Moore's Law vs. Murphy's Law: Who's Winning?

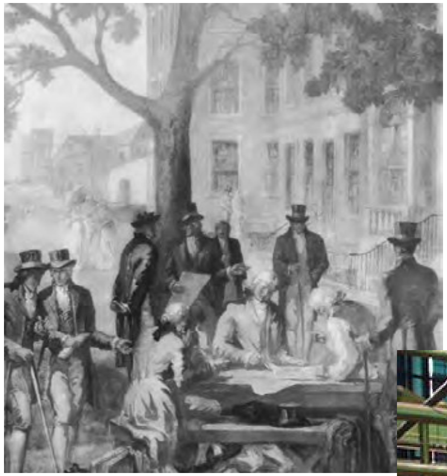
Andrew W. Lo, MIT

14th BIS Annual Conference

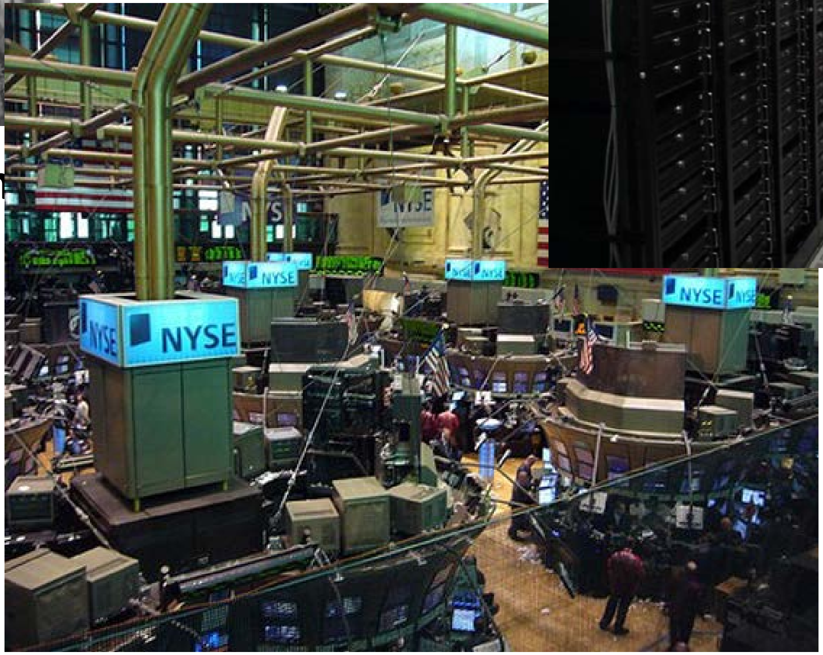
Towards a "New Normal" in Financial Markets?

MIT
Laboratory for
Financial Engineering

Technology and the Financial System



Buttonwood Agreem
May 17, 1792



Technology and the Financial System

Moore's Law vs. Murphy's Law

- Capital requirements are harder to implement
- New risks to financial stability have been created
- Fairness and privacy issues have emerged
- Speed of financial innovation has increased; speed of regulatory innovation has not kept pace, e.g., HFT
- Complexity has increased
- We need better "regulatory technology": (1) risk transparency; (2) adaptive regulations; (3) framework for financial regulation

The Challenge of Technology

"Nobody Knows Anything"

The May 6, 2010 "Flash Crash"

Friday, May 7, 2010 New York 55° | 38°

THE WALL STREET JOURNAL. | TODAY'S MARKETS

U.S. Edition Home ▾ | Today's Paper ▪ Video ▪ Blogs ▪ Journal Community

MARKET TUMULT: FROM A SLIGHT GAIN, TO DOWN 998.50 POINTS, TO DOWN 347.80

Dow Takes a Harrowing 1,010.14-Point Trip

Biggest Point Fall, Before a Snapback; Glitch Makes Thing Worse

Article | Video | Slideshow | Stock Quotes | Comments (248)

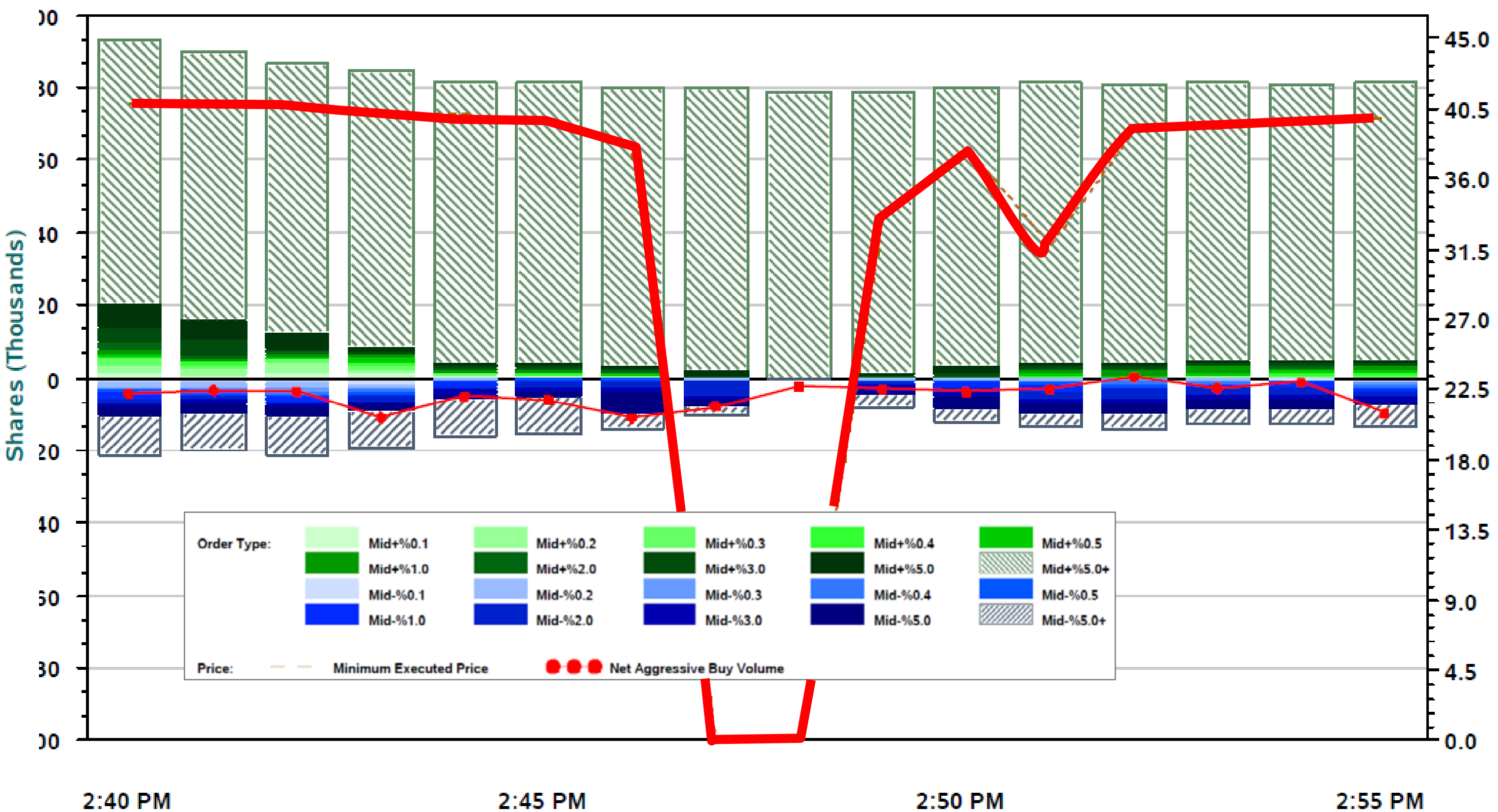
Email Print Save This Like 3K More + Text −

By TOM LAURICELLA And PETER A. MCKAY

“Nobody Knows Anything”

Accenture plc, Market Depth, Aggressive Buys, and Price

2:40pm - 2:55pm



"Nobody Knows Anything"

September 30, 2010

FINDINGS REGARD

THE MARKET EVE

OF MAY 6, 2010

SEPTEMBER 30, 2010

FINANCIAL TIM

April 3, 2013 5:43 pm

Flash crash exp

By Philip Stafford

THE W

This copy is fo
http://www.djre

http://www.wsj

BUSINES

The

Waddell

THE WALL STREET JOURNAL.

