On the Problems of the Student

by

Matthew S. Chiasson

An essay submitted to the Department of Economics
in Partial fulfillment of the requirements for
the degree of Master of Arts

Queen's University
Kingston, Ontario, Canada

July 2008

Abstract

In this paper, I model several aspects of Student decision making and classroom competition, within a game-theoretic framework. Specifically, in Part 1, I develop a Two-Period Learning Model (TPLM) to show how uncertainty over course difficulty prior to registration can lead to problems of overconfidence or underconfidence in equilibrium (the former case, providing a partial explanation for the problem of dropping-out). In Part 2, I consider the ways in which a classroom resembles a First-Price-All-Pay Auction and develop a Winner-Take-All Model (WTAM) to show how competition for the 'top-mark' affects effort in equilibrium. Within this competitive framework I show that increasing the class size has an *encouraging-effect* on the outcome attainment of high ability students, and a *discouraging-effect* on the outcome attainment of low-medium ability students. It turns out that, on average, the second effect dominates so that if we are concerned with maximizing the *average* educational outcome of the classroom then class-sizes should be kept small to reduce competitive discouragement; whereas, if we are only interested in maximizing the educational outcomes of the most elite, then we ought to increase class sizes to induce greater competition. Furthermore, I show that equilibrium effort is, at first, decreasing in ability (extra productivity is simply traded for leisure) but ultimately it increases in ability as students become more in contention for the 'top-mark'. I call these two effects the *leisure-substitution effect* and the *competitive effect*, respectively. Finally, I introduce several other settings that possess the essential features of classroom competition and discuss how my results may be carried over. As usual, I conclude with a brief discussion of some of the possible areas in which future research could be directed.

Acknowledgments

"I can no other answer make but thanks, And thanks, and ever thanks."

- William Shakespeare, *Twelfth Night*

First of all, I would like to thank my supervisor, Dr. Ruqu Wang, for his helpful advice, support and words of encouragement - always urging me to "keep up the good work" when I was struggling. I would also like to thank him for his time and patience in helping me *write* this paper, seeing it through many preliminary drafts. Without his prompt email responses this paper would not have been completed before August 1st. Lastly, I'd like to thank him for introducing me to graduate Microeconomic Theory in ECON 810. It is there that I learned about games of Imperfect Information and it was in his final exam that I was first introduced to the idea of a First-Price All Pay Auction.

I'd also like to thank the many friends I've made this year for their helpful conversations and suggestions - they were a great sounding board for ideas, but even more importantly, when the time was appropriate they provided an escape from Economics. Thank you to Queen's University, and in particular the Queen's Economics Department, for providing me with funding, online resources, and an office in which to work.

On a lighter note, I'd like to thank Mountain Equipment Co-op for their design of a lightweight, comfortable sleeping mat and blanket; the guys at the 'chip truck' for keeping me nourished; the ladies at Tim Horton's for keeping me caffeinated; and the many users of LaTeX for their online tutorials and cheat sheets.

Lastly, I'd like to thank all of the great teachers in the world for the work that they do each and every day. In particular, I'd like to thank those that helped me along the way, especially Mr. Chris Aube, Dr. Cathy Baker, and Dr. Niels Anthonisen.

Contents

List of Figures

List of Tables

0 Introduction

> "Begin at the beginning and go on till you come to the end; then stop."
>
> Lewis Carroll, *Alice in Wonderland*

When one dedicates themself to the study of any particular discipline for an extended period of time, they invariably undergo a period of indoctrination whereby they train themself to view everything in the world around them as being simple manifestations of the principles underlying their discipline. It is this sort of repeated exercise that leads the Physics student to exclaim, "Everything is Physics!", and the Philosophy student to ask, "What *isn't* Philosophy?". This is none more true than in the study of Economics, the social science which finds applications in most every nuance of human and societal decision making. Just as the Psychology student is eager to diagnose their friend with an obscure mental disorder at the most subtle suggestion of a symptom, the Economics student quickly attempts to assign rationality (or lack thereof) to any social interaction, no matter how insignificant. Just as the young Medical student worries that his/her flu-like symptoms are really the early signs of a rare, incurable infection, the student of Economics worries that his/her behaviour is not a 'best response' given the complex strategies chosen by everyone else! Learning the language and paradigms of the discipline, and becoming proficient at applying them to everyday life is almost a rite of passage for the young student of Economics. Inevitably, this activity leads one to try to reconcile their *own* behaviour with what the prevailing theory would predict, and hopefully eventually, reconcile the prevailing theory with one's own observed behaviour. It is this very type of critical introspection, relating the Principles of Economics to my own life, that has lead me to ask the following question, "Am I a rational student?", and perhaps more importantly, "What does it even mean to be a rational student?"

For years, students of Economics carefully grind a lens through which to view the world around them; after all, the more powerful the lens, the deeper the student can penetrate the problems that they encounter in Research and everyday life. While a good lens is great for peering deep into outer space, it is easily forgotten that some of the most interesting problems are in the foreground, and can be examined with the naked eye. Too often we forget that the lens can also be used as a mirror, for which to examine ourselves. It is in this spirit that I write my Master's essay on the topic of the Student, an occupation I have held for the past 17 years, but which I have yet to encounter as an Economic application outside the realm of job-market signaling.[1] Are their other salient features of Student behaviour that can be modeled using the standard Economic toolset, and if so, why is it that the first Economics-related result to the Google Scholar search query 'The Student's Problem' is an obscure paper from 1947 entitled "The Generalization of "Student's" Problem when Several Different Population Variances are Involved"? These remarks are admittedly facetious, and there is indeed innumerous volumes (both empirical and theoretical) that have been used to describe various aspects of individual decision making (which subsumes mere 'Student' behaviour). However, the problem still remains, can we usefully *apply* well known game-theoretic tools and results to the specific setting of the classroom, and if so, does this lend us any interesting policy implications for administrators?

In this paper I study two such applications. In Part 1, I consider one of the most common problems faced by the student, and that is, the decision of whether or not to register for a particular course, and subsequently, how much effort to allocate to studying, when the difficulty of the course is not known prior to completion of a preliminary task (midterm). In other words, the student must make the registration and effort-allocation decisions with only imperfect information regarding course difficulty. To address this problem, I develop a Two-Period Learning Model (TPLM) in Section 1.1. Specifically, I assume that the student enters period 1 with a prior

[1] Which reduces the classroom to nothing more than a costly exercise that the student must participate in to convey that they are of high-ability.

belief on their ability (relative to the course difficulty) and this determines whether they will register for the given course or not. If they register, they choose an effort level on a preliminary task (midterm) and achieve an outcome that is a deterministic function of their effort and their *true* ability. In period 2, based on the realized outcome from this preliminary task they can infer (learn) their true ability, and based on this knowledge, they decide whether they want to continue with the course and provide effort for the final task, or whether they would prefer to cut their losses and drop-out. If the student's final grade, which is a weighted average of the outcomes of the two tasks, exceeds some predetermined level, then the student 'passes' and obtains a fixed reward; otherwise, they 'fail' and obtain nothing. In this way, I capture the essential problem faced by the student that they only truly learn what they have 'gotten themself into' after they write a midterm and learn how difficult the course truly is. In Section 1.2, I formally derive the Subgame Perfect Nash Equilibrium of this model by backwards induction. This analysis leads to a discussion of scenarios of overconfidence, underconfidence, and perfect knowledge, in Sections 1.3, 1.4 and 1.5 respectively.[2] Next, I discuss the policy implications of this model in Section 1.6, namely, the practical ways in which Student welfare can be improved by reduction of uncertainty. In Section 1.7, I analyze the problem faced by the benevolent professor (social planner) in choosing the difficulty of the course, and lastly, I discuss some of the important drawbacks/limitations of the model in Section 1.8.

In Part 2, I switch gears and consider the ways in which classroom competition resembles a First-Price All Pay Auction. I argue that the typical classroom incentive structure is one that rewards both absolute performance (achieving some absolute outcome level, as in the previous model) and also relative performance (high class rank). In this way, I use Section 2.1 to model the classroom as a Winner-Take-All Market (WTAM) where all students are rewarded for passing, but only the student

[2] The first scenario provides a partial explanation for the phenomena of 'dropping-out' which I consider briefly.

with the top-mark receives additional rewards. Specifically, I consider the classroom to be an auction where abilities are private information, and effort levels are chosen to create outcomes - effectively bids - for the top mark. Regardless of whether the student achieves the top mark (i.e. wins the 'auction') they still incur the costs of effort required to produce their bid, and it is in this sense that the 'auction' is first-price *all-pay*. After a careful derivation of the equilibrium of this model in Section 2.2, the symmetric equilibrium outcome and effort functions are analyzed in Sections 2.3 and 2.5 respectively. In Section 2.3, I show that increasing class size has an *encouraging-effect* on the outcome attainment of high ability students, and a *discouraging-effect* on the outcome attainment of low-medium ability students. In Section 2.4, I consider the average outcome attainment of the class and show that is unambiguously decreasing in class-size. I then use this result to weigh-in on the infamous 'small class-size debate', arguing that if we are concerned with maximizing the average educational outcome of the classroom then class-sizes should be kept small to reduce competitive discouragement; whereas, if we are only interested in maximizing the educational outcomes of the most elite, then we ought to increase class sizes to induce greater competition. In Section 2.5, I show that equilibrium effort is, at first, decreasing in ability (the extra productivity is simply traded for leisure) but ultimately it increases in ability as students become more in contention for the 'top-mark'. I call these two effects the *leisure-substitution effect* and the *competitive effect*, respectively. In Section 2.6, I consider the robustness of these results to different types of density function specifications (i.e. different, more realistic, distributions of classroom ability). This leads to a discussion of *superstar effects* which arise when the density function is bell-shaped, and which leads students of extremely high ability to exert less effort ("rest on their laurels") since they face little challenge from their peers. In Section 2.7, I discuss some of the key limitations of this model and suggest some areas for improvement. Then, in Section 2.8, I introduce several other settings that possess the essential features of classroom competition and discuss how my results carry over. Finally, in Section 3, I offer some concluding remarks on each of the two models presented and discuss some areas for future work.

1 Part 1: The Course Registration and Effort-Allocation Problem

Undoubtably, two of the most important problems that a Student must confront are deciding which courses to register for, and how much effort to exert in those courses. The problem is non-trivial since registration and effort are costly, and different courses offer different educational rewards. For some, these decisions are easy and the result is a distribution of courses that provide them with the perfect balance of rewards and costs. For others, these decisions prove to be difficult, and the result is a timetable that is either unmanageably hard (in the extreme case leading to failure or withdrawal), or so trivially easy that the educational rewards don't outweigh the costs of registration. If the relative costs and benefits of taking a course could be perfectly forecasted, then such a problem would cease to exist; everyone could (and presumably would) choose the most optimal line-up of courses, subject to the various graduation requirements imposed in the Academic Calendar. This leads naturally to the question of *why* students can't form such perfect forecasts? The answer, it seems to me, lies in the fact that, prior to registration, the student only has imperfect information concerning their ability, relative to the course difficulty. After all, in most situations, the student's information set regarding a particular course is limited to a few, relatively vague lines of description from a course calendar. There is, of course, casual information sharing that occurs between students who have already experienced certain courses and certain professors, and there are often well-defined periods of 'course shopping' whereby students can collect additional information by attending preliminary lectures (for instance, picking up the course syllabus or reading list). Even still, such things only resolve small bits of uncertainty regarding course difficulty, and the problem remains. We characterize this observation with the following assumption,

Assumption 1.1 *Before registering for a course, Students only have imperfect information over their ability relative to the course difficulty.*

Thus, any model of the course-registration decision must incorporate this element of uncertainty. Now then, what can we say about the effort-allocation problem? Students receive rewards based on their academic achievement (eg. pass/fail rewards, distinctions for exceeding a predetermined cut-off grade such as the Dean's List, etc.) but how exactly does effort translate into achievement? There are obviously many ways to model this type of interaction between effort and outcome, the most obvious being the production function approach which views the educational outcome (aka. the 'mark' or 'grade') as being a function of a variety of factors or inputs (eg. raw talent, time spent studying, difficulty of the course, etc.). Here, for parsimony we make the following assumption,

Assumption 1.2 *The outcome achieved on a task (test) is a multiplicative function of effort and ability, where both factors exhibit diminishing returns. In other words, $Outcome = f(effort) \times g(ability)$ where f and g are both increasing concave functions.*[3]

This assumption, I think, is sufficiently general to encompass many plausible types of outcome functions but yet it is specific enough to give us something to work with for purposes of analysis. Notice that ability and effort have remained essentially undefined here. By 'effort' we mean any costly action taken to increase ones propensity to do well on the task (memorizing notes, rehearsing old exam questions, attending lectures, etc.), and by 'ability' we will mean any previously acquired, temporarily immutable quality that tends to increase one's propensity to perform well (high IQ, previous experience with the subject, mathematical training, ability to cope with

[3]When we speak of 'ability' we will always mean ability relative to the difficulty of the course.

stress, test-writing skills, etc.). Thus, for our purposes effort will be a choice variable and ability will be unchangeable. With these assumptions in hand, we are now prepared to introduce a model of the course-registration and effort-allocation decision.

