The CG Equation: A Probabilistic Approach to Predict Doctoral Success Likelihood

Sourojeet Chakraborty (sourojeet.chakraborty@mail.utoronto.ca)
University of Toronto
https://orcid.org/0000-0002-9461-880X

Daniela Galatro
University of Toronto

Research Article

Keywords: Doctoral Attrition (DA), Graduate Studies, Doctoral Success Likelihood (DSL), Time Taken to Graduate (TTG), Higher Education (HE), Mental Health

Posted Date: September 19th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1936442/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License
The CG Equation: A Probabilistic Approach to Predict Doctoral Success Likelihood

Sourojeet Chakraborty¹ (ORCID: 0000-0002-9461-880X), Daniela Galatro¹(ORCID: 0000-0003-0934-7492)

¹Department of Chemical Engineering & Applied Chemistry, University of Toronto, ON M5S 3E5 Canada
sourojeet.chakraborty@mail.utoronto.ca, daniela.galatro@utoronto.ca

Abstract: Doctoral attrition (DA) is a phenomenon of graduate students choosing to discontinue graduate studies and is universally encountered across all academic disciplines. Key parameters, that are typically perceived as valuable by Ph.D. students, are identified from a systematic literature review; and the Chakraborty-Galatro (CG) probabilistic equation is formulated to predict the likelihood of a successful Ph.D. experience, called the Doctoral Success Likelihood (DSL), thus minimizing; possibly eliminating DA. Our model provides prospective/novice graduates with a novel framework to self-assess and predict the success likelihood of their Ph.D. journey. Such a framework enables the graduate student to judiciously self-assess and make a rationally informed decision about their career, rather than taking a blind leap of faith. Our equation also accommodates force-majeure circumstances (such as a pandemic, the bereavement of a loved one, mental health issues, etc.), which may significantly impact the time taken to graduate (TTD); leading to a candidate choosing to drop out. Such circumstances typically derail/delay doctoral progress, and can push an initially feasible set of probabilities, into an undesired “infeasibility triangle”. Higher the net probability values obtained from our equation, stronger the likelihood of an enriching Ph.D. experience. When periodically tracked, our proposed equation can also help students identify and calibrate their own doctoral experience, while capturing tangible feedback and perspectives for both students, and supervisors. One author presents his own doctoral journey, applying the CG equation to evaluate DSL values for his Ph.D., over a three-year period.

Keywords: Doctoral Attrition (DA), Graduate Studies, Doctoral Success Likelihood (DSL), Time Taken to Graduate (TTG), Higher Education (HE), Mental Health.

Introduction: Present day graduate research is a far cry from how it was perceived and performed, even a century ago. From being compatriots towards a truth-seeking journey, the relationship between most supervisors and students has today become almost perfunctory, towards solely pursuing a publication rat race (Bedeian, 1996), with a publish-or-perish (De Rond and Miller, 2005), extremely competitive attitude. Additionally, there is very little element of mentorship in most supervisor-student relationships today, which adds to the ongoing mental health crisis among graduate students (Evans et al., 2018; Békouche et al., 2022). Most doctoral students today, depending on the discipline of interest and the choice of an appropriate research topic, spend anywhere between 3-11 years earning their Ph.D. and face intense pressure to publish early (Stoilescu and McDougall, 2010). These early but impactful years are extremely difficult to survive, and often lead to high amount of doctoral attrition (DA) in the absence of an adequate support system (Jacobsson & Gillström, 2006; Pyhältö et al., 2015). A large majority of those who successfully endure the ordeal, frequently fall victims to depression, anxiety, imposter syndrome, inferiority complex and/or other associated symptoms of deteriorated physical/mental health such as bipolar, borderline, and similar Cluster B personality disorders. In the US, most lifetime mental disorders show their first occurrence by 24 years, more so in college students compared to non-students of the same age (Hunt & Eisenberg, 2010). If such mental health struggles remain unaddressed or untreated, they may strongly impact doctoral/academic success. Some key aspects that Ph.D. students must simultaneously juggle during this demanding, strenuous journey include culture shocks, financial insecurity, strained supervisor / student relationship, rollercoaster of emotions (stress, depression, homesickness etc.), minimal to almost non-existent social life, severely compromised to non-existent romantic life, parental pressure, academic / emotional burnout, immense pressure from supervisor to work longer hours, to publish peer-reviewed articles in order to timely graduate, etc. Ph.D. students are also expected to pro-actively network and secure new opportunities at conferences, which may eventually lead to them securing jobs / postdocs / tenure-track faculty appointments later.

