Big Data, Cloud Computing, Macroimplications and EU privacy regulation

Prof. Federico Etro
Ca' Foscari University of Venice

Paris, June 16, 2014

Plan of the talk:

- the role of big data in the economy
- the macroeconomic implications of online data and cloud computing
- the impact of policy on investment in online computing
- Implications of privacy regulation

Big Data:

The recent increase in the creation and storage of larger and larger quantities of data, referred to as “big data” is:

- expected to enhance productivity of the global economy (accumulation of data as a new factor of production)
- create huge opportunities for data analysis and forecasting based on econometric techniques
- change the way firms, academic institutions and consumers do business, do research and interact with each other

We are talking of both structured data (databases) and unstructured data, that lack a predefined model (search queries on Google, posts on Twitter, 50 billion photos on Facebook)

Big Data:

Estimate (2013): the world generates 2 exabytes (=quintillion bytes..) of unstructured data every year and organizations generate and store even more exabytes of structured data every year

The use of cloud services to accumulate data implies that the volume of new data increases exponentially (velocity of collection increases)

- Cisco forecasts that the annual volume of global IP traffic will exceed a zettabyte (=1000 exabytes) next year
- 90% of the volume of global data was generated in the last two years
- Data creation is expected to increase 44 times between 2010 and 2020

Big Data:

What to do with all these data?

- consumers would need to know how their data are processed and protected,
- organizations need to develop tools to process, store and transmit these data (often export or import them: international trade dimension)
- and institutions must regulate (coordinate) data protection

Big Data:

Questions:

- What is the economic importance of big data for the global economy?
- What is the macroeconomic impact of cloud computing?
- What is the proper way to protect data and privacy when data are traded and potentially used globally?

Big Data:

The economy is going to benefit from the accumulation of big data and cloud computing thanks to creation of new business, but also to improved marketing and pricing possibilities and to the role biga data and the cloud as catalysts of innovations. This is possible thanks to:

-A technological factor: novel computing methodologies help firms to understand and use data by means of machine learning and analytical tools

-An economic factor: the hardware needed to store and process data has become incredibly inexpensive

Big Data:

How big data are going to *directly* benefit the economy?

-More efficient marketing (think of behavioral advertising addressing “customer heterogeneity”, increasing the return on marketing investment)

-More efficient pricing (think of pricing to market and price discrimination: first degree?) Think of complex auctions for ads (Varian), varying prices based on the outlet through which the product is sold (Priceline vs Expedia), or on the origin country (Mango online sales)

-More targeted product development & new business creation (which is also a source of more competition and efficiency)

Big Data:

Big data are going to benefit the economy *indirectly*:

-For instance, the use of personal data for diagnostic purposes and to avoid duplicative testing in the healthcare sector produces a real economic benefit, which however does not involve measurable market transactions; using cloud to improve public services gives similar positive externalities

-As another example, moving data storing to specific (cold, infrastructured) countries improves efficiency with cost benefits

All this makes it hard to measure the macroeconomic benefits of big data. Attempts have been limtied to parts of it, as cloud computing; it also points out privacy issues

Cloud Computing :

An Internet based GPT through which big data are stored in servers, processed and provided as a service on demand in a pay-per-use way:

1) Iaas (Infrastructure as a Service): renting virtual machines for data storage and management services (computer servers)
2) SaaS (Software as a Service): renting software as for email and ‘web based’ applications
3) PaaS (Platform as a Service): renting essentially an operating system on which applications can run in the cloud

- New Economics:
  - Pay for what you use
  - Lower and predictable costs
  - Shift from capex to opex
  - Rapid elasticity of usage
- Reduced Management cost
  - No maintenance
  - Robust multi-layered security
- Higher Productivity
  - Latest software for users
  - Resource pooling
  - Anywhere access

According to Gartner (2010), by this year 80 % of Fortune 1000 enterprises will be using some cloud computing services, 20 % of business will own no IT assets

