

Kto i dlaczego wytwarza oprogramowanie open source

Tomasz Plata-Przechlewski

Warszawa 2015

## Zagadnienia

- Specyfika produkcji/wytwarzania oprogramowania OSS
- Czynniki motywacji (krótki przegląd teorii)
- Empiryczna weryfikacja teorii autodeterminacji oraz charakterystyka demograficzno-społeczna twórców OSS
- Empiryczna weryfikacja teorii sygnalizacji
- Motywacja a zaangażowanie
- Podsumowanie
- Literatura

## Specyfika produkcji/wytwarzania oprogramowania OSS

Istotą licencji *Open Source* jest zgoda twórcy na zwielenie i zmianę/modyfikowanie programu (a nie to, że twórca udostępnia program „za darmo”). Copyleft vs copyright:

*Copyrights exist in order to protect authors from unauthorized copying or selling of their work. A Copyleft, on the other hand, provides a method for software [...] to be modified, and distributed back to the community, provided it remains Libre.*

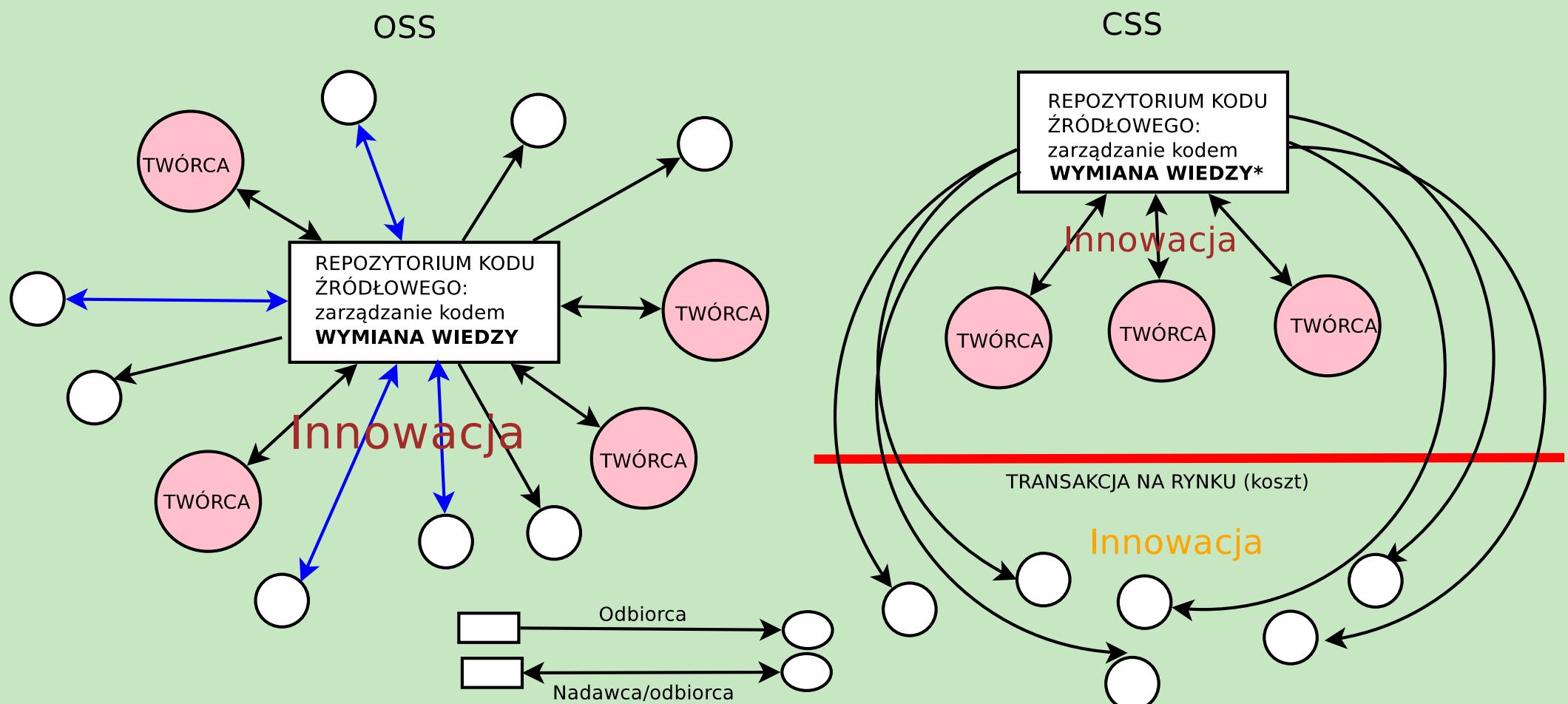
brak wyłącznych praw własności

(right to exclusive use of intellectual property)