This copy is for your personal, non-commercial use only. To order presentation-ready copies for distribution to your colleagues, clients or customers v
http://www.djreprints.com.

http://www.wsj.com/articles/u-k-man-arrested-on-charges-tied-to-may-2010-flash-crash-1429636758

MARKETS

April 21, 2015

'Flash Crash' Charges Filed

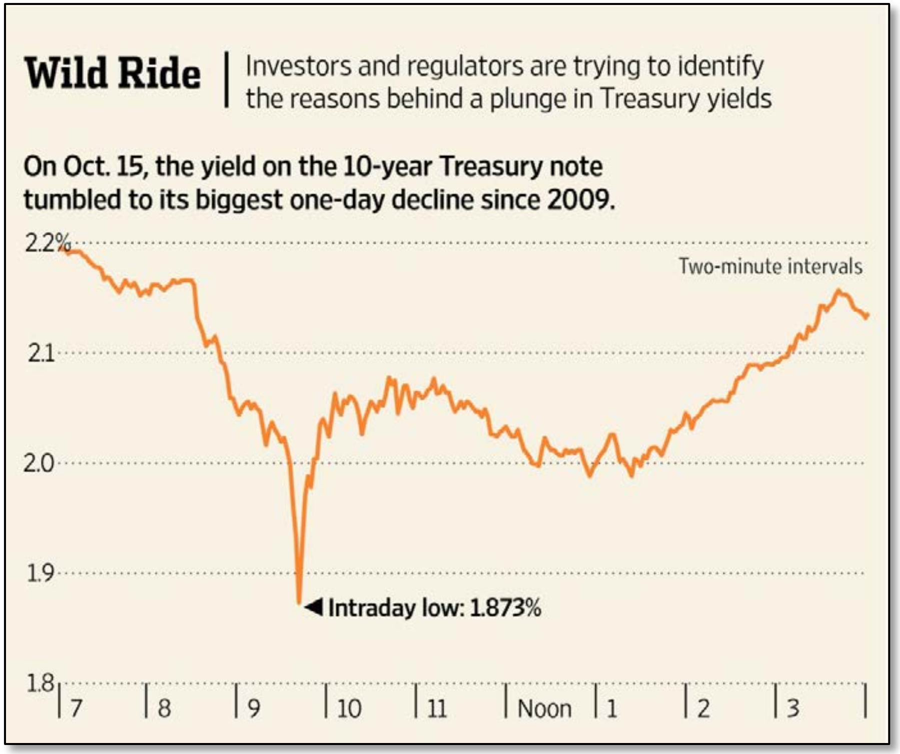
Authorities say a trader in the U.K. helped trigger wild swings that shook markets in May 2010

The final numbers of the day's trading is shown on a board on the floor of the New York Stock Exchange on May 6, 2010.
PHOTO: LUCAS JACKSON/REUTERS

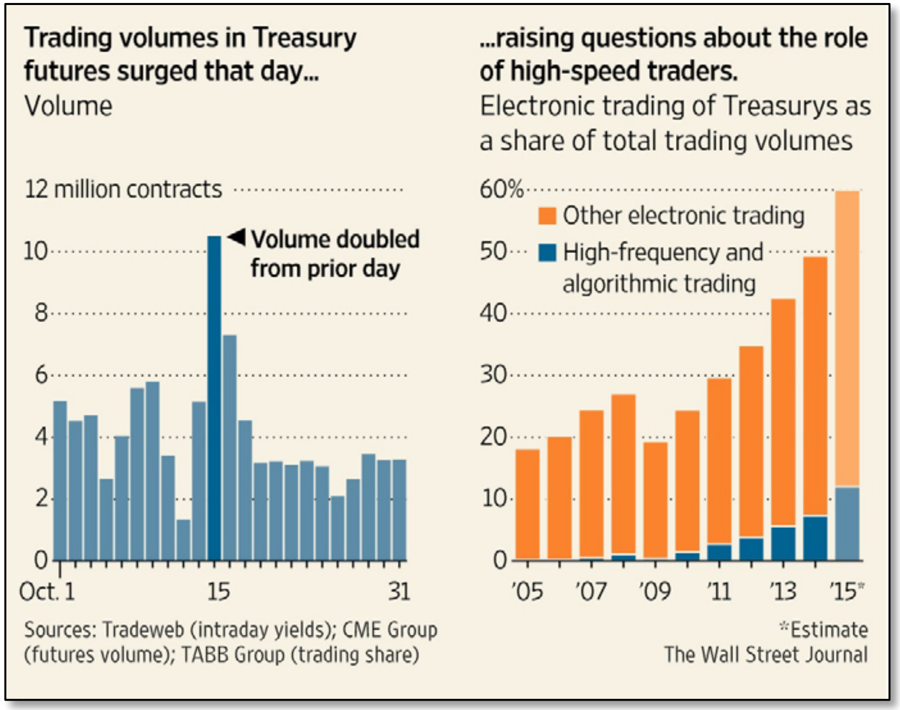
By **ARUNA VISWANATHA** in Washington,
BRADLEY HOPE in New York and **JENNY STRASBURG** in London
Updated April 21, 2015 7:59 p.m. ET

http://sec.gov/news/studies/2010/marketevents-report.pdf

"Nobody Knows Anything"



