Sweden is calling: What shapes the delay in the Nobel Prize discoveries? A research note

Polemis, Michael and Stengos, Thanasis

University of Piraeus, Department of Economics, Piraeus, Greece, University of Guelph, Ontario, Canada

1 January 2021

Sweden is calling: What shapes the delay in the Nobel Prize discoveries? A research note

Michael L. Polemis[a,b,c,] and Thanasis Stengos[d]

[a] *University of Piraeus, Department of Economics, Piraeus, Greece*

[b] *Hellenic Competition Commission, Athens, Greece.*

[c] *Hellenic Open University, Patras, Greece*

[d] *University of Guelph, Ontario, Canada*

Abstract

The scope of this study is to shed light on the determinants of the time gap between the publication of a Nobel discovery and the bestowment of the prize across three science fields (Physics, Chemistry, and Medicine). The econometric evidence supports that the delay gap is inversely related to the age of the Laureate when the Nobel-worthy contribution was published in Physics and Chemistry but not in Medicine. An increase of the age of the researcher by one year leads to a reduction of the Nobel delay by almost three months on average while sharing the Prize for the same research delays the award by approximately 1.2 years. Lastly, important theoretical discoveries increase the delay by 3.3 years on average, while obtaining the last education degree a year later delays the Nobel Prize by 4.7 months on average.

Keywords: Nobel Prize; Delay gap; Sciences; Contest

JEL Codes: C23; C80; D02

1. Introduction

It is widely perceived that scientists who publish breakthrough discoveries are, on average, waiting longer for a Nobel prize than ever before (Fortunato, 2014; Becattini et al, 2014). This is also expressed in Professor and Chairman of the Nobel Committee for Chemistry Sven Lidin's own words: *"We want to make sure that we award those who open the first door into new scientific insight. This means that naturally there is a delay. Typically, it takes about 20 years before the initial door opener has matured into a Nobel prize"*.[1]

This happens to be the case in many instances.[2] To give an example of this lengthy process, it is reported that the Higgs boson ("*God particle*"), whose discovery was awarded a Nobel prize in physics in 2013, was theoretically developed half a century ago back in 1964. However, at the time of the award one of the three original contributors, Dr. Robert Brout, had died (Becattini et al, 2014). The long history of the Nobel prize that originates back in 1901, provides sufficient ground to investigate the possible reasons for the time lag between a breakthrough discovery and a Nobel prize award.

Despite the crucial importance of the dissemination of academic research, few studies have tried to provide plausible answers. Becattini et al, (2014) discuss the causes for the Nobel prize delay. They argue that the increasing number of scientists, the increasing life expectancy, the changing research, and career policies, along with the increasing training time justify some of the reasons for the delay. However, they do not conduct a quantitative analysis to exemplify and weigh the impact of all these determinants on the Nobel Prize contest. In a similar vein, Fortunato (2014), recognizes that the time lag between the scientific discovery and the Nobel prize has been

[1] https://www.thelocal.se/20141005/nobel-delay-risks-making-prize-irrelevant-scientists.
[2] For some excellent examples, see Stephan and Levin, (1993).

increased since 1985 crossing the "*threshold*" of 20 years on average, without investigating the main causes for this outcome.

In a different context, Baffes and Vamvakidis (2011) forge a statistically significant correlation between the age of the Laureate when the Nobel-worthy contribution was published and the age of the scientist when the prize was awarded. Based on their estimates, they argue that delaying the contribution by one year, retards the Nobel Prize award by six (Physics) to eight months (Medicine). Jones (2010) argues that the mean age at great scientific advances for Nobel Prize winners rose by about six years over the twentieth century as a result of declining innovative output in the early life cycle.

This study lends support to pooled regression analysis to unravel the driving forces that determine the delay in the Nobel Prize award across sciences (Physics, Chemistry, and Medicine). The paper argues that the Nobel Committee does favor older nominees since the delay gap decreases (increases) as the age of the Laureate when the scientific achievement was brought to light increases (decreases). Other parameters including the age at which the Nobel Laureate received the last education degree, the number of recipients who share the award for the same research along with certain demographic factors (geographic origin and gender of the Laurate) also affect the delay process.