An extremely challenging aspect that continues to afflict doctoral students is their sheer inability to monitor/assess individual doctoral progress, especially during their impressionable, formative years. Novice graduate students often view their doctoral journey through a rose-tinted lens, and thus fail to recognize the red flags. Indeed, for most students, a Ph.D. is almost a leap of faith into an unknown abyss; the final success of which heavily relies on a series of factors that must cohesively align at the opportune moment and time. Therefore, it becomes even more imperative that a
prospective graduate student chooses the correct Ph.D. program and supervisor, to minimise the time taken to graduate (TTG) and have a meaningful, fulfilling research experience. In this work, we propose a probabilistic mathematical equation, to evaluate the likelihood of a doctoral student succeeding in their program \textit{a priori}, thus providing a tangible quantitative metric to objectively assess their doctoral feasibility. We have structured this article into two major sections – the first performs an exhaustive literature review, to recognize some common design hurdles that all students enrolled in a Ph.D. program encounter (the doctoral completion rate, challenges towards graduation etc.) and highlights recent advancements to understand these aspects qualitatively and quantitatively. This, in turn, sets the tone for the second section, where the Chakraborty-Galatro (CG) equation is proposed a method to gauge individual doctoral feasibility. Individual probabilities that contribute towards the CG equation are simply mathematical representations of key parameters contributing towards doctoral success, as identified from reviewing the scientific literature (in the previous section). We hope that this tool will enable graduate students and guide them towards better identifying to choosing their Ph.D. pathway and their supervisor(s) more consciously.

**Literature Review:** Doctoral research has piqued considerable interest in recent years (Jones, 2013) and continues to be explored and understood, in an increasingly competitive global landscape. As labour markets expand, the demand for skilled doctorates across all fields continues to surge. The doctoral degree or Ph.D. is perceived by academic institutions as the pinnacle of academic success (Park, 2005; Jairam & Kahl Jr., 2012) and, in turn, trains future leaders in industry and academia alike (Millet & Nettles, 2006; Davis \textit{et al.}, 2006). Additionally, doctoral students act as mediators, serving as a critical bridge (Thune, 2009) by linking industry (and its cautious pragmatism) to academia (and its research driven curiosity). However, due to an increasingly competitive academic market and diminishing job prospects, proper choice of a doctoral program becomes even more crucial for a prospective graduate student, to reap maximum benefits out of pursuing higher education (HE). Earning a doctorate is almost a ceremonially “rite-of-passage” and involves successfully navigating several difficult targets. The path towards doctoral graduation and achieving success is often solitary, physically/emotionally draining, socially confining and involves making several substantial sacrifices/compromises; therefore, choosing a capable and emotionally supportive supervisor is pivotal to successfully survive this intensely demanding marathon. Sadly, students pursuing doctoral studies today suffer even more, as they frantically struggle to maintain peak research productivity and academic output, while simultaneously avoiding the urge to consider dropping out, in a global landscape increasing affected with emotional/mental health crises (Nogueira-Martins \textit{et al.}, 2004; Jardon & Choi, 2022; Faisal \textit{et al.}, 2022).