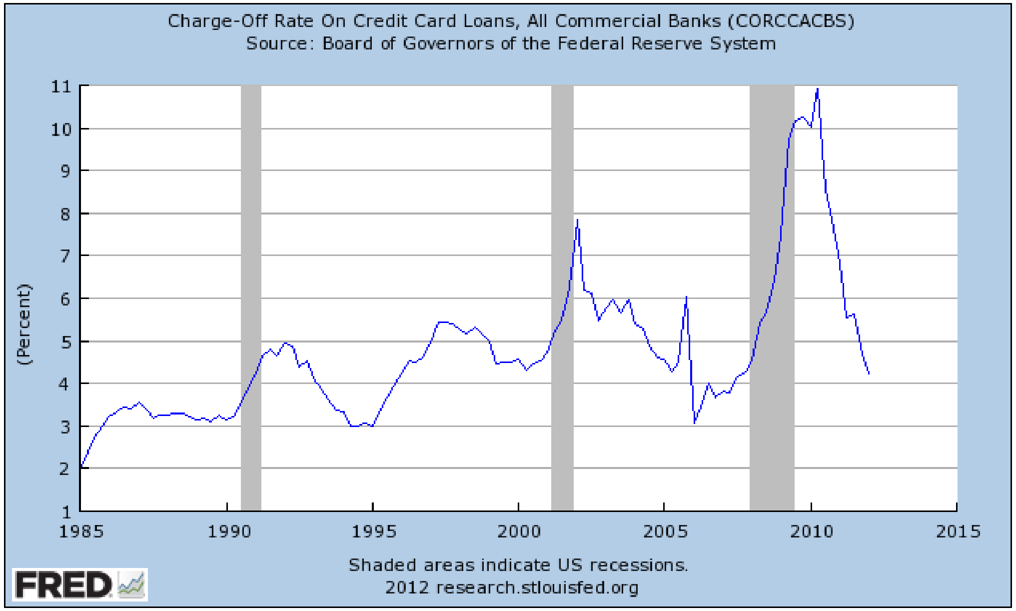
October 15, 2014



The Promise of Technology

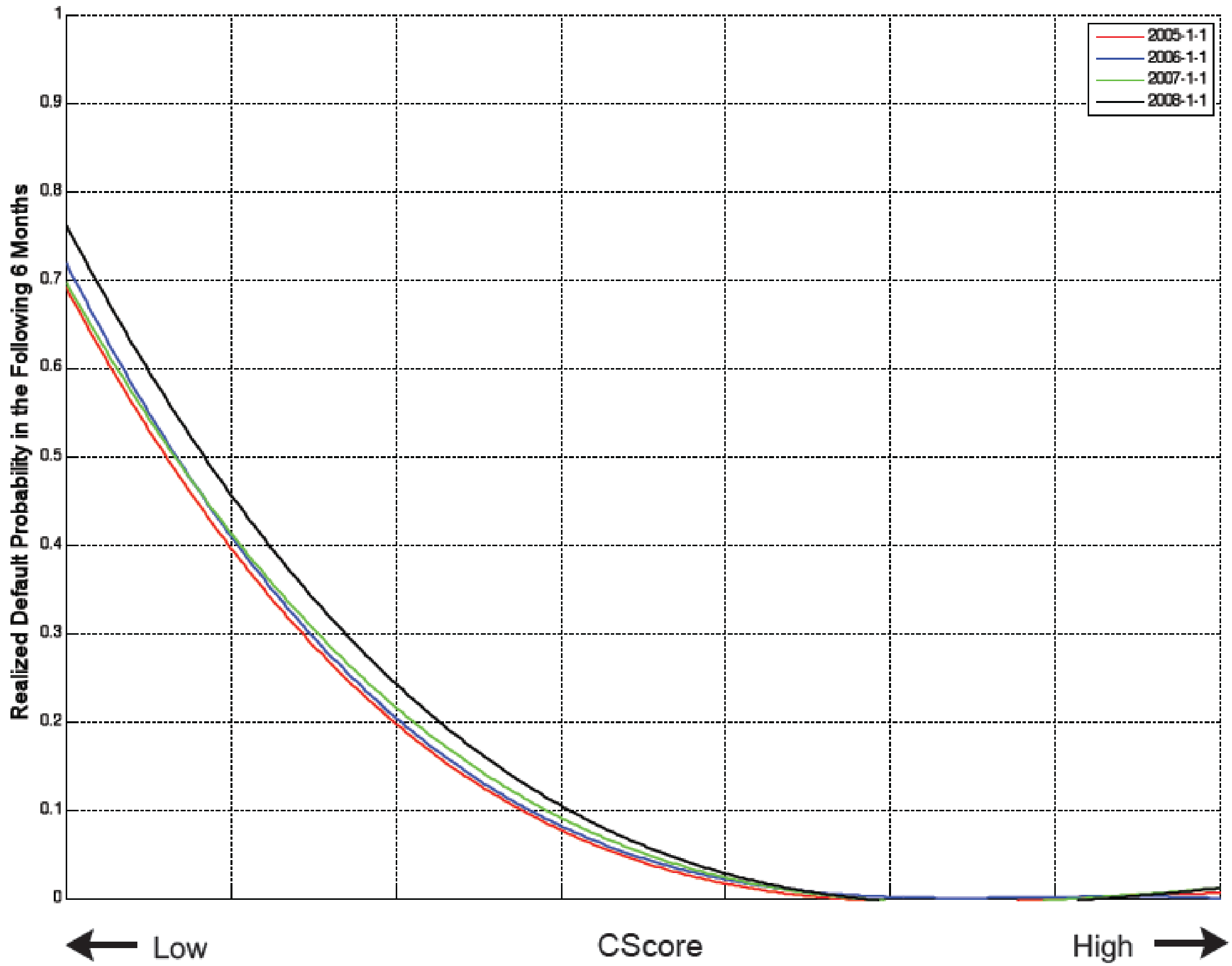
Big Data for Consumer Credit

- $3.4T of consumer credit outstanding as of Mar 2015
- $889B of revolving consumer credit outstanding as of Mar 2015
- 38% of households carry positive credit card balance in 2013 ($5,700)



Big Data for Consumer Credit

Standard Credit Scores Are Too Insensitive



Big Data for Consumer Credit

Anonymized Data from Large U.S. Commercial Bank

Transaction Data

Transaction Count
Total Inflow
Total Outflow

By Channel:

ACH (Count, Inflow and Outflow)
ATM (Count, Inflow and Outflow)
BPY (Count, Inflow and Outflow)
CC (Count, Inflow and Outflow)
DC (Count, Inflow and Outflow)
INT (Count, Inflow and Outflow)
WIR (Count, Inflow and Outflow)

By Category		
Mortgage payment	Hotel expenses	Bar Expenses
Credit car payment	Travel expenses	Fast Food Expenses
Auto loan payment	Recreation (golf	Total Rest/Bars/Fast-Food
Student loan payment	Department Stores Expenses	Healthcare related expenses
All other types of loan payment	Retail Stores Expenses	Health insurance
Other line of credit payments	Clothing expenses	Gas stations expenses
Brokerage net flow	Discount Store Expenses	Vehicle expenses
Dividends net flow	Big Box Store Expenses	Car and other insurance
Utilities Payments	Education Expenses	Drug stores expenses
TV	Total Food Expenses	Government
Phone	Grocery Expenses	Treasury (eg. tax refunds)
Internet	Restaurant Expenses	Pension Inflow
Collection Agencies	Unemployment Inflow	Collection Agencies

Balance Data

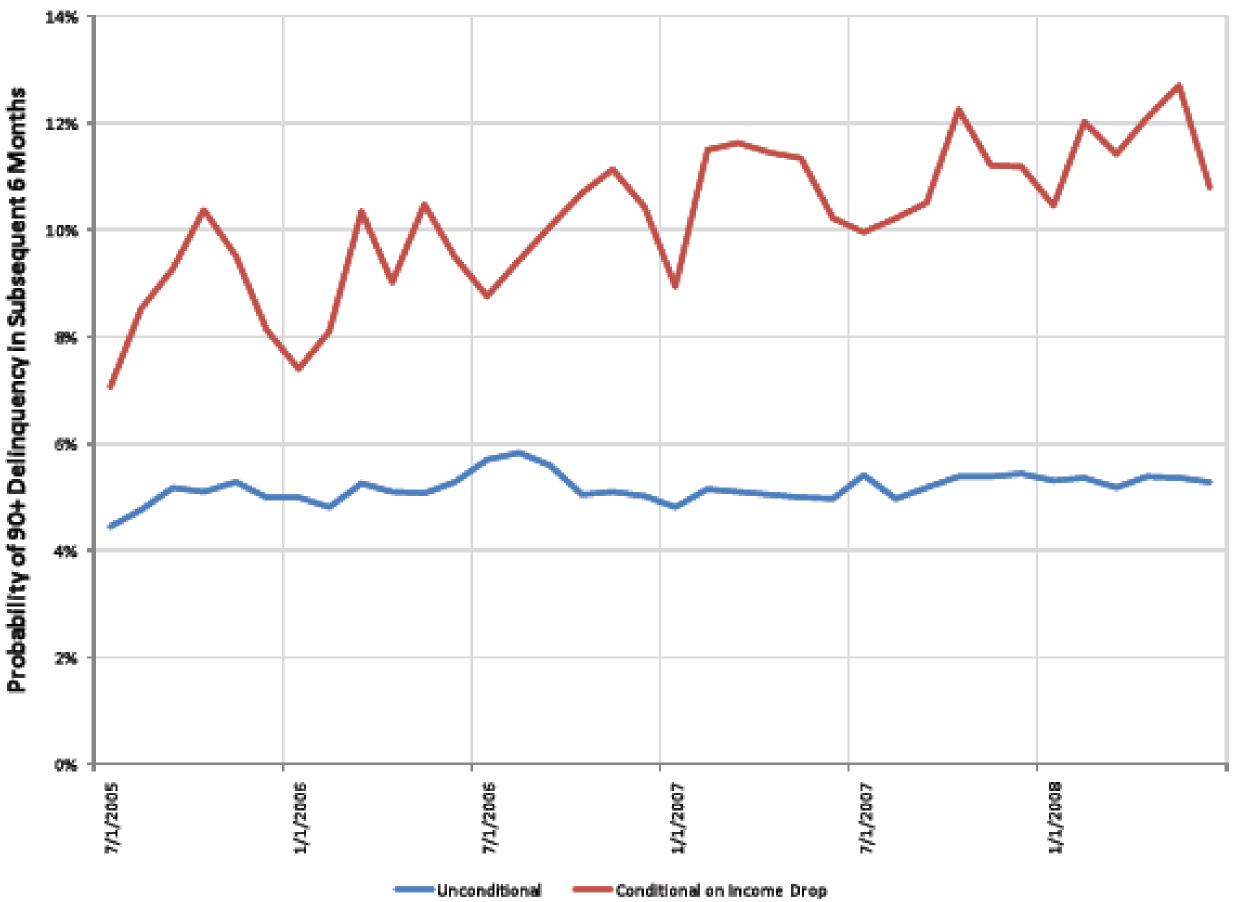
Checking Account Balance
Brokerage Account Balance
Saving Account Balance
CD Account Balance
IRA Account Balance

