The Pirate Bay & Box Office Buccaneers

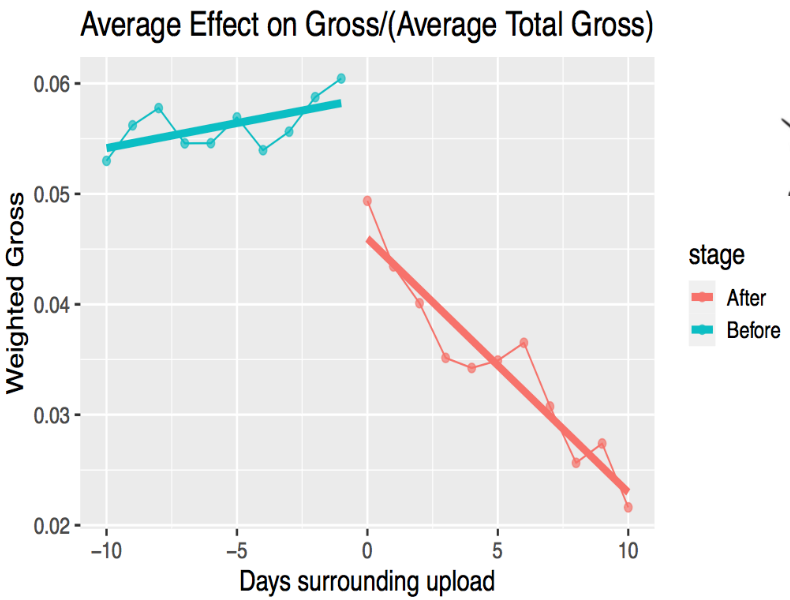

Figure 7: 10 days before and after the arrival of a good quality torrent

Klaus Ackermann
klaus.ackermann@monash.edu
Department of Econometrics and Business Statistics, SoDa Labs, Monash Business School
Monash University

Wendy Bradley
Department of Strategy
Cox School of Business
SMU

Jack Francis Cameron
Department of Econometrics and Business Statistics, SoDa Labs,
Monash Business School
Monash University

SMU | COX

SoDa Labs
Social [science] insights from alternative data.

MONASH University

The Challenge: Measuring digital piracy

Digital Piracy

- Digital piracy is the act of distributing or consuming content online without permission or payment
- A torrent is a link to a downloadable file through Peer-to-Peer technology; The Pirate Bay (TPB) is a catalogue of torrents
- Piracy has forced major change in media industries
- TPB is and has been the leading torrent catalogue on the Internet
- The web archive has an archive of all their files + scripts to update it to the most recent ones

To what extent does piracy displace legitimate sales, and do all firms suffer equally ?

Literature

- Revenue displacement in movies: Danaher and Waldfogel (2012), Smith and Telling 2012, Danaher and Smith (2014) e.g Megaupload
- Revenue increase in movies: Lu et al. (2019)
- Mixed results in music: Piracy is/was widespread with music, some studies find a displacement effect (e.g Hui and Png (2003) or Peitz and Waelbroeck(2004) others do not (e.g., Aguiar and Martens (2016))
- Digital goods vs physical goods: Qian (2014) investigates the effect of counterfeit products by quality tier in the fashion industry.