Earning a Ph.D. involves unwavering commitment towards a discipline, a supervisor and a tremendous investment of energy and time towards a research topic for several years (which is easier said than done). Consequently, a substantial proportion of doctoral students end up dropping out of their degree programs (this is termed as doctoral attrition (DA)); likewise, a significant number of those who choose to remain, fail to complete their PhDs. on time. Doctoral graduation timelines across America universities have steadily risen across all fields since 1967; the median time invested to earn a doctorate has climbed up to 7.1 years in 1993, from 6.6 years in 1983 (Ferrer de Valero, 2001). During that same period, the median time escaped between a bachelor’s to a doctoral degree rose from 9.8 to 10.5 years (Thurgood & Clarke, 1995). Doctoral attrition rates are seen to vary substantially, from a modest 10-20% to a moderate 33% recorded from 1996-2006 within the School of Physics and Chemistry at the University of Adelaide (Jiranek, 2010), to a staggeringly high value of 85% (Bourke \textit{et al.}, 2004). Contrary to previous belief, doctoral attrition is not completely gender-neutral (Seagram \textit{et al.}, 1998), and females are known to take almost 11% longer time than males to graduate. While there appears to be no sex bias in hiring for Australia (Jackson and Michelson, 2015), some female students in China fear gender discrimination during job-search (Wang \textit{et al.}, 2019). Likewise, citizenship and visa status play a key (albeit minor) role, with international students finishing their Ph.D.’s 10 months prior than their domestic counterparts, on average. Typically speaking, the sciences tend to exhibit a doctoral attrition rate of 30-40%, while the humanities are at 45-51% (Martin \textit{et al.}, 1999; Wright & Cochrane, 2000; Elgar, 2003). Other research places this number between 33-70% (Ivankova & Stick, 2007; Kim & Otts, 2010; Jiranek, 2010; Gardner & Gopaul, 2012) and notes that several students choose to quit the program within the first year itself (Lott \textit{et al.}, 2010; Jairam & Kahl Jr., 2012). In more recent years, “matching” prospective students to supervisors has perhaps been the only feasible solution proposed in an ever-increasingly online world, to combat and minimize DA rates (Orellana \textit{et al.}, 2016). Working style, role and academic skills of the supervisor are also key factors that heavily influence the extent of success of a Ph.D. candidate (Brown & Atkins, 1988; Quinn, 1988; Cullen \textit{et al.}, 1994; Latona & Browne, 2001; Pearson & Brew, 2002; Buttery \textit{et al.}, 2005; Taylor & Beasley, 2005; Vilkinas & Cartan, 2001, 2006; Vilkinas, 2007). Some key predictors of doctoral success and consequent timely graduation are typically reported to be high grades (Wright & Cochrane, 2000) / entrance scores (Siegel, 2005), strong supportive mentoring, a commensurate financial aid package, a personal student-advisor relationship with minimal conflict (Earl-Novell, 2006), a minimized sense of
self-isolation (Ali & Kohun, 2006; 2007), younger age, full-time enrollment in the program (Martin et al., 1999) and citizenship status (i.e., international over domestic) (Jiranek, 2010). This is by no means a complete checklist, and university-specific factors are acknowledged to also play strong roles in ultimately influencing doctoral attrition trends (Martin et al., 1999), resulting in completion rates ranging anywhere from between 20-66% (these values are recorded for Australian institutions of higher education). The cumulative result of all these design parameters is an effective lengthening of the time to attain a doctorate (TTD) (Wao & Onwuegbuzie, 2011). The TTD is reported to increase significantly in Education, as compared to other fields (Bowen & Rudenstine, 1992), with a median value of 10.7 to 12.7 years in Education vs. 7.7 to 7.9 years in other fields including Education (Hoffer et al., 2007). Increase in the TTD has sparked several studies to identify the parameters responsible; but almost all these studies are completely qualitative and somewhat abstract (Nerad & Cerny, 1993; Bair, 1999; Ferrer de Valero, 2001; Maher, Ford & Thompson, 2004; Stolzenberg, 2006).

There are some typical roadblocks all graduate students undergo, regardless of the field, that makes the whole process challenging. Identifying an existing research gap is usually the first step towards a Ph.D.; this often takes significant time, investment, commitment and is iterative in nature (Luse et al., 2012; Kuhn, 1962). Once a potential research gap is identified, a detailed literature review is conducted (Bourgeois, 1979; Ellis & Levy, 2006) to formally postulate a research hypothesis, that can answer an unaddressed knowledge gap and therefore, qualify as novel, impactful progress (Davis, 1971) worthy of being awarded a doctorate. It must be stressed here that the student needs to ensure they are not inattentive or perceive the problem in a unidirectional, monotonous, routinized fashion (Garfinkel, 1967). This creative alertness is extremely vital to ensure that a student eventually contributes tangibly to a field, by increasing understanding (Creswell, 2005; Leedy & Ormrod, 2005). A significant milestone for a graduate student is their first peer-reviewed publication. The ease with which this is accomplished depends (among other aspects) crucially on the supervisor-student relationship dynamics. Other factors like the academic institute one is publishing from (prestige bias), the global demand and perception of the field of study (subject bias) and the reputation of the supervisor within the academic community in the specific field (reputation bias).

All the points discussed above also relies on one inherent assumption, that the Ph.D. experience remains undisrupted by any unforeseen circumstances, which may not always be true. More recently, the covid-19 pandemic has augmented the already existent severe mental health crisis of the global population, by inducing anxiety, depression, and other psychological symptoms (Xiong et al., 2020, Cullen et al., 2020, Pfeiferbaun & North, 2020; Usher et al., 2020, Dong & Bouey, 2020; Talevi et al., 2020; Vindegaard & Benros, 2020; Moreno et al., 2020) such as Post-Traumatic Stress Disorder (PTSD). Graduate students have been especially impacted during the pandemic, with Ph.D. graduation timelines getting significantly delayed. Major concerns of Ph.D. students during the ongoing pandemic include physical/mental health issues (Zhai & Du, 2020), fear of contracting covid-19 themselves, fear of bereavement of a
loved one, fear due to disruption of experimental work hampering the TTD, etc. Specifically for the US college landscape (Texas A&M University), out of a dataset of 2031 participants comprising undergraduate and graduate students, only 43.25% felt capable of adequately coping with stress, 48.14% exhibited moderate-to-severe depression, 38.48% demonstrated moderate-to-severe levels of anxiety and 18.04% contemplated suicide (Wang et al., 2020). At another US school (University of California, Berkeley), 32% graduate and professional students and 35% undergraduates screened positive for major depressive disorder, and 39% of all three student classes screened positive for general anxiety disorder (Igor et al., 2020). A stark result of this research is that the likelihood of graduate and professional students being diagnosed with major depressive disorder is 2 times, and generalized anxiety disorder 1.5 times more in 2020, as compared to 2019. At Brazil’s Federal University of Sao Paulo (Nogueira-Martins et al., 2004), depressive traits were detected in over 42% of a study group (over 45% for anxiety and depression mixed episodes) of graduate students (majorly women, 146 students, 99 were masters and 47 doctoral students). In fact, depression and anxiety appear to be the most encountered disorders. Graduate students in the humanities (arts, design) are more susceptible to mental health problems, while engineering and business report the lowest mental health treatment rates (Lipson et al., 2016). Under such circumstances, it is but natural for graduate students to contemplate DA, especially when alternative pathways also appear bleak. It is only prudent to assume that there may also arise circumstances beyond a student’s control, which may significantly delay the course of a doctorate. A model that aims to accurately capture the probability of a successful doctorate must also account for unpredictable, force majeure circumstances.