The move from on premises servers to the private and public cloud

- Push the economy:
  - Shift fixed IT costs into (lower) variable costs in all sectors
  - Therefore induce enrtry in traditional sectors
  - Create new jobs (especially SMEs)
  - Strenghten competition and boost production

- Promote a healthy online ecosystem:
  - Consumerization of IT
  - Online privacy with a trusted, safe, secure internet
  - Innovative technology offers

- Adress societal challenges:
  - Supporting delivery of better healthcare
  - Supporting delivery of better education
  - Energy and environment improvements
  - Cost reduction/improvement productivity of government

Literature on the economic impact of online computing

- Fershtman and Gandal, 2012, Migration to the Cloud Ecosystem: Ushering in a New Generation of Platform Competition (CEPR 8907): *Despite the major changes in technology in IaaS services, estimates indicate that more than 90% of the cloud computing market (in terms of revenues) will involve (virtual) operating systems and applications software services (i.e., PaaS and SaaS services.)*

- Wynne Lam, 2013, Cloud Computing: Investment, Competition, and Demand Correlation, Toulouse S.E.: theoretical model on CC

- Etro, 2009, The Economic Impact of Cloud Computing on Business Creation, Employment and Output in Europe, *Review of Business and Economics*

- Lerner, Christensen et al., 2012, The Impact of Policy Changes on Investment in Cloud Computing Companies, Harvard University

- Etro, 2012, The Economics of Cloud Computing, Ch. 17 in "Cloud Computing Service and Deployment Models"

- Etro et al., 2013, The Impact of the Data Protection Regulation in the E.U., *European Financial Review*

- Etro and Colciago, 2013, Structural Change, Job Creation and Cloud Computing, in "Broadband in Latin America", United Nations

- Etro and Christensen, 2013, Big Data, the Cloud and the EU Regulation on Data Protection, *Intereconomics. Review of European Economic Policy*

The Economic Impact of Big Data management and Cloud Computing on macroeconomics (2013-2018)

Methodology: Dynamic Stochastic General Equilibrium model with Endogenous Market Structures

Background data: Eurostat (and eBusiness Reports)

Background data:

- *Eurostat data on EU -27 + Norway*
- *Aggregate sectors*
  - *manufacturing*
  - *wholesale & retail trade*
  - *hotels & restaurants*
  - *transport storage & communication*
  - *real estate and other business activities and services*

This image cannot currently be displayed.

Average share of ICT budget as % of total costs (by sector)

Impact on GDP growth (conservative estimates)

Short term (after one year):

+ 0.05-0.15 %

Medium term (after five years):

+ 0.1-0.3 %

Policy implications:

- *Part of the positive effects of cloud computing are going to be positively related to the speed of adoption of the new technology.*
- *Therefore, policymakers should promote as much as possible a rapid adoption of cloud computing. Concrete possibilities include:*
  - *– fiscal incentives and a specific promotion of cloud computing in particular dynamic sectors - for instance, governments could finance, up to a limit, the variable costs of computing*
  - *– business-friendly rules for the treatment and movement of data between their country and foreign countries.*

The Impact of Policy Changes on Investment in Companies active in Big Data management and Cloud Computing

Lerner, Christensen et al., 2012, The Impact of Policy Changes on Investment in Cloud Computing Companies, Harvard University

Overview

Objective: Understand the impact of copyright scope changes on investment by analyzing the effects of the court decisions on copyright infringements by businesses on **Remote Storage Digital Video Recorders** in US, France and Germany, on Venture Capital investment.

Methodology: Using a difference-in-difference approach, analyze whether investment in venture-backed U.S. and EU cloud companies (controlling for other factors) was affected by court rulings.

Results:

– VC investment increased significantly in the U.S. relative to the EU after the *Cablevision* decision (favorable to CC), with incremental investment in U.S. cloud computing companies estimated to be between $728 million and $1.3 billion. This may be the equivalent of $2 to $5 billion in traditional R&D investment.

– VC investment decreased significantly in France and Germany, relative to the rest of the EU, after the French and German court rulings (against CC), with decreased investment totaling approximately $87 million.

