IPEN 5140 · Quantitative Analysis and Empirical Methods

Zhanghao Chen (50023662)

Yanlin Chen (50030029)

Junxiang Lai (50029991)

Empirical Exercises 1: ***The Hidden Costs of Securing Innovation: The Manifold Impacts of Compulsory Invention Secrecy. Management Science (2022)***

This empirical exercise examines the impact of compulsory invention secrecy on the subsequent innovations of the same firms. In this report, we primarily use Python and Stata to conduct the analysis. Python is utilized for performing descriptive statistics while Stata is employed to replicate the main results of the original article by implementing the authors' provided code. The source code used in this study is available at https://github.com/Zhanghao25/5140_empirical_exercise.

Task1: Effect of compulsory secrecy on follow-up innovations by the same firms

We begin by examining the structure of the dataset. The file *data__assignee__nber_cat.dta* contains 42 columns and 392,717 unique firms. The missing values are primarily concentrated in the columns *self_total_secrate*, with 1,965,555 missing entries, and *self_incat_secrate*, with 2,213,585 missing entries.

Table 1 gives the results of two t-tests. It can be observed that there is a significant difference in the number of patents between OSRD and non-OSRD companies both before and after the war (p=0). Moreover, the t-statistic is higher in the post-war period, indicating that the difference is more pronounced after the war. The t-statistic for the average number of secret patents is 49, which also suggests a significant difference between OSRD and non-OSRD companies.

Table 1: T-statistics and p-values for the number of patents across different time periods

Variable	t-statistic	p-value
Number of Patents (Period 0: 1930-1934)	92.6898	0.0000
Number of Patents (Period 1: 1935-1939)	109.3195	0.0000
Number of Patents (Period 2: War time)	nan	nan
Number of Patents (Period 3: 1946-1950)	138.8149	0.0000
Number of Patents (Period 4: 1951-1955)	149.2752	0.0000
Number of Patents (Period 5: 1956-1960)	145.8186	0.0000
Average Number of Secret Patents	49.1760	0.0000

Figure 1 illustrates the relationship between the number of secret patents and the total number of patent

classes across firms. It is clear that most firms have no secret patents, and these firms also tend to have a lower number of patent classes. As the number of secret patents increases (from 1 to 4), the number of patent classes also increases, which may indicate that firms capable of filing secret patents are generally stronger and file more patents overall. However, when the number of secret patents exceeds 5 or 6, the total number of patent classes drops significantly. This sharp decline may be due to restrictions imposed by secrecy orders, which could limit the scope of patent filings for these firms.

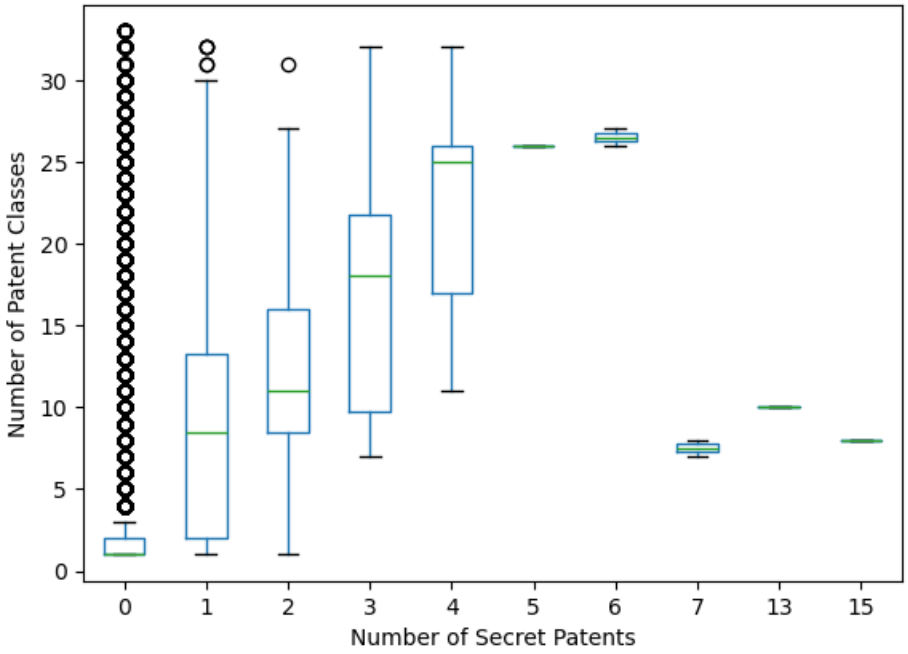

Figure 1: Boxplot of patent classes by secret patents.