1.1 The Two-Period Learning Model

Consider the following Two-Period Learning Model (TPLM):

Period 0: A prior on ability, θ, is formed exogenously. Here we will assume that the prior, which generally takes the form of a cdf $F(\theta)$, is simply a point $\hat{\theta}$, or equivalently, that $F(\cdot)$ is simply a one-point distribution that places all of the mass on $\hat{\theta}$. This assumption is for analytical tractability, however, in the appendix we relax it and derive a condition analogous to the one we achieve below.

Period 1: Based on their prior, the Student decides whether to Register for the course or not. Registration costs R, not registering costs 0. If the student registers, they choose effort level e_1 on a preliminary task (midterm) and achieve outcome u_1 given by the following deterministic outcome function:

$$u_1 = f(e_1) \cdot g(\theta) \tag{1}$$

where $f(\cdot)$ and $g(\cdot)$ are both concave, strictly increasing, passing through the origin and bounded above.[4] Based on the realized outcome from the task the student can infer their *true* ability θ,

$$\theta = g^{-1}\left(\frac{u_1}{f(e_1)}\right) \tag{2}$$

[4]For convenience we will assume that f and g are bounded above by 1 so that $u \in (0, 1)$ and $100 \cdot u$ can be thought of as a percentage grade.

Finally, effort costs $c(e)$ where $c' > 0, c'' > 0$, and $c(0) = 0$.[5]

Period 2: Here the Student decides whether to Continue with the course or Drop-Out. Dropping-out is costless, but the student still bears the registration/effort costs borne in the previous period. If the Student continues then they choose effort level e_2 which determines outcome u_2 on the final task (exam). Based on the two outcomes, a final mark is calculated according to:

$$u = \alpha \cdot u_1 + (1 - \alpha) \cdot u_2 \qquad \alpha \in (0, 1) \tag{3}$$

Here α represents the weight placed on the preliminary task (which the reader can assume to be less than or equal to $\frac{1}{2}$ in our interpretation). If the student's final mark, u, exceeds a predetermined level $\overline{u}$, the passing grade, then the student gets a payoff V. Otherwise, they fail and get payoff 0.

It is assumed that $V > R$ so that at least some students, with sufficiently large ability, will find it worthwhile to register and put in effort to pass. Finally, we assume no discounting ($\delta = 1$) for simplicity. For clarity, we illustrate the game in time-line form and present a payoff summary below.

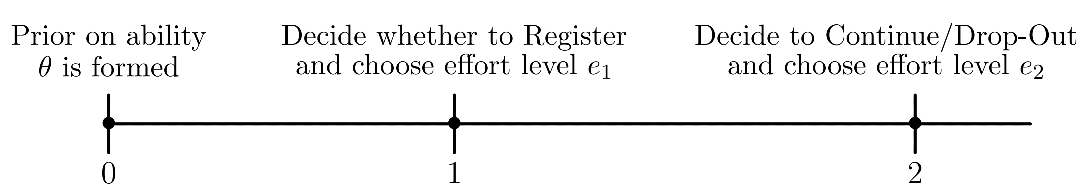


Figure 1: Timeline of TPLM

[5]There is no doubt that, in some cases, utility can be derived from the pure process of studying so that pure effort leads to a utility gain. Rather than incorporate this into the payoff function, we simply reinterpret $c(\cdot)$ to be the net cost of studying, and allow it to include the opportunity cost of not engaging in other welfare-improving activities.

Don't Register	$\Longrightarrow$	0	
Register then Drop-Out	$\Longrightarrow$	$0 - R - c(e_1)$	
Register then Continue	$\Longrightarrow$	V - R - c(e_1) $- c(e_2)$	if $u \geq \overline{u}$
		0 - R - c(e_1) $- c(e_2)$	if $u < \overline{u}$

Table 1: Payoff Summary

1.2 Equilibrium of the model

To derive the Subgame Perfect Nash Equilibrium of this model we start in Period 2 and proceed by backwards induction. Given u_1 observed at the end of Period 1, the student can extract their ability θ perfectly according to (2). To "pass" the students needs to obtain an outcome u_2 which satisfies $\alpha \cdot u_1 + (1-\alpha) \cdot u_2 \geq \overline{u}$. This requires,

$$\begin{aligned} &\Longrightarrow \quad u_2 \geq \tfrac{\overline{u}-\alpha \cdot u_1}{(1-\alpha)} \\ &\Longrightarrow \quad f(e_2) \cdot g(\theta) \geq \tfrac{\overline{u}-\alpha \cdot u_1}{(1-\alpha)} \\ &\Longrightarrow \quad e_2 \geq f^{-1}\left(\tfrac{\overline{u}-\alpha \cdot u_1}{g(\theta) \cdot (1-\alpha)}\right) \end{aligned} \tag{4}$$

Since effort is costly and the outcome of the task is deterministic the student will, at best, set e_2 to satisfy (4) with equality. Let $\underline{e_2}$ be this effort level - it represents the minimum amount of effort that the student can put in to "pass" in the second stage. We assume here that $\underline{e_2}$ exists and is positive so that the student cannot guarantee a "passing grade" after the preliminary task (this is obviously a reasonable assumption in most cases).[6] If $V \geq c(\underline{e_2})$ then the student will indeed choose to Continue-On with the course and supply effort level $\underline{e_2}$. On the other hand, if $V < c(\underline{e_2})$ then the student will find it optimal to Drop-Out. Notice that from the perspective of Period 2, both R and $c(e_1)$ are regarded as sunk. We now consider Period 1.

[6]Note that if $g(\theta) < \frac{\overline{u}-\alpha \cdot u_1}{1-\alpha}$ then $\underline{e_2}$ will not exist because there is no chance of passing even with $e_2 = \infty$. In this case we can just define $\underline{e_2} = \infty$ so that $c(\underline{e_2}) = \infty$ and the rest of the arguments proceed as usual.

Here, the student decides whether or not to Register. In contemplating the costs and benefits of Registration the student would optimally choose effort to solve the following planning problem (taking $\hat{\theta}$ to be their perceived ability).

$$\max_{e_1,e_2} \; V - c(e_1) - c(e_2) \qquad \text{subject to:} \quad \begin{aligned} \alpha \cdot u_1 + (1-\alpha) \cdot u_2 &\geq \overline{u} \\ e_1, e_2 &\geq 0 \end{aligned}$$

It is easy to verify that the FOC of this problem is given by,

$$\frac{\alpha}{1-\alpha} \cdot \frac{\frac{\partial u_1}{\partial e_1}}{\frac{\partial u_2}{\partial e_2}} = \frac{c'(e_1)}{c'(e_2)} \tag{5}$$

This, of course, just says that e_1 and e_2 will be chosen so that $MRS_{e_1,e_2} = MRT_{e_1,e_2}$ For simplicity, we will now assume that $\alpha = \frac{1}{2}$ so that $\frac{\alpha}{1-\alpha} = 1$, although we will relax this later when analyzing the equilibrium. Subbing in the deterministic outcome function and rearranging we find that (5) reduces to,

$$\underbrace{\left(\frac{1}{f'(e_1)}\right) \cdot c'(e_1)}_{h(e_1)} = \underbrace{\left(\frac{1}{f'(e_2)}\right) \cdot c'(e_2)}_{h(e_2)} \tag{6}$$

Evidently, $h(\cdot)$ is one-to-one since it is the product of two continuous, monotone increasing, functions.[7] Hence (6) is satisfied iff $e_1 = e_2$,

$$\begin{aligned} \Longrightarrow \quad e_1^* &= e_2^* \; = \; e^* \\ \Longrightarrow \quad u_1^* &= u_2^* \; = \; \overline{u} \\ \Longrightarrow \quad e^* &= f^{-1}\left(\frac{\overline{u}}{g(\hat{\theta})}\right) \end{aligned} \tag{7}$$

To ensure that e^* exists (and is positive) we require $\overline{u} < g(\hat{\theta})$ which simply says that it it possible for someone with ability $\hat{\theta} > 0$ to pass by providing sufficiently large effort

[7]Note that $f'' < 0$ implies $\frac{1}{f'(\cdot)}$ is increasing

(i.e. by making $f(e)$ sufficiently close to 1). This seems like a reasonable assumption to make, most people of even modest ability can pass a course provided they work hard enough. So, if the student decides that they want to Register and "pass" they will want to set $e_1^* = e_2^* = e^*$ according to (7) above. Therefore, they will only register if,

$$\underbrace{c(e_1^*) + c(e_1^*) + R}_{\text{Expected Cost of Passing}} \quad \leq \quad \underbrace{V}_{\text{Expected Benefit of Passing}} \tag{8}$$

$$\begin{aligned}
\Longrightarrow \quad 2 \cdot c(e^*) &\leq V - R \\
\Longrightarrow \quad e^* &\leq c^{-1}\left(\frac{V-R}{2}\right) \\
\Longrightarrow \quad f^{-1}\left(\frac{\overline{u}}{g(\hat{\theta})}\right) &\leq c^{-1}\left(\frac{V-R}{2}\right) \\
\Longrightarrow \quad \hat{\theta} &\geq g^{-1}\left(\frac{\overline{u}}{f\left(c^{-1}\left(\frac{V-R}{2}\right)\right)}\right) \equiv \theta_c
\end{aligned} \tag{9}$$

Thus, provided the student is sufficiently confident so that $\hat{\theta} \geq \theta_c$, the student will decide to register for the course; otherwise, they will not Register. Note, however, that if the 'true' θ were known then the student would only register if $\theta \geq \theta_c$. Thus, we may now state our first Theorem and consider several different scenarios:

Theorem 1.1 The unique subgame perfect Nash Equilibrium of (TPLM) is the following strategy: In Period 1, the student Registers iff $\hat{\theta} \geq \theta_c$ and supplies effort $e_1 = e^*$ as in (7). In Period 2, the student Continues-On with the course and supplies effort $e_2 = \underline{e_2}$ as in (4) iff $V \geq c(\underline{e_2})$; otherwise, they Drop-Out.

1.3 Overconfidence: $\hat{\theta} \geq \theta_c > \theta$

> "Well, I think we tried very hard not to be overconfident, because when you get overconfident, that's when something snaps up and bites you."
>
> -Neil Armstrong

Here, the student overestimates their raw ability (or alternatively underestimates the course difficulty) and registers when they shouldn't. They supply e_1^* in Period 1 and achieve an outcome u_1 which is less than what they expected (i.e. less than $\overline{u}$). Thus, after the preliminary task, the student learns their true ability θ and discovers that they are currently 'failing'. To pass, additional effort must be supplied in Period 2. If $V \geq c(\underline{e_2})$ then they will Continue (it is worthwhile for them to increase effort to pass given that R and $c(e_1)$ are now sunk); otherwise, they will decide to Drop-Out.

Application: Dropping-Out

The result above offers a partial explanation for the attrition phenomena observed in high school and University and accords with results found in empirical studies and surveys of drop-outs. Lloyd (1978) discovered that as early as the third grade, dropouts differed significantly from graduates in IQ level, marks received in course work, parent's occupational and educational level, and tested reading, arithmetic, and language skill achievement - suggesting that students who eventually drop-out significantly differ from their peers in terms of 'ability' (i.e. they have significantly lower θ's). Eckstein and Wolpin (1999) developed and structurally estimated a sequential model of high school attendance and work decisions. The model's estimates implied that youths who dropped-out of high school had different *traits* than those who graduated - they had lower school ability and/or motivation, they had lower expectations about the rewards from graduation, they placed a higher value on leisure and had a lower consumption value of school attendance.

In a comprehensive report released by FEDA (Further Education Development Agency) entitled "9,000 voices: student persistence and drop-out in further education" the authors surveyed 9,000 students from 31 colleges across the UK and identified a variety of drop-out predictors. In particular, they found that withdrawn students were much more likely than current students to believe that they have not been placed on the most appropriate course (In fact, this was the best predictor of student drop-out). Moreover, they found that student evaluations, which related ultimately to their study skills or to their confidence in their study skills, were predictors of dropouts. The authors did, however, acknowledge that these 'factors' were not operating in isolation,

> "Previous research shows that the reasons for drop-out (and persistence) are complex, multiple and inter-related. Students continually weigh the costs and benefits of completion and this process starts even before they enroll. If the scales tip too far toward the costs, they will withdraw."[8]

Finally, in an article entitled "Dropouts Give Reasons", 500 high-school dropouts, ages 16-25, were interviewed, and they gave the following reasons for leaving school:

- 47% said classes were boring
- 45% said they entered high school poorly prepared by their earlier schooling
- 69% said they were not motivated to work hard
- 32% said they left to get a job
- 35% said they were failing

By allowing ability θ to incorporate IQ, previous academic preparation, intrinsic motivation, interest in school, ability to handle stress, etc., this simple two-period model appears to offer a theoretical explanation of the mechanism behind these empirical findings. Indeed, failure in preliminary tasks (midterm) is a precursor to dropping-out

[8]In fact, this is the entire underpinning of my model, that students are 'rational'.

in this model, and the idea of 'dropping-out of school to get a job' indicates that continuing with school is no longer "worth it" from the student's perspective ($V < c(\underline{e_2})$, where we allow the marginal cost of supplying effort to include the opportunity cost of not working a job).