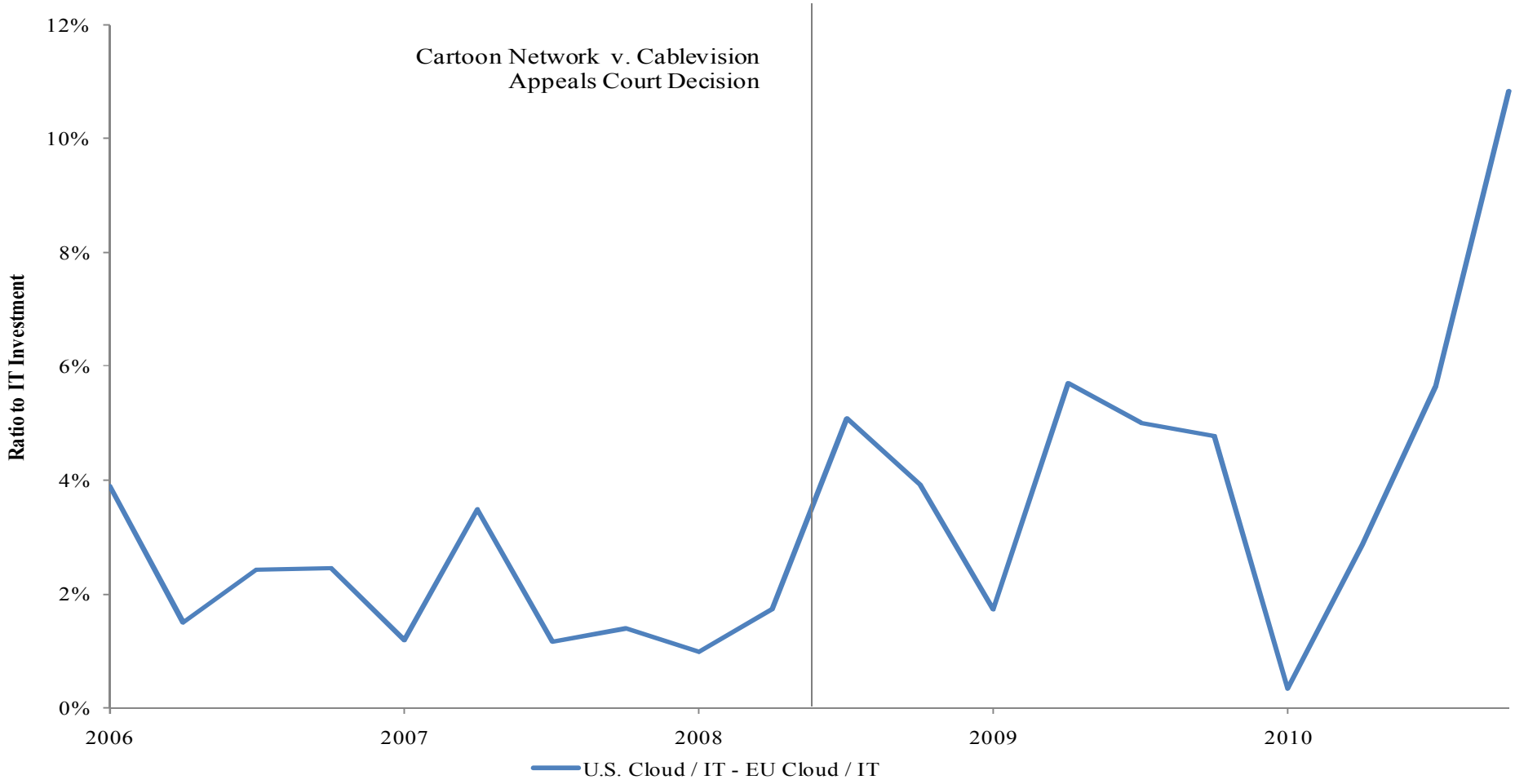
The *Cartoon Network v. Cablevision* Decision (August 2008) Reduced Uncertainty in the Cloud Computing Sector