We finally attempt to replicate Table 4 using a clogit combined with MLE. However, the log-likelihood function does not converge in our simulation, which indicates issues such as model complexity or potential multicollinearity within the data. To address this issue, we modify our approach by using a standard logit model instead. The logit model proves to be more stable for this replication, and the following code is used to implement the logit model:

```
eststo self_incat___cont___a'a'_i'i': ///
logit any_patent self_incat_secrate_t* self_incat_secrate period* 'ctrls1a' if ///
'cond_incumbency' & /// incumbency condition
self_incat_midwar>=1 & /// had midwar patents in given cat
self_incat_secrate>0 & /// had midwar secrecy orders in given cat
(year<=1939 | year>=1946) & 'subsample', ///
cluster(assignee)
```

Listing 1: Logit Regression with Clustered Standard Errors

Overall, the results from Table 2 demonstrate that wartime secrecy orders have a significant and disproportionate impact on different types of firms. Non-OSRD firms, small firms, and new market entrants face the greatest challenges, with their patent filings significantly reduced during the wartime periods. In contrast, OSRD firms seem to be largely unaffected, likely due to their privileged access to government projects and additional support. These findings suggest that wartime secrecy orders create an uneven playing field, where smaller and less connected firms experience greater constraints on innovation. This conclusion is consistent with the findings from Table 4 in the original paper.

Table 2: Logit regression results of the effect of secrecy order on patenting activity for different firm types.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	All	Non-OSRD	OSRD	Small	Large	Entrant	Entrant
Wartime Sec. Rate * (1935-39)	-1.094*	-1.541**	-0.898	-2.159**	-1.335		
	(0.574)	(0.732)	(1.062)	(0.974)	(0.833)		
Wartime Sec. Rate * (1946-50)	-1.959***	-2.236***	-1.070	-3.028***	-1.823***	-3.030***	-3.540***
	(0.420)	(0.544)	(0.779)	(0.683)	(0.626)	(0.306)	(0.444)
Wartime Sec. Rate * (1951-55)	-1.647***	-2.264***	-0.237	-3.593***	-0.829	-3.269***	-3.837***
	(0.438)	(0.596)	(0.855)	(0.807)	(0.632)	(0.338)	(0.528)
Wartime Sec. Rate * (1956-60)	-1.177***	-1.535***	-0.171	-2.323***	-0.578	-2.950***	-2.960***
	(0.417)	(0.557)	(0.815)	(0.805)	(0.597)	(0.340)	(0.514)
N	3635	1960	1675	905	2730	1809	1071
R^2							
Assignee x NBER Cat FEs	Y	Y	Y	Y	Y		
NBER Cat FEs						Y	Y

Task2: Effect of compulsory secrecy on follow-on innovation

In this task, we first present the kernel density estimation (KDE) of forward citations in Figure 2, grouped by secrecy status. Notably, there are two datasets labeled as ”paper_citation” in both the regression and descriptive folders. We choose the one with the larger values from the descriptive folder for plotting. Table 3 presents summary statistics for forward citations.