Of course, I must emphasize that this is still only a very partial analysis of the school attrition problem; students drop out for a variety of complex reasons that can't be subsumed into a single parameter. These include financial constraints, problems with social/academic integration, problems at home, more attractive educational opportunities, etc. Such problems can lead students of even high 'ability' to drop out, regardless of whether they are failing or excelling, and regardless of whether they are overconfident or under-confident. Even still, the very fact that dropping-out can be explained as the equilibrium consequence of overconfidence is somewhat interesting.

1.4 Under-Confidence: $\hat{\theta} \leq \theta_c < \theta$

> "Self pity is our worst enemy and if we yield to it, we can never do anything wise in the world."
>
> -Helen Keller

Here, the student underestimates their ability relative to the difficulty of the course and fails to register, even though it would be optimal for them to do so. Since they do not Register, they do not attempt the preliminary task, and thus, they do not ever learn their *true* ability relative to the course difficulty. In this way, Under-Confidence would appear to be a persistent trait in the sense that beliefs are never updated appropriately. This result is in accord with well-known results in the Psychology Literature on self-esteem. Josephs, Bosson, Jacobs (2003) write,

> Once established, self-esteem tends to remain stable across time. Research has shed light on many of the mechanisms that promote maintenance of high self-esteem but much remains to be learned about the

> mechanisms that promote stability of low self-esteem. Our conclusion in this article is that low self-esteem may maintain itself, in part, by making the individual immune to the beneficial effects of certain forms of esteem-enhancing feedback.

Thus, by failing to engage in activities that would provide esteem-enhancing feedback (i.e. not registering for the course), or by simply ignoring the feedback from such an activity, the low confidence individual does not 'learn' their true ability, and does not give themself the chance to update their beliefs accordingly.[9]

1.5 Perfect Knowledge of Ability: $\hat{\theta} = \theta$

> "The most difficult thing in life is to know yourself."
>
> -Thales
>
> "A man who knows he is a fool is not a great fool."
>
> -Chuang-tzu

In this case, the student registers iff it is actually in their best interest to do so. Furthermore, when they register, they put in the welfare maximizing allocation of effort across time-periods. This is clearly the ideal situation and any action or policy instrument that can be used to increase the extent to which students understand their capabilities will be welfare improving. Even when it is optimal for a student to register, overconfidence will cause them to underwork on the preliminary task, and they will need to work even harder on the final task in order to pass; conversely, underconfidence causes a student to overwork on their preliminary task and they will have wished they'd allocated more time to other welfare-improving activities (leisure, extracurriculars, etc.). Each of the various scenarios is depicted in Figure 2 below:

[9] It is worth noting, however, that if effort were allowed to positively influence future ability (which it almost surely does) then low self-confidence may actually be a self-fulfilling prophecy. Low self-confidence would lead to little experimentation with new challenges which would then lead to an *actual* lack of development relative to peers, and the start of a downward spiral of confirmed expectations. In this case, underconfidence would not be persistent in the sense that you would consistently underestimate your true ability, but rather, it would be a self-confirming belief. In any case, the problem of underconfidence is surely a deleterious one.

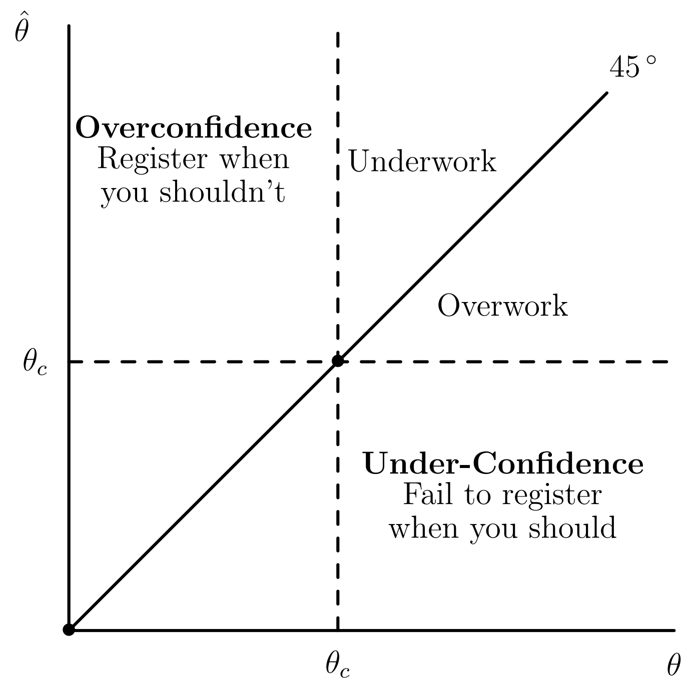


Figure 2: Scenario Diagram for TPLM

Given these scenarios, we can now analyze some ways of remedying the overconfidence/underconfidence problem.

1.6 Resolving Uncertainty over Ability: Reducing $|\hat{\theta} - \theta|$

As mentioned above, any way that a student can reduce $|\hat{\theta} - \theta|$ will be welfare improving since it will decrease the extent to which they end up in any of the undesirable scenarios previously described. Of course, in this model we made the formation of the prior $\hat{\theta}$ an exogenous activity, sweeping the whole belief formation process under the rug. In reality, the formation of the prior will be determined by a variety of complex factors (including previous experiences, information gathered on the course, etc.). Here, we intend to examine some of the current practices and services offered to help the student form the most accurate prior that they can.

The problem of uncertainty over ability is no doubt most prevalent in students who are new to a program or institution and thus unfamiliar with the expectations and course work demands. This is particularly true in introductory Mathematics or language courses at the University level, where students from a wide-range of backgrounds and varying degrees of academic preparedness are asked to register for courses for which they have only very little information (again, sometimes only a few relatively ambiguous lines of description from a course calendar). Thus, what is often observed in most universities is a “Placement Test” or “Entrance Exam”, which can be thought of as a relatively costless exercise (or preliminary task) taken in Period 0 to better inform the student of their ability relative to the difficulty of the class. Often the results of these Placement Tests will segment the population into appropriate levels of course difficulties (eg. placing ‘unprepared’ students in Pre-calculus or more basic Introductory Language classes, as oppose to their advanced counterparts which are reserved for the high-achievers). Indeed, many schools have perfected the design of such placement test, and their success suggests that they ought to be used more frequently. In the 1998 FEDA report discussed earlier, they identified the following list of solutions to the problems of “Incorrect or inappropriate expectations on the part of the student” and “Incorrect course placement”:

- Improved college and course publicity
- Pre-course briefings and Taster sessions
- Presentation of course overviews and expectations during induction process
- Early formative evaluations
- Clear entry criteria coupled with early screening and diagnostic assessments

Other, admittedly tenuous, ways of reducing uncertainty include such online services as RateMyProfessor.com or Pick-A-Prof.com which are websites that allow college and university students to anonymously (and freely) assign ratings to professors and courses of American, Canadian, British, New Zealand, and Australian Institutions. Assessing the credibility of posts found on these websites is obviously a subjective exercise (and there are certainly some adverse selection problems regarding the quality of reviews and reviewers attracted to these sites), but their extreme popularity suggests that there is at least some redeeming quality to be found from logging in. Indeed, if they are successful (at least on average) at resolving bits of uncertainty regarding the difficulty of a course/Professor then this will better inform the decision of whether to register, and will lead to a welfare improvement in general. In schools where students have many course options and are free to choose many electives or even design their entire educational program, this sort of information is extremely valuable in making educated decisions about which courses to take. In fact, from this perspective, it would seem worthwhile - at least in theory - for Universities to consider investing in more formal types of course evaluations, that would allow students to rate courses/Professors on a variety of dimensions and would allow future students to access these ratings prior to registering. Such a system, if implemented carefully, could be very informative.[10] In any case, the advent of Placement Tests and information sharing websites, suggests that the problem of over-confidence and underconfidence is a real one, that is acknowledged (if only tacitly) by students and administrators alike. Their continued use, I think, suggests that they remedy the uncertainty problem and, in doing so, constitute a welfare improvement.

1.7 The Social Planner's instruments: What can a teacher or institution do to improve student welfare?

We now consider the extent to which a Social Planner can influence θ_c by varying the parameters of the model V, R, α, and $\overline{u}$. Recall that the critical level of ability that determines whether a student will register for a course is given by $\theta_c = g^{-1}\left(\frac{\overline{u}}{f\left(c^{-1}\left(\frac{V-R}{2}\right)\right)}\right)$. Since $g^{-1}(\cdot)$, $f(\cdot)$ and $c^{-1}(\cdot)$ are all monotone increasing, we can readily see that,

[10] This, of course, ignores the ethical dilemma of publishing information that may harm the reputation of a Professor, and hinder there ability to advance successfully within the profession.

- $V \uparrow \Longrightarrow \theta_c \downarrow$
- $R \uparrow \Longrightarrow \theta_c \uparrow$
- $\overline{u} \uparrow \Longrightarrow \theta_c \uparrow$

The first two observations merely say that students require greater confidence over their ability to register for a course when either the benefits of taking that course are low or the costs of taking the course are high. Intuitively, this result makes sense because a greater confidence over one's ability reduces the perceived cost of effort required to pass. The last observation expresses the obvious fact that when the passing grade is higher, one needs to have a higher perception of their ability to induce registration into that course.

Recall that in the derivation of the equilibrium of the model we assumed $\alpha = \frac{1}{2}$ for simplicity. At this point it is not difficult to interpret the effect of α on θ_c by simply considering how behaviour would change if $\alpha \approx 0$. In this case, since the preliminary task does not make up a large component of the overall mark, the student will set $e_1 \approx 0$ and $e_2 \approx e^*$ as before. Thus, the optimal total amount of effort expended throughout the course will be $e_1 + e_2 \approx e^*$. Now, compare this scenario to the case where $\alpha = \frac{1}{2}$. Here, we found that the optimal total effort expenditure was $e_1 + e_2 = 2 \cdot e^*$ and, hence, we can see that approximately twice as much effort is expended to pass when the preliminary task carries 50% of the weight in the final mark! This seemingly odd result relies on several tacit assumptions: firstly, that effort in the preliminary task does not carry over in the final task (i.e. studying hard for the midterm does not impact your ability to perform on the final); two, that learning is perfect for any level of effort on the preliminary task (i.e. the student perfectly learns θ by exercising any positive amount of effort); and finally, that the outcome function does not depend on the weights assigned to the tasks (e.g. the midterm is just as difficult as the final exam even if the coverage/weights are completely disproportionate). Obviously, these assumptions are not realistically achieved in practice, but one can imagine that this result will still partially go through as long as the assumptions are approximately valid. In any case, it is not difficult to show that in this model total effort expenditure will be maximized for $\alpha = \frac{1}{2}$ and any deviation from this will strictly decrease the total effort required to pass for any given level of ability.[11] Thus, by increasing $|\alpha - \frac{1}{2}|$ the professor will decrease the expected cost of passing and therefore will decrease θ_c.

We are now prepared to discuss several conclusions from the comparative statics analyzed above. First of all, we consider the inherent trade-off that the *benevolent professor* faces when designing a course.

The Inherent Trade-Off faced by the Benevolent Professor

We begin by making three assumptions:

1. Professors can control the difficulty of their course by increasing the effort required to pass (by varying α and $\overline{u}$).

2. Professors want to maximize the knowledge gained by their students, and they can control this by varying the difficulty of the course.

3. Professors want to minimize the extent to which students end up in either the Over-confident or Underconfident regions (i.e. they want don't want students to have to drop the course, and they don't want to exclude competent people from registering on the grounds that the course appears too difficult).

Now, for such a benevolent professor, in order to attain the first objective listed above they would want to set α close to 50% and they would want to make $\overline{u}$ relatively large.

[11] The model presented here simply considered two periods with a midterm and final. But this could be easily generalized to k tests, with total effort being maximized by setting equal weights on each test. In fact, for a professor concerned with maximizing effort they would want to have many tests all worth a small and equal percentage of the final grade. As will be discussed below this creates an inherent tradeoff for a benevolent professor.