While these prior studies are commendable, exhaustive and acknowledge the possibility of several reasons that may lead towards DA, all these works are more qualitative and inherently assume incoming graduate students to be “sufficiently informed and aware” when choosing a Ph.D. in their respective fields. Reality is far from this; more commonly than not, it is a blind “leap of faith” that a student takes when choosing to pursue a Ph.D., often without little to almost no information about how the next few years are likely shape out. There have been no attempts to comprehensively understand, quantify or track these exact initial factors (using any mathematical frameworks) that a typical graduate student ponders upon, before accepting a Ph.D. offer. One therefore asks the obvious question: are prospective graduate school applicants even aware of what they are essentially “signing up” for? It appears from the literature that a large majority of graduate students are not, which is perhaps the reason why the fire to perform impactful research is often seen to steadily fade away with time, after an initial “honeymoon period” during the Ph.D. journey. With an aim to fill this lacuna, the next section proposes a probabilistic equation to capture these details, so that a graduate student may be able to impartially assess their own doctoral fit and feasibility in a doctoral program, and accordingly, decide on the best career pathway(s).

**Mathematical Model:** In this section, we present the Chakraborty-Galatro (CG) probability equation, that assimilates all key factors identified previously, to predict the Doctoral Success Likelihood (DSL). The DSL is simply an overall probability that incorporates individual probability contributions of the previously identified factors. Mathematically, we may write,

\[
DSL = (P_{SR} \times P_{IR} \times P_{SU} \times P_{S-S} \times P_{T}) - \mathcal{H} \times P_{FM} \quad \text{where} \quad 0 < DSL \leq 1. \tag{i}
\]

Here, \(P_{SR}\) is the subject reputation that indicates how favorably or unfavorably the Ph.D. subject is perceived globally, \(P_{IR}\) is the reputation of a specific academic institute within that subject discipline, \(P_{SR}\) is the supervisor’s actual/perceived reputation, as assessed through relevant metrics (h-index, citations, online talks, grants, awards etc.), \(P_{S-S}\) is a probability factor that captures actual/perceived supervisor-student relationship dynamics, \(P_{T}\) is the average graduation probability for the intended supervisor’s lab (defined in detail later), \(\mathcal{H}\) is the Heaviside step function and \(P_{FM}\) is the probability accounting for any unprecedented, force-majeure circumstance. Each of these probabilities are now described qualitatively, and numerical values are assigned below. While this model is robust, further research on each of these probability factors is needed to create a more thorough scoring scale that succinctly maps real scenarios – this is recognized as future work.

**Subject Reputation (\(P_{SR}\)):** The perceived market reputation of the subject domain strongly depends on the existent socio-political/economic trends that drive market growth (Nerad, 2004). A field in “demand” has a high perceived reputation and is scored \(P_{SR} = 1\), in contrast to a field that is in lesser perceived demand, which is scored \(P_{SR} = 0.9\). For instance, in the last few years, fields like biomedical engineering, data analytics, machine learning and artificial intelligence have seen an unprecedented boom; graduate studies in these fields would therefore be scored \(P_{SR} = 1\). In contrast, fields like religion, women, and gender studies, despite their immense potential are “perceived” in lower market demand, have lower full-time employment opportunities and prospects, and are scored \(P_{SR} = 0.9\). Fields in moderate perceived demand like economics, business, engineering etc. are scored \(P_{SR} = 0.95\). Theoretically speaking,
the $P_{SR}$ value may drop to even lower values, however, $P_{SR} \neq 0$ (as there wouldn’t be the need for a subject if no demand exists).