- In 2006, Cablevision announced the development of a Remote Storage DVR.
- In response, a consortium of U.S. television and copyright holders sued Cablevision over alleged copyright infringement in May 2006. The case was litigated and in August 2008, the Appellate Court ruled that the DVR did not violate copyright law.
- The *Cablevision* decision was perceived as likely to positively impact cloud computing:

  "[A] rule holding Cablevision liable merely because it housed and maintained the servers in this case could imperil a wide variety of innovative business models that rely on the use of remote computing, ranging from examples like Internet-enabled self-service photo processing and printing, to cloud computing services offered by companies like Amazon, Apple and Google." (Kwun, 2008)

Investment Increased in U.S. Cloud Computing after the Final Ruling in *Cartoon Network v. Cablevision* (August, 2008)

Difference in the Ratio of Investment in Cloud Computing Companies to Investment in all IT Companies in the U.S. and EU



Source: Private Equity Investment data Jan 2006 - Dec 2010 from Thomson ONE.

French and German Court Rulings Were Seen as Negatively Affecting Cloud Computing

- In May 2008, Wizzgo launched the first online DVR platform in France. In response, a consortium of French television and copyright holders filed complaints against Wizzgo over alleged copyright infringement. On November 2008, the Tribunal de Grande Instance de Paris declared a final set of summary judgments against Wizzgo.
- In 2005, Shift.tv, and in 2006, Save.tv, were founded in Germany; both companies are subscription-based services that allow customers to select and store television content on servers. Two German television channels filed complaints over alleged copyright infringement. Although the litigation continues today, a number of court rulings, which were favorable to the plaintiffs, were made in 2006, 2007, and 2009.
- The French and German rulings were perceived as likely to negatively impact cloud computing in these countries

Results

- VC investment increased significantly in the U.S. relative to the EU after the *Cablevision* decision, with incremental investment in U.S. cloud computing companies estimated to be between $728 million and $1.3 billion. This may be the equivalent of $2 to $5 billion in traditional R&D investment.

– VC investment decreased significantly in France and Germany, relative to the rest of the EU, after the French and German court rulings, with decreased investment totaling approximately $87 million.

Cloud Computing Regression Analysis: *Cablevision* Specification

$$\textbf{VC Ratio}_{r,t} = \beta_0 + \beta_1(\textbf{US Dummy})_r + \beta_2(\textbf{2008 Dummy})_t + \beta_3(\textbf{Effect of } \textit{Cablevision} \textbf{ Decision in US})_{r,t} + \varepsilon_{r,t}$$

where:

- *r* indexes region (either the U.S. or EU);
- *t* indexes quarter of year (Q1 1995 – Q3 2011 and Q1 2006 – Q3 2011);
- VC $\text{Ratio}_{r,t}$ is equal to VC dollars invested in the cloud computing sector in region *r* at quarter *t* divided by total IT VC dollars invested in region *r* at quarter *t*;
- US Dummy is equal to 1 for the U.S. cloud computing sector and 0 for the EU cloud computing sector;
- 2008 Dummy is a dummy variable for the *Cablevision* Appellate Court decision and is equal to zero for all quarters before Q3 2008 and one for all quarters including and after Q3 2008;
- Effect of *Cablevision* Decision in US is equal to the interaction between the US Dummy and 2008 Dummy; and,
- The unit of observation is quarterly investment in cloud computing by region.

Cloud Computing Regression Results: U.S. vs. EU[1,2]
Dependent Variable: Ratio of Cloud Computing VC Dollars to Total IT VC Dollars

U.S. Indicator	0.0202***	0.0129***
	(0.0048)	*(0.0045)*
2008 Dummy[3]	0.0059	-0.0094
	(0.0080)	*(0.0090)*
Effect of Cablevision on U.S. VC Investment	0.0257**	0.0256**
	(0.0114)	*(0.0095)*
Percent Change in GDP		0.0093***
		(0.0030)
Broadband Penetration Rate		0.3754***
		(0.0900)
Constant	0.0117***	-0.0629***
	(0.0038)	*(0.0167)*
Observations	40	40
Adjusted R-Squared	0.544	0.699
Implied Increase in U.S. Cloud VC Investment ($ Millions)	**$730**	**$728**
Length of Time Period	1Q 2006 - 4Q 2010	1Q 2006 - 4Q 2010

Notes:

[1] Robust standard errors (by quarter) are provided under the point estimates in italics.