However, in doing this they will simultaneously be maximizing θ_c which as we can see in Figure 2 will maximize the area of both the overconfident and underconfident regions and will undermine the extent to which the professor wants to adhere to their second objective. Conversely, by making θ_c small in an effort to attain their second objective, they will necessarily be making their course 'easier' to pass (in the sense that progressively weaker students will be capable of achieving a passing grade) which will work to minimize effort expenditure, and will decrease the knowledge attained by their pupils. Thus, their is an inherent trade-off faced by the benevolent professor when designing the course curriculum and syllabus. The optimal difficulty level chosen will necessarily depend on the relative weights placed on the two objectives. If the professor's ideology is like that of the shepherd who strives to keep their entire flock intact, not valuing any individual sheep more than the other, then they will most likely adhere to the second objective - at the unfortunate expense of not challenging the strongest students in the group. If on the other hand, the Professor identifies more with the winemaker who cares only about the quality of the finest berries on the vine and hand-picks them to prevent inferior quality fruit from contaminating the lot - then they will most likely adhere to the first objective and maximize the educational outcome attainment of the most talented students by making their course extremely difficult. This, of course, is the age-old problem of whether to cater to the 'average' or to the 'elite'.[12]

Scholarships, and Tutoring

Another way of improving welfare, that avoids the inherent trade-off described above, is to reduce the Expected Cost of Passing through tutoring, help sessions, and Scholarships. Tutoring and help sessions lower the time cost of studying for struggling

[12]The interested reader will appreciate that this debate goes back at least as far as Plato's "Laws" where he writes "No better would be the result with pilots or generals, or householders or statesmen, or any other such class, if they neglected the small and regarded only the great; as the builders say, the larger stones do not lie well without the lesser."

students, which acts to lower θ_c without actually making the course any 'easier'.[13] This will help to prevent the deleterious affects of dropping-out without watering down the course material to cater to low ability students. On the other hand, Scholarships help to defray the cost of registration R which will remedy the problem of underconfidence by making the net benefit of registration more attractive.

1.8 Limitations of the Model

There are, of course, several limitations to this simple model. For one, the model predicts that everyone, regardless of ability, will attain a bare minimum passing mark, $\overline{u}$. While this implication does seem to resonate well with the observation made by author Max Forman that "Education seems to be, in America the only commodity of which the customer tries to get as little as he can for his money.", we feel that is still not nearly compatible with the behaviour that we observe in the real world. This result is primarily due to the assumption of a deterministic outcome function. In reality, the outcome function has a degree of unpredictable 'noise' and the risk-averse student will no doubt put in more than the bare minimum of effort to achieve a passing grade. This type of 'noisy' outcome function could also account for the phenomena whereby students continue-on with a course, and yet still achieve a failing grade because of an uncharacteristically poor exam day. Moreover, it would account for the fact that learning is not perfect; instead of learning their ability exactly, students would only be able to update their prior beliefs in the Bayesian sense, incorporating the (presumably) known distribution of the noise.

Another way of making this model more realistic would simply be to assume a stratified set of rewards $V_1 > V_2 > \ldots > V_n$ corresponding to outcomes $u_1 > u_2 > \ldots > u_n$. In this way, people of varying ability levels will sort themselves into different levels on the outcome ladder (high level students will shoot for the high rewards, and low

[13]In the 1998 FEDA report, the authors identified Tutoring, peer support/coaching/mentoring as ways of remedying "feelings of isolation and not belonging" which are large predictors of drop-outs

ability students will stick to the low rewards).[14] This story makes sense in schools that use letter grades (or some similar stratified reward system) but we are still left with the implication that students with abilities on a continuum will cluster themselves into discrete outcome levels.

The second limitation has already been suggested in the discussion above, and that is, the model does not allow for complimentarities between effort in Period 1 and ability in Period 2. In module-based courses dealing with several relatively disjoint subjects (such as those found in Medical School or Pharmacy programs) this may not be a huge issue, since exams tend to cover disjoint areas (i.e. they are non-cumulative). In these courses, time spent studying for early tests does not necessarily help you perform better in later tests. However, for the great majority of courses, especially in Math and Economics, the trend is to 'build up' a set of tools/models with each new concept depending on (or relating to) the previous one. In these courses, cumulative finals are the norm, and there is no doubt that hard-work expended early in the course, will lessen the load in the latter parts of the course. In any event, such an interplay between effort and ability is an added complexity that would be interesting to study as an extension of this model. However, in the interest of time and space, I will not be considering this here.

The third limitation is that this model ignores the competitive effects observed in most classrooms, whereby students not only care about passing, but also about their rank-order because 'top marks' lend themselves to even greater rewards. With competitive effects we would expect to see higher ability students encouraged to work for more than just a passing grade (which would serve to address the first limitation). In the following section, we will consider this problem in great detail.

[14] From a signaling perspective, this would induce the kind of separating equilibrium that employers desire.

2 Part 2: The Classroom as a First-Price All Pay Auction

To begin with, it is useful to introduce the notion of an All-Pay auction in general, then motivate its usefulness in describing competitive behaviour. In a First-Price all-pay auction, each bidder ($i = 1, 2 \ldots n$) submits a (non-negative) sealed bid, x_i, for an item valued by player i at v_i. All players forfeit their bids, but only the highest bidder wins the item. In this way, the all-pay auction is similar to a standard (winner pay) first-price auction, except that losers must also pay the auctioneer their bids.[15] Thus, in the general set-up we can write the payoff function to player i as follows,

$$u_i(x_1, \ldots, x_n) = \begin{cases} -x_i & \text{if } \exists\ j \text{ such that } x_i < x_j, \\ v_i - x_i & \text{if } x_i > x_j \ \forall j \neq i. \end{cases} \tag{10}$$

In some circumstance, one must specify a tie breaking rule (usually ties are broken randomly), but we will not worry about this additional complexity here. What is important to note, however, is that in an all-pay auction one can interpret differences in the valuations v_i's as arising from differences in abilities. To see this, consider Baye (1996),

> ...suppose the utility to player i of winning a prize of W by putting forth effort x_i is $u_i^* = U_i(W) - \beta_i x_i$, where x_i is effort, and β_i is the marginal cost to player i of effort. Since behavior is invariant to affine transformations, we may just as well write the utility function as $u_i = \frac{u_i^*}{\beta_i} = v_i - x_i$, where $v_i \equiv \frac{U_i(W)}{\beta_i}$. Thus, differences in the v_i's may be due to differences in valuations or differences in the abilities of players to convert an entry into a prize: players with higher v_i's can be thought of as 'stronger' players.

It is this key observation that has contributed to the now widespread use of the all pay auction in economics because it captures the essential elements of "contests" [Dixit (1987)]. It has been used to model technological competition and R&D races [Dasgupta (1986)]; tournaments and job promotions [Lazear and Rosen (1981), Narasimha

[15] This difference is crucial. Just imagine if eBay was organized on the basis of a sealed bid all-pay auction and ask yourself how that would impact your bid for an item you really valued?

(1988)] as well as a host of other situations including political campaigns, wars of attrition, etc. Essentially, these economic problems boil down to a contest that is an all-pay auction in effort; the player putting forth the greatest effort wins the prize, while the efforts of other contestants go essentially unrewarded. It is in this spirit that we view the "classroom" as a manifestation of a first-price all pay auction, although our development will be a slight departure from the 'greatest effort wins the prize' view, since we will incorporate the dual impact of effort and ability on outcome achievement. Students with varying abilities (that are private information), compete for the top-mark by their choice of effort, which together with their ability determines their outcome, or bid. Students get rewards for passing the course (attaining some critical outcome level) and the student with the 'top-mark' (i.e. highest bid) gets an additional reward W for out-doing each of his/her peers. In this way, we capture two elements: one, students value passing the course (exceeding some absolute outcome level) and two, students value their rank-order in the class. Of course, in reality there is probably a stratified set of rewards $V_1 > V_2 > \ldots > V_k$ each corresponding to a different absolute outcome on the ladder (think letter grades A+, A, A−, B, C, etc.) as mentioned in the previous section. Also, in reality their is some, albeit smaller, reward to be had for placing second or third in the class. In principle, these additional complexities could be included in our model; there are, after all, such things as multi-prize all-pay auctions with heterogeneous prizes [Barut (1998)]. However, in spirit of Occam's Razor we will focus our attention on the most basic model that captures the effects we are interested in - the Winner-Take-All Model. We begin with two assumptions,

Assumption 2.1 *Students derive additional rewards from achieving high grades relative to their peers. This leads to rank-order competition amongst students for the top mark.*

Although this assumption seems rather intuitive, it is deserving of at least some words of motivation. First of all there is - I think undeniably - a fundamental human quality that makes us want to outperform our peers in competitive tasks. After all, it is undoubtable that Tiger Woods would rather be 10 over par and win the British Open, then be 10 under par and place second! Perhaps this motivation is routed in a sense of pride, an intrinsic feeling of accomplishment for outachieving others; perhaps it is just an evolutionary trait passed down from a bygone era where an inner ambition to 'outdo' was selected for; or perhaps, as in many cases, the reward is simply extrinsic (after all, the difference between first and second place in this year's British Open is approximately $600,000!). Whatever the case may be, relative performance matters greatly, especially for individuals at the top of the pack (case in point: the difference between placing 55 and 56 is roughly $200 at this year's British Open). Markets such as these, in which small differences in performance translate into extremely large differences in reward, are characteristic of what are called Winner-Take-All Markets, a term coined by Robert Frank and Philip Cook in their book "The Winner-Take-All Society". In these special markets, sellers whose goods are viewed as the best (or are thought of as the top performers) reap far greater rewards and capture a disproportionately large share of the market compared to other sellers whose goods are also of high quality. Sellers immediately below the top may perform *almost* as well as those at the top, but gain far fewer rewards. In Frank and Cook's example, a handful of top opera tenors command high fees for performing and recording, while artists who are technically close to them remain obscure, are paid far less, and are rarely asked to record. Similar behavior is found among top athletes, as well as among law firms, top consulting companies, and authors. Sports and Entertainment represent the prototypical Winner-Take-All Markets because here performance is easily observed and demand for the most-elite is heavily sought after. Recently, however, higher education has been likened to this type of market as well. Frank (1999) writes,

> The market for higher education is fundamentally different from the typical markets portrayed in economics textbooks. When excess demand

> arises in the market for an ordinary good or service, it is almost always fleeting: producers rush to fill the void, or prices rise so much that the market quickly clears. Not so in the upper reaches of the academic market. Despite the persistence of excess demand, elite colleges and universities continue to turn away thousands of well qualified applicants, while charging those they admit only about one-third of what it costs to serve them. Why dont elite universities simply raise their prices? Because a universitys status depends heavily on the average intellectual ability of its students, elite universities need top students every bit as much as top students feel the need to attend elite institutions. This co-dependence creates multiple positive-feedback loops that amplify the rewards for a university that succeeds in its efforts to recruit top students and faculty. The result is a quintessential winner-take-all market, in which success breeds success and failure breeds failure.

Thus, there appears to be a significant premium to attending a top-ranking program, and in turn, there appears to be a significant advantage given to top-ranking students entering and exiting these programs. Bowen and Bok (1998) found that graduation from the most selective colleges was a great advantage in terms of entry to elite professional schools and long-run posteducational incomes, net of a student's initial academic performance or skills (SAT and GPA). Thus, they conclude that the economic payoff to a given academic ability is greatly enhanced by entry to a highly selective college. So, then the next natural question would be "What does it take to get into a top-tiered program?"

Hernandez (1997), the assistant director of admissions at Dartmouth College, detailed the "Blue Book" algorithm used there and in other Ivy League colleges to calculate an academic index (AI) for comparing students' applications. She reported that there is a high degree of agreement between admissions decisions using this method and decisions made by other highly selective colleges that use the same basic inputs but in a slightly different way. The formula calculates an AI by combining three components: SAT I scores, SAT II scores, and the student's *class rank* in his or her particular high school. The last is called a "converted rank score" or CRS and it drops off steeply as one moves down from the head of the class. In simulations using the admission's algorithm, it was shown that using class rank in combination with SAT scores creates a dramatic *valedictorian effect*. Specifically, being at the top of one's class was worth about 70 points on the SAT I plus 60 points on the SAT II in terms of the final index score. However, this class-rank benefit was only extended to the top handful in a high school class. Therefore, rank-order performance was shown to be an extremely important factor in gaining entry into top programs.

Thus, if entry into a top program delivers extremely high rewards, and if high-rank order performance is extremely important to gain entry into these schools (and not to mention, to gain scholarships to attend such schools) then it would seem like students (at least at the top of the ability spectrum) would have a great extrinsic incentive to work hard for the top-mark.[16] On the flip side of this, we would expect that Students below the top of their class, in terms of ability, would have less incentive to attain higher-grades, since they are at a substantial disadvantage from attaining the top-mark required (it is too costly in terms of the effort expenditure). Later I will refer to these two effects as the *encouragement-effect* and *discouragement-effect* respectively. In light of this observation we are now ready for our second assumption,

Assumption 2.2 *In general, higher ability students (higher θ's) achieve higher grades. That is, the outcome function $u(\theta)$ is strictly increasing.*

This assumption will be needed when deriving the symmetric equilibrium outcome function (i.e. the bidding function) of our model. Notice that in this framework higher grades do not necessarily translate into higher effort levels since, for us, the stronger students do not need to work as hard to achieve a given outcome. Thus, Assumption 2.2 will not necessarily imply that the equilibrium effort function $e(\theta)$ is strictly increasing. In fact, in light of the discouragement and encouragement effect hypothesized above, we should expect that, at first, effort will be decreasing

[16] Add in the intrinsic rewards discussed earlier, and this effect is only magnified further.

in ability as students are still discouraged from attaining the top mark and simply trade their additional productivity for leisure, but then, at the high end of the ability spectrum students begin to work harder as ability increases, buying an additional competitive edge with their increased productivities. We will later refer to this as the *leisure-substitution effect* and *competitive-effect* respectively and will take great care in analyzing each of them. We are now ready to introduce our model.