Credit Bureau Data

File Age	Type (CC, MTG, AUT, etc)
Credit Score	Age of Account
Open/Closed Flag & Date of Closure	Balance
Bankruptcy (Date & Code)	Limit if applicable
MSA & Zip	Payment Status
	48-Month Payment Status History

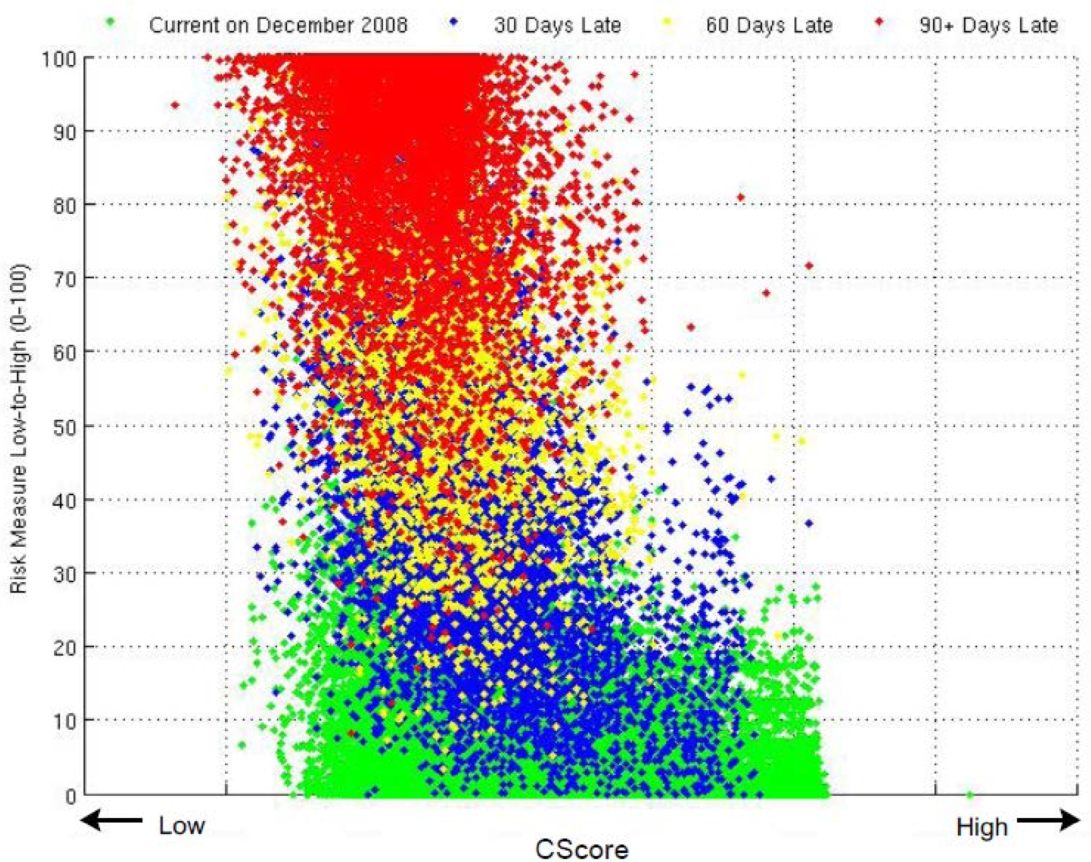
1% Sample = 10 Tb!