⇒ brak oprogramowania

Oprogramowanie własnościowe (CSS = *Closed Source Software*) vs OSS

Konsekwencje (braku) wyłącznych praw własności w przypadku wytwarzania oprogramowania:



## OSS w liczbach na przykładzie repozytorium sourceforge.net

Liczba projektów OSS w (największym) repozytorium kodu sourceforge.net (styczeń 2010) – 160,141.

Rozkład projektów ze względu na liczbę twórców (sourceforge.net styczeń 2010)

Liczba twórców	Liczba projektów	%
1	110,622	82,62
2	15,412	11,51
3	4,529	3,38
4 i więcej	3,329	2,49
Razem	133,892	100,0

Rozkład projektów ze względu na liczbę pobrań (sourceforge.net styczeń 2010)

Liczba pobrań	Liczba projektów	%
0	55,986	41,61
1-999	43,619	32,42
1000-9999	24,438	18,16
10000 i więcej	10,506	7,81
Razem	153,730	100,0

Źródło: Emanuel i inni (2010), *Statistical analysis of the Open Source Software Projects from sourceforge.net* Int. Conference on Computer and Mathematical Sciences

## OSS w liczbach na przykładzie repozytorium sourceforge.net cd.

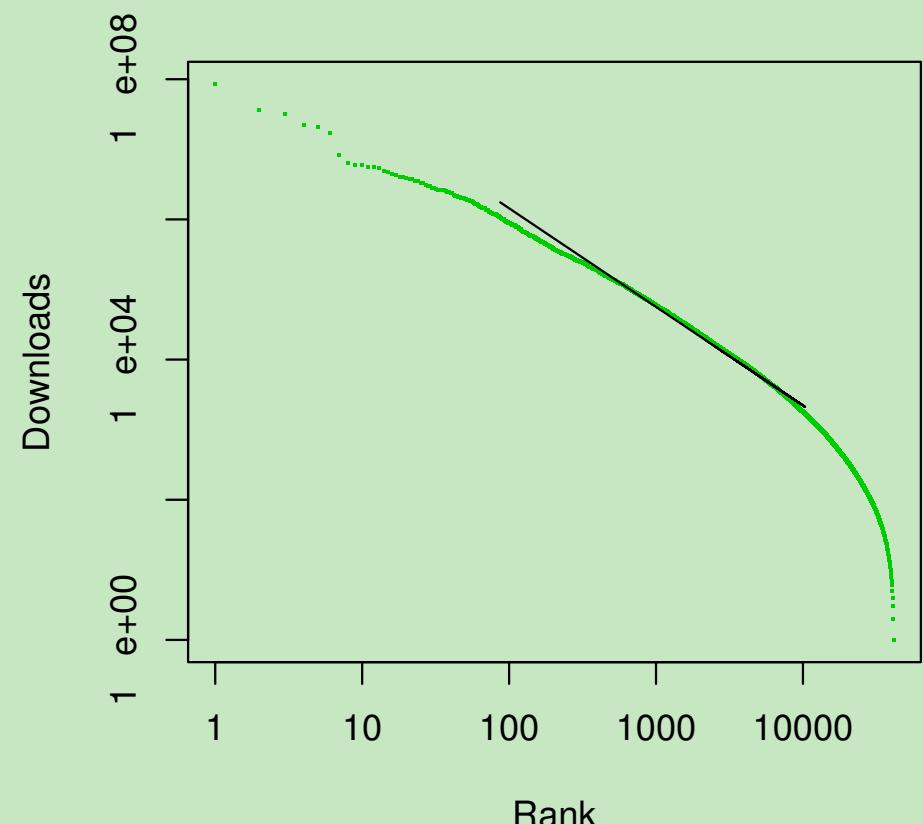
Analiza sourceforge.net w 2005 r. (Feitelson i Heller (2006)\*)

Łączna liczba projektów – 122,205. Liczba projektów pobranych co najmniej raz – 41,608 (34%).