2.1 The Winner-Take-All model

Consider the following two-period Winner-Take-All model (WTAM) with m students:

Period 0: Each student learns their *true* θ (exogenously). θ is private information, but the distribution of ability is known to be uniform on the interval $(0, 1)$. In this way, Student's can form expectations of where they stand relative to their peers.[17]

Period 1: Based on their level of ability, the Student decides whether to Register for the course or not. Registration costs R, not registering costs 0.

Period 2: If the student registers, they choose effort level e and achieve outcome (mark) u given by the following deterministic outcome function:[18]

$$u = e \cdot \theta \tag{11}$$

If the student's mark, u, exceeds a predetermined level $\overline{u}$, the passing grade, then the student gets a payoff V. Otherwise, they fail and get payoff 0. If, in addition

[17]Note that we are now relaxing the assumption of imperfect information over ability. In the original conception of this model ability was only known imperfectly, as in the previous model, however this quickly became intractable. In any case, we have already considered the deleterious effects of imperfect information, and now we wish to focus attention on a new problem: to study the effects of competition in the classroom when abilities are certain and private.

[18]Here we choose to specify a particular functional form for analytical tractability. The reader will quickly notice that we've abandoned our boundedness assumption on $f(\cdot)$ and $g(\cdot)$. In lieu of this $\overline{u}$ can no longer be interpreted as a percentage grade, and we no longer need to worry about it being impossible for Students to pass.

to passing, the student achieves the 'top' grade (i.e. $u \geq u_i \ \forall \ i = 1, 2 \ldots n$) then they obtain an additional reward W.[19] As usual effort is costly, and the cost of effort is given by $c(e)$, as before. Given the form of the outcome function the reader may interpret e as the sum total of all effort expended over the duration of the course.

As usual, it is assumed that $V > R$ so that at least some students, with sufficiently large ability, may find it worthwhile to register and put in effort to pass (even if they do not put themselves in much contention for the top mark). Again, we assume no discounting ($\delta = 1$) for simplicity.

2.2 Equilibrium of the model

To begin with, we will assume that there exists a cut-off type, $\theta_L \in (0, 1)$, who finds it optimal to register and exert the bare minimum amount effort to 'pass' (i.e. they set $u(\theta_L) = \overline{u}$). Anticipating a strictly increasing symmetric equilibrium outcome function, for type θ_L, the cost of passing will simply equal the benefit of passing: [20]

$$c(e_L) = V - R \qquad\qquad \text{where } e_L = \frac{\overline{u}}{\theta_L} \tag{12}$$

and this implies that $\theta_L = \frac{\overline{u}}{c^{-1}(V-R)}$. All $\theta < \theta_L$ won't register, and conversely, all $\theta \geq \theta_L$ will register and will set $u \geq \overline{u}$. Thus, everyone who registers will 'pass' in Period 2. Since the costs and benefits of registering are known to all students (even the cost of effort function $c(\cdot)$ is assumed to be known), all students can compute θ_L. From this knowledge, students who are registered in the class will no longer consider the abilities of their classmates to be independent random draws from the unit interval, but rather, from the interval $(\theta_L, 1)$. Therefore, in Period 2 the payoff

[19] Note that we do not include an additional option of 'dropping-out' after seeing the results of a 'preliminary' test since this will be redundant in the case of perfect knowledge over ability and a deterministic outcome function. Since students can perfectly forecast the costs and benefits of their decisions,they will register iff it is optimal for them to do so. Also, we do not specify a tie-breaking rule because, as it will turn out, the probability of two-students sharing the top mark will be zero. Lastly, note that we use $n \leq m$ as the number of 'registered' students.

[20] Since, in equilibrium, they give themself a zero probability of having the "top mark" we do not need to incorporate the probability of obtaining W. This will be justified in the analysis that follows.

function to a representative student i of type θ_i (who has registered) can be written as follows:

$$\Pi(e,\theta_i) = V - R + W \cdot Pr(e \cdot \theta_i \text{ is the top mark}) - c(e) \tag{13}$$

As in standard auction theory, we will assume a symmetric equilibrium bidding function, which in this case will be an equilibrium outcome function $u(\theta)$ mapping abilities to final grades. We assume that $u(\theta)$ is monotone increasing (as in Assumption 2.2) and, as noted before, that $u(\theta_L) = \overline{u}$. This equilibrium outcome function will implicitly define a symmetric equilibrium effort function given by $e(\theta) = \frac{u(\theta)}{\theta}$. We will take great care in analyzing this later. With this assumption we may now rewrite (13) above as:

$$\begin{aligned}
\Pi(e,\theta_i) &= V - R + W \cdot Pr(e \cdot \theta_i \geq u(\theta_j) \;\; \forall\, j \;\neq i) - c(e) \\
&= V - R + W \cdot Pr(\theta_j \leq u^{-1}(e \cdot \theta_i) \quad \forall\, j \neq i) - c(e) \\
&= V - R + W \cdot \left(\frac{u^{-1}(e \cdot \theta_i) - \theta_L}{1-\theta_L}\right)^{n-1} - c(e) \qquad (14)
\end{aligned}$$

The first step makes use of the fact that $u(\cdot)$ is monotone (invertible) and the second step makes use of the fact that θ's are independent random draws from the interval $(\theta_L, 1)$. We want students to choose the optimal amount of effort given the outcome function chosen by others, so we now differentiate (14) with respect e and set it equal to 0.

$$\frac{d\Pi}{de} = W \cdot (n-1) \cdot \left(\frac{u^{-1}(e \cdot \theta) - \theta_L}{1-\theta_L}\right)^{n-2} \cdot \frac{1}{1-\theta_L} \cdot \frac{1}{u'(u^{-1}(e \cdot \theta))} \cdot \theta - c'(e) = 0 \tag{15}$$

Of course, since we want $u(\theta)$ to be a *symmetric* equilibrium function, we need (15) to be satisfied by $e(\theta) = \frac{u(\theta)}{\theta}$.[21] Imposing this equilibrium condition, (15) reduces to:

[21] That is, given that everyone else is using the bidding function $u(\theta)$, it must also be optimal for *you* to use it.

$$W \cdot (n-1) \cdot \left(\frac{\theta - \theta_L}{1-\theta_L}\right)^{n-2} \cdot \frac{1}{1-\theta_L} \cdot \frac{1}{u'(\theta)} \cdot \theta - c'(e(\theta)) = 0 \tag{16}$$

At this point it is useful to assume a particular functional form for our cost function. For simplicity we take the cost of effort to be linear $c(e) = a \cdot e$ where $a > 0$ represents the marginal cost of effort. We could equally well assume cost to be quadratic in effort, but at the unfortunate expense of making the solution to the resulting differential equation analytically complex. Since very few additional insights are gained from a more realistic cost function, we choose to stick with linearity. Subbing in our cost function and rearranging, (16) becomes,

$$\begin{aligned} u'(\theta) &= \frac{W}{a} \cdot (n-1) \cdot \left(\frac{\theta - \theta_L}{1-\theta_L}\right)^{n-2} \cdot \frac{\theta}{1-\theta_L} && (17) \\ &= \frac{W}{a} \cdot (n-1) \cdot \left[\left(\frac{\theta - \theta_L}{1-\theta_L}\right)^{n-1} + \left(\frac{\theta - \theta_L}{1-\theta_L}\right)^{n-2} \cdot \frac{\theta_L}{1-\theta_L}\right] && (18) \end{aligned}$$

We are now prepared to solve for the equilibrium outcome function $u(\theta)$ by integrating,

$$\begin{aligned} u(\theta) &= \int_{\theta_L}^{\theta} u'(t)dt + u(\theta_L) \\ &= \frac{W}{a}(n-1) \left[\frac{1-\theta_L}{n}\left(\frac{\theta - \theta_L}{1-\theta_L}\right)^{n} + \frac{\theta_L}{n-1}\left(\frac{\theta - \theta_L}{1-\theta_L}\right)^{n-1}\right] + \ u(\theta_L) \\ &= \frac{W}{a}\left(\frac{n-1}{n}\right)\left(\frac{\theta - \theta_L}{1-\theta_L}\right)^{n-1}\left(\theta + \frac{1}{n-1}\theta_L\right) + \ \overline{u} && (19) \end{aligned}$$

Observing that $u(\theta)$ is increasing in θ (justifying our initial assumption) our derivation of the equilibrium is now complete and we are ready to state our second Theorem.

Theorem 2.1 With linear cost of effort there exists a subgame perfect Nash Equilibrium of WTOM in which: in Period 1 each student registers iff $\theta \geq \theta_L = \frac{\overline{u}}{c^{-1}(V-R)}$, and in Period 2 students choose effort according to the symmetric effort function $e(\theta) = \frac{u(\theta)}{\theta}$ where $u(\theta)$ is given by (19).

Now that we have derived an expression for the equilibrium outcome function (and hence, also the equilibrium effort function), we are ready to study its properties in detail.

2.3 The Equilibrium Outcome Function, $u(\theta)$

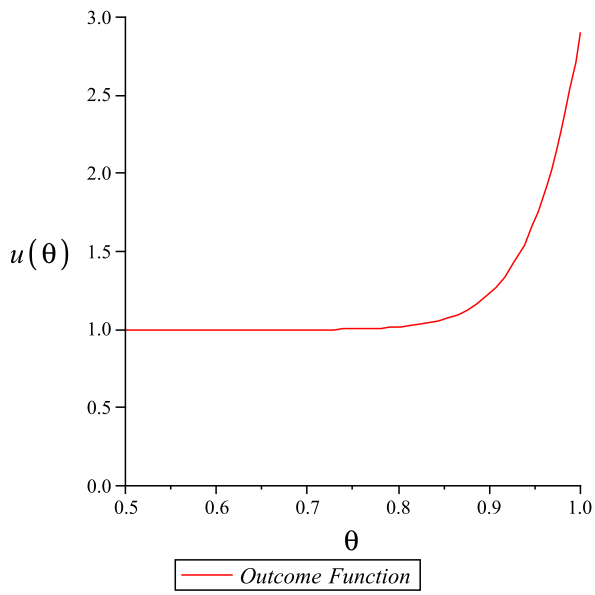
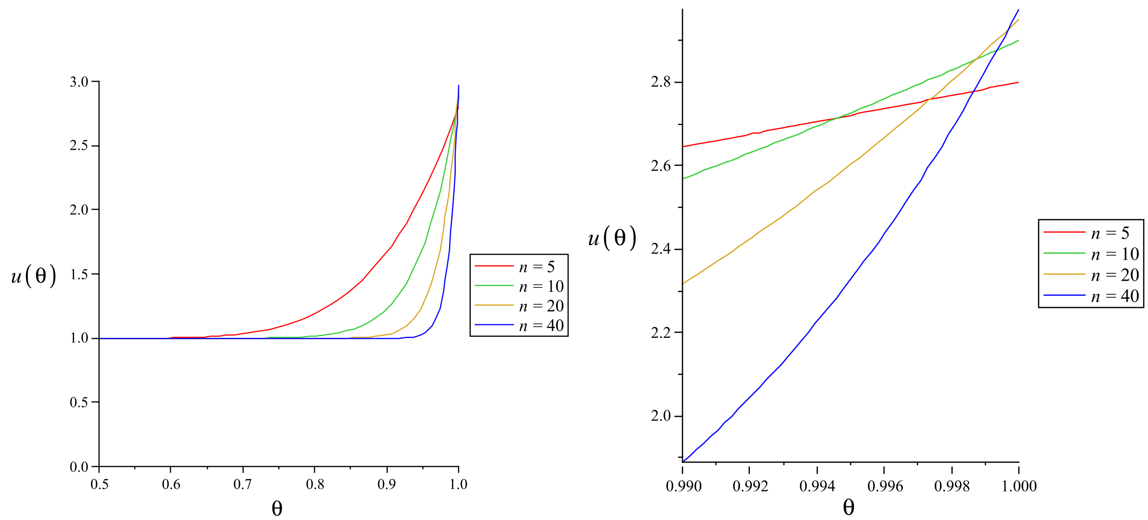


Figure 3: $u(\theta)$ for $W = 2, a = 1, \overline{u} = 1, \theta_L = 0.5, n = 10$

Pictured in Figure 3 is our equilibrium outcome function (suitably parametrized). The first thing to notice about it is that it is only defined for $\theta \geq \theta_L$ and it is relatively flat for all θ close to θ_L. The basic intuition behind this trait is that registered students at the lowest ends of the ability spectrum have little to shoot for beyond a passing grade since supplying sufficient effort to attain the 'top mark' would be far too costly for them. On the other hand, as θ moves closer and closer to the upper end of the ability spectrum the student finds themself more and more capable of attaining the 'top mark' and, as a result, they strive for higher outcomes. This, in part, justifies our initial hypothesis of a discouragement and encouragement region on the ability spectrum. Note that in the extreme case when $\theta = 1$ the student's outcome is numerically equivalent to their amount of effort. In this case we see that the student



(a) All Abilities (b) High Ability

Figure 4: Outcome Function for varying class sizes

with $\theta = 1$ would be willing to work just under 2 units more the passing level of effort since this would almost guarantee them the top-mark and a reward of $W = 2$. Now, let's consider the impact of class size on the outcome function, by allowing n to vary.