Big Data for Consumer Credit



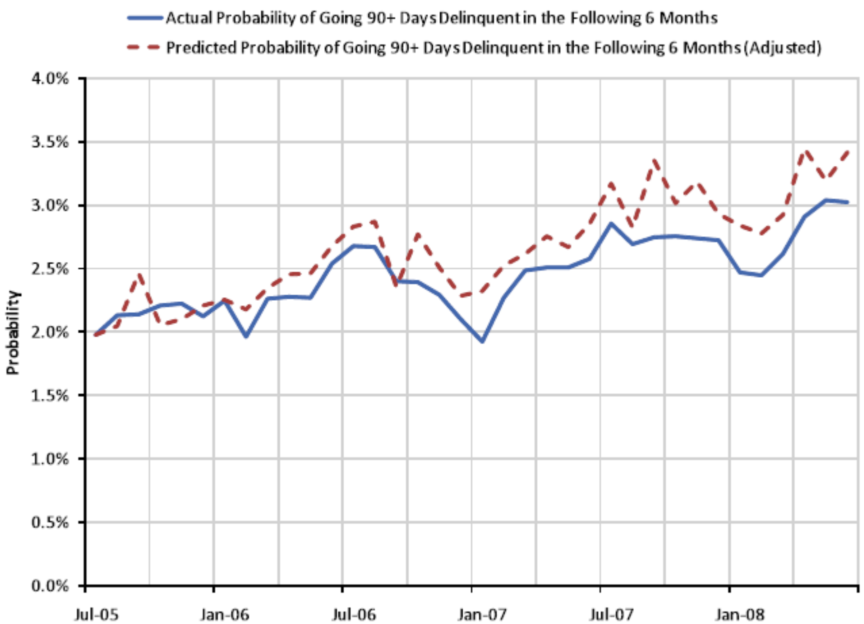
Big Data for Consumer Credit

- Khandani, Kim, and Lo (2010)
- 600,000 credit cards per month; 40-hour runtime

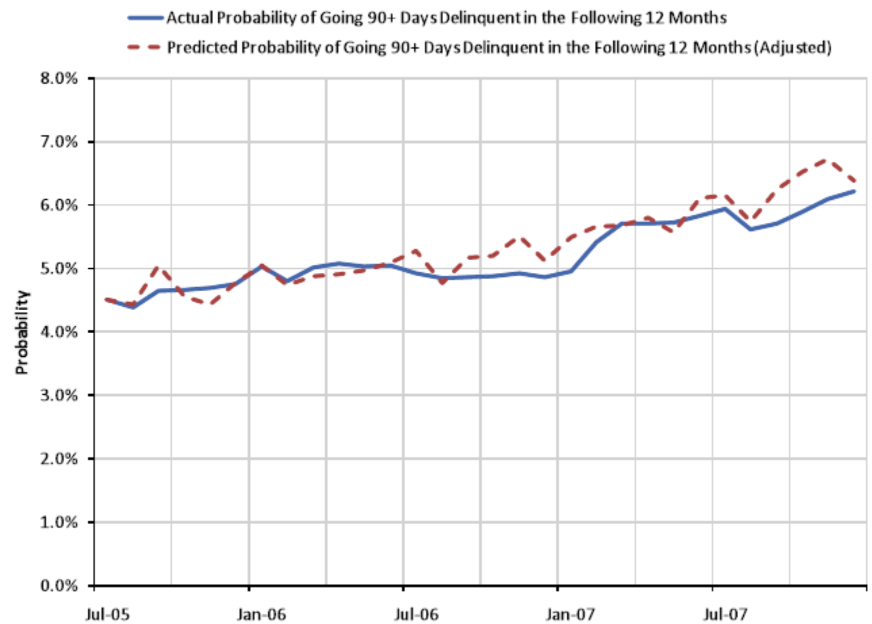


Big Data for Consumer Credit

Credit Forecasts Over Time



(a) Time series of actual and predicted 90-days-or-more delinquency rates (6-month)



(b)Time series of actual and predicted 90-days-or-more delinquency rates (12-month)

Current Research

Risk and Risk Management in the Credit Card Industry*

Florentin Butaru[1], Qingqing Chen[1], Brian Clark[1,4], Sanmay Das[2], Andrew W. Lo[3], Akhtar Siddique[1]

This Revision: 14 June 2015

Abstract

Using account level credit-card data from six major commercial banks from January 2009 to December 2013, we apply machine-learning techniques to combined consumer-tradeline, credit-bureau, and macroeconomic variables to predict delinquency. In addition to providing accurate measures of loss probabilities and credit risk, our models can also be used to analyze and compare risk management practices and the drivers of delinquency across the banks. We find substantial heterogeneity in risk factors, sensitivities, and predictability of delinquency across banks, implying that no single model applies to all six institutions. We measure the efficacy of a bank's risk-management process by the percentage of delinquent accounts that a bank manages effectively, and find that efficacy also varies widely across institutions. These results suggest the need for a more customized approached to the supervision and regulation of financial institutions, in which capital ratios, loss reserves, and other parameters are specified individually for each institution according to its credit-risk model exposures and forecasts.

* We thank Michael Carhill, Jayna Cummings, Misha Dobrolioubov , Dennis Glennon, Amir Khandani, Adlar Kim, Mark Levonian, David Nebhut, Til Schuerman, Michael Sullivan and seminar participants at the Consortium for Systemic Risk Analysis, the Consumer Finance Protection Bureau, the MIT Computer Science and Artificial Intelligence Laboratory (CSAIL), the Office of the Comptroller of the Currency, and the Philadelphia Fed's Risk Quantification Forum for useful comments and discussion. The views and opinions expressed in this article are those of the authors only, and do not necessarily represent the views and opinions of any institution or agency, any of their affiliates or employees, or any of the individuals acknowledged above. Research support from the MIT CSAIL Big Data program, the MIT Laboratory for Financial Engineering, and the Office of the Comptroller of the Currency is gratefully acknowledged.

[1] U.S. Department of the Treasury, Office of the Comptroller of the Currency, Enterprise Risk Analysis Division.

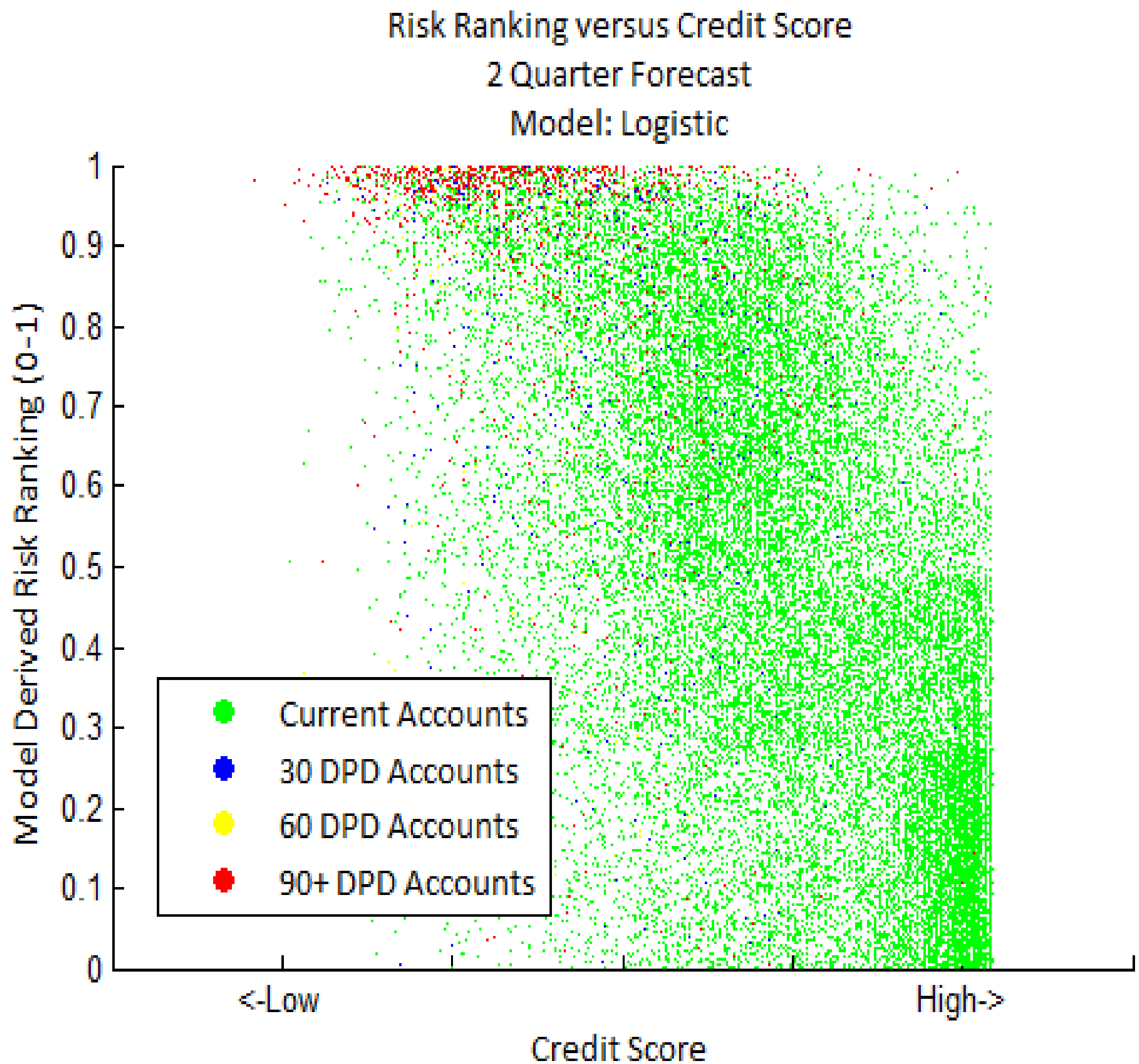
[2] Washington University in St. Louis, Department of Computer Science & Engineering.

[3] Massachusetts Institute of Technology, Sloan School of Management, Computer Science and Artificial Intelligence Laboratory, Electrical Engineering and Computer Science; AlphaSimplex Group, LLC.

[4] Rensselaer Polytechnic Institute (RPI), Lally School of Management.