Identification strategy exploits the random arrival of a good quality copy

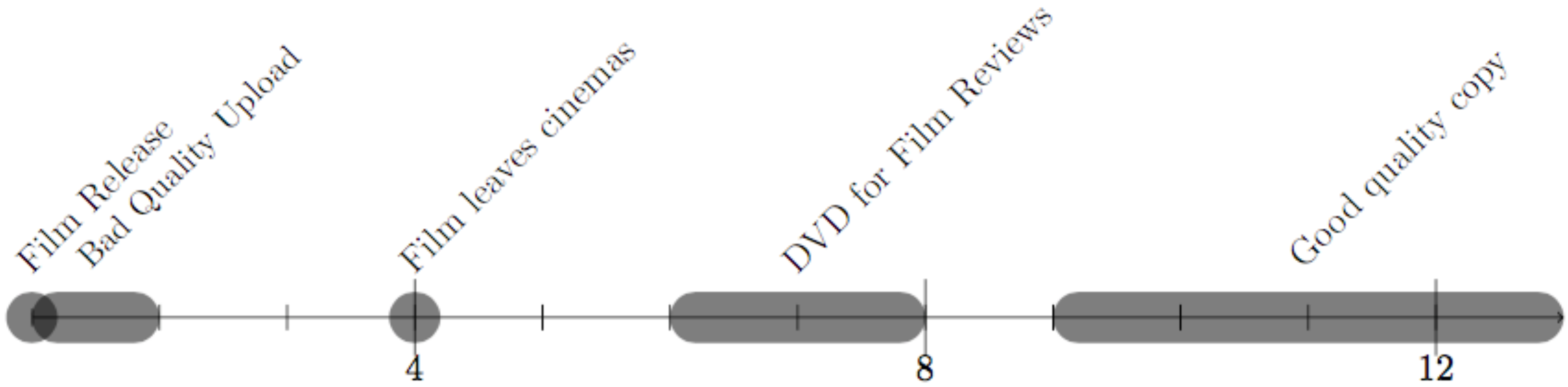

Figure 1: Theoretical timeline for the expected upload of a torrent

Table 1: Average days to upload by torrent quality

	Good Quality	Bad Quality
Average Days	29	17
Median Days	15	9

Identification strategy exploits the random arrival of a good quality copy

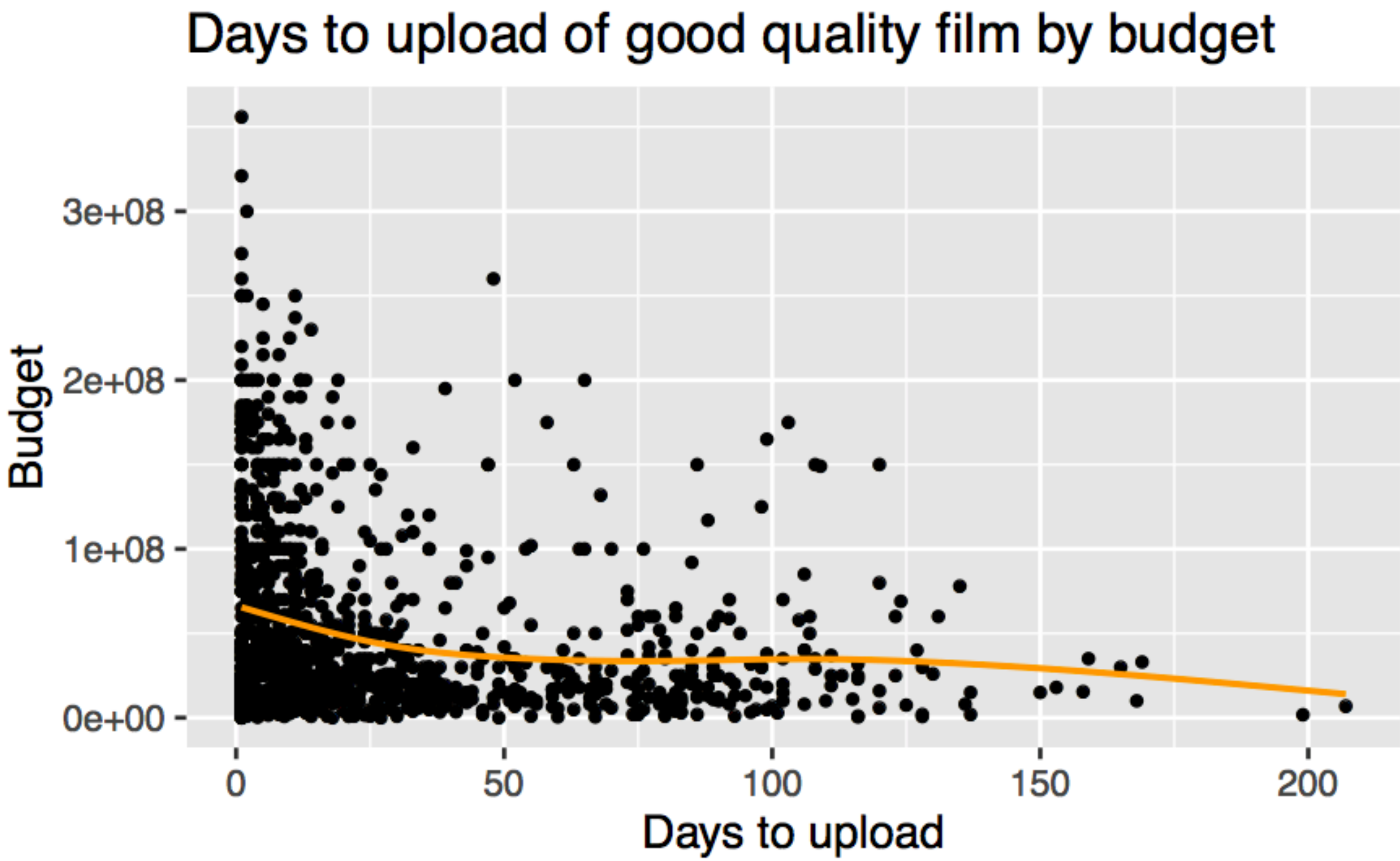

Figure 2: Average days to good quality uploads by budget

Identification strategy exploits the random arrival of a good quality copy

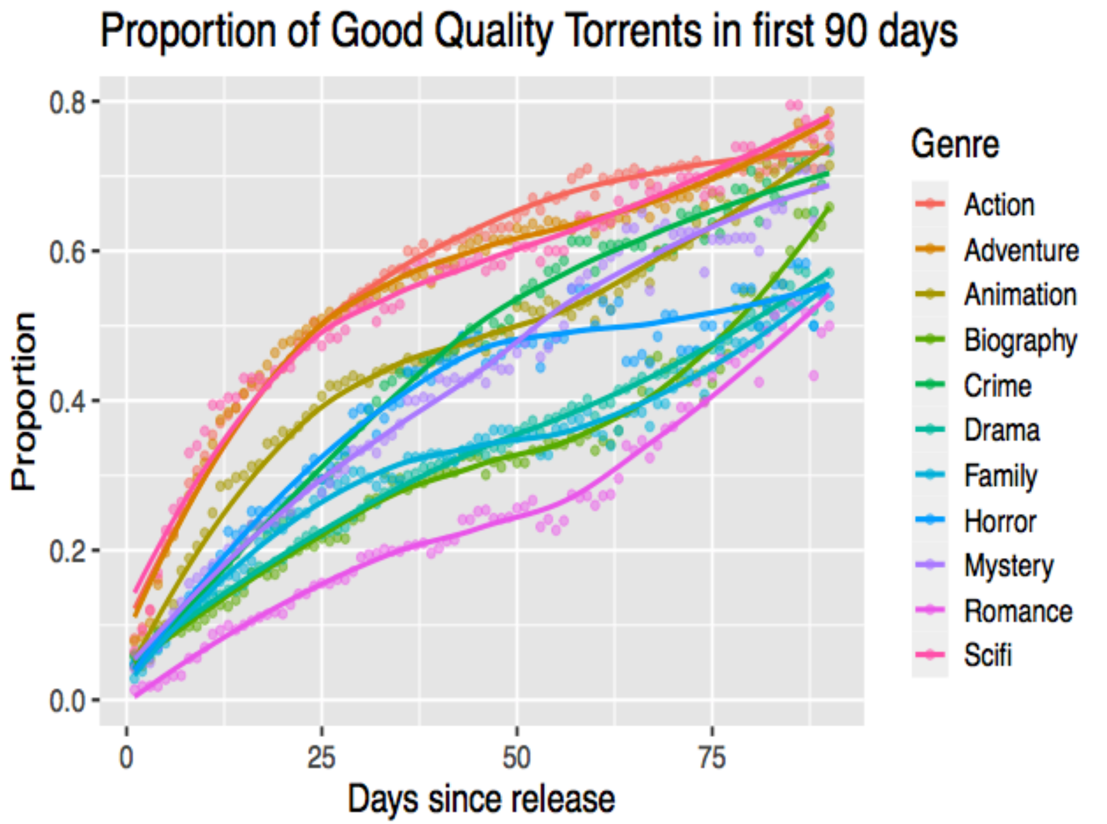