As can be seen from Figure 4, the impact of an increase in class size serves to "discourage" the low-medium ability types, and tends to "encourage" the very high ability types (as mentioned earlier we call these two phenomena the *discouragement effect* and *encouragement effect* respectively). The intuition behind this result is that high-ability types are 'threatened' by the increase in class-size and are capable of rising to the occasion since they see larger returns to effort than low-medium ability types. In fact, by fixing θ and allowing n to vary we can pinpoint the regions where these students are encouraged by additional competition, and where they ultimately become discouraged. Evidently, we can see from Figure 5, for low ability types the discouragement effect kicks in almost immediately, whereas, for the high ability students the encouragement effect dominates up until a critical threshold, beyond which the discouragement effect takes over. The higher the ability of the student, the larger the class size needs to be before they eventually become discouraged.

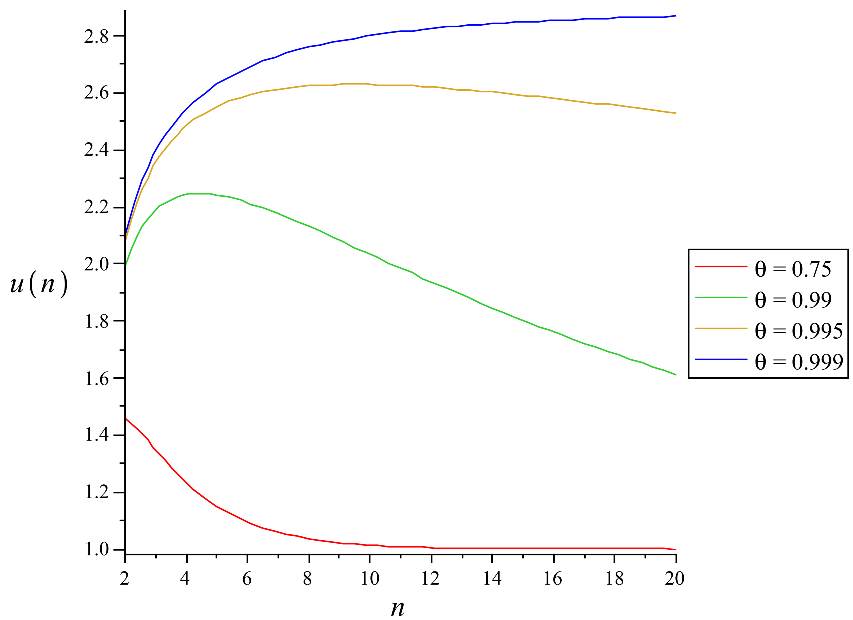


Figure 5: Regions of Encouragement and Discouragement for Various θ's

2.4 The Average Outcome and the Class-Size Debate

Recall that in Period 2, θ's are uniformly distributed on the interval $(\theta_L, 1)$. Thus, we can easily compute the average outcome, for a given set of parameters, by working out the expected value of $u(\theta)$,

$$
\begin{aligned}
E[u(\theta)] &= \int_{\theta_L}^{1} u(\theta) \cdot \left(\frac{1}{1-\theta_L}\right) d\theta \\
&= \frac{W}{a}\left[\frac{n-1}{n}\int_{\theta_L}^{1}\left(\frac{\theta-\theta_L}{1-\theta_L}\right)^{n} d\theta \; + \frac{\theta_L}{1-\theta_L}\int_{\theta_L}^{1}\left(\frac{\theta-\theta_L}{1-\theta_L}\right)^{n-1} d\theta\right] + \; \overline{u} \\
&= \frac{W}{a}\left[\frac{n-1}{n}\cdot\frac{1}{n+1}(1-\theta_L) \; + \; \frac{\theta_L}{n}\right] + \overline{u} \\
&= \frac{W}{a \cdot n(n+1)}\left[(n-1) + 2\cdot\theta_L\right] + \overline{u} \\
&= \frac{W}{a \cdot n(n+1)}\left[(n-1) + \frac{2\,\overline{u}}{\left(\frac{V-R}{a}\right)}\right] + \overline{u} \qquad (20)
\end{aligned}
$$

Now, we can analyze the effects that our parameters have on the class average,

- $V \uparrow$ or $R \downarrow \Longrightarrow \theta_L \downarrow \Longrightarrow E[u(\theta)] \downarrow$ (class more diluted with low ability students)
- $n \uparrow \Longrightarrow E[u(\theta)] \downarrow$ (since probability of achieving top mark $\downarrow$ for **most** students)
- $\overline{u} \uparrow \Longrightarrow E[u(\theta)] \uparrow$ (since more effort is required to pass)

The results above are both intuitive and obvious in this framework. What is not immediately obvious, however, is how an increase in a will affect the class average. On the one hand the greater the marginal cost of effort, the less likely it is for a student to want to strive hard for the top mark (holding the reward W fixed); but on the other hand, the greater the marginal cost of effort, the more prohibitive the class is to register for, and thus the larger is θ_L (the ability of the weakest student in the class). These two effects work in opposite directions, and ultimately, the effect of increasing a would seem to be ambiguous. However, taking a partial derivative of $E[u(\theta)]$ with respect to a and simplifying we find that,

$$\frac{\partial E[u(\theta)]}{\partial a} = -\frac{(n-1)\,W}{(n+1)\,na^2} < 0 \tag{21}$$

and so, it is unambiguously clear that the first effect dominates.

Application: The Class-Size Debate

One thing that is worth pointing out, is that within this competitive framework, increasing class-size tends to decrease the overall classroom achievement (as measured by the class average). The issue of class size and the idea of class size reduction is probably the most popular and most funded school improvement policy in Canada and the United States.[22] By way of example, for the 2007-08 school year, the Ontario Ministry of Education committed a total investment of \$406 million in ongoing

[22] For an excellent summary of this debate, featuring a collection of papers from several eminent economists, see Mishel & Rothstein's *The Class Size Debate*

funding to support over 5,100 additional primary teachers in the province in hopes of reducing class sizes. As well, they are providing \$700 million in funding to support capital projects that will create more than 1,900 classrooms to accommodate the additional space required for these smaller class sizes, with the ultimate goal of reducing all primary classes to 23 students or fewer.[23] Hoxby (1997) writes,

> Class size reductions are enacted often because they are popular with nearly every constituency interested in schools. Parents like smaller classes because their personal experience suggests that they themselves give more to each child when they have fewer children to handle. Even if parents in a school disagree bitterly about educational methods, they can agree that class size reduction is good: smaller classes give teachers the opportunity to practice more of each parents favored educational method. Teachers, teachers unions, and administrators like smaller classes for the same reasons parents do, but they may also like smaller classes for reasons that spring from self-interest. Teachers may like smaller classes because they reduce the effort that they must expend in order to deliver instruction. Teachers unions may like class size reductions because they increase the demand for teachers. Administrators may like class size reductions because they increase the size of their domain. As a result of the policys popularity, the twentieth century has been a period of continuous decline in class size.

Proponents of smaller class sizes, the Ontario Ministry of Education included, argue that with a smaller pupil-teacher ratio, students can get more one-on-one attention, have less distractions, feel more engaged, are more willing to ask questions, and are more likely to succeed in high school and beyond. All of these effects, they say, serve to increase a child's educational outcome and, hence, their cognitive development. While, our model has nothing to say about these effects, it does add one more argument to the small class-size camp. In our model, with smaller class sizes, the class is encouraged to work harder on average. However, this conclusion must be taken with a grain of salt, as one must do whenever talking about an effect that is only true 'on average'. In our model, the smaller class size encourages the low-medium ability students to work harder (they are less discouraged by competition), but on the other

[23]Figures taken from Ontario Ministry of Education webpage, `www.edu.gov.on.ca`

hand, it encourages the high-ability students to work less (they are less threatened by competition). If, as society, we wish to maximize the average outcome attainment of our students, then this model predicts that we should have smaller class sizes - but this will have the unfortunate consequence of holding back the most gifted students from reaching their greatest potential.[24] Conversely, if we wish to maximize the outcome attainment of our most gifted students, then we ought to keep class sizes large to encourage competition - but at the unfortunate expense of discouraging the majority. This expresses the essential ideological dichotomy of whether to view the classroom as a herd or as a jungle (and whether to view the teacher as a shepherd or winemaker).

2.5 The Equilibrium Effort Function, $e(\theta)$

Having considered various aspects of the Outcome function $u(\theta)$ we are now ready to examine the properties of the equilibrium effort function $e(\theta) = \frac{u(\theta)}{\theta}$, pictured in Figure 6. The first thing to notice is that while the outcome function is strictly increasing, the effort function first decreases for low values of θ and then increases as θ becomes closer to 1. The intuition here is that for low values of θ there is only a small probability of achieving the 'top mark', and so, for small increases in ability the student substitutes their labour for leisure (effectively, they buy themselves leisure time with their extra-productivity). However, as θ gets large, the student is more and more capable of attaining the 'top mark' so they begin to use their extra productivity as a competitive advantage, by working even harder. As foreshadowed earlier, we will call these two phenomena, the *leisure-substitution effect* and the *competitive effect* respectively. We can highlight each of these phenomena by allowing W and n to vary as in Figure 7.

As we can see from Figure 7, by allowing the class-size to increase we make it in-

[24]It is, I think, this type of philosophy that lead Col. Robert Ingersoll to say that "Colleges are places where pebbles are polished and diamonds are dimmed."

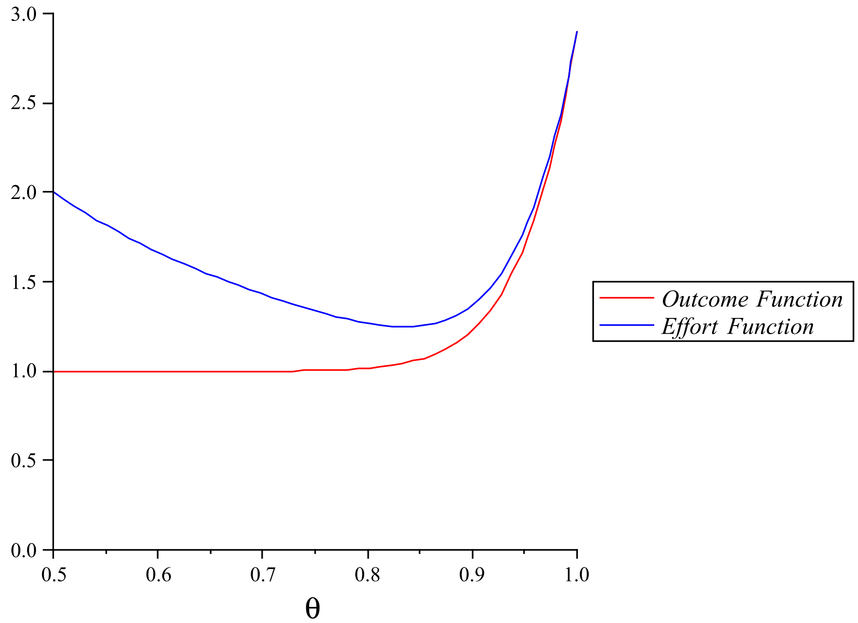
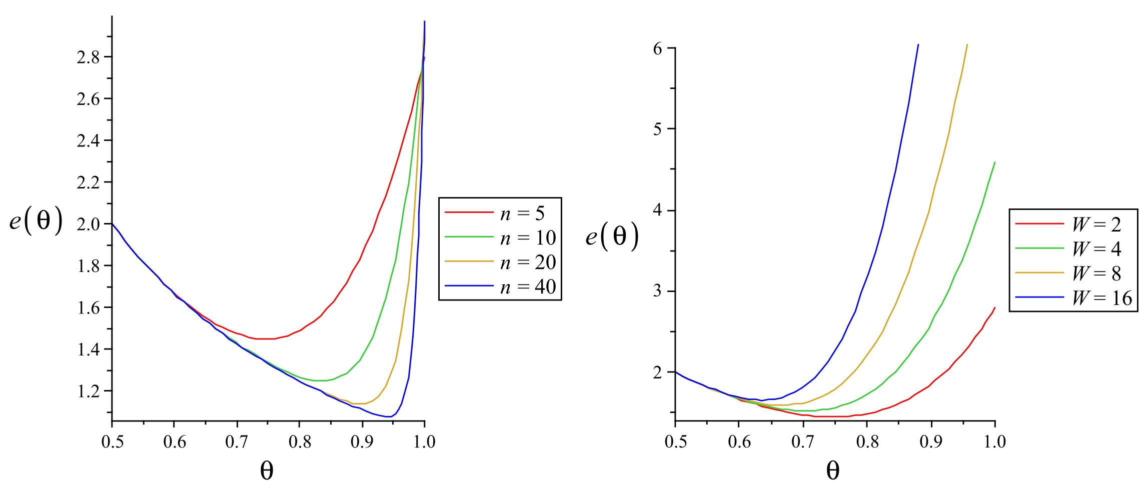


Figure 6: Effort Function vs. Outcome Function



(a) Varying class size

(b) Varying reward for top-mark

Figure 7: Effort Function for Various Class Sizes and Top-Mark Rewards

creasingly difficult for student's to achieve the top mark, and so they will tend to use their additional productivity (i.e. higher θ's) to buy additional leisure even when they are already of relatively high ability. In other words, the *leisure-substitution effect* is more pronounced when class sizes are large. This, of course, is just another name for the "discouragement" effect analyzed earlier. On the other hand, by increasing W we increase the attractiveness of attaining the top-mark, and so the *competitive effect* tends to take over earlier.