**Institute Reputation ($P_{IR}$):** It is well-established that the global reputation of the institution one earns their PhD. in significantly influences a graduate’s academic / industrial prospects (Webb et al., 1997). Graduates from high-ranked universities typically enjoy more benefits in terms of reputation by university association, global exposure, career opportunities, and strategic networking. Academic reputation of the institution plays a key role for prospective graduate student (more so for international students choosing an overseas Ph.D. program), a pattern that has been observed in China (Li et al., 2021), is one of the highest ranked factors in the US (Kim et al., 2018), Germany, and the UK (Abbas et al., 2021). The global top 50 institutes in a chosen field globally are scored $P_{IR} = 1$, the next 50 are scored $P_{IR} = 0.95$ and any institutes lower than these are scored $P_{IR} = 0.9$. The top 50 institutes in any field may be easily obtained through online subject rankings such as the QS Subject Rankings, Times Higher Education Rankings, Forbes World University Rankings etc., although there is debate on how “unbiased” the rankings truly are. Nevertheless, a graduate student can judiciously assign a number to this probability index, except $P_{IR} \neq 0$ as every academic institute has some intrinsic reputation.

**Supervisor Reputation ($P_{SU}$):** This probability index accounts for the supervisor’s perceived reputation in their own field (by their research colleagues), with whom the student wishes to work to earn their PhD. It is evident that graduate students are trained differently under different supervisors; and being under a prolific supervisor can often “make-or-break” a student’s prospects, especially for tenure-track jobs in academia (Kim et al., 2018). A supervisor’s success may be decently estimated by looking at some key indicators: published journal articles / conference proceedings, h-index, and citations (excluding self-citations), the career trajectory of past students advised, etc. One may even contact the supervisor’s doctoral / postdoctoral advisors to seek impartial feedback. A study performed at the University of Grenoble; France (Mangematin, 2000) proves that the success of prior students in academia/industry drives a prospective student to work with a supervisor. A supervisor who ideally meets these criteria is scored $P_{SU} = 1$, who somewhat meets these requirements is scored $P_{SU} = 0.95$ and who fails to significantly meet these requirements is scored $P_{SU} = 0.9$. It is of course, more difficult to evaluate newly hired tenure-track faculty based on such objective parameters, and the assessing student should be mindful of this. Academia is a competitive field, and credit must be attributed to anyone who has successfully made through the tenure track system. Most supervisors are very good at excelling at the tenure-track game (otherwise they wouldn’t be there), hence the score for a supervisor’s reputation is recommended not go below 0.9. Thus, $0.9 \leq P_{SU} \leq 1$ are recommended as more realistic bounds.

**Supervisor-Student Relationship ($P_{S-S}$):** This is perhaps the most sensitive and impactful of all the contributing probabilities when it comes to determining the DSL. The supervisor-student relationship is known to be the most impactful factor towards shaping the overall success of a student’s Ph.D. journey (Sverdlik et al., 2018; McAlpine et al., 2020). The benefits of a “good” supervisor-student relationship are immense – it is intellectually rewarding, makes the candidate pro-active, inculcates in them a proper research mindset, and significantly improves their mental health and perceived well-being over their doctoral years. On the contrary, a less conducive supervisor-student relationship can significantly deteriorate doctoral progress, trigger DA, and implant in a graduate student chronic issues like impostor syndrome, inferiority complex and emotional burnout. Evidence shows that there is a prevalence of depression, self-harm, anxiety, and suicidal tendency among graduate students (Lipson et al., 2016). Good supervisors recognize their graduate student as assets and are empathic, supportive, strategic, efficient at conflict resolution, act as an efficient mentor (Creighton et al., 2010) and lead by example. A good exercise for a potential grad student is to do a little background research on the supervisor and how are they perceived by their own prior/current graduate students, as well as within the department. Obtaining multiple (and detailed) feedback/crosstalk from individuals is strongly encouraged, as it tends to portray the supervisor in a more neutral, objective fashion. It is strongly encouraged that significant effort be invested to assess this relationship through the techniques listed above. A relationship that is perceived (or is) conducive towards a good graduate research experience is scored $P_{S-S} = 1$, a relationship perceived as moderate is scored $P_{S-S} = 0.95$ and a bad relationship scored $P_{S-S} = 0.9$.

**Timely Graduation ($P_T$):** Graduating on time continues to be a vital concern affecting the mental health of graduate students (Jiranek, 2010; Pitchforth et al., 2012; Ellenbecker et al., 2017). It is common to hear about supervisors who intentionally delay/extend graduation timelines of their own graduate students, for various motives such as getting more work done. Good supervisors recognize the sacrifice students undertake to earn their Ph.D., as they are essentially alienating themselves from the job market for several years. Students graduating beyond the normal academic timeline of a program almost always have a hard time finding full-time employment immediately, because
of how a delayed timeline is perceived. Most employers view a delayed timeline as failure on the student’s part; while even fewer realize that the graduation timeline is, almost exclusively controlled by the supervisor. Academics have tried to map out this behaviour using Bayesian networks (Pitchforth et al., 2012), and universities have employed strategies to accelerate timely Ph.D. graduation (Gasson, 2015). Calculating this probability requires us to have two data: the average time a doctoral student takes to graduate from the department ($\bar{T}_D$), and the average time a doctoral student takes to graduate from the supervisor’s group ($\bar{T}_S$). The probability $P_T$ is defined as follows:

$$P_T = \begin{cases} 
1 & \text{when } \bar{T}_S \leq \bar{T}_D \vspace{1mm} \\
1 - 0.1 \times (\bar{T}_S - \bar{T}_D) & \text{when } \bar{T}_S > \bar{T}_D
\end{cases}$$

