The Digital Commons: Tragedy or Opportunity?
The Effect of Crowdsourced Digital Goods on Economic Growth

Frank Nagle

Marshall School of Business

University of Southern California

ARNIC

10/13/16

Introduction

- The Tragedy of the Commons (Hardin, 1968) – overuse of a public good
  - Digital goods are non-rival and essentially infinitely abundant
- The Tragedy of the Digital Commons – free crowdsourced digital goods destroy existing business models, without replacing them in an economically measurable way
  - E.g., the encyclopedia industry
  - Creative destruction (Schumpeter, 1942), but the destructive impact is more easily measured than the creative impact

Introduction

- The Opportunity of the Digital Commons
  - May not contribute to economic value directly, but contribute to productivity and general quality of life
  - Knowledge repositories
  - User review sites
  - Open source software (OSS)

- Three papers that examine the contribution of OSS to the economy and firm production

Paper 1 - Digital Dark Matter and GDP
(w/ Shane Greenstein)

- Digital Dark Matter - Digital goods that are non-pecuniary and effectively limitless
  - Important inputs into production
  - Systematically undercounted in productivity measures
- Open Source Software (OSS) is an example of Digital Dark Matter that is widely used throughout the economy
  - Others include Wikipedia, Yelp, YouTube, digitized 3D blueprints, and many more
- The Apache Web Server is an important example of this phenomenon – non-pecuniary, widely used, originally government funded R&D

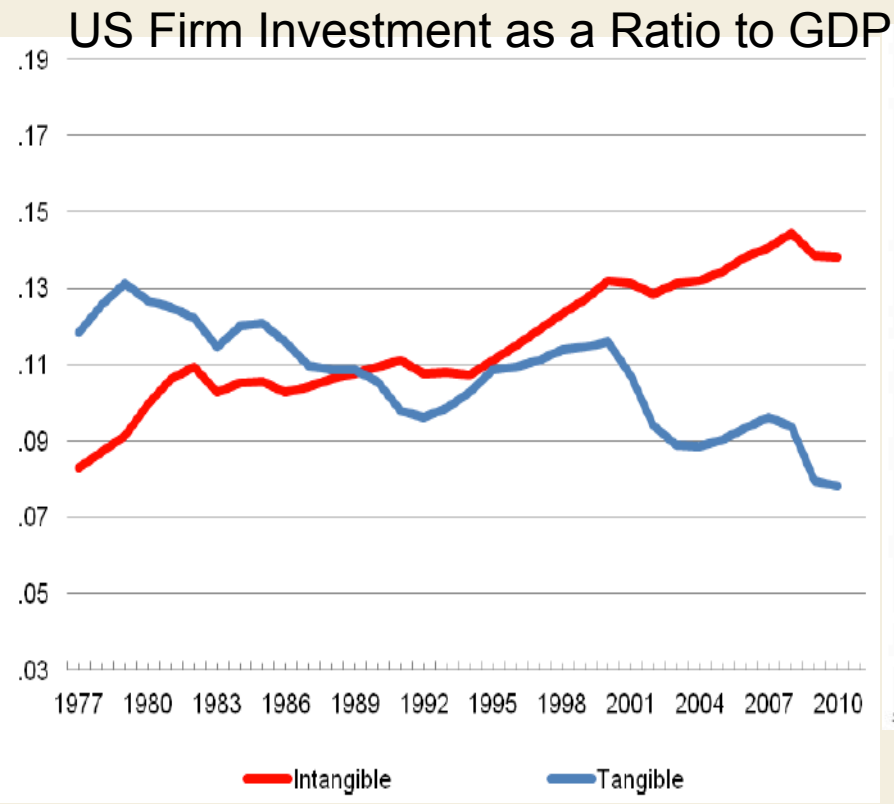
Measuring Digital Dark Matter and Apache

- Scanned 1% of the 1.5 billion IPv4 addresses in the US
  - Found ~200,000 web servers, 23% were running Apache
  - This leads to an estimate of ~4 million Apache servers in the US