It is interesting to note that the hardest workers in this model, are those on the tails of the distribution. The low ability students work hard so that they can just scrape by with a passing grade, and the high ability workers work hard so that they can potentially achieve the top mark. In between, students have very little to motivate them; realistically, they can't vie for the rewards of top marks, and so there is little more to strive for beyond a passing grade, which they can achieve with relatively little effort. This story, as simplistic as it is, seems to accord with evidence from casual empiricism. Students work hard because they struggle, or they work hard because they are competitive, and in the middle, they are simply apathetic.

2.6 A generalization to other types of probability distributions over θ

Before we begin to discuss some of the limitations of this model, it is useful to explore the robustness of these results to different functional form specifications. In particular, we may consider alternative distributions of θ on the unit interval. The shrewd reader would surely have cringed at the assumption of a *uniform* distribution over abilities. There is indeed much evidence that suggests that distributions over intelligence (and other talents) tend to resemble the Gaussian or Normal Distribution. Whether or not this is precisely true[25], few would be willing to argue that talent is uniform. There

[25] See "Is intelligence distributed normally" by C. Burt - British Journal of Mathematical & Statistical Psychology, 1963

are relatively few people who possess exceptional high or low intelligence, whereas there are very many individuals clustered somewhere around the middle. In lieu of this, we now consider a family of distributions that more closely resemble the normal distribution in shape and form. Of course, we could simply use the normal distribution itself, so long as we were willing to accept negative abilities, or if we were willing to retool the outcome function, but this would come at the unfortunate expense at not being able to compare the new results to the old results in a meaningful way. Consequently, we suggest the following distribution, $\sigma(\theta)$, indexed by parameter k:[26]

$$\sigma(\theta) = \left(\frac{\Gamma\left(3/2 + k\right) 2^{k+1}}{\sqrt{\pi} \cdot \Gamma\left(k+1\right)} \right) \left(2\,\theta\left(1-\theta\right)\right)^{k} \tag{22}$$

where $\Gamma(z) = \int_0^\infty t^{z-1} e^{-t}\, dt$ is an extension of the familiar factorial function to real and complex numbers (i.e. $\Gamma(n) = (n-1)!$ when n is a positive integer). We illustrate this distribution for various values of k in Figure 8.

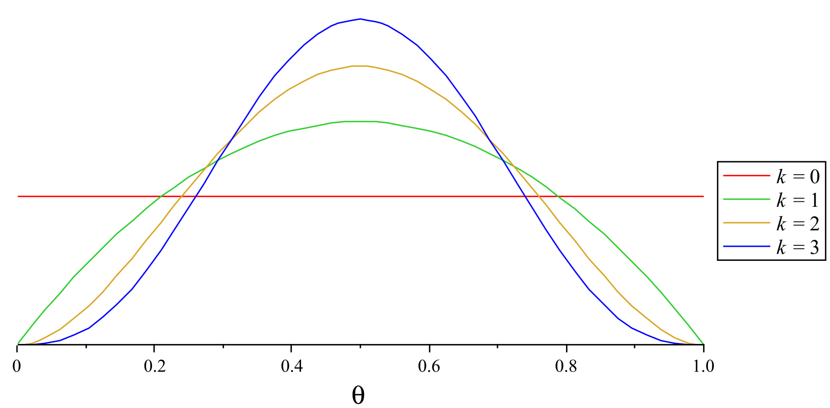


Figure 8: Distribution for different values of k

The first thing to notice is that the distribution is symmetric around $\theta = 0.5$, and when $k = 0$ we reduce to the familiar uniform distribution studied earlier. Increasing k has the effect of squeezing the distribution toward the center, so that less and less

[26] As far as I know, there is no name for this distribution. I invented it for the sole purpose of having a bell-shaped distribution over the unit interval, that would reduce to the uniform distribution as a special case. The complex first term in the expression is required to make the pdf integrate to 1.

students are located at the extremes. In the limit as k tends to infinity the distribution tends to the degenerate distribution that places all the probability mass at $\theta = 0.5$. Thus we will interpret k as an index of the clusteredness of ability. We are now prepared to see how this specification will effect our equilibrium outcome and effort functions for various levels of k.

We begin by noting that for a general pdf $\sigma(\theta)$ over ability, and with corresponding cdf $\Sigma(\theta)$, we can rewrite (14) as,

$$\Pi(e, \theta_i) = V - R + W \cdot \left[\Sigma\left(u^{-1}(e \cdot \theta_i)\right)\right]^{n-1} - c(e) \tag{23}$$

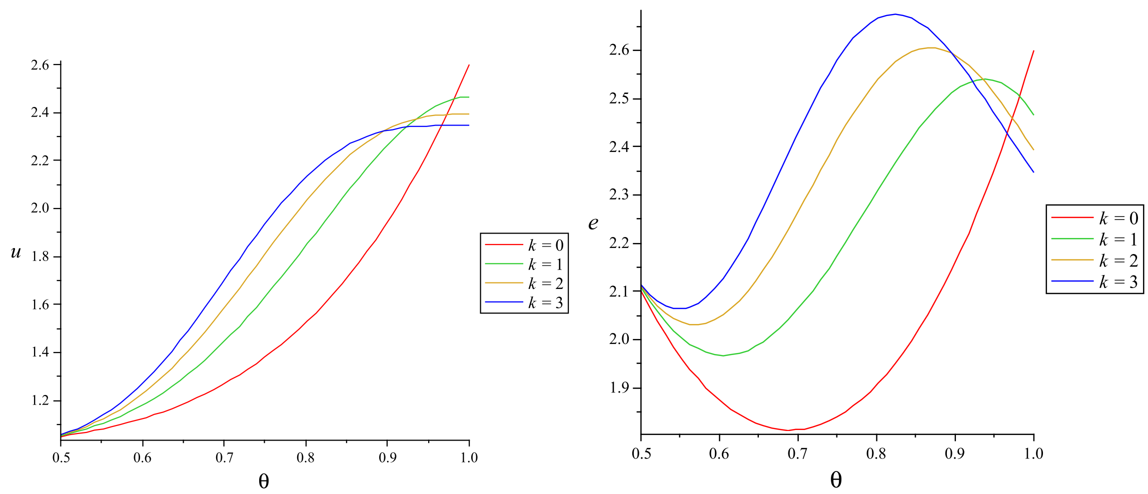
and we can rewrite the corresponding FOC (18) as

$$u'(\theta) = \frac{W}{a}(n-1) \cdot [\Sigma(\theta)]^{n-1} \cdot \sigma(\theta) \cdot \theta \tag{24}$$

Due to the analytical complexity of $\sigma(\cdot)$ we cannot derive a closed form expression for the equilibrium solution $u(\theta)$. We can, however, use MAPLE to perform the integration for various values of k and plot the resulting functions on a graph.[27]

As can be seen from Figure 9, when $k = 0$ our distribution reduces to the uniform distribution and we get the familiar pictures for the outcome and effort functions derived earlier. However, when k is increased, we obtain a new feature on both the outcome and effort functions, and that is, an increasingly prominent s-shape. The intuition here is that for large values of k, there is less and less probability of having exceptionally high-ability students in the classroom (abilities are more clustered). Thus, the rare high-ability student will find themself in a 'league of their own' and will not likely face competition from students of similar ability. These students use their extremely high ability to their advantage by working relatively less, and that

[27]Unless otherwise stated we will use the same parametrization as before, $W = 2$, $a = 1$, $\theta_L = 0.5$, and $n = 5$

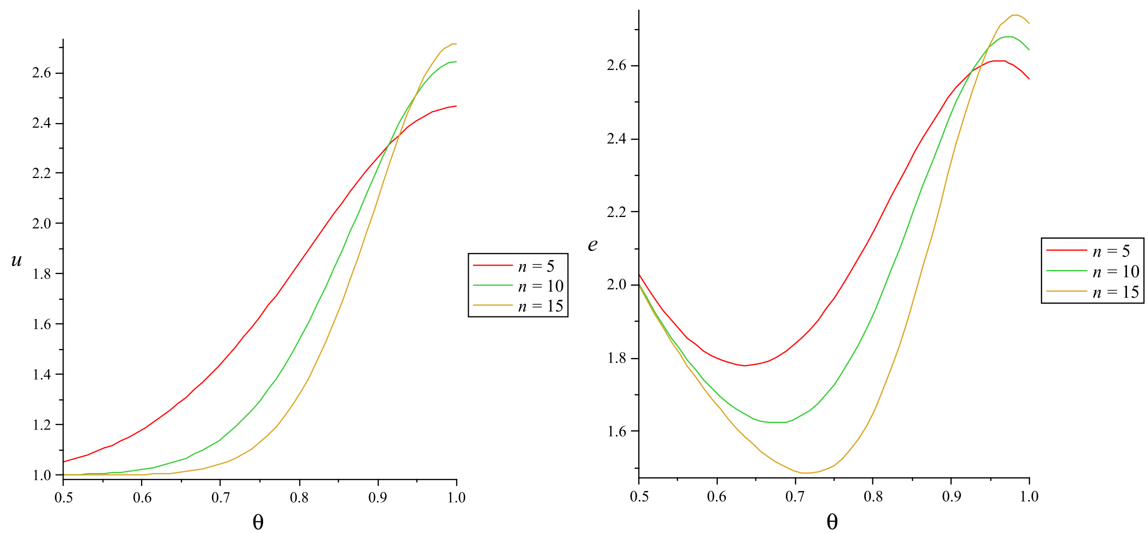


(a) Outcome Functions (b) Effort Functions

Figure 9: Outcome and Effort Functions for various k's

explains why the effort function dips down near the highest end of the ability spectrum (in turn, this accounts for the change in concavity of the outcome function). We still see the same *leisure-substitution* effect and *competitive effect* discussed earlier but now we must add in a third effect, the *superstar effect*. The superstar effect kicks in at the point where a student is so extremely talented that they are unlikely to face competition from anyone of comparable ability, and this encourages them to work less (i.e. exploit there uncharacteristically high productivity and simply rest on their laurels). As we increase k, the superstar effect starts earlier because students are more clustered around average ability, so that even small deviations from the average are 'extreme'). Looking at Figure 10, which fixes $k = 2$ and allows n to vary, we see that same crossing properties are retained. Thus, the encouragement and discouragement effects are robust. We also see that the superstar effect is mitigated to the extent that class size is increased. This makes sense because for larger class sizes there is more chance of having many extremely talented students competing for the top mark (i.e. it's too risky to rest on your laurels in a large class).

Another interesting conclusion is that as we increase k, the outcome and effort func-



(a) Outcome Function (b) Effort Function

Figure 10: Outcome and Effort Functions for various class sizes

tions are seen to shift upwards for the majority of ability types. In other words, by trying to cluster students of similar abilities in the same classroom we increase the achievement and effort levels of the majority of students (only the highest ability students are encouraged to work less). The intuition here is that with greater clustering, students are less discouraged by the prospects of having much more talented students in the classroom with them, and this increases their estimation of the probability of attaining the top mark. Therefore, if we are concerned with maximizing the educational attainment of the classroom, we would like to increase the extent to which students of similar ability levels are grouped together. Interestingly, this is often done by the population segmenting Placement Tests discussed earlier! This also offers a justification for grouping high ability students together in 'advanced courses' in grade school.

2.7 Limitations of the model

We are now prepared to discuss some of the important drawbacks and limitations of this model. For one, the model assumes a particular functional form for the outcome function and cost function. Although these assumptions are not thought to be driving the main results (i.e. we believe that these results are robust to many different functional form specifications) there is no way of testing this hypothesis since the model becomes very complicated to analyze when different forms are imposed, even for the computer. Apart from this, we still have the unrealistic implication that no student will ever 'fail' a course, and this is primarily due to the deterministic outcome function as well as the assumption of certainty over ability. In reality, the classroom more closely resembles a hybrid of our first model and the current model, and indeed, this was originally the model I explored. However, for purposes of analytical tractability, it was not possible to study these two models together and that is why they were eventually broken up. Another problem with this model is that it predicts a very strange non-symmetric grade distribution, one that places the vast majority of students at low grade levels and only a small handful at the top. Such a left-skewed grade distribution does not accord with the familiar bell-curved distribution that most teachers seek to attain. A more realistic model, that would partially get around this problem would be one that included a stratified reward set $V_1 > \ldots > V_k$ corresponding to different absolute outcome levels. Such an additional complexity would not change the character of our main results, but it would tend to replicate the behaviour we observe more closely by spreading out the grade distribution.