- 6 large banks from Jan 2009 to Dec 2013
- Macro and institution-specific factors (137)
- 25 Tb of data
- Used to gauge quality of risk management across institutions
- Models vary greatly

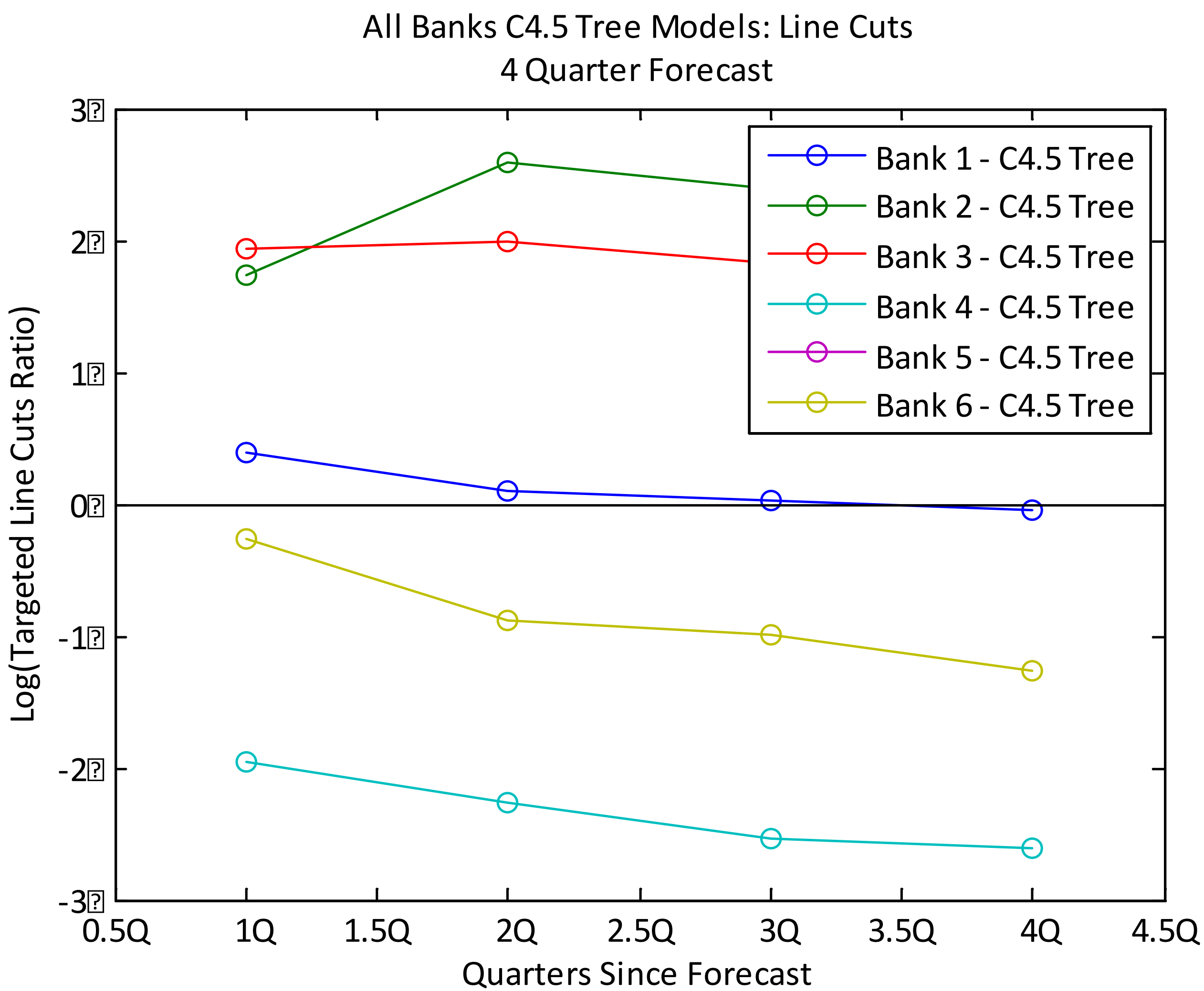
Current Research



Current Research

Category	Attribute	Bank 1	Bank 2	Bank 3	Bank 4	Bank 5	Bank 6
Utilization	MonthUtilization1MoChange	6.2	6.5	1.5	4.5	4.9	5.3
Utilization	CycleUtilization1MoChange	1.8		5.4		1.4	1.6
Utilization	MonthUtilization					1.0	
Utilization	MonthUtilization3MoChange		2.5				0.0
Utilization	CycleUtilization						1.3
Utilization	Dum1IfTotBal_TotLmtAllOpenBankCardAcctsEQ0			0.1			
delinquency Status	Dum1IfGT0Acct60DPD	5.6	2.3	4.8	5.2	0.8	3.4
Delinquency Status	DaysPastDue	5.4	5.7		5.5	3.8	2.2
Delinquency Status	Dum1IfGT0Acct90DPD	4.2	4.5	4.4	3.2		1.4
Delinquency Status	NumOfAcc60DPD	4.1	2.8	2.5	4.3	0.2	-0.6
Delinquency Status	Dum1IfGT0Acct30DPD	3.8	4.3		2.4		1.5
Delinquency Status	NumAccts30DPD	3.0	2.6		2.4		
Delinquency Status	NumOfAcc90DPD	2.4		2.9	1.8		
Delinquency Status	TotNumAcc60DaysPastDue12MoVerif			-0.1		-0.7	
Delinquency Status	TotNumOpenBankCard60DPD12MoVer					-0.2	
Delinquency Status	Dum1IfGT0BankCardAcct60DPD12MoVer		2.9	-0.5		0.2	
Borrower Payment behavior	ActualPmtAmt_TotPmtDue	5.0	4.0	3.8	4.9	2.0	0.6
Borrower Payment behavior	PaymentEqDueLast3MoFlag	3.9	1.7	3.3	2.3	0.7	-0.8
CardCharacteristics	CurrentCreditLimit	2.4			3.9	0.1	0.8
CardCharacteristics	MonthEndBalance	2.2	2.6	0.1		-0.6	1.7
CardCharacteristics	ProductType	1.8					
CardCharacteristics	CycleEndBalance			0.3	6.5	0.9	2.2
CardCharacteristics	TotNumberOfAccounts			-0.5			
CardCharacteristics	TotNumberGoodAccounts		3.1		2.9		
CardCharacteristics	TotNonMortgBalAllAcc12MoVerif						-0.6
CardCharacteristics	MaxTotAmt60DPDAllAcctsOrTotBalOpenBankCards60DPD		5.7			1.5	
CardCharacteristics	TotCredLmtBankCardAccts			-0.2			
CardCharacteristics	Dum1IfTotCredlmtAllRvlvgAcctsGT012MoVer		2.3	-0.2			
CardCharacteristics	CreditCardType				3.8		
BorrowerCharacteristics	3MoChangeRefreshedFICO	3.5		-0.4			
BorrowerCharacteristics	BehevScore	2.3	3.1	0.7	4.6	1.9	2.9
BorrowerCharacteristics	RefreshedFICO	1.9	1.6		1.7	0.4	1.9
BorrowerCharacteristics	6MoChangeBehevScore						-0.9
AccountStatus	chg1Mo_LineFrozenFlag_0	2.4			1.5	1.8	
AccountStatus	LineFrozenFlag	2.4	1.5				
AccountStatus	LineDecreaseFlag				3.5		
AccountStatus	TotalPaymentDue				2.1	-0.4	2.0
AccountStatus	OverLimitLast3MoFlag					0.4	
Macro	MACROd3hrs_wkly_private	1.5	2.9	0.6	2.7		
Macro	MACROd3num_total_private_nsa		2.5				
Macro	MACROl12hrs_wkly_leisure						0.0
Macro	MACROd12index_sa			-0.3			