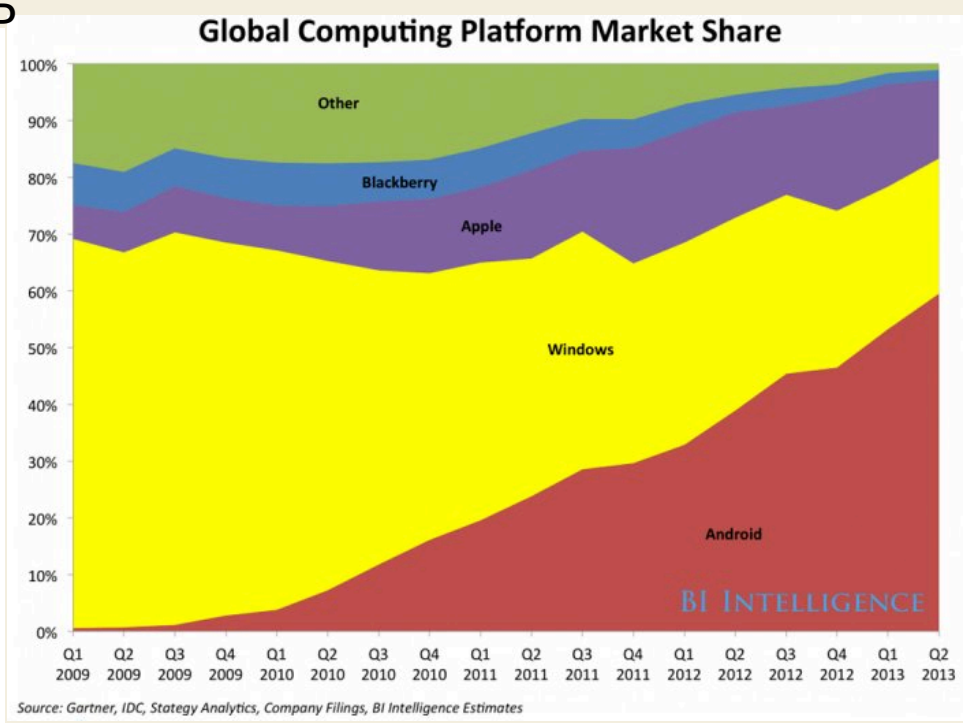
Measuring Digital Dark Matter and Apache

- Follow Nordhaus (2006) and impute the price of Apache based on a comparable market good – Microsoft IIS
  - Value of Apache is between **$2 billion** and **$12 billion**
  - Equivalent of 1.3% to 8.7% of the value of all prepackaged software investments
  - Represents a 17% to 19% rate of return, if Apache was the *only* good to come out of all of the US government's NSF investment in super-computing centers from 1985-1995

Paper 2 – OSS and Firm Productivity



Source: Corrado and Hulten (2013)



What is the impact of crowdsourced digital goods on firm-level productivity?

Related Literature

User Innovation

- Important topic in management as early as von Hippel (1986)
- OSS is a frequently studied phenomenon within this literature
- Lerner and Tirole, 2002; Lakhani and von Hippel, 2003; von Hippel and von Krogh, 2003; West and Lakhani, 2008; Lerner and Schankerman, 2010; Casadesus-Masanell and Llanes, 2011; many others
- **Predominately focused on the supply side, not the productivity of usage**

Productivity of IT

- IT investment contributes to both firm and national growth
- Brynjolfsson and Hitt, 1996; Jorgenson, Ho, and Stiroh, 2005; Syverson, 2011; Tambe, Hitt, and Brynjolfsson, 2011; Huang, Ceccagnoli, Forman, and Wu, 2013; many others
- **IT is measured via expense on IT HW, SW, or labor, which does not account for IT without a price**

Institutional Context – OSS Operating Systems

- GNU Project and Free Software Foundation established in mid-1980's
  - First effort to create a free and open operating system and related software
- Linux Kernel created in 1991 and added to GNU
  - Many, many flavors of Linux built on top of the kernel
  - Linux is developed and maintained by a community of mostly unpaid contributors
- Many pecuniary systems are built on non-pecuniary OSS (e.g., RedHat Linux, IBM HTTP Server)

