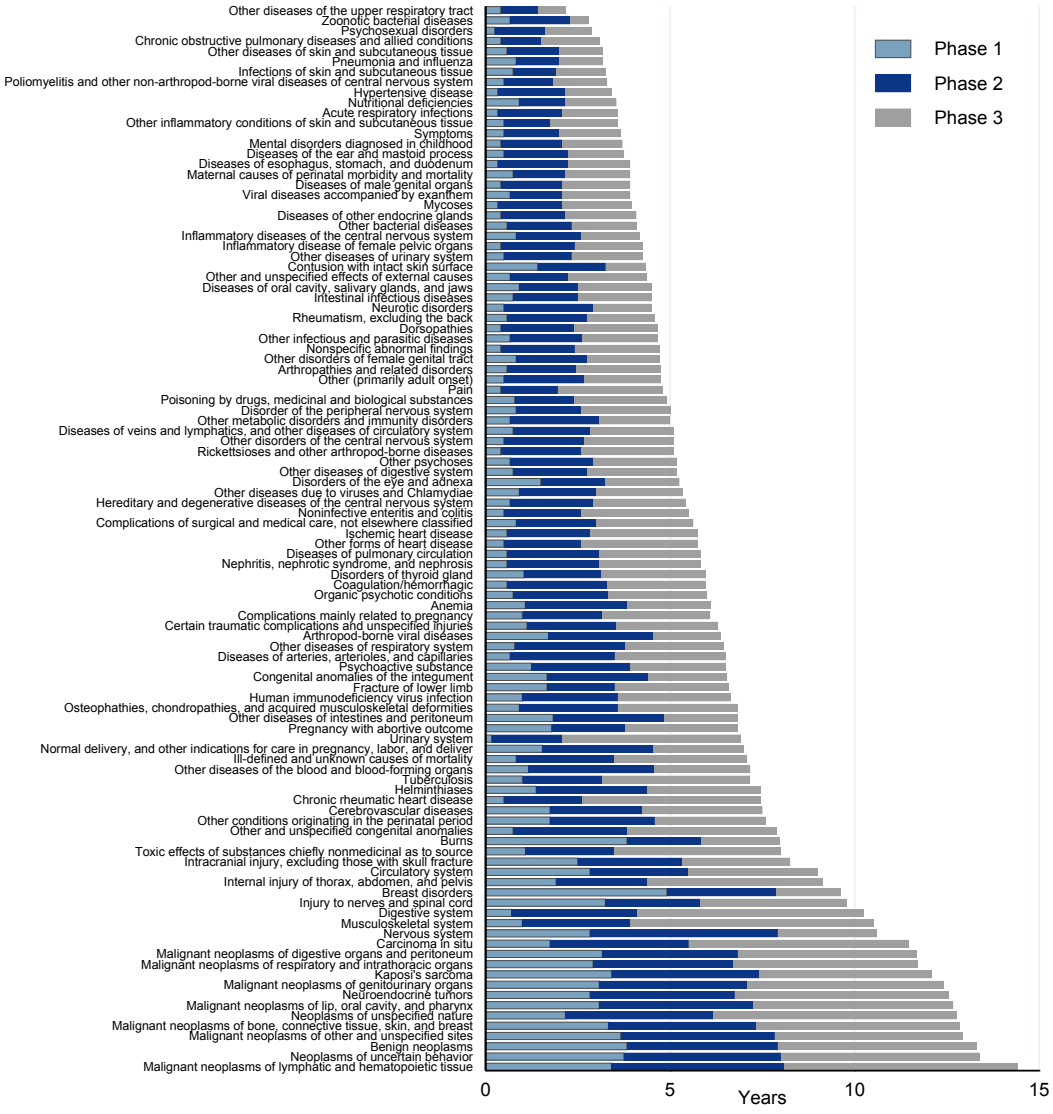
For Online Publication

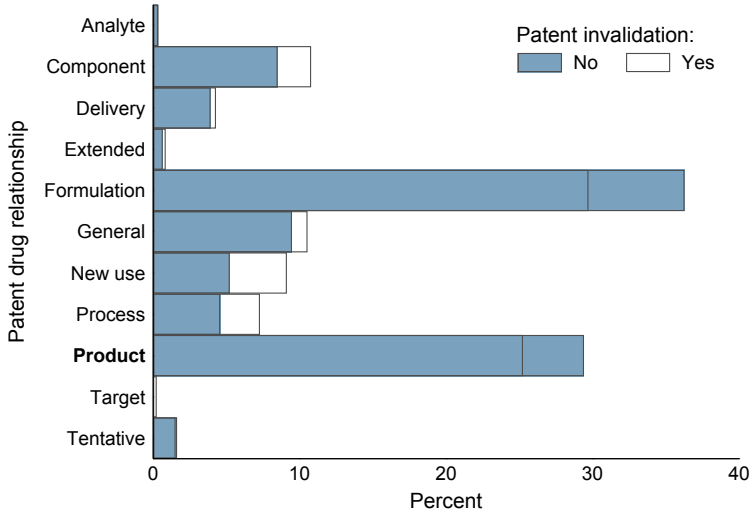
A Appendix: Figures

Figure A-1: Distribution of clinical trials lengths



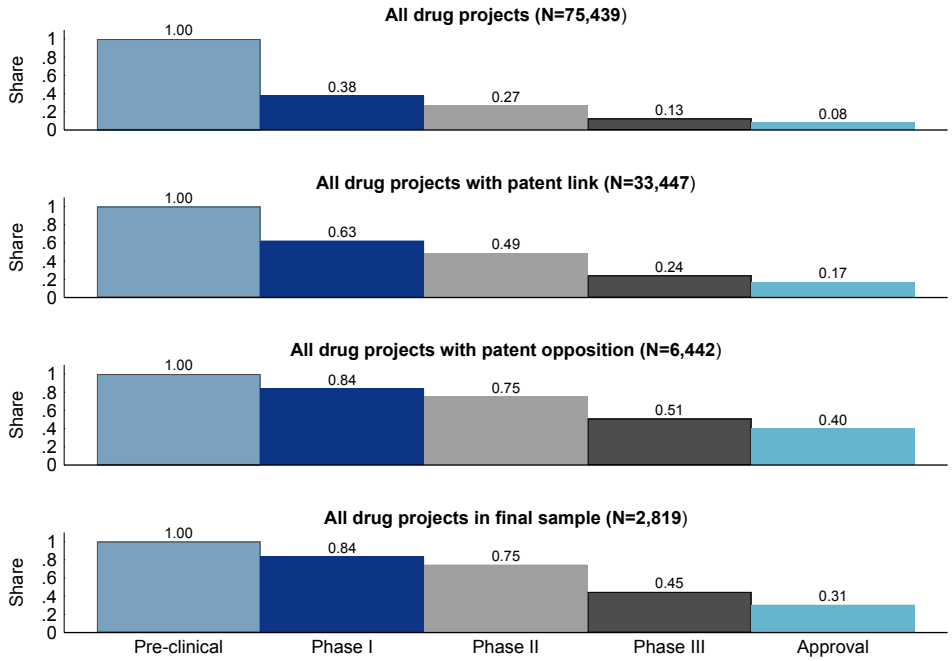
Notes: This figure shows the distribution of lengths of clinical trials by phase at aggregated indication level. Observations are at the aggregated indication level. Lengths calculated as the medians of the full sample of clinical trials in Cortellis. Only aggregated indications that are part of the final sample included.

Figure A-2: Distribution of patent-drug relationships by opposition outcome



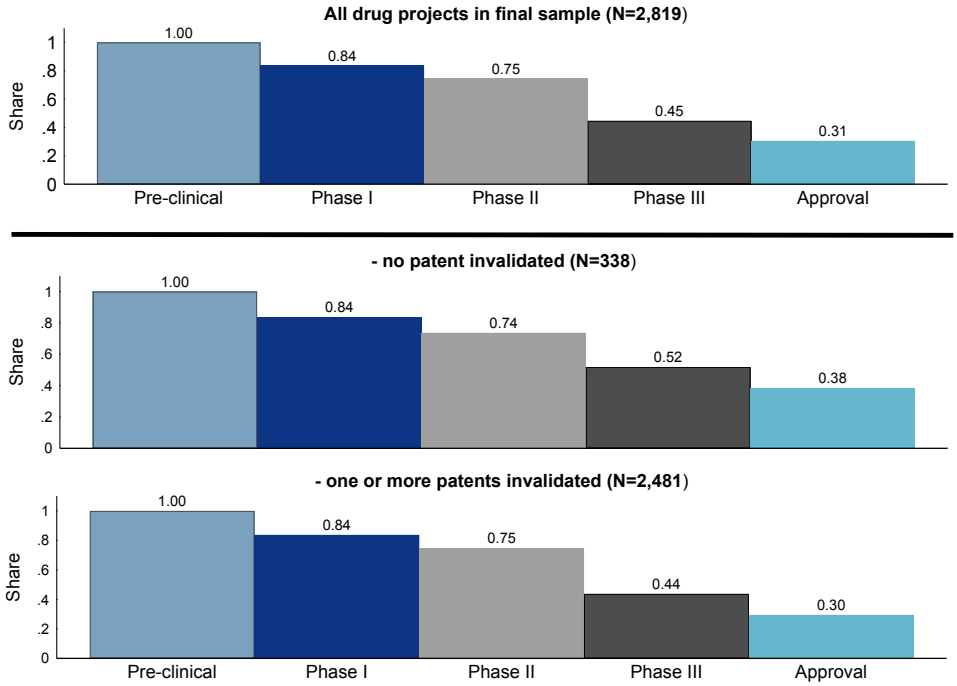
Notes: This figure shows the distribution of patent-drug relationships by opposition outcome. Patent-drug relationship as stated in Cortellis. Observations are at the drug-patent level. Bold relationships indicate "primary patent status". The null-hypothesis that the two distributions are equal can be rejected with a p-value of 0.124 ($\text{Chi}^2 = 15.218$).