Current Research



The Threat of Technology

Privacy vs. Transparency

Predicting Social Security numbers from public data

Alessandro Acquisti[1] and Ralph Gross

Carnegie Mellon University, Pittsburgh, PA 15213

PNAS 106 (July 2009)

Communicated by Stephen E. Fienberg, Carnegie Mellon University, Pittsburgh, PA, May 5, 2009 (received for review January 18, 2009)

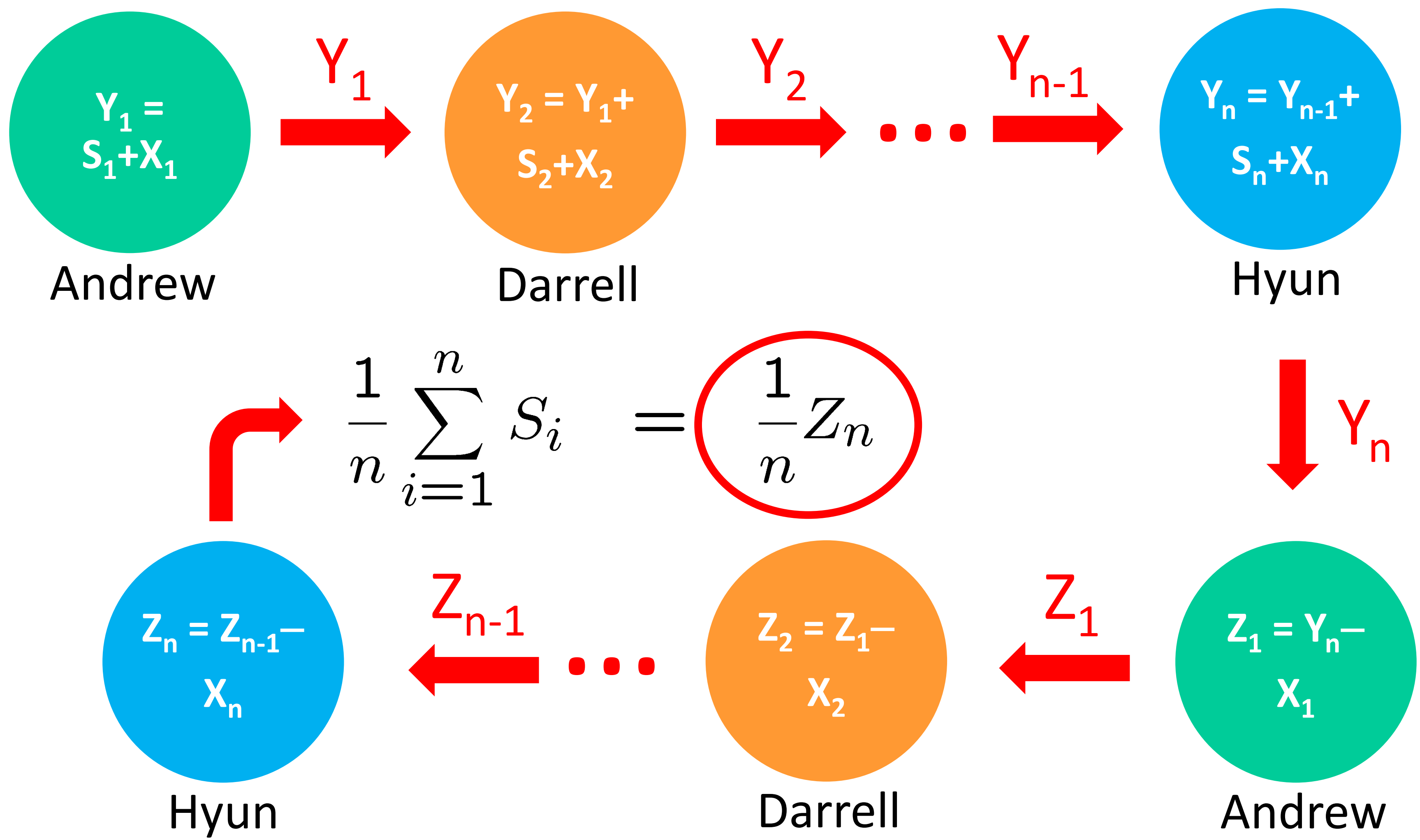
Information about an individual's place and date of birth can be exploited to predict his or her Social Security number (SSN). Using only publicly available information, we observed a correlation between individuals' SSNs and their birth data and found that for younger cohorts the correlation allows statistical inference of private SSNs. The inferences are made possible by the public availability of the Social Security Administration's Death Master File and the widespread accessibility of personal information from multiple sources, such as data brokers or profiles on social networking sites. Our results highlight the unexpected privacy consequences of the complex interactions among multiple data sources in modern information economies and quantify privacy risks associated with information revelation in public forums.

identity theft | online social networks | privacy | statistical reidentification

number (SN). The SSA openly provides information about the process through which ANs, GNs, and SNs are issued (1). ANs are currently assigned based on the zipcode of the mailing address provided in the SSN application form [RM00201.030] (1). Low-population states and certain U.S. possessions are allocated 1 AN each, whereas other states are allocated sets of ANs (for instance, an individual applying from a zipcode within New York state may be assigned any of 85 possible first 3 SSN digits). Within each SSA area, GNs are assigned in a precise but nonconsecutive order between 01 and 99 [RM00201.030] (1). Both the sets of ANs assigned to different states and the sequence of GNs are publicly available (see www.socialsecurity.gov/employer/stateweb.htm and www.ssa.gov/history/ssn/geocard.html). Finally, within each GN, SNs are assigned "consecutively from 0001 through 9999" (13) (see also [RM00201.030], ref. 1.)

Is There A Compromise Between Data Privacy and Transparency?