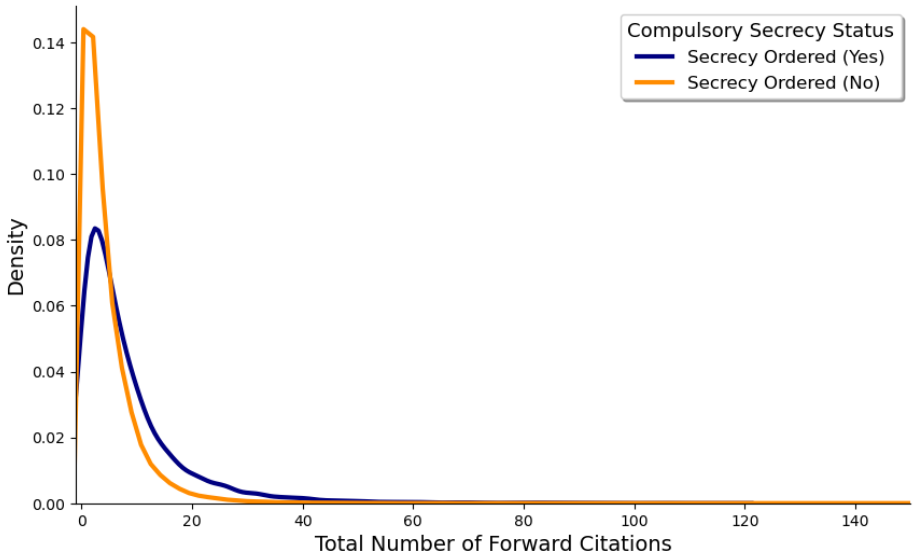

Figure 2: Kernel density estimation of forward citations by secrecy status.

Table 3: Summary statistics of forward citations by secrecy status.

Statistic	Secrecy Ordered (Yes)	Secrecy Ordered (No)
Count	6,242	236,372
Mean	8.42	4.53
Standard Deviation	9.96	5.51
Min	0.00	0.00
25th Percentile	2.00	1.00
Median (50th Percentile)	5.00	3.00
75th Percentile	11.00	6.00
Max	116.00	343.00

Combining the insights from Figure 2 and Table 3, it is evident that patents under secrecy orders tend to have a higher number of forward citations compared to those without secrecy restrictions. The KDE plot shows a rightward shift in the distribution for secrecy-ordered patents, indicating a higher density of citations. Meanwhile, the summary statistics in Table 3 confirm this observation, with a mean of 8.42 forward citations for secrecy patents, compared to 4.53 for non-secrecy patents. The t-test results further reinforce this significant difference, with a t-statistic of 30.72 and an extremely small p-value (2.49e-193), indicating that the difference in forward citations between the two groups is statistically significant.

In the next step, we replicate the results reported in original Table 6, replacing the dependent variable with the total number of forward citations received. We then estimate the same equation using the maximum

likelihood estimator (MLE) based on Poisson regressions. The results are shown in Table 4.

Compared to Table 6 in the original paper, many point estimates become statistically insignificant when the dependent variable is replaced with the total number of forward citations using the MLE estimator based on Poisson regressions. All regressions include grant year and patent class-year fixed effects to control for unobserved impacts.

When using the total number of forward citations as the dependent variable, no significant correlation exists between secrecy orders and the number of forward citations, except for the year 1942 (0.187, $p<0.1$) for all patents ordered secret. This contrasts with the original Table 6, where the probability of receiving any forward citation was significantly affected, indicating that secrecy orders and OSRD status do not influence citation counts.

For patents evaluated for secrecy but not ordered secret, results are more consistent with the original Table 6. For non-OSRD patents evaluated but not ordered secret, both the probability (0.036, $p<0.05$) and the number of forward citations (0.167, $p<0.05$) show statistically significant positive estimates, similar to the original table. However, for OSRD-evaluated patents, forward citations increase significantly from 1941 to 1944 (0.170 to 0.211, $p<0.05$) and 1945 (0.237, $p<0.1$), suggesting that OSRD patents may have broader post-war applications despite initially being focused on wartime efforts.

In contrast, non-OSRD estimates by year remain close to zero, showing little impact on forward citation counts, further indicating the different nature and influence of OSRD and non-OSRD patents on forward citations.

Task 3: Comparison of Forward Citations of Secret and Nonsecret Patents (1939-1945)

In Task 3, we analyze the forward citations of secret and nonsecret patents during the period from 1939 to 1945. To investigate the potential impact of secrecy orders on patent influence, we conduct a series of linear regressions where the number of forward citations is used as the dependent variable. The key independent variables include whether a patent was subject to a secrecy order (denoted as *secret*) and whether it was evaluated for secrecy but not ordered secret (*secrecy_eval*). We also control for grant year and patent class-year fixed effects to account for temporal and technological variations.