To demonstrate the calculation of this probability, let us consider two cases. Case 1: Let us assume that a supervisor that graduates his students typically in 8 years, as against a departmental average of 6 years ($\bar{T}_S = 8$, $\bar{T}_D = 6$). The probability is then calculated to be $P_T = 1 - 0.1 \times (8 - 6) = 1 - 0.2 = 0.8$. Case 2: Let us assume that a supervisor makes their students graduate on time, or even faster than the department’s average graduation timeline. In this case, the default value of the probability is $P_T = 1$.

The Heaviside / Unit Step Function ($\mathcal{H}$): The Heaviside step function (or the unit step function) behaves as a switch and is set to a default value of $\mathcal{H} = 0$, and only assumes the value of $\mathcal{H} = 1$ when unforeseen circumstances delaying doctoral progress arise. Thus, $\mathcal{H}$ only serves to include or exclude the probability accounting for force majeure, $P_{FM}$.

Force Majeure Circumstances($P_{FM}$): Unforeseen circumstances may significantly alter the PhD. pathway for a graduate student. A classic recent example is the covid-19 pandemic that severely impacted the mental health (Jardon et al., 2022) and graduation timelines of graduate students since 2020. Laboratory/library access was severely restricted for graduate students, which in turn extended doctoral trajectories and the TTD. Also, mass layoffs and hiring freezes continue to occur globally in a post-pandemic economy, fueling even more uncertainty among current doctoral researchers who are at advanced stages. Other examples of force majeure may include – sudden bereavement of family and loved ones, diagnosis of a physical/mental disorder, bankruptcy, natural disasters, unexpected legislative action, lockdowns, slowdowns, strikes, sudden illness/death of the supervisor, etc. An unforeseen circumstance that impacts a student’s doctoral progress strongly is scored $P_{FM} = 0.3$, one that impacts moderately is scored $P_{FM} = 0.2$ and one that impacts weakly (but is strong enough to not be completely ignored) is scored $P_{FM} = 0.1$.

Results: The results of our proposed model may now be evaluated and summarized for three typical cases – strong, moderate, and weak (Table I). Except for the probability $P_T$, only a simple multiplication needs to be performed for each of the individual probabilities. We also highlight here that the procedure used to propose the CG equation is generic, and therefore, can be used universally across all fields to predict DSL. Typical values are calculated for two representative cases (timely graduation (Table II) vs. a one-year delay (Table III)), primarily to compare scenarios. The relative shading in the Table II and Table III represents the desirableness of outcomes – darker the shade, more desirable is the outcome (as higher the DSL), and more likely is the probability of having a fulfilling doctoral experience. We also define an “infeasibility triangle”, represented by the pink triangle, as the region in the table with the least shading, where the DSL drops to a value of 0.55 or lower ($DSL \leq 0.55$). The choice of this cut-off probability is purely individual, and different risk-averse students may define their own infeasibility triangles/regimes by choosing a different cut-off value, depending on individual risk-reward dynamics. With a one-year delay arising from unforeseen circumstances ($\bar{T}_S = 6$, $\bar{T}_D = 5$, $P_T = 0.9$), the area covered by the infeasibility triangle increases (refer to Table III), as there are now more probabilistic outcomes for the Ph.D. experience to fail. The final decision is of course, a purely individual one, but we hope that the procedure outlined here can provide concrete insights in an objective, unbiased fashion for both prospective and currently enrolled doctoral candidates. Future studies may be performed with a sufficiently large dataset of graduate students ($N \sim 1000$) who initially self-assess their DSL using the CG equation, and then track it with time over several years, during their doctorate journey. Future researchers may use the CG equation to not only validate our model but will also help identify typical limits of the infeasibility triangle, across various fields (and perhaps, across different countries). Our approach may also be employed to track attrition rates and success likelihoods for undergraduate studies (Triventi & Trivellato, 2009).
As a personal example, Sourjeet Chakraborty (SC) is presently a doctoral candidate enrolled at the University of Toronto’s Department of Chemical Engineering & Applied Chemistry. He employed this framework to keep track of his own doctoral progress, and anticipate the TTD, from 2019-2022. He is scheduled to graduate in December 2022; and his individual results are summarized below.