Secure Multi-Party Computation



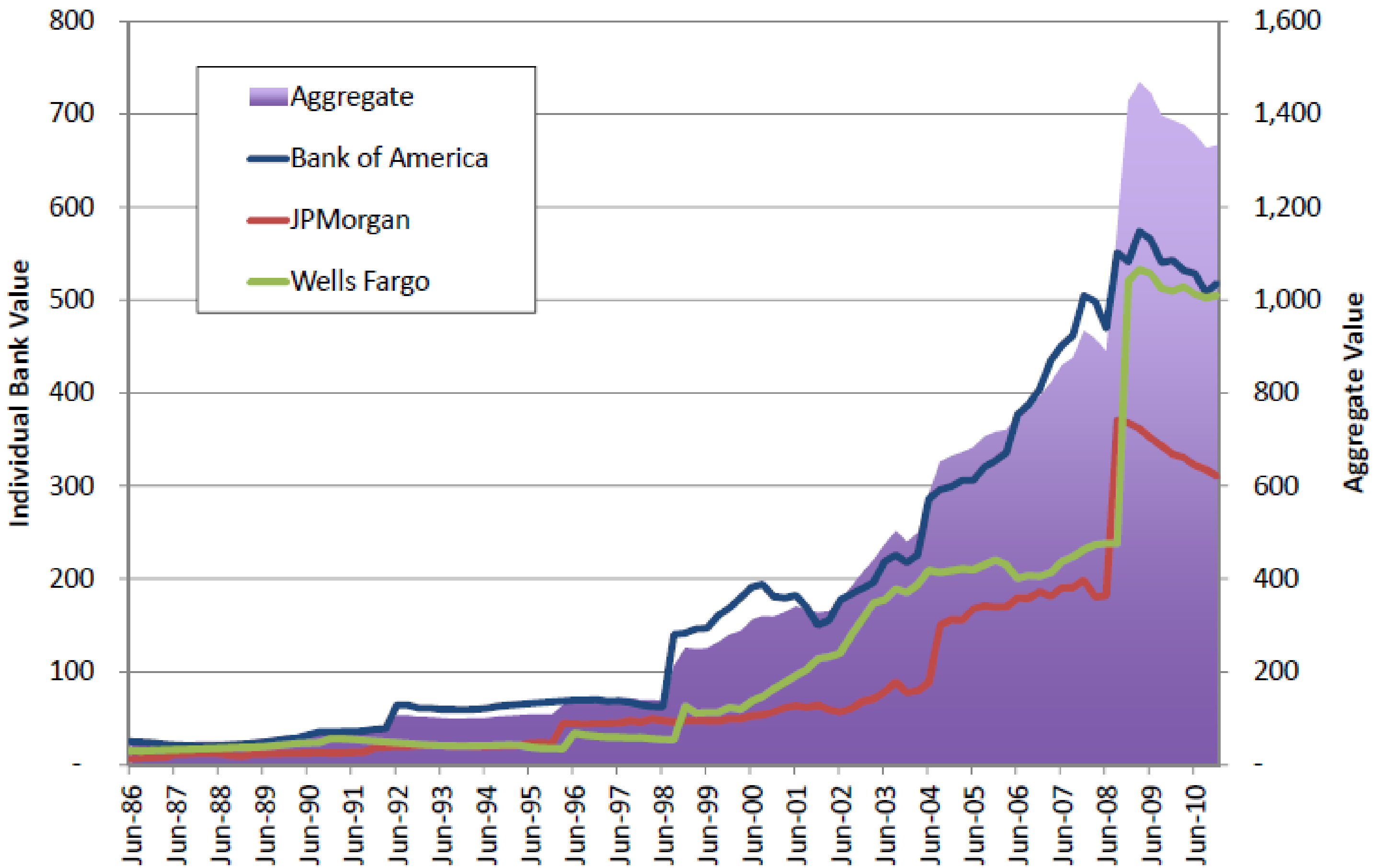
Privacy and Transparency

Transparency and Privacy Can <u>Both</u> Be Achieved

- Abbe, Khandani, and Lo (2012, 2015)
- Individual data is kept private, e.g., RSA
- Encryption algorithms are "collusion-robust"
- Aggregate risk statistics can be computed using encrypted data
  - Means, variances, correlations, percentiles, Herfindahl indexes, VaR, CoVaR, MES, etc.
- Privacy is preserved, no need for raw data!

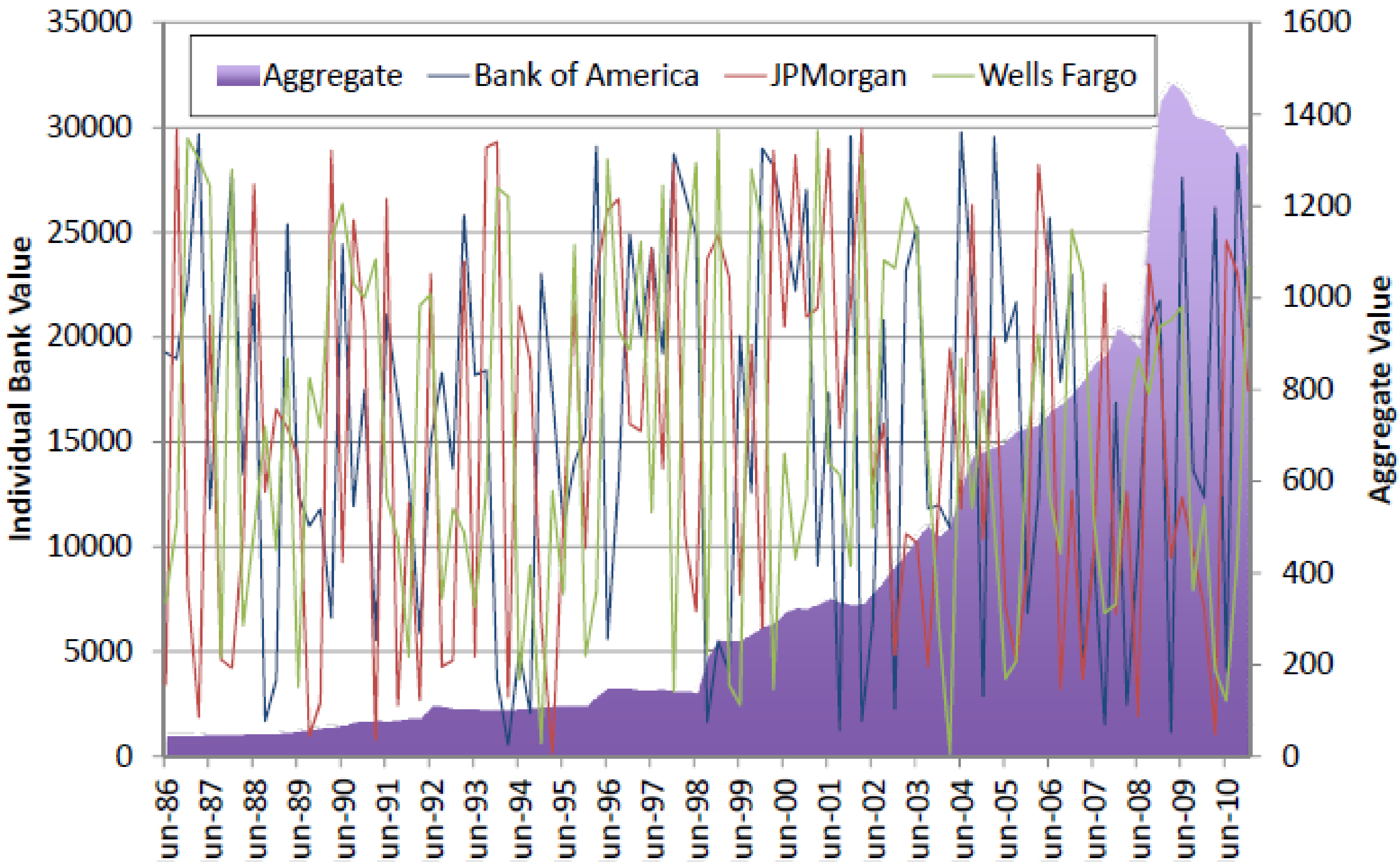
Privacy and Transparency

Real Estate Loans Outstanding



Privacy and Transparency

Real Estate Loans Outstanding



Privacy and Transparency

OFFICE OF FINANCIAL RESEARCH
U.S. DEPARTMENT OF THE TREASURY

Office of Financial Research
Working Paper #0011
September 4, 2013

Cryptography and the Economics of Supervisory Information: Balancing Transparency and Confidentiality

Mark Flood,[1] Jonathan Katz,[2] Stephen Ong,[3] and Adam Smith[4]

[1] Office of Financial Research, mark.flood@treasury.gov
[2] University of Maryland, jkatz@cs.umd.edu
[3] Federal Reserve Bank of Cleveland, stephen.j.ong@clev.frb.org
[4] Pennsylvania State University and Office of Financial Research, asmith@cse.psu.edu

The Office of Financial Research (OFR) Working Paper Series allows staff and their co-authors to disseminate preliminary research findings in a format intended to generate discussion and critical comments. Papers in the OFR Working Paper Series are works in progress and subject to revision.

Views and opinions expressed are those of the authors and do not necessarily represent official OFR or Treasury positions or policy. Comments are welcome as are suggestions for improvements, and should be directed to the authors. OFR Working Papers may be quoted without additional permission.

www.treasury.gov/ofr

Conclusion

- Technology has transformed everything!
- Financial markets are vastly better off
- But new challenges have emerged
- We can do better
- We have to do better

- Regulation has to account for technology and how it interacts with human behavior
- Regulators, industry, and academia must collaborate to create the Financial System 2.0

Thank You!