Figure A-3: Project advancement at the drug-indication level by sample



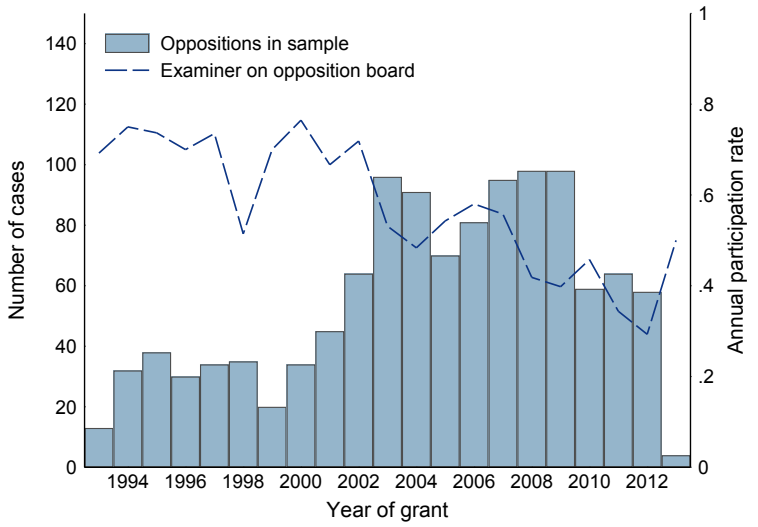
Notes: Observations are at the drug-indication level. "All drugs" refers to all non-pending drug projects with development information in the Cortellis database. "All drugs with patent link" refers to all drug projects where a link to a patent family was identified. "All drugs with patent opposition" refers to all drug projects with at least one patent challenged in opposition proceedings. "All drugs in final sample" refers to all drug projects that are part of our final sample of analysis. Higher approval rates for drug projects with opposed patents compared to our final sample are due to the fact that companies adapt their innovation efforts only conditional on opposition outcome.

Figure A-4: Project advancement at the drug-indication level by opposition outcome



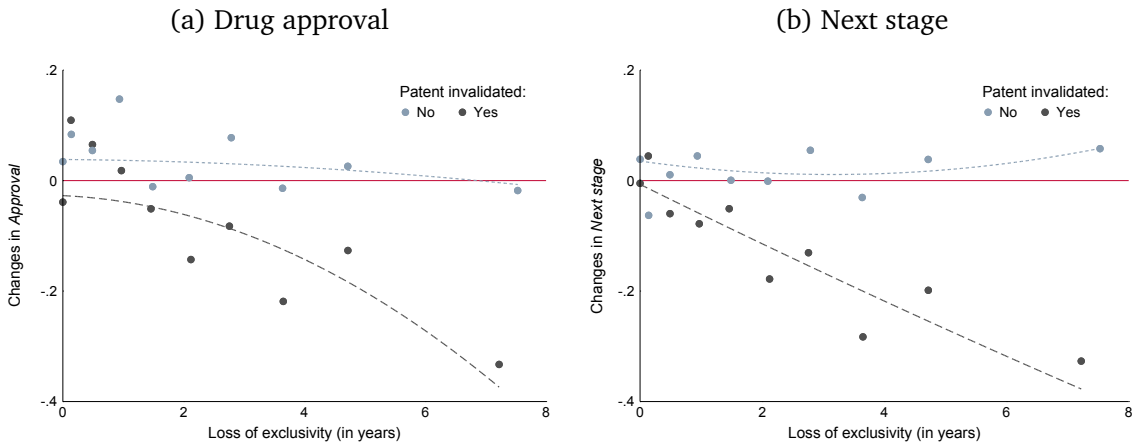
Notes: Observations are at the drug-indication level. "All drug projects in the final sample" refers to all drug projects that are part of our final sample of analysis.

Figure A-5: Annual number of opposed patents and rate of examiner participation



Notes: This graph includes all opposition proceedings (at the patent level) in the regression sample. Based on the same sample, the figure also shows the annual rate of examiner participation in opposition proceedings.

Figure A-6: Effects of market exclusivity loss (in years) on drug development



Notes: These figures plot market exclusivity loss (in years) on the change in drug development outcome (approval or continuation of drug development). To construct this binned scatterplot, we predict the residuals from the estimated regressions in column 4 and column 8 in Table 3, and plot these by the instrumented variable *Loss of exclusivity (in years)*.

B Appendix: Tables

Table B-1: Groups of control variables

Group name	**Variables in group**
Patent characteristics*	Dummy for PCT application Dummy for accelerated examination Dummy for examination in Munich branch Dummies for publication language: ENG – FRA – GER Size of docdb family Number of IPC classes Number of claims Number of inventors log(1 + Number of patent references) log(1 + Number of patent 3yrs forward citations)
Patent examination characteristics*	Duration of examination Duration of wait until examination
Patent age effects*	Dummies for age (in yrs)
Year effects*	Dummies for patent grant year Dummies for opposition outcome year
Drug characteristics†	Number of indications Dummy for prior approval Number of patent families Dummy for orphan drug status Dummy for pediatric use status
Development characteristics†	Time since discovery (in yrs) Dummies for current development stage
Disease effects†	Dummies for aggregated ICD-9 levels (19)
Technology effects†	Dummies for technology class (34) Dummy for biologics
Originator characteristics‡	Dummy for originator European Dummy for originator corporation Dummies for originator portfolio size: small – large
Opponent characteristics‡	Number of opponents Dummy for opponent European Dummy for opponent corporation

Notes: Variables with * capture heterogeneity in patent value and describe the examination process at the EPO. They (primarily) have explanatory value for the first stage (patent) regressions. Variables with † capture characteristics of the underlying drug (project) and control for the timing of the opposition outcome relative to the development process. They (primarily) have explanatory value for the second stage (drug project) regressions. Variables with ‡ capture characteristics of the drug originator and the opponent. They have explanatory value for the first and second stage regressions.

Table B-2: Summary statistics

Drug-indication level	All drug projects in final sample				
	Mean	Median	Std. Dev	Min	Max
Patent characteristics					
PCT application (d)	0.75	1.00	0.43	0.00	1.00
Publication language GER (d)	0.05	0.00	0.23	0.00	1.00
Publication language ENG (d)	0.92	1.00	0.26	0.00	1.00
DOCDB family size	29.91	25.00	21.39	1.00	187.00
Number of IPC classes	4.50	3.00	3.93	1.00	32.00
Number of claims	21.41	16.00	19.03	1.00	249.00
Number of inventors	3.45	3.00	2.43	1.00	22.00
Backward citations	1.40	1.39	0.66	0.00	3.81
Forward citations within 3 yrs	5.64	4.00	6.76	0.00	113.00
Primary patent (d)	0.31	0.00	0.46	0.00	1.00
Patent examination characteristics					
Accelerated examination (d)	0.24	0.00	0.43	0.00	1.00
Examination in Munich office (d)	0.79	1.00	0.41	0.00	1.00
Examination time (in yrs)	4.87	4.57	2.24	0.71	15.17
Pre-examination time (in yrs)	2.67	2.50	1.63	0.00	18.59
Patent age effects					
Age of patent (in yrs)	10.98	11.00	3.07	4.00	22.00
Year effects					
Patent grant (yr)	2002.99	2004.00	5.43	1992.00	2013.00
Opposition outcome (yr)	2006.39	2008.00	4.99	1994.00	2015.00
Drug characteristics					
# Indications	10.54	9.00	11.97	1.00	62.00
Prior approval	0.41	1.00	0.49	0.00	1.00
# Patent families	27.14	21.00	53.60	1.00	1037.00
Orphan Drug Designation (d)	0.48	1.00	0.50	0.00	1.00
Paediatric Investigation Plan (d)	0.06	0.00	0.24	0.00	1.00
Biologic (d)	0.47	0.00	0.50	0.00	1.00
Development characteristics					
Time betw/ drug discovery and opposition outcome	12.07	13.00	6.29	0.00	61.00
Development stage at opposition outcome:					
Pre-clinical (d)	0.52	1.00	0.50	0.00	1.00
Phase I (d)	0.07	0.00	0.26	0.00	1.00
Phase II (d)	0.23	0.00	0.42	0.00	1.00
Phase III (d)	0.18	0.00	0.39	0.00	1.00
Originator characteristics					
Originator European (d)	0.33	0.00	0.47	0.00	1.00
Originator corporation (d)	0.97	1.00	0.18	0.00	1.00
Originator small (d)	0.41	0.00	0.49	0.00	1.00
Originator large (d)	0.59	1.00	0.49	0.00	1.00
Opponent characteristics					
# Opponents	2.02	1.00	2.54	1.00	18.00
Opponent European (d)	0.71	1.00	0.46	0.00	1.00
Opponent Corporation (d)	0.94	1.00	0.23	0.00	1.00

Notes: Observations are weighted by the inverse frequency at drug-indication-patent level. "All drug projects in final sample" refers to all drug projects that are part of our final sample of analysis. Patent families follow the INPADOC patent family definition. Dummies for aggregated ICD-9 levels omitted.

Table B-3: Impact of patent invalidation on drug development by timing as well as portfolio size

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep var	Approval				Next Stage			
Sample	Small		Large		Small		Large	
	Pre-clinic/Phase I	Phase II/III	Pre-clinic/Phase I	Phase II/III	Pre-clinic/Phase I	Phase II/III	Pre-clinic/Phase I	Phase II/III
Dep var mean	0.310	0.289	0.603	0.334	0.352	0.266	0.671	0.318
Invalidated	−0.376**	0.324	0.028	0.184	−0.259	0.314	0.220	0.016
	(0.170)	(0.372)	(0.385)	(0.196)	(0.177)	(0.458)	(0.335)	(0.204)
Invalidated × Potential LoE (in yrs)	−0.007	−0.090	−0.077**	−0.041	−0.044	−0.071	−0.091***	−0.031
	(0.038)	(0.068)	(0.036)	(0.032)	(0.036)	(0.066)	(0.029)	(0.031)
Potential LoE (in yrs)	0.042	0.047	0.095***	0.030	0.033	0.018	0.058**	0.006
	(0.027)	(0.057)	(0.029)	(0.025)	(0.026)	(0.053)	(0.023)	(0.023)
Drug characteristics	Yes***	Yes	Yes**	Yes	Yes***	Yes**	Yes***	Yes
Development characteristics	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Patent characteristics	Yes***	Yes***	Yes	Yes***	Yes**	Yes***	Yes**	Yes***
Technology effects	Yes***	Yes**	Yes***	Yes	Yes***	Yes**	Yes***	Yes
Disease effects	Yes***	Yes***	Yes*	Yes***	Yes***	Yes***	Yes***	Yes***
Examination characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age effects	Yes***	Yes**	Yes***	Yes***	Yes*	Yes**	Yes	Yes
Originator characteristics	Yes**	Yes*	Yes	Yes	Yes**	Yes*	Yes	Yes
Opponent characteristics	Yes***	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year effects	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Underidentification test	20.2	9.9	9.8	15.5	20.2	9.9	9.8	15.5
Weak identification test	18.8	6.9	5.9	10.5	18.8	6.9	5.9	10.5
Observations	1,769	695	2,242	1,268	1,769	695	2,242	1,268
Observations (weighted)	739	418	862	722	739	418	862	722

Notes: Columns (1) to (8) show 2SLS regressions for the impact of invalidation on drug development on subsamples defined by the development stage of the drug project at the time of opposition outcome and the size of the originator. The samples used in columns (1), (3), (5) and (7) include drug projects only if the patent opposition outcome occurred during the preclinical phase or in Phase I clinical trials. Likewise, the samples used in columns (2), (4), (6) and (8) include drug projects only if the patent opposition outcome occurred during Phase II or III clinical trials. The samples used in columns (1), (2), (5) and (6) include drug projects only if the drug originates from small entities. Likewise, the samples used in columns (3), (4), (7) and (8) include drug projects only if the drug originates from large entities. In each 2SLS regression the *Invalidated* dummy is instrumented with the corresponding probability predicted by a probit regression on the *Examiner participation* dummy and all other exogenous variables. The underidentification and weak identification tests are the heteroskedasticity-robust Kleibergen and Paap (2006) rk LM and Wald F statistics, respectively, as reported by Stata's ivreg2 command (Baum et al., 2010). A comprehensive list of the control variables contained in the indicated groups can be found in Table 3 in the Appendix. Standard errors are two-way clustered by drug and indication. Observations are weighted by the inverse frequency at drug-indication-patent level. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Table B-4: Impact of patent invalidation on drug development (binary distinction of "Potential LoE")

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Estimation method	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Dep var	Approval						Next Stage					
Sample	All		Potential LoE = 0		Potential LoE > 0		All		Potential LoE = 0		Potential LoE > 0	
Dep var mean	0.308	0.308	0.248	0.248	0.431	0.431	0.509	0.509	0.520	0.520	0.510	0.510
Invalidated	0.014	0.072	0.003	−0.010	−0.015	−0.488***	0.033	0.058	−0.005	−0.115	0.026	−0.330**
	(0.032)	(0.124)	(0.033)	(0.103)	(0.039)	(0.163)	(0.036)	(0.132)	(0.034)	(0.110)	(0.041)	(0.158)
Invalidated × Potential LoE > 0	−0.095*	−0.401***					−0.057	−0.388***				
	(0.048)	(0.102)					(0.052)	(0.107)				
Potential LoE > 0	0.213***	0.457***					0.025	0.288***				
	(0.046)	(0.088)					(0.048)	(0.089)				
Drug characteristics	Yes***	Yes***	Yes*	Yes*	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Development characteristics	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Patent characteristics	Yes**	Yes***	Yes*	Yes*	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes**
Technology effects	Yes***	Yes***	Yes	Yes*	Yes***	Yes***	Yes	Yes**	Yes	Yes	Yes***	Yes***
Disease effects	Yes**	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes**	Yes***
Examination characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes*	Yes
Age effects	Yes	Yes**	Yes***	Yes***	Yes***	Yes**	Yes	Yes	Yes*	Yes**	Yes***	Yes***
Originator characteristics	Yes**	Yes**	Yes*	Yes*	Yes	Yes	Yes***	Yes***	Yes**	Yes**	Yes	Yes
Opponent characteristics	Yes	Yes	Yes*	Yes*	Yes***	Yes	Yes	Yes	Yes	Yes	Yes***	Yes*
Year effects	Yes***	Yes***	Yes***	Yes**	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Underidentification test		34.1		36.5		36.6		34.1		36.5		36.6
Weak identification test		24.6		68.6		42.1		24.6		68.6		42.1
Observations	5,974	5,974	4,135	4,087	1,839	1,806	5,974	5,974	4,135	4,087	1,839	1,806
Observations (weighted)	2,741	2,741	1,762	1,762	1,254	1,254	2,741	2,741	1,762	1,762	1,254	1,254

Notes: Columns (1) and (2) as well as (5) and (6) provide a comparison between the OLS and the 2SLS regressions for the impact of invalidation on drug development (approval or continuation to next stage) when accounting for the actual loss of exclusivity. Columns (3) and (4) as well (7) and (8) include an interaction term capturing the loss of exclusivity in linear form. In each 2SLS regression the *Invalidated* dummy is instrumented with the corresponding probability predicted by a probit regression on the *Examiner participation* dummy and all other exogenous variables. The underidentification and weak identification tests are the heteroskedasticity-robust Kleibergen and Paap (2006) rk LM and Wald F statistics, respectively, as reported by Stata's ivreg2 command (Baum et al., 2010). A comprehensive list of the control variables contained in the indicated groups can be found in Table 3 in the Appendix. Standard errors are two-way clustered by drug and indication. Observations are weighted by the inverse frequency at drug-indication-patent-level. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Table B-5: Impact of patent invalidation on drug development (alternative operationalization of "Invalidated")

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Estimation method	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Dep var	Approval				Next Stage			
Dep var mean	0.308	0.308	0.308	0.308	0.308	0.509	0.509	0.509
Invalidated	−0.026	−0.039	0.008	0.059	0.007	−0.026	0.032	0.058
	(0.023)	(0.118)	(0.024)	(0.120)	(0.025)	(0.112)	(0.027)	(0.116)
Invalidated × Potential LoE (in yrs)			−0.022***	−0.050***			−0.016**	−0.041***
			(0.009)	(0.016)			(0.008)	(0.014)
Potential LoE (in yrs)			0.034***	0.051***			0.001	0.017*
			(0.007)	(0.011)			(0.006)	(0.009)
Drug characteristics	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Development characteristics	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Patent characteristics	Yes**	Yes**	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Technology effects	Yes***	Yes***	Yes***	Yes***	Yes	Yes**	Yes	Yes***
Age effects	Yes	Yes*	Yes	Yes**	Yes	Yes	Yes	Yes
Examination characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Originator characteristics	Yes*	Yes**	Yes**	Yes**	Yes***	Yes***	Yes***	Yes***
Opponent characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Disease effects	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Year effects	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Underidentification test		25.8		25.7		25.8		25.7
Weak identification test		37.3		18.3		37.3		18.3
Observations	5,974	5,974	5,974	5,974	5,974	5,974	5,974	5,974
Observations (weighted)	2,741	2,741	2,741	2,741	2,741	2,741	2,741	2,741

Notes: Columns (1) and (2) as well as (5) and (6) provide a comparison between the OLS and the 2SLS regressions for the impact of invalidation on drug development (approval or continuation to next stage) when accounting for the actual loss of exclusivity. Columns (3) and (4) as well (7) and (8) include an interaction term capturing the loss of exclusivity in linear form. In each 2SLS regression the *Invalidated* dummy is instrumented with the corresponding probability predicted by a probit regression on the *Examiner participation* dummy and all other exogenous variables. The underidentification and weak identification tests are the heteroskedasticity-robust Kleibergen and Paap (2006) rk LM and Wald F statistics, respectively, as reported by Stata's ivreg2 command (Baum et al., 2010). A comprehensive list of the control variables contained in the indicated groups can be found in Table 3 in the Appendix. Standard errors are two-way clustered by drug and indication. Observations are weighted by the inverse frequency at drug-indication-patent level. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Table B-6: Impact of patent invalidation on drug development – first indication only

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Estimation method	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Dep var	Approval				Next Stage			
Dep var mean	0.422	0.422	0.422	0.422	0.563	0.563	0.563	0.563
Invalidated	−0.058	−0.184	−0.022	0.017	−0.001	−0.034	0.034	0.197
	(0.039)	(0.144)	(0.048)	(0.170)	(0.040)	(0.141)	(0.045)	(0.171)
Invalidated × Potential LoE (in yrs)			−0.012	−0.059**			−0.016	−0.079***
			(0.013)	(0.027)			(0.012)	(0.027)
Potential LoE (in yrs)			0.025**	0.061***			0.001	0.050**
			(0.012)	(0.021)			(0.012)	(0.021)
Drug characteristics	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Development characteristics	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Patent characteristics	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Technology effects	Yes***	Yes***	Yes***	Yes***	Yes	Yes***	Yes	Yes***
Disease effects	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Examination characteristics	Yes	Yes	Yes	Yes	Yes**	Yes**	Yes**	Yes**
Age effects	Yes	Yes	Yes	Yes*	Yes	Yes**	Yes	Yes***
Originator characteristics	Yes**	Yes**	Yes**	Yes**	Yes***	Yes***	Yes***	Yes***
Opponent characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year effects	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Underidentification test		38.0		34.8		38.0		34.8
Weak identification test		45.1		21.7		45.1		21.7
Observations	1,583	1,583	1,583	1,583	1,583	1,583	1,583	1,583
Observations (weighted)	913	913	913	913	913	913	913	913

Notes: Columns (1) and (2) as well as (5) and (6) provide a comparison between the OLS and the 2SLS regressions for the impact of invalidation on drug development (approval or continuation to next stage) when accounting for the actual loss of exclusivity. Columns (3) and (4) as well (7) and (8) include an interaction term capturing the loss of exclusivity in linear form. First indication refers to the drug-indication project that was first initiated for a particular compound. In each 2SLS regression the *Invalidated* dummy is instrumented with the corresponding probability predicted by a probit regression on the *Examiner participation* dummy and all other exogenous variables. The underidentification and weak identification tests are the heteroskedasticity-robust Kleibergen and Paap (2006) rk LM and Wald F statistics, respectively, as reported by Stata's ivreg2 command (Baum et al., 2010). A comprehensive list of the control variables contained in the indicated groups can be found in Table 3 in the Appendix. Standard errors are two-way clustered by drug and indication. Observations are weighted by the inverse frequency at drug-indication-patent level. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Table B-7: Impact of patent invalidation on drug development – primary patent only

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Estimation method	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Dep var	Approval				Next Stage			
Dep var mean	0.311	0.311	0.311	0.311	0.548	0.548	0.548	0.548
Invalidated	−0.038	−0.219	−0.008	−0.126	0.006	0.090	0.012	0.153
	(0.039)	(0.135)	(0.043)	(0.143)	(0.036)	(0.126)	(0.039)	(0.136)
Invalidated × Potential LoE (in yrs)			−0.016	−0.044**			−0.005	−0.039**
			(0.011)	(0.019)			(0.010)	(0.017)
Potential LoE (in yrs)			0.033***	0.054***			−0.007	0.020
			(0.010)	(0.016)			(0.008)	(0.014)
Drug characteristics	Yes**	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Development characteristics	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Patent characteristics	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Technology effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Disease effects	Yes***	Yes***	Yes**	Yes**	Yes***	Yes***	Yes***	Yes***
Examination characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age effects	Yes	Yes**	Yes	Yes***	Yes	Yes	Yes	Yes*
Originator characteristics	Yes*	Yes	Yes*	Yes	Yes**	Yes**	Yes**	Yes**
Opponent characteristics	Yes**	Yes*	Yes*	Yes	Yes	Yes	Yes*	Yes*
Year effects	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Underidentification test		32.1		30.3		32.1		30.3
Weak identification test		42.8		19.9		42.8		19.9
Observations	2,643	2,643	2,643	2,643	2,643	2,643	2,643	2,643
Observations (weighted)	2,643	2,643	2,643	2,643	2,643	2,643	2,643	2,643

Notes: Columns (1) and (2) as well as (5) and (6) provide a comparison between the OLS and the 2SLS regressions for the impact of invalidation on drug development (approval or continuation to next stage) when accounting for the actual loss of exclusivity. Columns (3) and (4) as well (7) and (8) include an interaction term capturing the loss of exclusivity in linear form. Primary patent refers to patents that protect the formulation or product of the underlying drug. In each 2SLS regression the *Invalidated* dummy is instrumented with the corresponding probability predicted by a probit regression on the *Examiner participation* dummy and all other exogenous variables. The underidentification and weak identification tests are the heteroskedasticity-robust Kleibergen and Paap (2006) rk LM and Wald F statistics, respectively, as reported by Stata's ivreg2 command (Baum et al., 2010). A comprehensive list of the control variables contained in the indicated groups can be found in Table 3 in the Appendix. Standard errors are two-way clustered by drug and indication. Observations are weighted by the inverse frequency at drug-indication-patent level. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Table B-8: Impact of patent invalidation on drug development – approval in Europe

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Estimation method	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Dep var	Approval				Next Stage			
Dep var mean	0.188	0.188	0.188	0.188	0.478	0.478	0.478	0.478
Invalidated	0.010	−0.095	0.034	−0.042	0.011	−0.089	0.039	0.005
	(0.020)	(0.103)	(0.021)	(0.103)	(0.029)	(0.115)	(0.031)	(0.125)
Invalidated × Potential LoE (in yrs)			−0.012	−0.025			−0.016**	−0.041***
			(0.008)	(0.016)			(0.008)	(0.015)
Potential LoE (in yrs)			0.025***	0.034**			0.006	0.024**
			(0.008)	(0.013)			(0.007)	(0.012)
Drug characteristics	Yes***	Yes**	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Development characteristics	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Patent characteristics	Yes**	Yes**	Yes**	Yes**	Yes**	Yes**	Yes**	Yes**
Technology effects	Yes	Yes***	Yes*	Yes***	Yes	Yes***	Yes*	Yes***
Disease effects	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Examination characteristics	Yes	Yes	Yes*	Yes*	Yes	Yes	Yes	Yes
Age effects	Yes	Yes*	Yes	Yes**	Yes	Yes	Yes	Yes
Originator characteristics	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Opponent characteristics	Yes**	Yes	Yes*	Yes	Yes	Yes	Yes	Yes
Year effects	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Underidentification test		36.5		33.9		36.5		33.9
Weak identification test		53.2		24.2		53.2		24.2
Observations	5,974	5,974	5,974	5,974	5,974	5,974	5,974	5,974
Observations (weighted)	2,741	2,741	2,741	2,741	2,741	2,741	2,741	2,741

Notes: Columns (1) and (2) as well as (5) and (6) provide a comparison between the OLS and the 2SLS regressions for the impact of invalidation on drug development (approval or continuation to next stage) when accounting for the actual loss of exclusivity. Columns (3) and (4) as well (7) and (8) include an interaction term capturing the loss of exclusivity in linear form. Drug projects that have not reached approval in Europe have the dependent variable set to 0. In each 2SLS regression the *Invalidated* dummy is instrumented with the corresponding probability predicted by a probit regression on the *Examiner participation* dummy and all other exogenous variables. The underidentification and weak identification tests are the heteroskedasticity-robust Kleibergen and Paap (2006) rk LM and Wald F statistics, respectively, as reported by Stata's ivreg2 command (Baum et al., 2010). A comprehensive list of the control variables contained in the indicated groups can be found in Table 3 in the Appendix. Standard errors are two-way clustered by drug and indication. Observations are weighted by the inverse frequency at drug-indication-patent level. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Table B-9: Impact of patent invalidation on drug development – biological drugs only

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Estimation method	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Dep var	Approval				Next Stage			
Dep var mean	0.261	0.261	0.261	0.261	0.468	0.468	0.468	0.468
Invalidated	0.017	−0.101	0.047	−0.045	0.081*	−0.062	0.100*	0.045
	(0.037)	(0.117)	(0.042)	(0.128)	(0.049)	(0.134)	(0.055)	(0.144)
Invalidated × Potential LoE (in yrs)			−0.025	−0.061*			−0.017	−0.095**
			(0.019)	(0.036)			(0.019)	(0.045)
Potential LoE (in yrs)			0.039**	0.070**			−0.005	0.063
			(0.019)	(0.033)			(0.018)	(0.041)
Drug characteristics	Yes**	Yes	Yes***	Yes*	Yes***	Yes***	Yes***	Yes***
Development characteristics	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Patent characteristics	Yes**	Yes**	Yes**	Yes**	Yes***	Yes***	Yes***	Yes***
Technology effects	Yes	Yes**	Yes	Yes**	Yes	Yes	Yes	Yes
Disease effects	Yes***	Yes***	Yes**	Yes***	Yes**	Yes***	Yes**	Yes***
Examination characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Originator characteristics	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Opponent characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year effects	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Underidentification test		29.1		29.2		29.1		29.2
Weak identification test		47.8		23.8		47.8		23.8
Observations	2,357	2,357	2,357	2,357	2,357	2,357	2,357	2,357
Observations (weighted)	1,158	1,158	1,158	1,158	1,158	1,158	1,158	1,158

Notes: Columns (1) and (2) as well as (5) and (6) provide a comparison between the OLS and the 2SLS regressions for the impact of invalidation on drug development (approval or continuation to next stage) when accounting for the actual loss of exclusivity. Columns (3) and (4) as well (7) and (8) include an interaction term capturing the loss of exclusivity in linear form. The sample includes only drug projects on large molecule drugs ("biologics"). In each 2SLS regression the *Invalidated* dummy is instrumented with the corresponding probability predicted by a probit regression on the *Examiner participation* dummy and all other exogenous variables. The underidentification and weak identification tests are the heteroskedasticity-robust Kleibergen and Paap (2006) rk LM and Wald F statistics, respectively, as reported by Stata's ivreg2 command (Baum et al., 2010). A comprehensive list of the control variables contained in the indicated groups can be found in Table 3 in the Appendix. Standard errors are two-way clustered by drug and indication. Observations are weighted by the inverse frequency at drug-indication-patent level. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Table B-10: Impact of patent invalidation on drug development – biological drugs from 2004 onward only

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Estimation method	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Dep var	Approval				Next Stage			
Dep var mean	0.241	0.241	0.241	0.241	0.428	0.428	0.428	0.428
Invalidated	0.001	−0.081	−0.018	−0.063	0.086	0.032	0.090	0.128
	(0.039)	(0.100)	(0.046)	(0.106)	(0.055)	(0.103)	(0.063)	(0.105)
Invalidated × Potential LoE (in yrs)			0.021	−0.032			−0.009	−0.141*
			(0.020)	(0.052)			(0.025)	(0.079)
Potential LoE (in yrs)			−0.025	0.022			−0.029	0.091
			(0.020)	(0.048)			(0.025)	(0.076)
Drug characteristics	Yes*	Yes	Yes*	Yes	Yes***	Yes***	Yes***	Yes***
Development characteristics	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Patent characteristics	Yes	Yes	Yes	Yes	Yes***	Yes***	Yes***	Yes***
Technology effects	Yes	Yes*	Yes	Yes*	Yes	Yes	Yes	Yes
Disease effects	Yes***	Yes***	Yes***	Yes***	Yes	Yes*	Yes**	Yes**
Examination characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Originator characteristics	Yes**	Yes**	Yes**	Yes**	Yes***	Yes***	Yes***	Yes***
Opponent characteristics	Yes	Yes	Yes	Yes	Yes	Yes*	Yes*	Yes**
Year effects	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Underidentification test		25.2		11.2		25.2		11.2
Weak identification test		41.1		2.7		41.1		2.7
Observations	1,893	1,893	1,893	1,893	1,893	1,893	1,893	1,893
Observations (weighted)	868	868	868	868	868	868	868	868

Notes: Columns (1) and (2) as well as (5) and (6) provide a comparison between the OLS and the 2SLS regressions for the impact of invalidation on drug development (approval or continuation to next stage) when accounting for the actual loss of exclusivity. Columns (3) and (4) as well (7) and (8) include an interaction term capturing the loss of exclusivity in linear form. The sample includes only drug projects on large molecule drugs ("biologics"). In each 2SLS regression the *Invalidated* dummy is instrumented with the corresponding probability predicted by a probit regression on the *Examiner participation* dummy and all other exogenous variables. The underidentification and weak identification tests are the heteroskedasticity-robust Kleibergen and Paap (2006) rk LM and Wald F statistics, respectively, as reported by Stata's ivreg2 command (Baum et al., 2010). A comprehensive list of the control variables contained in the indicated groups can be found in Table 3 in the Appendix. Standard errors are two-way clustered by drug and indication. Observations are weighted by the inverse frequency at drug-indication-patent level. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Table B-11: Impact of patent invalidation on drug development – relative loss of exclusivity

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Estimation method	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Dep var	Approval				Next Stage			
Dep var mean	0.308	0.308	0.308	0.308	0.508	0.508	0.508	0.508
Invalidated	−0.029	−0.133	0.016	0.017	0.008	−0.142	0.038	0.018
	(0.030)	(0.121)	(0.030)	(0.125)	(0.031)	(0.116)	(0.033)	(0.132)
Invalidated × Potential LoE (relative)			−0.324**	−0.878***			−0.229*	−0.889***
			(0.129)	(0.263)			(0.127)	(0.254)
Potential LoE (in yrs)			0.575***	0.997***			0.042	0.545***
			(0.116)	(0.213)			(0.113)	(0.208)
Drug characteristics	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Development characteristics	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Patent characteristics	Yes**	Yes**	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Technology effects	Yes***	Yes***	Yes***	Yes***	Yes	Yes**	Yes	Yes***
Disease effects	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Examination characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age effects	Yes	Yes*	Yes	Yes***	Yes	Yes	Yes	Yes*
Originator characteristics	Yes*	Yes*	Yes**	Yes**	Yes***	Yes**	Yes***	Yes***
Opponent characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year effects	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Underidentification test		36.5		33.5		36.5		33.5
Weak identification test		53.2		23.9		53.2		23.9
Observations	5,974	5,974	5,974	5,974	5,974	5,974	5,974	5,974
Observations (weighted)	2,741	2,741	2,741	2,741	2,741	2,741	2,741	2,741

Notes: Columns (1-8) provide a comparison between the OLS and the 2SLS regressions for the impact of invalidation on drug development (approval or continuation to next stage) when accounting for the actual loss of exclusivity as a share of the total market exclusivity. Columns (3) and (4) as well (7) and (8) include an interaction term capturing the relative loss of exclusivity in linear form. In each 2SLS regression the *Invalidated* dummy is instrumented with the corresponding probability predicted by a probit regression on the *Examiner participation* dummy and all other exogenous variables. The underidentification and weak identification tests are the heteroskedasticity-robust Kleibergen and Paap (2006) rk LM and Wald F statistics, respectively, as reported by Stata's ivreg2 command (Baum et al., 2010). A comprehensive list of the control variables contained in the indicated groups can be found in Table 3 in the Appendix. Standard errors are two-way clustered by drug and indication. Observations are weighted by the inverse frequency at drug-indication-patent level. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

C Appendix: Incentives to invest in innovation

We present the incentives to invest into R&D in a highly stylized model which has its roots in the model by Nordhaus (1969). In this model, investments in R&D need to be made which stochastically yield successful innovations before the innovator can earn profits conditional on having been successful in innovating. We hereby analyze how changes in the duration of exclusivity an innovator enjoys before imitation can occur affect incentives to invest in R&D.

Investments in R&D: We denote the amount of risky R&D as I that has to be invested in $t = 0$ and the probability of successfully innovating in $t = 0$ as p. p is a concave function of a firm's investments I with $p'(I) > 0$ and $p''(I) < 0$. (This implies that the firm's cost function $I(p)$ is strictly convex with $I'(p) > 0$ and $I''(p) > 0$). We acknowledge that the assumption of investments and innovation outcomes happening instantaneously in $t = 0$ is an abstraction from the lengthy R&D process in the pharmaceutical industry that we describe in Section 2 of this paper. Budish et al. (2015) provide a more nuanced model that takes the commercialization lag (i.e., the time between discovery and drug approval) into account; regarding the incentive effect of the duration of exclusivity, it leads to the same conclusions as the simple model presented here.

Returns from R&D: In case of successful innovation, the profits of an innovator depend on the level of exclusivity of the innovation. In case of full exclusivity (no imitation is possible for competitors), the innovator earns monopoly profits assumed to equal π^m. If (partial) imitation is possible, the innovator earns some competitive return π^c with $0 \leq \pi^c < \pi^m$. The duration of exclusivity in case of successful innovation is T. As described in Section 2, T is derived from patent protection as well as data exclusivity and depends on the project specific speed of commercialization. Abstracting from the obsolescence of an innovation (a drug might be replaced by a more efficient drug in which case profits would be reduced to zero; see Budish et al., 2015) and assuming a discount rate of r, a successful innovator therefore earns a total profit $\Pi(T)$ of

$$\Pi(T) = \int_0^T e^{-rt}\pi^m dt + \int_T^\infty e^{-rt}\pi^c dt.$$

Private incentives to invest in R&D: The firm's problem is to choose p so as to maximize the expected payoff from investing into R&D $p\Pi(T) - I(p)$. The first order condition (FOC) to this problem is given by $I'(p^*) = \Pi(T)$. The optimal likelihood of successful innovation $p^*(T)$ clearly increases in the duration of exclusivity T during which monopoly profits π^m can be earned. Since $\pi^c < \pi^m$, $\Pi(T)$ increases in T with $\frac{\partial \Pi(T)}{\partial T} > 0$. The longer the period of exclusivity, the higher the innovator's return and, hence, its incentive to invest into R&D. Since

$p(I)$ is concave, the FOC holds for increasing T only for rising I. The opposite holds as well. A loss of exclusivity will reduce a firm's incentives to invest in R&D.

References

Baum, J. a. C., R. Cowan, and N. Jonard (2010). Network-Independent Partner Selection and the Evolution of Innovation Networks. *Management Science 56*(11), 2094–2110.

Budish, E., B. Roin, and H. Williams (2015). Do Firms Underinvest in Long-Term Research? Evidence from Cancer Clinical Trials. *American Economic Review 105*(7), 2044–2085.

Kleibergen, F. and R. Paap (2006). Generalized Reduced Rank Tests Using the Singular Value Decomposition. *Journal of Econometrics 133*(1), 97 – 126.

Nordhaus, W. D. (1969). *Invention, Growth and Welfare: A Theoretical Treatment of Technological Change*. MIT press.