Kryterium sukcesu jest liczba pobrań. Autorzy zakładają *per analogiam*, że rozkład liczby pobrań powinien być zbliżony do rozkładu Zipfa

*Zipf distributions are encountered in many other fields [...] Number of pages in a website, number of visitors to a website, number of links to other sites*

[...]



Status	% projektów	Liczba pobrań	Udział pobrań
Super-projekty	0,2	1,1mln–85mln	59,4
Udane	24,7	1,681–1,08mln	39,1
Nieudane	75,1	1–1,681	1,5

Łącznie projekty super+udane = 10358/122205 = 8,5%

Źródło: Feitelson D.G., Heller G.Z., Schach S.R. (2006) *An Empirically-Based Criterion for Determining the Success of an Open-Source Project*, (ASWEC'06), 363–368

Inne analizy statystyczne projektów OSS: Krishna i Srinivasa (2011) – 76% twórców rozwija tylko jeden projekt, 90% zarejestrowanych użytkowników nie rozwija żadnego projektu, 70% projektów rozwija 1 twórcę.

## Jakim cudem to działa?

Duża część twórców OSS to nieopłacani wolontariusze. Dlaczego poświęcają swój (wolny) czas na produkcje oprogramowania?

**Why should thousands of top-notch programmers contribute freely to the provision of a public good – Lerner and Tirole (2002)**

**Who can afford to do professional work for nothing? What hobbyist can put 3-man years into programming, finding all bugs, documenting his product and distribute for free?**

**– William Henry Gates III (1976)**

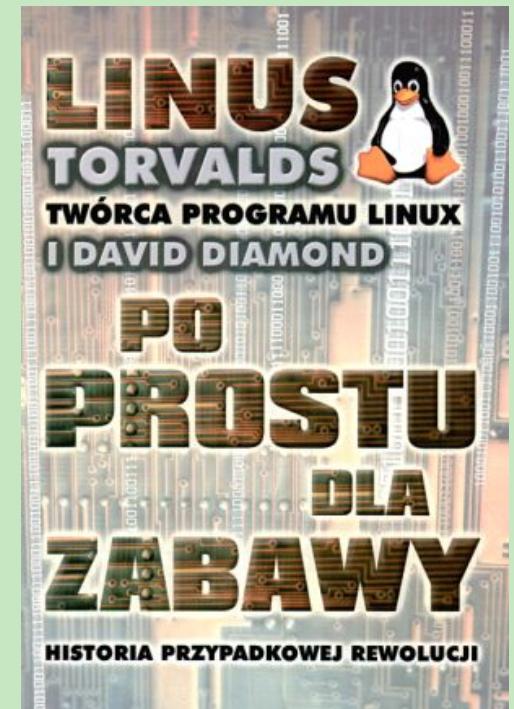
**Dwa modele innowacji:**

Prywatna inwestycja (*private investment*): zwrot z inwestycji jest możliwy dzięki skutecznej ochronie własności intelektualnej.

„Działanie wspólne” (*collective action*): w sytuacji zawodności rynku, innowatorzy współpracują w celu wytworzenia *dobra publicznego* (problem jazdy na gapę).

Hippel i von Krogh (2003) proponują „prywatno-wspólny” model, w którym (przy pewnych warunkach) *prywatne korzyści* dla producentów są wyższe niż dla tych, którzy chcieliby jechać na gapę. Przykładowo rezultatem zaangażowania w tworzenie OSS jest korzyść w postaci zwiększenia umiejętności i/lub satysfakcji/przyjemności, niedostępna dla gapowicza.

Kiedy koszty związane z udostępnieniem/odebraniem informacji są niskie, nawet bardzo małe korzyści mogą skutkować decyzją o upublicznieniu.



## Czynniki motywacji

Teoria autodeterminacji (SDT, (Ryan i Deci, 2000)) dzieli czynniki motywacji na: wewnętrzne i zewnętrzne.

Motywacja wewnętrzna to zaspokojenie potrzeby poprzez określona czynność (np. programowanie).

→ ideologiczne, altruizm, radość tworzenia (uniesienie, *flow* (Csíkszentmihályi))

→ uwewnętrzniione zewnętrzne: reputacja, wzajemność, uczenie się (*human capital*), użytek własny

**Kapitał ludzki:** kwalifikacje i umiejętności podnoszą wartość pracownika na rynku pracy.

**Kapitał społeczny:** wzajemne relacje społeczne oraz zaufanie jednostek, przy pomocy których możliwe jest pomnażanie osiąganych korzyści.