The remainder of the paper unfolds as follows. Section 2 describes the data and estimation strategy. Section 3 presents and discusses the empirical findings of this study, while Section 4 concludes the paper.

2. Data and Methodology

2.1 Sample and variables

The sample was obtained by Jones and Weinberg, (2011) and is based on publicly available data provided by the Nobel Foundation's official website (nobelprize.org).[3] Table 1 presents descriptive statistics. As it is evident, the dependent variable (Delay) exhibits the highest variation, which is to be expected, while the average time lag between scientific discovery and recognition in all the disciplines exceeds 16 years. Most of the variables are positively skewed, following a leptokurtic distribution.

<Table 1>

The Nobel Prize delay can be illustrated in Figure 1. The latter plots the time difference (in years) between the discovery and the awarding of the Nobel prize, versus the year when the award is received. As it is evident, there is an increasing trend (solid black line) in the time between the discovery and the Nobel award for the three disciplines, with rates for Physics, Chemistry, and Medicine, equal to 17, 16.4, and 17.2 years on average.

<Figure 1>

Lastly, Figure 2 presents the kernel density of the Nobel Prize delay for the sample scientific fields. As it is evident, there is a large variation in years that makes the distribution non-normal (left-skewed). The largest mass of the delay refers to 17 years, but a substantial amount is also concentrated around 15 years, while the rest is related beyond the 20 years.

<Figure 2>

[3] https://data.nber.org/data-appendix/w11359/

2.2 Identification strategy

Following the spirit of Baffes and Vamvakidis (2011), this study estimates the following panel regression equation for the period 1901–2008 for all three sciences, as well as by discipline:

$$Delay_{i,t} = \beta_0 + \beta_1 Contribution_{i,t} + \beta_2 Share_{i,t} + \beta_3 Europe_{i,t} + \beta_4 Gender_{i,t} +$$

$$+\beta_5 Last\_Degree_{i,t} + \beta_6 Theory_{i,t} + \beta_7 Trend_t + \gamma_i + \delta_t + \varepsilon_{i,t} \quad (1)$$

where Delay denotes the time lag between the discovery and the Nobel recognition for discipline i=1,2,3 at time t = 1,2,…,190. Contribution is the age of the Laureate when the Nobel-worthy contribution was published. Share denotes the number of Laurates who share the award for the same research and takes the value of 1, 2, or 3. Europe is a dummy variable taking the value of one when the recipient was born in a European country and zero otherwise. Gender takes the value of one for male and zero for female recipients. Last_Degree is the age at which the Nobel Laureate received the last education degree (e.g., Ph.D. or Post-Doc).

Theory is a dichotomous variable that takes the value of one (zero) for discovery with an important theoretical (empirical) component. Trend denotes the linear time trend used as a proxy for increasing life expectancy and the expanded pool of Nobel candidates and subfields of research (Baffes and Vamvakidis, 2011). The model controls for field and time fixed effects (γ_i and δ_t respectively) to allow unobserved heterogeneity. Lastly, $\varepsilon_{i,t}$ denotes the i.i.d error term.

3. Results and discussion

Table 2 displays the pooled specification results for all sciences. Columns 1-4 report the OLS estimates with or without the inclusion of the quadratic term (Contribution-squared). As it is observed, nearly all the covariates (except for the Europe dummy), are statistically significant exhibiting anticipated signs.

The estimate of the age of the Laureate when the Nobel discovery was published (Contribution) is significantly less than one in the linear specifications (see Columns 1&2), while the magnitude dictates that delaying the contribution by a year, "*quickens*" the award by only two months.