2019 – 2020: He rated $P_{SR} = 0.95$ because chemical engineering is still perceived as moderate to highly desirable in industry/academia alike. $P_{IR} = 1$ as the University of Toronto is one of Canada’s finest academic institutions, and consistently ranks among the top 50 universities of the world. His supervisor is fairly well-known in his field, and thus, he ranked $P_{SU} = 0.95$. He scored $P_{S-S} = 0.95$ as he qualified his relationship with his supervisor as moderate. For Ph.D. candidates in his specific research lab, students typically graduate in 5.5 to 6 years ($\bar{T}_S = 6$), as against a departmental average of 5.6 years ($\bar{T}_D = 5.5$); this number is obtained from the department’s self-study report, published in 2020. Thus, $P_T = 1 - 0.1 \times (\bar{T}_S - \bar{T}_D) = 1 - 0.1 \times (6 - 5.5) = 0.95$. In 2019, the covid-19 pandemic was just beginning, and a very weak case of force majeure ($H = 1$) was identified, thus $P_{FM} = 0.1$. Thus,

$$DSL_{2019–2020} = (P_{SR} \times P_{IR} \times P_{SU} \times P_{S-S} \times P_T) - P_{FM} = (0.95 \times 1 \times 0.95 \times 0.95 \times 0.95) - 0.1 = 0.715$$

This is a high DSL value (71.5%) and lies above the infeasibility probability of 55%. Therefore, SC continued with graduate school, steadily working towards completion.

<table>
<thead>
<tr>
<th>Contributing Probability</th>
<th>Weak Case</th>
<th>Moderate Case</th>
<th>Strong Case</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_{SR}$</td>
<td>0.9</td>
<td>0.95</td>
<td>1</td>
</tr>
<tr>
<td>$P_{IR}$</td>
<td>0.9</td>
<td>0.95</td>
<td>1</td>
</tr>
<tr>
<td>$P_{SU}$</td>
<td>0.9</td>
<td>0.95</td>
<td>1</td>
</tr>
<tr>
<td>$P_{S-S}$</td>
<td>0.9</td>
<td>0.95</td>
<td>1</td>
</tr>
<tr>
<td>$P_{FM}$</td>
<td>0.3</td>
<td>0.2</td>
<td>0.1</td>
</tr>
</tbody>
</table>

Table II: DSL values for different force-majeure conditions, assuming timely graduation for a PhD. student.

<table>
<thead>
<tr>
<th>DSL</th>
<th>Weak Case</th>
<th>Moderate Case</th>
<th>Strong Case</th>
</tr>
</thead>
<tbody>
<tr>
<td>No $P_{FM}$</td>
<td>0.656</td>
<td>0.815</td>
<td>1</td>
</tr>
<tr>
<td>Weak $P_{FM}$</td>
<td>0.556</td>
<td>0.715</td>
<td>0.9</td>
</tr>
<tr>
<td>Moderate $P_{FM}$</td>
<td>0.456</td>
<td>0.615</td>
<td>0.8</td>
</tr>
<tr>
<td>Strong $P_{FM}$</td>
<td>0.356</td>
<td>0.515</td>
<td>0.7</td>
</tr>
</tbody>
</table>

Table III: DSL values for different force-majeure conditions, with a 1-year graduation delay for a Ph.D. student.

<table>
<thead>
<tr>
<th>DSL</th>
<th>Weak Case</th>
<th>Moderate Case</th>
<th>Strong Case</th>
</tr>
</thead>
<tbody>
<tr>
<td>No $P_{FM}$</td>
<td>0.591</td>
<td>0.733</td>
<td>0.9</td>
</tr>
<tr>
<td>Weak $P_{FM}$</td>
<td>0.491</td>
<td>0.633</td>
<td>0.8</td>
</tr>
<tr>
<td>Moderate $P_{FM}$</td>
<td>0.391</td>
<td>0.533</td>
<td>0.7</td>
</tr>
<tr>
<td>Strong $P_{FM}$</td>
<td>0.291</td>
<td>0.433</td>
<td>0.6</td>
</tr>
</tbody>
</table>

2020 – 2021: This was the worst year globally for the covid-19 pandemic, with labs being shut down at the University of Toronto. SC’s experiments were put on hold, as the situation became especially challenging for international graduate students like him. However, he did not feel DA, but instead, focused his efforts on a computational work that
The revised values of these probabilities were \( P_{SR} = 0.95, P_{IR} = 1, P_{SU} = 0.95, P_{SR} = 1, P_T = 0.95 \) and \( P_{FM} = 0.3 \). His relationship with his supervisor improved more, because of regular Zoom meetings, and was helpful in helping him stay on track, even as the world was in a pandemic (\( H = 1 \)), which is reflected in the higher value of \( P_{FM} \). Therefore,

\[
DSL_{2020-2021} = (P_{SR} \times P_{IR} \times P_{SU} \times P_{3-s} \times P_T) - P_{FM} = (0.95 \times 1 \times 0.95 \times 1 \times 0.95) - 0.3 = 0.557
\]

The DSL value (55.7%) for 2020-2021 came almost close to the infeasibility limit (55%); but one recognizes that the reduction of the DSL arises primarily due to an increase in the \( P_{FM} \) value. SC decided to go ahead, and work towards his timely graduation, since the end seemed nearer.