Motywacja zewnętrzna to instrument do osiągnięcia konkretnego wyniku

→ kariera (sygnalizacja), wynagrodzenie (np. usługi komplementarne)

**Teoria sygnalizacji** (Spence 1973) postuluje, że pracodawca nie będąc w posiadaniu wszystkich informacji o pracownikach (asymetria), może wykorzystywać informacje o wykształceniu jako *sygnał* wysokich zdolności, pracowitości lub ambicji.

Ponieważ dla pracowników o niższych zdolnościach, osiągnięcie tego samego sygnału wiąże się z wyższym kosztem, sygnalizacja pozwala na rozróżnienie między lepszymi i gorszymi pracownikami i zatrudnienie odpowiedniej osoby i/lub zaproponowanie odpowiedniej stawki płacowej.

Sygnalizacja jest tym silniejsza im: 1. bardziej widoczne (dla relevantnego odbiorcy) są osiągnięte rezultaty 2. wysiłek ma większy wpływ na wynik

## Efekt wypychania

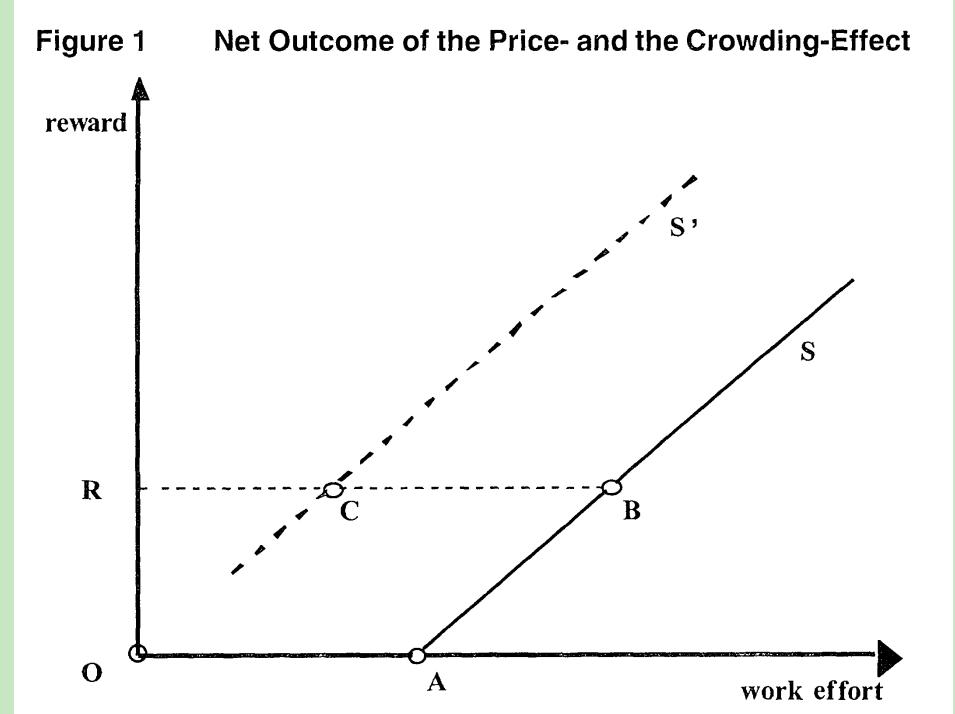
Teoria agencji postuluje, że czynniki motywacji wewnętrznej i zewnętrznej sumują się.

Wg teorii SDT zachęty pieniężne mogą wypychać (*crowd out*) czynniki motywacji wewnętrznej: *jeżeli nagroda (płaca) postrzegana jest jako forma kontroli, to osłabia ona motywację wewnętrzną*.

Zachęty/bodźce są słabymi stymulatorami działającym w krótkim okresie; w długim okresie okazują się destymulantami skłaniając do zachowań oportunistycznych (Bénabou i Tirole 2003).

Wyniki eksperymentu Ederera i Manso (2010): płaca za efekt ma ujemny wpływ na kreatywność i wynik.

Sauerman i Cohen (2010): empiryczna weryfikacja na próbie 11 tys pracowników działów R&D. Oba rodzaje motywacji mają wpływ na wysiłek (czas pracy) i wynik (liczba patentów), ale wpływ motywacji wewnętrznej jest większy.



Źródło: Osterloh M., Frey B.S. (2000)  
*Motivation, Knowledge Transfer, and Organizational Form* Organizational Science 11/5

## Czynniki motywacji – opinie twórców OSS



**Richard Stallman** (ideolog, 61 l.): podkreśla *moralne zobowiązania programisty* do dzielenia się z innymi. OSS jako słuszny wybór moralny.



**Eric Raymond** (pragmatyk, 56 l.): radość tworzenia, własny użytek, lepsza jakość.



**Linus Torvalds** (student, 44 l.) Stworzył Linuksa dla zabawy (a potem napisał o tym książkę).



**Donald Knuth** (profesor, 76 l.) stworzył TEXa na własny użytek, ponieważ nigdzie nie mógł kupić programu, którego potrzebował.

## Hars i Ou 2002 (H&O2002)

Badanie ankietowe twórców OSS. N=79 (389, 21,0%). Wiek: 50% mniej niż 30 lat, 85% mniej niż 40 lat. Ponad 70% posiada wykształcenie wyższe.

Czynnik	Hobbyści (42%)		Programiści (34%)		Kontrakt OSS (16%)	
	%	r*	%	r*	%	r*
<b>Radość tworzenia</b>	<b>82,0</b>	0,07	<b>92,6</b>	-0,30	61,5	0,22
Altruizm	24,2	0,19	11,1	0,06	7,7	-0,16
<i>Community identification</i>	36,4	0,12	18,5	-0,13	30,8	-0,31
Usługi komplementarne	6,1	0,36	3,7	0,49	<b>53,8</b>	<b>0,30</b>
<b>Kapitał ludzki</b>	<b>96,9</b>	<b>0,14</b>	<b>88,5</b>	<b>0,07</b>	<b>84,6</b>	<b>0,06</b>
Sygnalizacja/Autopromocja	33,3	0,31	29,6	0,20	<b>69,2</b>	<b>0,42</b>
Peer recognition	42,4	-0,02	48,1	-0,14	46,2	-0,18
Użytek własny	36,4	0,30	38,5	0,19	38,5	0,32

\* współczynnik korelacji z nakładem pracy (czas poświęcony wytwarzaniu OSS)

→ Czynniki wewnętrzne są bardziej istotne ale nie są powiązane z nakładem pracy

→ Radość tworzenia/samospełnienie ma wpływ na decyzję o partycypacji, ale nie na jej wielkość?

## Lakhani i Wolf 2005 (L&W2005)

Badanie ankietowe przeprowadzone w 2001–2002 r. 648 respondentów, 287 projektów wylosowanych z repozytorium sourceforge.net

97,5% – mężczyźni. Średnia wieku – 29,8 lat. Staż pracy 11,86 lat. „Straż” w społeczności OSS – 5,31 lat. Średnia liczba projektów – 4,95 / twórcę. Twórcy na kontrakcie OSS – 13%.

Czas poświęcony na OSS (Z): 14,3 godzin tydzień (odchylenie standardowe: 15,7), w tym opłacani 17,7 godzin tydzień, wolontariusze: 11,7 godzin tydzień.

Czynnik	%	wolontariusze	opłacani*	t
Programowanie jest intelektualnie stymulujące (intS)	<b>44,9</b>	<b>46,1</b>	<b>43,1</b>	ns
Lubię pracę w zespole OSS (team)	20,3	25,1	18,5	ns
Nauka programowania	<b>41,3</b>	<b>45,8</b>	33,2	3,56*
Potrzeba własna	<b>58,7</b>	—	—	—
Zwiększenie statusu zawodowego	17,5	13,9	22,8	3,01*
Kod źródłowy powinien być otwarty	33,1	34,8	30,6	ns
Potrzeba odwzajemnienia ponieważ używam OSS	28,6	29,6	26,9	ns
Nie lubię oprogramowania komercyjnego	11,3	11,5	11,1	ns
Zwiększenie reputacji w środowisku OSS (commR)	11,0	12,0	9,5	ns

\* Na kontrakcie OSS i tworzący OSS w czasie pracy (za zgodą pracodawcy)

$$\ln(Z) = 1,6 \cdot \text{intS} + 0,88 \cdot \text{paidS} + 0,84 \cdot \text{team} + 0,56 \cdot \text{commR} - 1,6 \cdot \text{altHr} - 0,6 \cdot \text{ITtraining}$$

## Sygnalizacja (Bitzer i inni 2009)

Bitzer i inni (2009) empirycznie zweryfikowali hipotezę o wpływie wytwarzania OSS na wielkość wynagrodzenia zaangażowanych na podstawie ankiety dołączonej do popularnego magazynu komputerowego. Wielkość próby 7115, z czego 1226 było zaangażowanych w tworzenie OSS (256 w roli kluczowego wytwórcy)

Hipoteza: tworzenie OSS w czasie wolnym skutkuje wyższymi zarobkami.

Oszacowano następujący wariant równania Mincera:

$$\ln P_i = \alpha + \beta Z_i + \theta D_u + \gamma W + \epsilon_i \quad (1)$$

gdzie:  $P$  – wynagrodzenie,  $Z$  – zaangażowanie w tworzenie OSS,  $D$  – zbiór zmiennych demograficznych (wiek, płeć, wykształcenie)  $W$  – zbiór zmiennych opisujących zatrudnienie (wielkość firmy, branża, stanowisko).

Jeżeli  $Z$  jest sygnałem dla nieobserwowanej produktywności, to musi być endogeniczne. Potwierdzeniem sygnalizacji jest nieistotność różnicy pomiędzy modelem KMNK oraz modelem z zmienną instrumentalną.

Oceniono łączny wpływ  $Z$  oraz wieku, wykształcenia i wielkości firmy. Jedyny statystycznie istotny współczynnik wystąpił w przypadku zmiennej interakcyjnej  $Z \times \text{braku-formalnego-wykształcenia}$ .

Zaangażowanie w wytwarzanie OSS **nie skutkuje premią wynagrodzenia** (60% ankietowanych wierzy, że jest inaczej) .

## Sygnalizacja (Il-Horn i inni 2004)

Badanie zrealizowane w latach 1999–2002 wśród 147 twórców *Apache Software Foundation* (panel) Średnia liczba przepracowanych godzin w roku 2122 w tym 206 na rzecz ASF (około 10%). Staż – 5,7 lat, zaangażowanie (LOC) – 24,7 (odchylenie standardowe 138,10)

Oszacowano następujący wariant równania Mincera (zastosowano estymator o efektach stałych):

$$\begin{aligned}\ln P_{it} = & \alpha_i + \beta_1 Z_{i,t-1} + \beta_2 Dev_{i,t-1} + \beta_3 Com_{i,t-1} + \beta_4 Pmc_{i,t-1} + \beta_5 Expr_{i,t-1} + \beta_6 Expr_{i,t-1}^2 \\ & + \beta_7 Ledu_{i,t-1} + \beta_8 Jswch_{i,t}^2 + \beta_9 Fpub_{i,t} + \beta_{10} Fswin_{i,t}^2 + \beta_{11} Stdnt_{i,t} \\ & + \beta_{12} Pdapc_{i,1}^2 + \beta_{13-15} Time_{i,t-1} + \xi_{i,t} \quad (i = 1, \dots, N, t = 1 \dots T)\end{aligned}$$

gdzie:  $P$  – wynagrodzenie,  $Z$  – zaangażowanie w wytwarzanie OSS mierzone łączną liczbą wierszy wytworzonych kodu (LOC);  $Dev-Pmc$  – formalny status odpowiednio: *Developer*, *Committer*, *PMC member*;  $Expr$  – staż (lata);  $Ledu$  – wykształcenie (lata ukończone);  $Stdnt$  – student;  $Pdapc$  – na kontrakcie z ASF;

Współczynnik  $\beta_1 = -0,0001$  ( $p = 0,07$ ) jest istotny statystycznie ale jego wielkość jest z praktycznego punktu widzenia równa zero. Alternatywny pomiar  $Z$  (liczba godzin) nie zmienił tego wyniku. Zaangażowanie w wytwarzanie OSS *per se nie skutkuje premią wynagrodzenia*.

Współczynnik przy zmiennych określających status formalny wynoszą odpowiednio:  $\beta_2 = -0,055$  ( $p = 0,20$ ),  $\beta_3 = 0,132$  ( $p = 0,04$ ),  $\beta_4 = 0,257$  ( $p = 0,01$ ), co oznacza że status *Developer* także **nie skutkuje premią wynagrodzenia**. Natomiast premia wynagrodzenia występuje w przypadku statusów *Committer* i *PMC member* (odpowiednio 14,11% i 29,32%).

$\beta_4 - \beta_3 = 0,125$ , ( $p < 0,001$ ) *PMC member* przeciętnie zarabia więcej 13,3% niż *Committer*.

## Motywacja a zaangażowanie (R2006)

Badanie zrealizowane w latach 1999–2002 wśród 288 twórców *Apache Software Foundation* (Roberts i inni 2006).

Czynniki motywacji: **Extrinsic** – liczba godzin/tydzień opłaconych za pracę na rzecz ASF; **Intrinsic, Status, UseValue** – zmienne ukryte mierzone za pomocą miar refleksyjnych (Likerta 1–5):

Skala czynnika UseValue ( $\alpha = 0,71$ ):

1 *I fixed a bug or problem that was causing me trouble*

2 *I added a feature I wanted or needed to use*

Skala czynnika Status ( $\alpha = 0,73$ ):

1 *It gives me the chance to attain a recognized qualification or skill*

2 *It gives me status at work*

3 *It increases my opportunity for a better job*

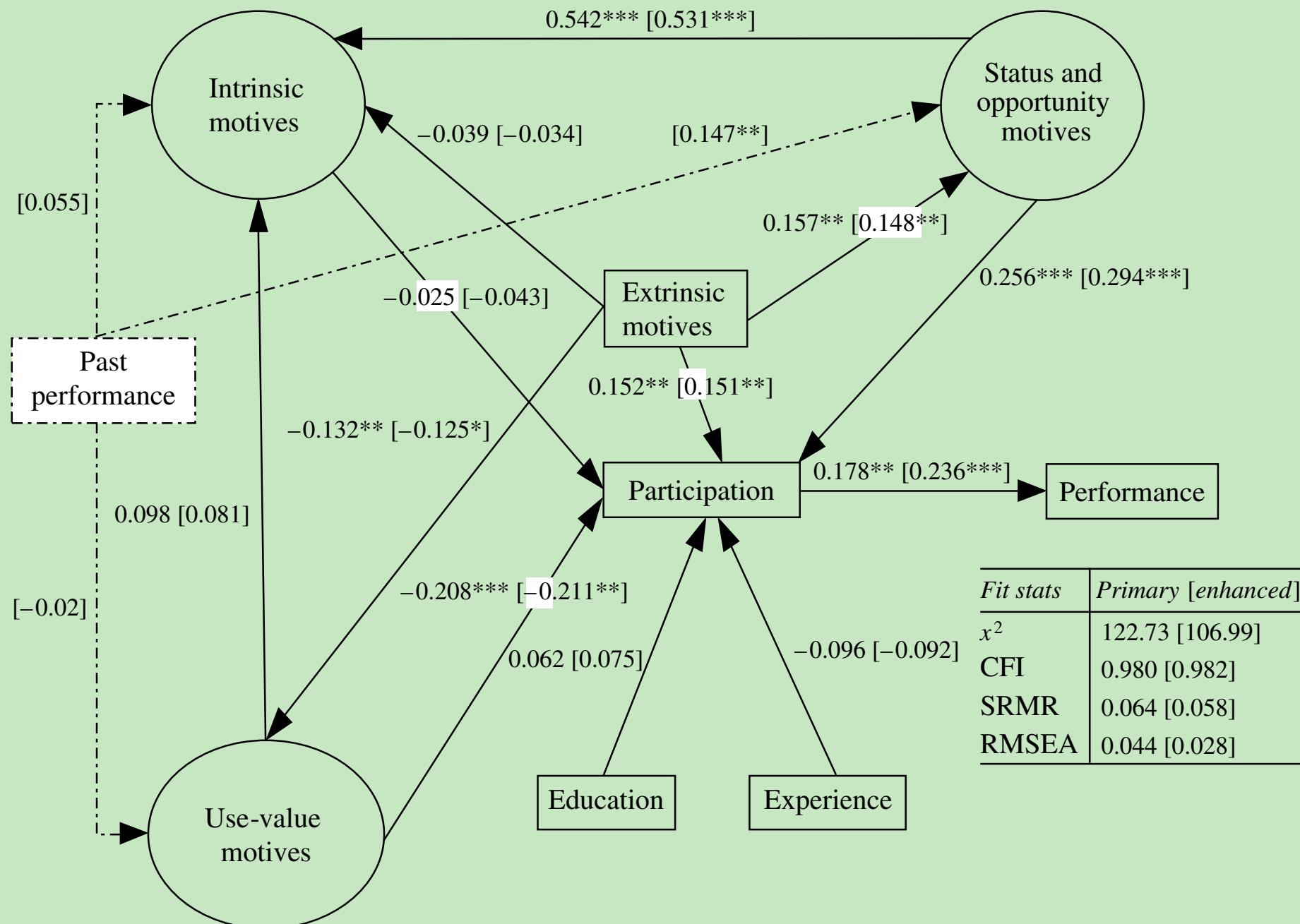
4 *It gives me status in ASF*

**Participation** – łączna liczba wierszy wytworzonego kodu (LOC);

**Performance** – zmiana formalnego statusu (*Developer, Commiter, PMC member*)

$$\begin{aligned} Participation_{t_2^3} &= \beta_0 + \beta_1 Intrinsic_{t_2^2} + \beta_2 Intrinsic_{t_2^2} \\ &+ \beta_3 UseValue_{t_2^2} + \beta_4 Status_{t_2^2} + \beta_5 Education_{t_2^2} + \beta_6 Experience_{t_2^2} + \epsilon_2 \end{aligned}$$

## Motywacja a zaangażowanie (oszacowany model ścieżkowy)



	Fit stats	Primary [enhanced]
$\chi^2$		122.73 [106.99]
CFI		0.980 [0.982]
SRMR		0.064 [0.058]
RMSEA		0.044 [0.028]

## Podsumowanie/zestawienie

Czynnik	H&O2002	D&S2008	L&W2005	R2006
Kapitał ludzki	●	○	●	
Radość tworzenia	●	×	●	×
Sygnalizacja	○		×	●
Czynniki ideologiczne		○	×	
Użytek własny		●	●	?

● – istotny; × – nieistotny; ○ – częściowo istotny

→ Radość tworzenia/samospełnienie – wyniki są sprzeczne; możliwe, że samospełnienie ma istotny wpływ na decyzję odnośnie partycypacji ale nie na jej wielkość.

→ Wpływ motywacji pieniężnej/wynagrodzenia na wynik/nakład pracy, weryfikacja efektu wypychania nie był podejmowany w literaturze.

→ W (licznej) grupie twórców produkujących OSS bez bezpośredniego wynagrodzenia za tą działalność (87% wg L&W2005) efekt wypychania może być istotny.

→ Premia wynagrodzenia związana z sygnalizacją dotyczy co najwyżej najlepszych.

## Ważniejsza literatura

- Bénabou R., Tirole J. (2003) *Intrinsic and Extrinsic Motivation* Review of Economic Studies 70, s. 489–520.
- Bitzer J., Geishecker I., Schroeder P. (2010) *Returns to Open Source Software Engagement: An Empirical Test of the Signaling Hypothesis* Oldenburg Working Papers.
- David P., Shapiro J. (2008) *Community-based production of open-source software*: What do we know about the developers who participate, Information Economics and Policy s. 364–398.
- Hars A., Ou S. (2002) *Working for Free? Motivations for Participating in Open-Source Projects* International Journal of Electronic Commerce 6/3, s. 25–39.
- Roberts J.A., Il-Horn H., Slaughter S. (2006) *Understanding the Motivations, Participation, and Performance of Open Source Software Developers: A Longitudinal Study of the Apache Projects* Manag Science, 52/7, s. 984–999.
- Il-Horn H., Roberts J.A., Slaughter S., Fielding R. (2004) *An empirical analysis of economic returns to open source participation* Carnegie-Mellon University.
- Krishna i Srinivasa (2011) *Analysis of Projects and Volunteer Participation in Large Scale Free and Open Source Software Ecosystems*, ACM SIGSOFT Software Eng. Notes, 36/2.
- Lakhani K.R., Wolf R.G. (2005) *Why Hackers Do What They Do: Understanding Motivation and Effort in Free/Open Source Software Projects w Perspectives on Free and Open Source Software*, MIT Press.
- Lerner J., Tirole J. (2002) *Some simple economics of open source*, Journal of Industrial Economics 50/2, s. 197–234.
- Ryan R.M., Deci E.L. (2000) *Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being*. American Psychologist, 55.
- Spence M. (1973) *Job market signaling* The Quarterly Journal of Economics, 87/3, s. 355–374.

Dziękuję za uwagę