Risks of Non-Pecuniary OSS

- Using non-pecuniary OSS can be risky:
  - No guaranteed technical support (Woods and Guliani, 2005)
  - No guaranteed technical path (Kogut and Metiu, 2001)
  - Security concerns (despite Linus's Law)
  - No contractual relationship (no one to sue)
- Free software is not truly free
  - Costs of software are **<** 10% of total cost of implementing software (MacCormack, 2003; Varian and Shapiro, 2003)
  - Long-term costs of open-source software are 5% to 20% higher than proprietary closed source (Giera and Brown, 2004)
- *"No one ever got fired for buying Microsoft."*

Benefits of Non-Pecuniary OSS

- Joy's Law: "No matter who you are, most of the smartest people work for someone else."

  Bill Joy, co-founder Sun Microsystems

  - Knowledge is distributed throughout society and cannot be fully aggregated in one central body (von Hayek, 1945)
  - The Linux kernel has 10,000 contributors
  - The Windows 8 kernel had one team of < 40 people (Sinofsky, 2011)

- Up-front cost savings

- Flexibility to alter and enhance code (Schwarz and Takhteyev, 2011)

Data

- Firm-level observations for 1566 public firms from 2000-2009
- Technology usage (Harte Hanks IT Survey)
  - Site-level survey of technology usage and employment
  - Includes information on desktop and server operating system usage, including free and paid OSS
- Financial performance (Compustat)

Variables

- VA_{it} = productive value-added output of firm *i* at time *t*
- ITK_{it} = IT capital stock
- ITL_{it} = IT labor
- K_{it} = non-IT capital stock
- L_{it} = non-IT labor
- Operating Systems: Count of the number of computers running a give type of OS

Estimation Strategy

- Cobb-Douglas production function including IT

$$VA_{it} = K_{it}^{\alpha} L_{it}^{\beta} IT_{it}^{\gamma} A_{it}$$

$$\ln(VA_{it}) = \alpha \ln K_{it} + \beta \ln L_{it} + \gamma \ln IT_{it} + \varepsilon_{it}$$

$$\ln(VA_{it}) = \alpha \ln K_{it} + \beta \ln L_{it} + \gamma_1 \ln ITK_{it} + \gamma_2 \ln ITL_{it} + \varepsilon_{it}$$

$$\ln(VA_{it}) = \alpha \ln K_{it} + \beta \ln L_{it} + \gamma_1 \ln ITK_{it} + \gamma_2 \ln ITL_{it} + \underline{\gamma_3 \ln OSS_{it}} + \varepsilon_{it}$$

Continuous Adoption of Non-Pecuniary OSS

DV: Value-Added (VA_{it})	1	2	3	4	5
Model	OLS	OLS – NAICS5	OLS FE	ABOND	BBOND
IT Capital (IT_{it})	0.054***	0.040***	0.026***	0.023***	0.023***
	(0.008)	(0.007)	(0.006)	(0.005)	(0.005)
Non-IT Capital (K_{it})	0.260***	0.226***	0.060	-0.154	-0.080
	(0.014)	(0.019)	(0.049)	(0.102)	(0.095)
Non-IT Labor (L_{it})	0.725***	0.768***	0.790***	0.698***	0.680***
	(0.017)	(0.022)	(0.040)	(0.054)	(0.051)
$non\_pecuniary\_OSS_{it}$	-0.004***	-0.001	-0.001	-0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Constant	0.131	0.291***	1.114***	0.019	-0.273
	(0.197)	(0.063)	(0.260)	(0.680)	(0.616)
Year fixed effect?	Y	Y	Y	Y	Y
Industry fixed effect	NAICS2	NAICS5	-	NAICS2	NAICS2
Number of firm/year observations	10355	10355	10355	7650	8988
Number of firms (groups)	1566	1566	1566	1424	1501
R^2 (within for panel)	0.930	0.952	0.444		

Adoption of NP OSS by IT-Producers

DV: Value-Added (VA_{it})	1	2	3	4	5
Model	OLS	ABOND	BBOND	ABOND	ABOND
IT Capital (IT_{it})	0.053***	0.022***	0.022***	0.031**	0.031**
	(0.008)	(0.005)	(0.005)	(0.014)	(0.013)
Non-IT Capital (K_{it})	0.259***	-0.157	-0.087	-0.242	-0.248
	(0.014)	(0.102)	(0.095)	(0.182)	(0.181)
Non-IT Labor (L_{it})	0.724***	0.698***	0.681***	0.806***	0.812***
	(0.017)	(0.054)	(0.052)	(0.163)	(0.161)
$non\_pecuniary\_OSS_{it}$	-0.006***	-0.001	-0.001	0.004**	0.005***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
$non\_pecuniary\_OSS_{it}$ x IT-producer	0.016***	0.004**	0.005**		
	(0.002)	(0.002)	(0.002)		
$non\_pecuniary\_OSS_{i,t-1}$					0.003*
					(0.002)
$non\_pecuniary\_OSS_{i,t-2}$					
$non\_pecuniary\_OSS_{i,t-3}$					
Constant	0.132	-0.022	-0.305	0.664	0.417
	(0.195)	(0.685)	(0.619)	(1.106)	(1.102)
Year FE?	Y	Y	Y	Y	Y
Industry FE	Y	NAICS2	NAICS2	NAICS2	NAICS2
Sample Restriction	-	-	-	IT-Producing Firms	IT-Producing Firms
# obs	10355	7650	8988	1070	1070
# firms	1566	1424	1501	237	237
R^2	0.932				

Robustness Checks

DV: Value-Added (VA_{it})	1	2	3	4	5	6	7	8	9
Model	ABOND – NAICS5	BBOND – NAICS5	LevPet GMM	IPW OLS	CEM OLS	2SLS (emp)	2SLS (rev)	ABOND – IV (emp)	ABOND – IV (rev)
IT Capital (IT_{it})	0.022***	0.022***	0.083***	0.051***	0.068***	0.035	0.067	0.023***	0.023***
	(0.005)	(0.005)	(0.007)	(0.012)	(0.014)	(0.135)	(0.132)	(0.005)	(0.005)
Non-IT Capital (K_{it})	-0.158	-0.158	0.504***	0.230***	0.221***	0.248***	0.256***	-0.181*	-0.180*
	(0.103)	(0.103)	(0.099)	(0.023)	(0.017)	(0.036)	(0.036)	(0.101)	(0.101)
Non-IT Labor (L_{it})	0.687***	0.687***	0.725***	0.754***	0.741***	0.706***	0.710***	0.695***	0.696***
	(0.057)	(0.057)	(0.021)	(0.027)	(0.023)	(0.030)	(0.030)	(0.055)	(0.055)
$non\_pecuniary\_OSS_{it}$	-0.001*	-0.001*	-0.009***	-0.005***	-0.007***	-0.022	-0.040	-0.004***	-0.004***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.079)	(0.077)	(0.001)	(0.001)
$non\_pecuniary\_OSS_{it}$ x IT-producer	0.004**	0.004**	0.022***	0.013***	0.029***	0.170***	0.159***	0.010*	0.010*
	(0.002)	(0.002)	(0.002)	(0.003)	(0.006)	(0.038)	(0.037)	(0.006)	(0.006)
Constant	0.369	-0.613	-	0.095	0.311*	0.255	-0.048	-0.597	-0.982
	(0.693)	(1.108)	-	(0.218)	(0.188)	(1.294)	(1.274)	(0.671)	(0.634)
Year FE?	Y	Y	Y	Y	Y	Y	Y	Y	Y
Industry FE	NAICS5	NAICS5	NAICS2	NAICS2	NAICS2	NAICS2	NAICS2	NAICS2	NAICS2
First Stage F-statistic	-	-	-	-	-	52.83	52.79	-	-
Cragg and Donald Min. Eigenvalue	-	-	-	-	-	1.815	1.835	-	-
# obs	7650	8988	10355	10355	6015	10355	10355	7650	7650
# firms	1424	1501	1566	1566	1414	1566	1566	1424	1424
R^2				0.932	0.910	0.809	0.823		

Additional Interactions

- Smaller firms get more from using NP OSS
- Firms with fewer IT employees get more from using NP OSS
- No interaction effects with ITK, R&D, Local IT Authority

Paper 3 – Free Ride vs. Contribute

- Profitability and success of firms is driven by competitive advantage
  - *Resources* or *capabilities* a firm has that it's competitors do not
  - How do public goods that all firms can use factor into competition?

- Some firms pay their own employees to contribute to the creation of public goods that their competitors can use for free. Why?

Theory and Hypotheses

- Learning by doing
  - Knowledge is a public good (Stiglitz, 1999)
  - Information is codified into knowledge via experience
  - Learning by doing has important implications for growth
    - At the economy level (Arrow, 1962; Romer, 1989)
    - At the organization-level (Herriott, Levinthal, and March, 1985; Levitt and March, 1988)
  - Transfer of knowledge via experience may be the reason firms exist (Huber, 1991; Kogut and Zander, 1992, 1996; Grant, 1996; Kogut, 2000)
  - Integration of external knowledge can be improved via investing in absorptive capacity (Cohen and Levinthal, 1989, 1990)

Theory and Hypotheses

- Learning by Contributing
  - Public physical goods are simple
    - Learning by doing is enough
  - Public information goods are complex
    - Learning by doing (free-riding) is not enough
    - Learning by contributing allows a deeper understanding of how to use the good for productive purposes

Variables

- VA_{it} = productive value-added output of firm *i* at time *t*
- IT_{it} = IT expenditure
- K_{it} = non-IT capital stock
- L_{it} = non-IT labor

- OSS_{it} = Number of OSS operating systems at the firm

- $Post_{it}$ = 1 after the firm (or it's match) contributes

- # Contributors = Number of contributors to Linux from firm this year
- # Changes = Number of lines contributed to Linux
- # Signoffs = Number of approvals of lines contributed to Linux

Estimation Models

$$Y_{it} = K_{it}^{\alpha} L_{it}^{\beta} IT_{it}^{\gamma} A_{it}$$

$$\ln(VA_{it}) = \alpha \ln K_{it} + \beta \ln L_{it} + \gamma \ln IT_{it} + \varepsilon_{it}$$

$$\ln(VA_{it}) = \alpha \ln K_{it} + \beta \ln L_{it} + \gamma_1 \ln IT_{it} + \gamma_2 \ln OSS_{it} + \varepsilon_{it}$$

$$\begin{aligned}\ln(VA_{it}) = {} & \alpha \ln K_{it} + \beta \ln L_{it} + \gamma_1 \ln IT_{it} + \gamma_2 \ln OSS_{it} + \gamma_3 Contrib_i + \gamma_4 Post_{it} \\ & + \gamma_5 Contrib_i * \ln OSS_{it} + \gamma_6 Contrib_i * Post_{it} + \gamma_7 \ln OSS_{it} * Post_{it} \\ & + \gamma_8 Contrib_i * \ln OSS_{it} * Post_{it} + \varepsilon_{it}\end{aligned} \quad (4)$$

$$\begin{aligned}\ln(VA_{it}) = {} & \alpha \ln K_{it} + \beta \ln L_{it} + \gamma_1 \ln IT_{it} + \gamma_2 \ln OSS_{it} + \gamma_3 Contrib_i + \gamma_4 Post_{it} \\ & + \gamma_5 Contrib_i * \ln IT_{it} + \gamma_6 Contrib_i * Post_{it} + \gamma_7 \ln IT_{it} * Post_{it} \\ & + \gamma_8 Contrib_i * \ln IT_{it} * Post_{it} + \varepsilon_{it}\end{aligned} \quad (5)$$

Results – Benefits of Contribution

DV: Value-Added (VA_{it})	1	2	3	4	5
$(Contrib_i * OSS_{it})$	0.004 (0.026)	-0.061 (0.040)	-0.016 (0.031)	-0.036 (0.027)	-0.087*** (0.028)
$(Contrib_i * Post_{it})$	0.059 (0.130)	-0.402* (0.229)	-0.286 (0.207)	-0.142 (0.172)	-0.316 (0.196)
$(OSS_{it} * Post_{it})$	-0.013 (0.022)	-0.060** (0.027)	-0.026 (0.024)	-0.034 (0.021)	-0.081*** (0.021))
$(Contrib_i * OSS_{it} * Post_{it})$		0.107** (0.041)	0.066* (0.034)	0.046* (0.027)	0.073** (0.031)
R&D Expense (RD_{it})			0.322*** (0.060)		
Standard Error	Clustered	Clustered	Clustered	Clustered	GMM
Robustness check			R&D Control	Firm FE	Arellano-Bond
N	681	681	611	681	604
R^2	0.907	0.908	0.932	0.563	-

***p<.01, **p<.05, *p<.1. All variables are the natural log of the underlying variable. All columns include controls for IT_{it}, K_{it}, L_{it}, OSS_{it}, $Contrib_i$, $Post_{it}$, and year as well as a constant not shown for space. The regressions include a lagged variable (not shown due to space constraints) for all variables related to the use of and contribution to OSS including *OSSit*, *Postit*, and any interaction variable including one of these two.

Results – Contribution Intensity

DV: Value-Added (VA_{it})	1	2	3	4	5	6
Contribution Intensity ($Contrib\_Int_{it}$)	-0.029* (0.015)	-0.022* (0.012)	-0.023* (0.013)	-0.020* (0.010)	-0.022* (0.013)	-0.018 (0.011)
($Contrib\_Int_{it}$ * OSS_{it})	0.004* (0.002)	0.003** (0.002)	0.003* (0.002)	0.003** (0.001)	0.003* (0.002)	0.003* (0.001)
Measure of Contribution Intensity	Number of Contributors	Number of Contributors	Number of Changes	Number of Changes	Number of Signoffs	Number of Signoffs
Model Type	OLS	Firm FE	OLS	Firm FE	OLS	Firm FE
N	307	307	307	307	307	307
R^2	0.934	0.540	0.934	0.540	0.934	0.537

***p<.01, **p<.05, *p<.1. All variables are the natural log of the underlying variable. All regressions are OLS models and use clustered standard errors at the firm level. All columns include controls for IT_{it}, K_{it}, L_{it}, OSS_{it}, and year as well as a constant not shown for space. The regressions include a lagged variable (not shown due to space constraints) for all variables related to the use of and contribution to OSS including OSS_{it}, $Post_{it}$, and any interaction variable including one of these two.

Results – Spillover to all IT

DV: Value-Added (VA_{it})	1	2	3	4
$(Contrib_i * OSS_{it})$		-0.051 (0.045)		
$(Contrib_i * Post_{it})$	-0.569* (0.308)	-0.651** (0.290)	-0.488 (0.321)	-0.268 (0.307)
$(OSS_{it} * Post_{it})$		-0.047 (0.032)		
$(Contrib_i * OSS_{it} * Post_{it})$		0.099** (0.046)		
$(Contrib_i * IT_{it})$	-0.100 (0.076)	-0.046 (0.079)	-0.029 (0.073)	-0.070 (0.078)
$(IT_{it} * Post_{it})$	-0.113** (0.049)	-0.067 (0.048)	-0.065 (0.054)	-0.062* (0.035)
$(Contrib_i * IT_{it} * Post_{it})$	0.159** (0.073)	0.072 (0.076)	0.132* (0.076)	0.083 (0.070)
R&D Expense (RD_{it})			0.357*** (0.064)	
Robustness Check			R&D Control	Firm FE
N	681	681	607	681
R^2	0.906	0.909	0.932	0.560

Conclusion

- Crowdsourced digital goods & digital dark matter are missed in GDP calculations
  - Apache alone accounts for between $2 billion and $12 billion that is missing from GDP
- OSS has a positive impact on firm productivity
  - For non-IT producing firms, it takes 1-6 years to occur
  - Bigger effect for smaller firms
- Contributing to OSS leads to higher productivity from using OSS
  - Contributors gain 11% more productivity from using OSS than free-riding peers

The Digital Commons: Tragedy or Opportunity?

The Effect of Crowdsourced Digital Goods on Economic Growth

Frank Nagle

Marshall School of Business

University of Southern California

frank.nagle@marshall.usc.edu

Twitter: @frank_nagle