For simplicity, we use OLS regression with clustered standard errors at the patent class level to ensure robust estimates. Table 5 presents the regression results comparing forward citations of secret and nonsecret patents filed between 1939 and 1945. The results show that patents ordered under secrecy received sig-

nificantly more forward citations, had a greater number of unique citers, and spanned more citing classes compared to nonsecret patents. This suggests that, despite being kept confidential, these patents had a substantial long-term impact on subsequent innovation once secrecy was lifted. Similarly, patents that were evaluated for secrecy also exhibited greater citation influence. The results indicate that compulsory secrecy did not diminish the eventual technological impact of these patents, suggesting that secrecy orders often targeted patents with high technical significance, resulting in high citation rates after their disclosure. All regressions control for grant year fixed effects and class-year fixed effects. Standard errors are clustered by patent class and reported in parentheses.

Table 5: Regression Results: Forward Citations of Secret and Nonsecret Patents (1939-1945)

	Outcomes:				
	Cites	Citers	Citing classes	Non-self	Self
Secrecy ordered	1.578***	1.232***	0.364***	1.422***	0.134***
	(0.349)	(0.275)	(0.080)	(0.321)	(0.039)
Secrecy evaluated	0.926***	0.615***	0.293***	0.743***	0.170***
	(0.111)	(0.077)	(0.036)	(0.100)	(0.018)
N	242614	242614	242614	242614	242614
R^2	0.09	0.09	0.12	0.09	0.06
Grant year FEs	Y	Y	Y	Y	Y
Class-year FEs	Y	Y	Y	Y	Y
Mean of DV	4.63	4.23	2.06	4.27	0.27
s.d. of DV	5.70	4.65	1.90	5.31	0.95

Conclusion

In this replication task, we examine the impact of compulsory secrecy orders on corporate innovation. In Task 1, we find significant differences between OSRD and non-OSRD companies in both total and secret patents. Logit regression results from Table 4 confirm that post-war patent filings decrease significantly, particularly among non-OSRD and smaller firms, largely due to secrecy orders. In Task 2, we observe that patents under secrecy orders receive more follow-up citations than nonsecret patents, suggesting greater impact. In Task 3, regression analysis from 1939 to 1945 shows that secrecy orders increase forward citations, with secret patents generally receiving more citations than nonsecret ones.

Table 4: Impact of Secrecy Orders on Patent Filings (1941-1945)

	All (1)	Non-OSRD (2)	OSRD (3)	Triple difference	
				All (4)	OSRD (rel. to all) (5)
Secrecy ordered	0.081	-0.004	0.263	0.051	0.091
	(0.112)	(0.133)	(0.164)	(0.123)	(0.172)
× filed in 1941	0.018	-0.061	-0.021	-0.105	0.181
	(0.112)	(0.149)	(0.145)	(0.144)	(0.174)
× filed in 1942	0.187*	0.195	0.048	0.129	0.045
	(0.105)	(0.136)	(0.160)	(0.121)	(0.172)
× filed in 1943	0.161	0.242	-0.027	0.175	-0.076
	(0.111)	(0.153)	(0.145)	(0.149)	(0.186)
× filed in 1944	0.128	0.180	-0.085	0.125	-0.070
	(0.109)	(0.198)	(0.151)	(0.184)	(0.218)
× filed in 1945	-0.002	0.431	-0.314	0.366	-0.546*
	(0.159)	(0.274)	(0.217)	(0.271)	(0.308)
Secrecy evaluated	0.092	0.167**	-0.043	0.182**	-0.219**
	(0.060)	(0.073)	(0.099)	(0.074)	(0.105)
× filed in 1941	0.128**	0.117	0.170*	0.117	0.017
	(0.064)	(0.088)	(0.100)	(0.089)	(0.122)
× filed in 1942	0.040	-0.003	0.070	-0.028	0.124
	(0.066)	(0.078)	(0.112)	(0.079)	(0.110)
× filed in 1943	0.068	-0.003	0.139	-0.037	0.188*
	(0.071)	(0.092)	(0.102)	(0.093)	(0.113)
× filed in 1944	0.082	-0.063	0.211**	-0.063	0.255**
	(0.066)	(0.099)	(0.098)	(0.098)	(0.125)
× filed in 1945	0.097	-0.038	0.237*	-0.058	0.264
	(0.096)	(0.132)	(0.144)	(0.132)	(0.172)
N	162548	125042	37263	162548	
Grant year FEs	Y	Y	Y	Y	
Class-year FEs	Y	Y	Y	Y	
Mean of DV	4.79	4.49	5.81	4.79	
s.d. of DV	5.77	5.28	7.09	5.77	