Sharing the Prize for the same research has also a statistically significant impact on the Nobel delay since adding one more recipient for the same research extends the time lag up to 1.135 years (Column 4). Male recipients (Gender) delay the Nobel Prize, even by nine years, whereas the time trend is negatively correlated with the dependent variable, implying that increasing life expectancy reduces the Nobel delay by a negligible rate (0,3 months on average). Obtaining the last education degree, a year later (Last_Degree), delays the Nobel Prize even by 6.7 months (see Column 1). Moreover, theoretical contributions delay further the Nobel award exceeding in some cases the three years timespan.

<Table 2>

It is noteworthy that the OLS estimates uncover a non-monotonic statistically significant (concave) pattern when a quadratic term is included in the estimation models (see Columns 3&4). This finding coincides with Jones and Weinberg, (2011), who unravel "*hump-shaped*" Kernel estimated patterns of the Physicist Nobel laureates before the ages of 30 and 40.

To test the robustness of our findings, we re-estimate our basic model by controlling for fixed effects (i.e., field and time dummies). Columns 5-6 present the relevant estimations. As it is evident, the inclusion of fixed effects does not change the validity and inference of the estimated parameters, though the magnitude of the coefficients is larger in this case.

The results by field emerge significant differences (see Table 3). The age in which the last degree was received is negative and statistically significant in Physics and Chemistry, though the

magnitude of the estimates is larger than the pool regression results (see Columns 1 and 2). However, this effect is not significant for Medicine as it seems that in this field any delay has to do with the nature of the discovery rather than its timing. The dummy variable for sharing the Prize has a negative but not statistically significant estimate in Physics, while the dummy variable for European laureates' research exerts a negative and statistically significant effect only in Physics (see Column 1). The male dummy variable retains its positive sign in all fields but is not statistically significant in Medicine possible due to the smaller number of male recipients compared to the other two sciences (see Column 3). Finally, the theoretical contributions delay the Nobel Prize by 4 (Chemistry) to 5.3 years (Physics), while the estimate of the time trend is negative and statistically significant only in Physics.

<Table 3>

4. Conclusion

It has been argued that the time lag between the publication of a Nobel discovery and the conferment of the prize (Nobel Prize delay) has been rapidly increasing for all disciplines. This note attempts to identify the key determinants of this Nobel Prize delay, an issue that has been nearly overlooked by the existing literature.

The sample used in the empirical analysis includes various characteristics of the Nobel Laureates including *inter alia* data on dates of birth, the year of Nobel prizes and year(s) of publication(s) of prize-winning work, the gender, and the nationality of the recipients, across the three scientific fields (Physics, Chemistry, and Medicine).

The basic econometric model relies on pooled regression analysis of the most likely drivers of the Nobel Prize delay. The empirical results support the notion that the Nobel Committee favor older nominees since the delay gap decreases (increases) as the age of the Laureate when the

scientific achievement was conducted increases (decreases). However, the field-by-field analysis unravels that this finding is stronger in Physics and Chemistry and nearly absent in Medicine. The econometric analysis has also studied the impact of other drivers stimulating the Nobel Prize delay. Specifically, we argue that sharing the Prize for the same research delays the award by approximately 1.2 years. However, this does not hold for Physics, despite the existence of a negative (though not statistically significant) relationship with the delay variable. Important theoretical contributions increase the delay gap by 3.3 years on average for all fields. The time lag between the discovery and the award, further increases by 4 years for Chemistry and 5.3 years for Physics, while this parameter does not impose a statistically significant effect in Medicine. It is also highlighted that obtaining the last education degree a year later delays the Nobel Prize by 4.7 months on average for all the three disciplines and 9.3 months for Physics only. Lastly, gender has a positive and statistically significant effect in Physics and Chemistry but not in Medicine as a result of the smaller number of male researchers compared to the other two sciences.

Based on the above, this note is not free from limitations, while it provides avenues for further research. First, expanding the database to include a richer set of explanatory variables, that may affect the Nobel delay as well as incorporating more information regarding the demographic and other socioeconomic characteristics of all candidates would provide us with more robust estimates. Moreover, the estimation of non-parametric or semi-parametric models to explore the non-linear pattern of the relationship between the age of the Laurate at the time of the discovery and the Nobel delay would also shed some light on this research direction in the near future.