\( 2021-2022: \) Analysis for this year is still undergoing, but based on present circumstances (where SC’s doctoral committee has approved him to write his dissertation and schedule a defense in December 2022), the new values assigned to the probabilities are \( P_{SR} = 0.95, P_{IR} = 1, P_{SU} = 0.95, P_{SR} = 0.9, P_T = 0.95 \) and \( P_{FM} = 0 \). Two of the values have changed now; \( P_{SR} = 0.9 \) because his doctoral advisor had a medical issue that minimized his ability to advise (\( H = 1 \)). This is also then reflected in \( P_{FM} = 0.1 \), as he has had to substantially rely on and seek feedback from his other Ph.D. committee members, to ensure timely graduation. The new DSL value is,

\[
DSL_{2021-2022} = (P_{SR} \times P_{IR} \times P_{SU} \times P_{3-s} \times P_T) - P_{FM} = (0.95 \times 1 \times 0.95 \times 0.9 \times 0.95) - 0.1 = 0.672
\]

The DSL value for this year (67.2%) increased, primarily due to the eventual end of the covid-19 pandemic (which led to labs being opened again at the University of Toronto). Thus, the CG equation serves as a quick quantitative metric to objectively assess an individual’s DSL. The metric may also potentially serve as impartial feedback which the candidate may use to improve on potential areas, to better enjoy the Ph.D. (and graduate school) experience.

Conclusion: The present work performs a deep dive into graduate school life and workplace dynamics of a typical doctoral candidate and identifies key design parameters influencing DA and the TTG. The issue highlighted is especially relevant during current times, with doctoral graduation timelines being delayed globally due to the covid-19 pandemic and a rise its associated repercussions among graduate students (mental health crisis, financial stability, career insecurities, etc.). The key factors crucial towards ensuring a high DSL are identified, and, building on this, the probabilistic CG equation is proposed as a self-assessment metric for the student. Evidence suggests that the key parameters identified are most sensitive towards determining the eventual success/failure of a graduate school undertaking. The DSL assumes a best-case value of 100%, and a worst-case value of 35.6%, in the absence of unanticipated circumstances. Should force majeure circumstances arise, these probabilities are reduced to 90% and 29.1% (for a 1-year delay) and 80% and 22.5% (for a 2-year delay) respectively. A subtle point to not is that the probability incorporating force majeure conditions \( P_{FM} \) is also likely to influence indirectly the probability of timely graduation \( P_T \), these probabilities are therefore, not purely independent. This observation is also intuitive; for instance, the recent covid-19 pandemic is a force majeure circumstance \( P_{FM} \), that has, and continues to delay global PhD graduation timelines \( P_T \). While a higher value of the DSL does predict a more favourable outcome, care must be taken to choose individual probability values rationally; for it is only too easy to fall into the trap of confirmation bias, by setting all values to unity, and feeling content, which may not necessarily mirror actual reality. Choosing a favorable vs. unfavorable outcome primarily relies on choosing an appropriate, practical cut-off probability; this choice may somewhat be circumvented if all other probabilities in the CG equation are at high values. In other words, if one chooses to pursue doctoral studies in a “desirable” field, within the top 50 universities of the world, under a globally recognized supervisor, who ensures that students timely graduate, and builds a meaningful mentorship/feedback-based relationship with them, one will likely receive maximum benefit from the Ph.D. experience. We are hopeful that the quantitative assessment tool presented here is adopted universally, across multiple disciplines, fields, across several countries; and will lead to a minimization of DA rates, and the TTD, for prospective/current graduate school students. This, coupled with other initiatives to catalyze student empowerment (Mousavi et al., 2018), will likely go a long way in transforming the doctoral experience of students, towards more positive pathways. Systematic tracking of such factors would also help in providing supervisors feedback about their advisory capacities and recognize aspects for improvement. For instance, several Ph.D. graduates in Australia report that they would prefer active industry-based mentoring (leading to knowledge and skills transfer) and interdisciplinary research opportunities during their doctoral journey, which would have helped them enter the job market sooner (Manathunga et al., 2009).
Statements and Declarations:
There are no declarations to be made.

Competing Interests:
The authors declare no competing interests.

References:


