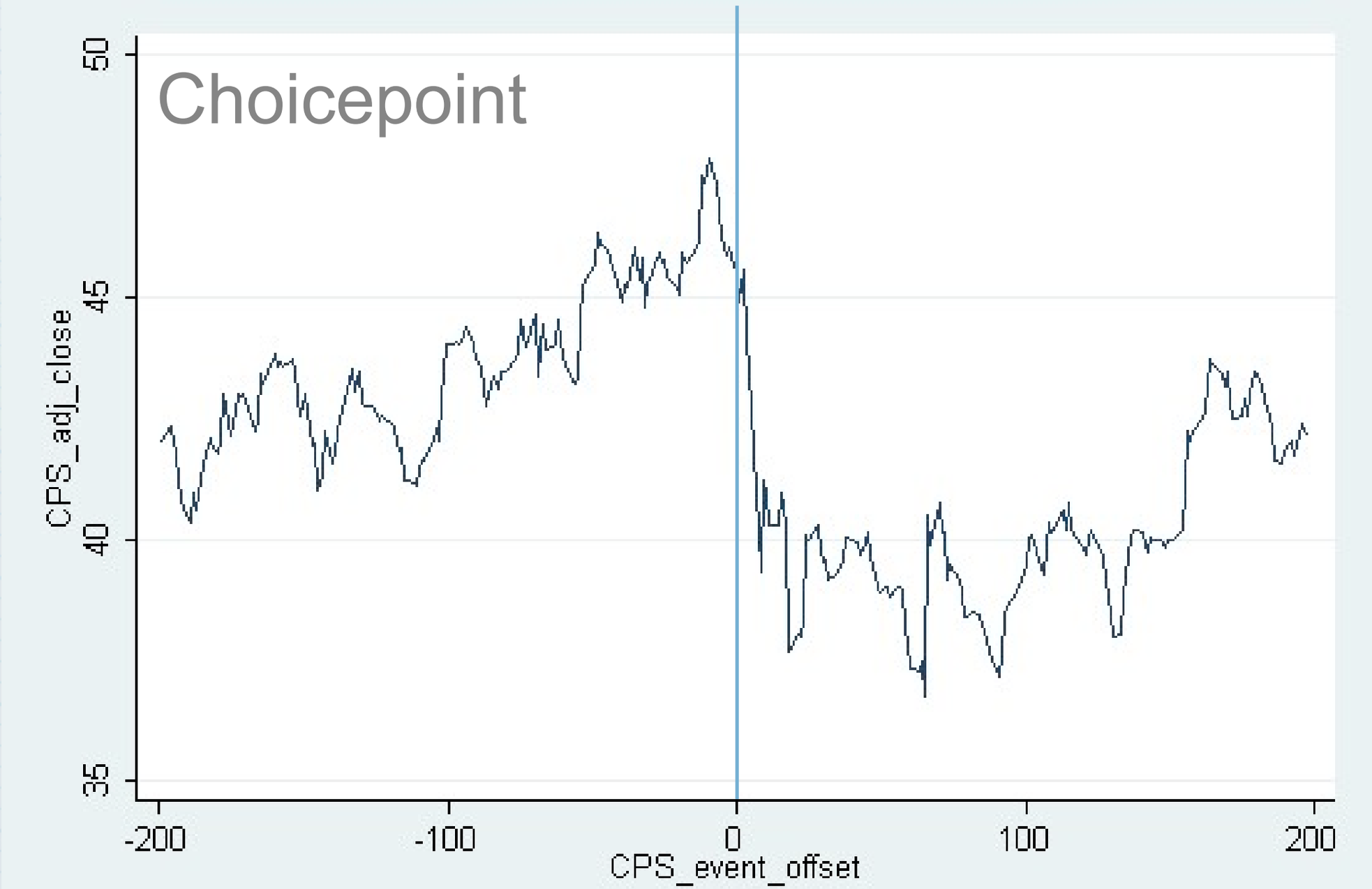
Aggregated metrics are trivially derived from any of the above.

There's some unit confusion: "records" vs. "people".

Event Studies

* Econometric estimates of "abnormal return" across a sample of firms subject to similar events
* English: On average, how much does a security breach decrease a company's stock price -- if at all?
* In pictures ...

Is this typical?



Data: Yahoo Finance

Is this?

Source: Google

Results

Study	Period Examined	Abnormal Return	N
Campbell, et. al., 2003	1997-2000	-5.4%	11
Cavusoglu, et. al., 2004	1996-2001	-2.1%	78
Acquisti, et. al., 2006	2000-2006	-0.6%	79

Hard to say which aspects of breaches contribute to losses -- confidentiality seems to matter (Campbell), but jury is out on other independent variables.

Chris the grad student sez:

Additional research is needed in this crucial area...

Actual 1991 Photo

More specifically...

Research Agenda

* Measure impact on govt, educational organizations
* Find independent variables affecting breach impact. Is time one of them? Is firm "frankness"?
* Is this an iceberg? How can we tell?
* Do we have enough info on breaches we know about?

* Validate model assumptions about investor attitude, using survey research

* Examine sampling issues in existing event studies -- has SB1386 improved data availability, added noise, or what?

* Look inside organizations -- do decision-makers act to minimize breach impact?

* Does behavior vary across organization types or governance structures?

* Can we integrate findings from fraud-detection 'sensor networks', honeynets, and monitoring of underground economy in PII to validate breach volume information?

* Replicate Campbell, et. al. with more recent data.

* Some non-US data would be nice!

Read me:

Acquisti, Alessandro, et. al., Is There a Cost to Privacy Breaches? An Event Study, [DRAFT -- URL omitted]

Anderson, Keith B., Identity Theft: Does the Risk Vary with Demographics?, http://www.ftc.gov/be/workpapers/wp279.pdf

Belva, Kenneth F., How It's Difficult to Ruin a Good Name: An Analysis of Reputational Risk, http://www.ftusecurity.com/pub/FiTechSummit_final_paper.pdf

Campbell, et. al., The economic cost of publicly announced information security breaches: empirical evidence from the stock market, http://iospress.metapress.com/link.asp?id=5nkxhffc775tuel9

Cavusoglu, et. al., The Effect of Security Breach Announcements on the Market Value of Breached Firms and Internet Security Developers, http://mesharpe.metapress.com/link.asp?id=mx6xwxy2rfx166ge

Ponemon, Larry, Lost Customer Information: What Does a Data Breach Cost Companies?

Thanks

Please see http://www.cwalsh.org/metricon/ for full citations, links to materials mentioned, and (real soon now) a more formal paper-length discussion of the issues raised here.