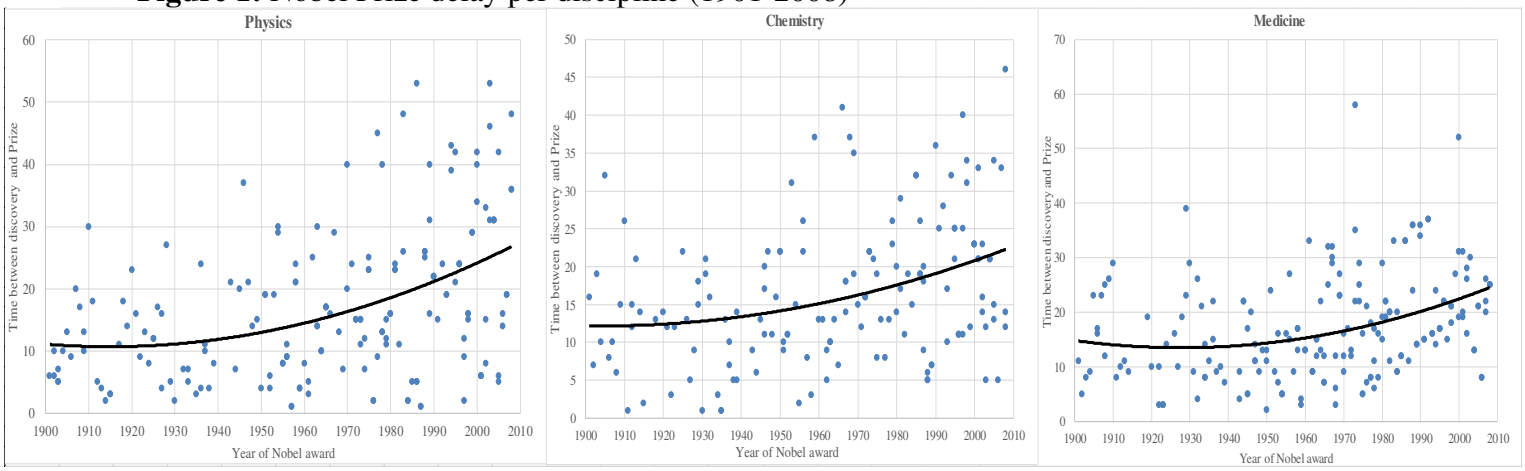
Tables & Figures

Table 1: Descriptive statistics

Variable	Mean	Median	Min	Max	Standard deviation	Skewness	Kurtosis
Delay *(in years)*	16.87	15	5	58	10.34	0.885	3.799
Contribution *(in years)*	39.04	38	20	80	8.545	0.696	4.079
Share *(in numbers)*	1.901	2	1	3	0.839	0.188	1.448
Europe *(binary indicator)*	0.514	1	0	1	0.500	-0.057	1.003
Gender *(binary indicator)*	0.975	1	0	1	0.156	-6.116	38.41
Last_Degree *(in years)*	26.10	26	14	43	3.417	1.004	5.310
Theory *(binary indicator)*	0.185	0	0	1	0.388	1.625	3.639

Notes: Delay denotes the time lag between the discovery and the Nobel recognition. Contribution is the age of the Laureate when the Nobel-worthy contribution was published. Share denotes the number of Laurates who share the award for the same research. Europe is a dummy variable taking the value of one when the recipient was born in a European country and zero otherwise. Gender takes the value of one for male and zero for female recipients. Last_Degree is the age at which the Nobel Laureate received the last education degree. Theory is a dichotomous variable that takes the value of one (zero) for discovery with an important theoretical (empirical) component. Trend denotes the linear time trend.

Figure 1: Nobel Prize delay per discipline (1901-2008)



Notes: Each plot shows the raw data corresponding to the time difference (in years) between the discovery and the awarding of the Nobel prize, versus the year when the award is received for Physics, Chemistry, and Medicine, respectively. The solid black line denotes the polynomial trend of degree.

Figure 2: Kernel density estimate for the Nobel Prize delay

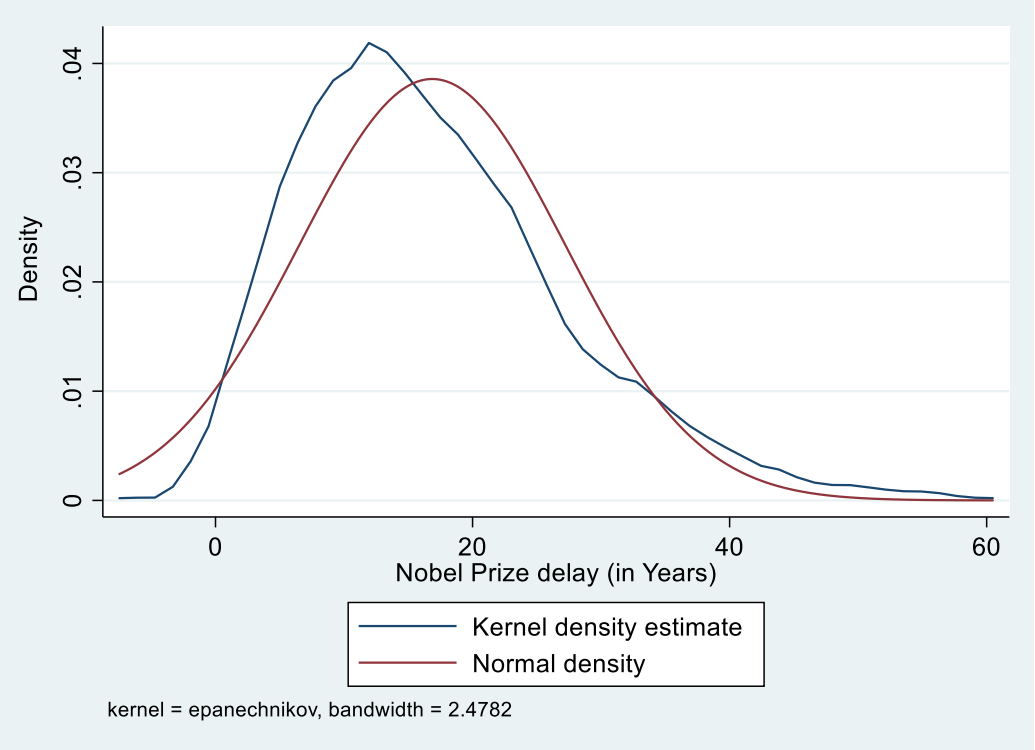


Table 2: Pooled regression results for all disciplines

Variable **Method**	**(1)** **OLS**	**(2)** **OLS**	**(3)** **OLS-NL**	**(4)** **OLS-NL**	**(5)** **FE**	**(6)** **FE**
Constant	-	-	-	-	17.37**	16.99*
					(6.762)	(5.835)
Contribution	-0.160***	-0.169***	0.630***	0.473**	-0.309***	-0.331**
	(0.0503)	(0.0506)	(0.191)	(0.189)	(0.0114)	(0.0340)
Contribution-squared	-	-	-0.0101***	-0.008***	-	-
			(0.0023)	(0.00233)		
Share	0.925*	1.202**	0.763	1.135**	1.232***	1.385***
	(0.556)	(0.553)	(0.549)	(0.547)	(0.254)	(0.136)
Europe	-0.592	-0.776	-1.027	-1.173	0.254	0.259
	(0.898)	(0.903)	(0.889)	(0.900)	(0.508)	(0.495)
Gender	8.928***	7.525***	4.033*	3.205	5.906*	6.405
	(2.354)	(2.326)	(2.583)	(2.609)	(4.165)	(4.184)
Last_Degree	0.555***	0.520***	0.222*	0.241*	0.0046	0.0216
	(0.102)	(0.102)	(0.127)	(0.128)	(0.157)	(0.178)
Theory	3.015**	2.621**	3.015***	2.528**	2.576*	3.334*
	(1.172)	(1.174)	(1.153)	(1.162)	(1.556)	(1.275)