[2] *** indicates significance at a 1 percent level, ** indicates significance at a 5 percent level, and * indicates significance at a 10 percent level.

[3] Decision by Appellate Court (judgment of District Court is reversed) (8/4/2008). The 2008 Dummy variable is set equal to one for all quarters after 2Q 2008.

Cloud Computing Regression Results: U.S. vs. EU[1,2]
Dependent Variable: For Each of IaaS, PaaS, and SaaS, Ratio of Cloud Computing VC Dollars to Total IT VC Dollars

	IaaS	PaaS	SaaS
U.S. Indicator	0.0026	0.0057***	0.0151***
	(0.0034)	*(0.0013)*	*(0.0037)*
2008 Dummy[3]	-0.0016	0.0000	0.0081
	(0.0028)	*(0.0005)*	*(0.0073)*
Effect of Cablevision on U.S. VC Investment	0.0117**	0.0061**	0.0089
	(0.0050)	*(0.0030)*	*(0.0087)*
Constant	0.0039	0.0003	0.0074
	(0.0026)	*(0.0003)*	*(0.0026)*
Observations	40	40	40
Adjusted R-Squared	0.389	0.546	0.448
Length of Time Period	1Q 2006 - 4Q 2010	1Q 2006 - 4Q 2010	1Q 2006 - 4Q 2010

Notes:

[1] Robust standard errors (by quarter) are provided under the point estimates in italics.

[2] *** indicates significance at a 1 percent level, ** indicates significance at a 5 percent level, and * indicates significance at a 10 percent level.

[3] Decision by Appellate Court (judgment of District Court is reversed) (8/4/2008). The 2008 Dummy variable is set equal to one for all quarters after 2Q 2008.

Debate on the EU Privacy Regulation

Christensen and Etro (*Intereconomics*, 2013)

- Impact of the EU Data Protection Regulation (as of 2013) on a wide number of EU SMEs

- High compliance costs. Examples:

- Costs for **Data Protection Officer** for more than 250 emploees or processing data as part of core activities

  - EU Impact Assessment estimates 4 hours/year for 0.1% of SMEs that are data controllers (50% of 18 million EU SMEs), for a total of 9000 SMEs
  - We estimate 100 hours/year and recalculated the number of SMEs involved

- Principle of Minimization of Data Processing and Right to object to profiling

  - No costs in EU Impact Assessment, but losses from lost revenues for eliminating **targeted advertising** and other forms of direct marketing
  - UK Direct Marketing Association estimated average losses of 94000 euros/year per firm in the UK

Debate on the EU Privacy Regulation

– Notification of personal **data breaches** to the supervisory authorities within 24 hours

- Extremely demanding, especially for non serious data breaches
- High cost of sanctions, especially without distinction between intentional and unintentional harm

– Data protection by design (of systems and procedures) and default (to ensure minimum processing of personal data) and **Data Protection Impact Assessment** prior to risky processing for the rights of data subjects

•The EU Impact Assessment provided three examples: small scale DPIA (commercial use of sensitive data) at 14000 euros, medium (location based data) at 34000, and large (security and biometrics) at 149000

•But the frequency is high and sums to a large cost

– Benefits:

•“One-stop principle”: regulation by a single supervisory authority (but not for companies that are both data controllers and processors, as cloud computing providers)

•Effort in promoting secure data transfers

Debate on the EU Privacy Regulation

Average costs for SMEs:

between 3000 and 7000 euros

between 18% and 40% of IT budget

plus additional Macroeconomic costs (business destruction)

Policy is key to
- promote the adoption of technologies that create growth (cloud computing),
- promoting innovation (businesses developed on the cloud) and
- optimizing the regulation of economic activity (data protection regulation)