Figure 3: Average proportion of good torrents over time by genre

Splitting the data, pre and post of the arrivals of a good quality copy

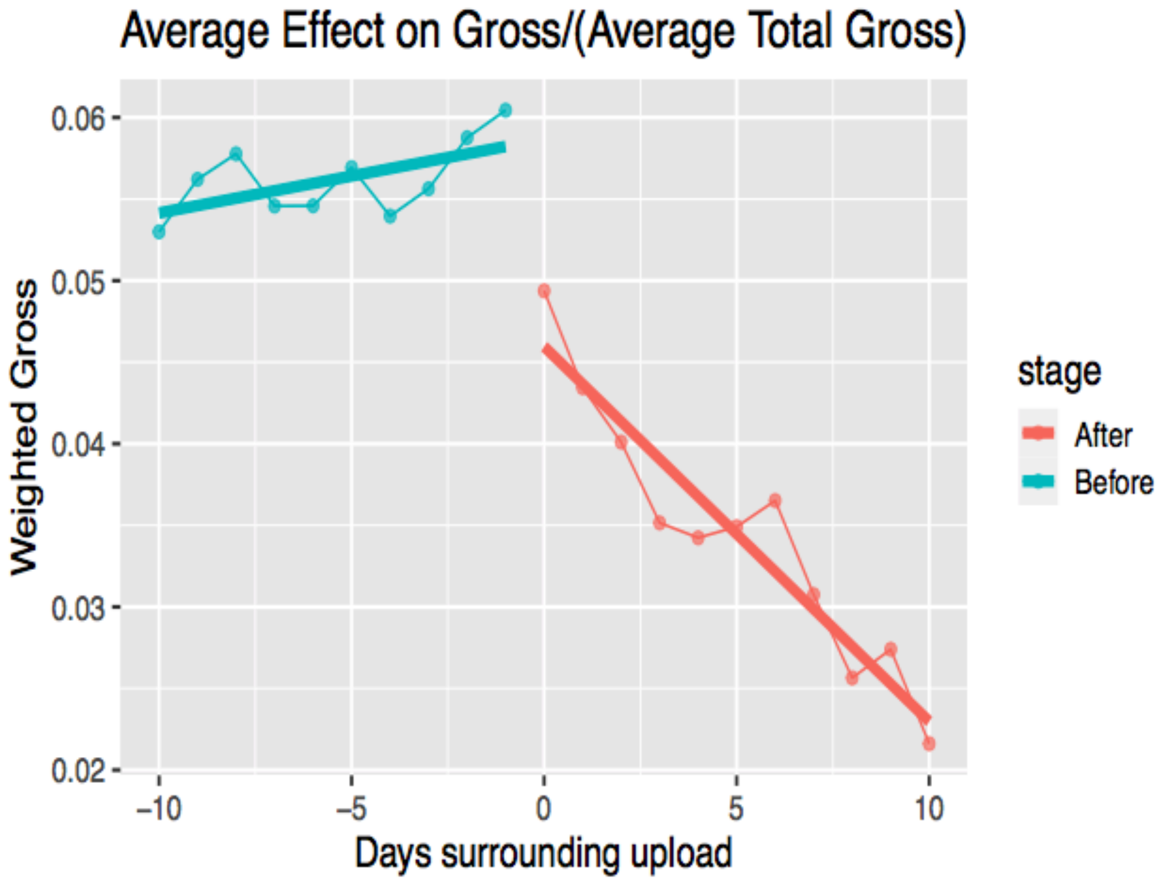

Figure 7: 10 days before and after the arrival of a good quality torrent

Estimation: Across 2,145 movies from 2004-2018 (release year)

$$InRevenue_{ijt} = \beta_1 Bad\_Quality_{ijt} + \beta_2 Good\_Quality_{ijt} + X_{ijt} + \mu_i + \Upsilon_j + \phi_t + \varepsilon_{jt} \quad (1)$$

	(1)
VARIABLES	All Movies
bad_quality_dummy	0.0320**
	(0.0153)
good_quality_dummy	-0.161***
	(0.0160)
Observations	110,792
R^2	0.869
Movie FE	Yes
Days since release FE	Yes
Year-Week FE	Yes
Number of movies	2145

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Estimation: Across 2,145 movies from 2004-2018 (release year)

Table 3: Budget subset

VARIABLES	(1) Movies <P25	(2) Movies IQ	(3) Movies >P75
bad_quality_dummy	-0.344***	0.027	0.149***
	(0.040)	(0.021)	(0.027)
good_quality_dummy	0.227***	-0.124***	-0.048**
	(0.044)	(0.022)	(0.022)
Observations	28,775	54,736	27,281
R^2	0.845	0.870	0.926
Movie FE	Yes	Yes	Yes
Days since release FE	Yes	Yes	Yes
Year-Week FE	Yes	Yes	Yes
Number of movies	712	1026	407

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Estimation: Heterogeneity by genre

Table 4: Genre subset 1

VARIABLES	(1) Action	(2) Adventure	(3) Animation	(4) Biography	(5) Comedy
bad_quality_dummy	0.063**	0.092***	-0.116**	0.493***	-0.106***
	(0.027)	(0.030)	(0.050)	(0.070)	(0.022)
good_quality_dummy	-0.050**	-0.046*	-0.005	-0.646***	-0.210***
	(0.021)	(0.026)	(0.054)	(0.078)	(0.024)
Observations	32,842	26,300	9,309	12,051	41,615
R^2	0.927	0.915	0.921	0.855	0.882
Movie FE	Yes	Yes	Yes	Yes	Yes
Days since release FE	Yes	Yes	Yes	Yes	Yes
Year-Week FE	Yes	Yes	Yes	Yes	Yes
Number of movies	589	412	143	220	770

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Negative impact on genres with focus on story

Estimation: Heterogeneity by genre

Table 5: Genre subset 2

VARIABLES	(1) Crime	(2) Drama	(3) Family	(4) Fantasy	(5) Horror
bad_quality_dummy	-0.020	0.014	0.081*	0.304***	-0.143***
	(0.041)	(0.025)	(0.046)	(0.060)	(0.035)
good_quality_dummy	0.018	-0.158***	-0.186***	-0.006	0.344***
	(0.039)	(0.028)	(0.061)	(0.056)	(0.037)
Observations	17,333	55,535	7,433	9,213	9,510
R^2	0.904	0.830	0.941	0.929	0.965
Movie FE	Yes	Yes	Yes	Yes	Yes
Days since release FE	Yes	Yes	Yes	Yes	Yes
Year-Week FE	Yes	Yes	Yes	Yes	Yes
Number of movies	377	1142	133	175	225

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Postive impact on genres with focus on experience

What is the effect on a movie's Box office revenue of the arrival of a good quality copy of a competing movie in the same genre released the same date ?

Table 8: Cohort Genre subset 1

VARIABLES	(1) Action	(2) Adventure	(3) Animation	(4) Biography	(5) Comedy
bad_quality_dummy	0.063**	0.093***	-0.109**	0.497***	-0.109***
	(0.027)	(0.030)	(0.050)	(0.070)	(0.022)
good_quality_dummy	0.028	-0.146***	-0.141**	-0.509***	-0.261***
	(0.027)	(0.033)	(0.059)	(0.090)	(0.029)
cohort_dummy	-0.135***	0.170***	0.348***	-0.209***	0.081***
	(0.027)	(0.033)	(0.073)	(0.070)	(0.028)
Observations	32,842	26,300	9,309	12,051	41,615
R^2	0.927	0.915	0.922	0.855	0.882
Movie FE	Yes	Yes	Yes	Yes	Yes
Days since release FE	Yes	Yes	Yes	Yes	Yes
Year-Week FE	Yes	Yes	Yes	Yes	Yes
Number of movies	589	412	143	220	770

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

- Are there implications for at-home release vs in theatres ?
- There maybe lost revenue. Is lost revenue retained by consumer or transferred to other film studios ?

Conclusion

Recap

- The timing of a good quality release as a random positive shock
- Strong evidence of an average negative impact, e.g. a loss of a fifth of the daily revenue
- Different genres have differentiated effects: Horror, Mystery or Scifi may be enhanced through a cinematic experience; story based movies suffer, e.g. comedy and drama
- Experiential film genres, a good quality movie is not an exact substitute; instead the appearance of a pirated movie acts as discovery version, thereby increasing the demand for the cinematic experience
- Low-budget movies benefit from the discovery effect

Implications

- Film studios could focus their enforcement of piracy on genres that are most affected
- Release strategies could take into account an optimal number of genres studios release (product mix) given piracy, and the timing of releases, given competing films
- Returns to low-budget, story-based films may be highest when released on streaming services
- Importance of quality tiers in piracy -- may also apply to video games, cooperative versus non-cooperative video games