Trend	-0.0237***	-	-0.0305***	-	-0.194***	-0.192***
	(0.0078)		(0.0079)		(0.0045)	(0.0117)
Observations	*525*	*525*	*525*	*525*	*525*	*525*
Adjusted R^2	*0.739*	*0.735*	*0.748*	*0.741*	*0.553*	*0.554*
Field fixed effects	*No*	*No*	*No*	*No*	*No*	*Yes*
Time fixed effects	*No*	*No*	*No*	*No*	*Yes*	*Yes*

Notes: Delay denotes the time lag between the discovery and the Nobel recognition. Contribution is the age of the Laureate when the Nobel-worthy contribution was published. Share denotes the number of Laurates who share the award for the same research. Europe is a dummy variable taking the value of one when the recipient was born in a European country and zero otherwise. Gender takes the value of one for male and zero for female recipients. Last_Degree is the age at which the Nobel Laureate received the last education degree. Theory is a dichotomous variable that takes the value of one (zero) for discovery with an important theoretical (empirical) component. Trend denotes the linear time trend. Time dummies are included but not reported. Robust standard errors clustered at three disciplines in parentheses. FE = Fixed effects, NL = Non-Linear. *** p<0.01, ** p<0.05, * p<0.1.

Table 3: OLS regression results by discipline

	(1) Physics	(2) Chemistry	(3) Medicine
Contribution	-0.322***	-0.207***	-0.012
	(0.094)	(0.079)	(0.088)
Share	-1.499	2.229**	1.653*
	(1.159)	(0.994)	(0.861)
Europe	-3.160*	-0.796	1.252
	(1.787)	(1.472)	(1.444)
Gender	16.637***	13.140***	2.738
	(5.491)	(4.098)	(3.214)
Last_Degree	0.772***	0.325**	0.461***
	(0.227)	(0.162)	(0.155)
Theory	4.016**	5.304***	-1.782
	(1.878)	(1.932)	(2.816)
Trend	-0.049***	-0.009	-0.017
	(0.015)	(0.013)	(0.013)
Adjusted R^2	*0.72*	*0.80*	*0.75*
Observations	*182*	*153*	*190*

Notes: Delay denotes the time lag between the discovery and the Nobel recognition. Contribution is the age of the Laureate when the Nobel-worthy contribution was published. Share denotes the number of Laurates who share the award for the same research. Europe is a dummy variable taking the value of one when the recipient was born in a European country and zero otherwise. Gender takes the value of one for male and zero for female recipients. Last_Degree is the age at which the Nobel Laureate received the last education degree. Theory is a dichotomous variable that takes the value of one (zero) for discovery with an important theoretical (empirical) component. Trend denotes the linear time trend. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

References

Baffes, J., and Vamvakidis, A. (2011). Are you too young for the Nobel Prize?, *Research Policy*, **40** (10), 1345-1353.

Becattini, F., Arnab, C., Fortunato, S., Mitrovic D.M., Pan, R.K., Parolo, P. (2014). The Nobel Prize delay. *Physics Today*.

Fortunato, S. (2014). Growing time lag threatens Nobels. *Nature* **508**, 186.

Jones, B., and Weinberg, B.A. (2011). Age dynamics in scientific creativity. *Proceedings of the National Academy of Sciences*, **108** (47), 18910-18914.

Jones, B. (2010). Age and Great Invention, *The Review of Economics and Statistics*, **92**(1), 1-14.

Karazija, R., and Momkauskaitė, A. (2004). The Nobel prize in physics - regularities and tendencies. *Scientometrics* **61**, 191–205.

Stephan, P.E., and Levin, S.G. (1993). Age and the Nobel prize revisited. *Scientometrics* **28,** 387–399