2.8 Other Applications

While I have dealt exclusively with the setting of the classroom in the formation and presentation of this model, there are actually many other settings that resemble the essential features of 'classroom' competitiveness. Take, for instance, the non-commissioned salesman or telemarketer who must sell a certain amount of products to avoid being fired (sales quota) but who gets a bonus or promotion for outselling all of the other salesman in the branch. Replace 'classroom' with 'firm', 'student' with 'salesman', and 'final grade' with 'total dollar sales' and essentially the same story carries over. On the other hand, here the Manager of the firm ('teacher') almost certainly cares most about maximizing average performance since this will maximize total revenue. Thus, in this setting, the shepherd vs. winemaker debate is moot. For the teacher, the objective function isn't as clearly defined, and so there is an inherent trade-off.

Another setting analogous to the classroom, in many ways, is the world of academia. Professors would like to work hard enough to attain tenure (a difficult task in itself), but beyond that, the next plateau of significant rewards would seem to go to the select few who attain star status for being at the top of their field. Replace 'student' with 'professor', 'final grade' with 'research output' and 'passing mark' with 'tenure' and the same basic story carries over. This parallel is probably only partially true, since professors would seem to get some type of 'commission' off of every publication, either in the form of pure intrinsic utility, or in the form of extrinsic rewards such as research grants, offers from other schools, etc. At least at a general level, however, it would seem that incentives are lined up in the same way as the classroom: rewards for attaining a specific bar of achievement, beyond which rewards are given out on the basis of relative performance. Ironically, this is the very way in which the incentives are aligned for the QED Master's Essays. There is some reward for submitting a paper that is basically acceptable (namely, a passing grade and diploma) but the next substantial extrinsic reward is reserved exclusively for the top paper (namely, the prestigious Scarthingmoor Prize). How strange and convenient it is when the very paper you write is an application of itself!

3 Concluding Remarks

Admittedly, the results obtained and methodology used herein are easily grasped by the Honours Undergraduate student. If there is anything novel to be found in this paper, it is in the application of well-known tools to a new and interesting area. Indeed, this paper is the first to consider the ways in which Auction Theory can be applied to the classroom setting! In fact, if one is willing to be creative, 'student' and 'classroom' can be placeholders for many other occupations and competitive settings, as we saw in the previous section. Even still, the classroom - it seems to me - is a very fruitful and important setting in and of itself, even if it is somewhat narrow. As Plato said, "The direction in which education starts a man will determine his future in life." Thus, understanding and predicting behaviour in the classroom, particularly in grade school, could prove quite instrumental in predicting behaviour later on in life, and in other settings of economic interest. Perhaps, as I have suggested in this paper, if we can understand the mechanisms behind Student decision making, we can choose policy instruments to ensure that our educational system sets youth in the *right* 'direction'. For example, if the student registration and effort allocation problem resembles the TPLM developed in Part 1, then that gives us some ideas of how to correct the problem of dropping-out (namely to reduce $|\hat{\theta} - \theta|$ through the various measures suggested in Section 1.6). Similarly, if classroom competition resembles the WTAM developed in Part 2, then that tells us how effort and outcome attainment will be affected by changing class sizes and by clustering students of similar ability. In the end, we are still left with important value judgements when choosing the optimal policy, as in the inherent trade-off faced by the benevolent professor, but at least we've reduced the problem to something that is manageable and somewhat quantifiable.

I believe that it is this type of theoretical modeling, coupled with empirical estimation and verification, that will help direct administrators in developing sound educational policies. Too often, educational studies are based on analysis of pure statistical correlation.[28]. While these are valuable and informative, they lack the explanatory power of a well crafted theoretical model which endogenizes the decision making of the

[28]Such as surveys of drop-outs, regressions of class size on student achievement, etc.

individual. Tinto (1975) writes,

> Despite the very extensive literature on dropout from higher education, much remains unknown about the nature of the dropout process. In large measure, the failure of past research to delineate more clearly the multiple characteristics of dropout can be traced to two major shortcomings; namely, inadequate attention given to questions of definition and to the development of theoretical models that seek to explain, not simply to describe, the processes that bring individuals to leave institutions of higher education...knowing, for instance, to what degree an individual's measured ability and social status relate to the probability of his leaving college does not mean knowing how these attributes affect the process of dropping out from college. Whereas the former requires little more than a simple comparison of the rates of dropout among individuals of differing ability and social status characteristics, the latter requires the development of a theoretical longitudinal model that links various individual and institutional characteristics to the process of dropping out from college.

As mentioned before, their are many limitations and drawbacks to the models presented herein. In principle, future work could be directed towards correcting these limitations so that the models are more accurate representations of reality. In particular, more thought should be given to the idea of a 'noisy' outcome function, a stratefied set of rewards, and an interplay between effort and ability. Also it seems to me that the phenomena of 'procrastination' ought to be accounted for in any multi-period model where effort must be exerted well before rewards are reaped.[29] Perhaps the inclusion of hyperbolic discounting into each model could fix this problem and lend us some interesting conclusions. In any case, I have just begun to scratch the surface of the Problems of the Student, and with these simple models I've uncovered many interesting implications. Along the way, I've tried to relate these implications to what we actually observe in the real world, but more work should be done to this end. In particular, the so called *leisure-substitution*, *competitive*, *encouragement*, *discouragement* and *superstar* effects seem to accord with what I've called 'casual empiricism'; which is to say, they agree with what I've observed throughout my 17 years as a student. However, future work should be directed towards measuring these

[29] As any student will certainly agree!

effects empirically to determine whether they actually exist to a significant degree.

In conclusion, I believe I have succeeded in answering my essential question, “Are their other salient features of Student behaviour that can be modeled using the standard Economic toolset?”. While my models are, admittedly, somewhat limited in range and scope, I hope that they have at least illuminated in some small way the vast and rich area of behavioural research that can be found in the setting of the classroom.

References

Barut, Y., and D. Kovenock (1998) ‘The symmetric multiple prize all-pay auction with complete information.’ *European Journal of Political Economy* 14, 627–644

Baye, M., Kovenock D. de Vries C. (1996) ‘The all-pay auction with complete information.’ *Economic Theory* 8(2), 291–305

Bowen, W., and D. Bok (1998) ‘The shape of the river: Long term consequences of considering race in college and university admissions.’ Princeton, NJ: Princeton University Press.

Chiasson, M. (2008) ‘On the problems of the student.’ *QED Master’s Essay*

Cook, Philip J., and Robert H. Frank. (1996) ‘The winner-take-all society: Why the few at the top get so much more than the rest of us.’ London: Penguin Books.

Dasgupta, P. (1986) ‘The theory of technological competition.’ In: J. E. Stiglitz and G. F. Mathewson (eds.) New developments in the analysis of market structure, pp. 519-547. Cambridge: MIT Press

Dixit, A. (1987) ‘Strategic behavior in contests.’ *American Economic Review* 77, 891–898

Eckstein, Z., and K. Wolpin (1999) ‘Why youths drop out of high school: The impact of preferences, opportunities, and abilities.’ *Econometrica* 67(6), 12951339

Frank, R.H. (1999) ‘Higher education: the ultimate winner-take-all market?’ Working Paper, Retrieved July 2009 from Cornell University, ILR School site: `http://digitalcommons.ilr.cornell.edu/cheri/2`

Hernandez, M. (1997) ‘A is for admission.’ New York: Warner Books.

Hoxby, C. (1997) ‘How the changing market structure of u.s. higher education explains college tuition.’ NBER Working Paper (W6323).

Josephs, R., Bosson J. Jacobs C. (2003) ‘Self-esteem maintenance processes: Why low self-esteem may be resistant to change.’ *Personality and Social Psychology Bulletin* 29(7), 920–933

Lazear, E. P., Rosen S. (1981) ‘Rank-order tournaments as optimum labor contracts.’ *Journal of Political Economy* 89, 341–364

Lloyd, N. (1978) ‘Prediction of school failure from third-grade data.’ *Educational and Psychological Measurement* 38(4), 1193–1200

Martinez, P., and F. Munday (1998) ‘9000 voices: Student persistence and dropout in further education.’ FEDA Report Vol. 2 No 7. London, UK: Further Education Development Agency.

Mishel, L., and R. Rothstein (1997) 'The class size debate.' Published by the Economic Policy Institute, Washington, DC

Narasimhan, C. (1988) 'Competitive promotional strategies.' *Journal of Business* 61, 427–449

Pytel, B. (2006) 'Dropouts give reasons.' Online Publication. www.educationalissues.suite101.com

Tinto, V. (1975) 'Dropout from higher education: A theoretical synthesis of recent research.' *Review of Educational Research* 45(1), 89–125

A TPLM with prior given by $\Gamma(\theta)$

Here we relax the assumption that the students prior on their ability is simply a point $\hat{\theta}$, and allow it to be a general distribution given by cdf $\Gamma(\theta)$.[30] As in our derivation of the equilibrium in Section 1.2 we begin in Period 2 and proceed by backwards induction:

Period 2: Given θ and e_1 the optimal behaviour is given by,

- Continue and set $e_2 = \underline{e_2}(e_1, \theta)$ if $V \geq c(\underline{e_2}(e_1, \theta))$
- Drop-Out if $V < c(\underline{e_2}(e_1, \theta))$

Where $\underline{e_2}(e_1, \theta) = f^{-1}\left(\frac{\overline{u} - \alpha \cdot g(\theta) f(e_1)}{g(\theta) \cdot (1-\alpha)}\right)$ is the minimum amount of effort required to pass, as before.

Let θ_c be the type that finds themself indifferent between Continuing and Dropping-Out after providing e_1 in Period 1 (i.e. $\theta_c(e_1)$ satisfies $V = c(\underline{e_2}(e_1, \theta_c(e_1)))$. We are now ready to examine the optimal behaviour in Period 1.

Period 1: Given the optimal behaviour described in Period 2, if the student decides to Register and puts in effort e_1 then they will receive,

- $-R - c(e_1)$ if $\theta < \theta_c(e_1)$ (they will Drop-Out)
- $V - R - c(e_1) - c(\underline{e_2}(e_1, \theta))$ if $\theta \geq \theta_c(e_1)$ (they will Continue)

Thus we can take an expectation of their payoff with respect to their prior over θ and we obtain,

$$\begin{aligned} EU(e_1) &= \int_0^{\theta_c(e_1)} (-R - c(e_1))\, d\Gamma(\theta) + \int_{\theta_c(e_1)}^{\infty} \big(V - R - c(e_1) - c(\underline{e_2}(e_1, \theta))\big)\, d\Gamma(\theta) \\ &= \underbrace{-R - c(e_1)}_{\text{Sunk cost of Registering}} + \underbrace{\int_{\theta_c(e_1)}^{\infty} \big(V - c(\underline{e_2}(e_1, \theta))\big)\, d\Gamma(\theta)}_{\text{Expected premium for Passing}} \qquad (25) \end{aligned}$$

If (25) is positive for some $e_1 > 0$ then the student will Register, otherwise they won't. Examining the two terms in (25) we see that the first term is decreasing in e_1 whereas the second term is increasing in e_1 (both the argument of the integral increases and the limits of integration are expanded). This expresses the basic trade-off faced by the student - by registering and working hard in Period 1, they incur a large sunk

[30] To avoid confusion, we use the greek letters Γ and γ to refer to the cdf and pdf over θ respectively. We could use $\mathcal{F}$ and f but this might confuse the reader with the f from the students outcome function $u(\theta) = f(e) \cdot g(\theta)$.

cost but simultaneously increase their chances of passing in Period 2 and obtaining benefit V. Another thing to notice is that the second term in (25) is larger when the student places much of the probability mass on large values of θ (e.g. when the pdf $\gamma(\theta)$ is heavily right-skewed). Thus, the greater the students confidence over their ability, the more likely that (25) will be positive for some level of effort, and the more likely that the student will Register. Conversely, the less confident that the student is over their ability, the less likely that they will find it optimal to Register. Although it is not easy to derive an exact condition on Γ that will characterize the registration decision, as we had done in Section 1.2, the basic idea still holds - and the unfavorable scenarios of overconfidence and underconfidence will still be present. Even though we can't derive a characterizing rule on Γ it is, however, still interesting to see how the optimal level of effort would be derived in this more general setting.

Let's suppose that (25) is positive for some $e_1 > 0$, to derive the optimal amount of effort we will first want to calculate $\frac{dEU}{de_1}$,

$$\begin{aligned}
\frac{dEU}{de_1} &= -c'(e_1) + \frac{\partial}{\partial e_1} \int_{\theta_c(e_1)}^{\infty} \big(V - c(\underline{e_2}(e_1, \theta))\big)\, \gamma(\theta) d\theta \\
&= -c'(e_1) + \big[V - c(\underline{e_2}(e_1, \theta_c(e_1))\big]\, \gamma(\theta_c) \cdot \frac{d\theta_c}{de_1} - \int_{\theta_c(e_1)}^{\infty} c'(\underline{e_2}(e_1, \theta)) \left(\frac{\partial \underline{e_2}}{\partial e_1}\right) \gamma(\theta) d\theta \\
&= -c'(e_1) + \int_{\theta_c(e_1)}^{\infty} c'(\underline{e_2}(e_1, \theta)) \left(-\frac{\partial \underline{e_2}}{\partial e_1}\right) \gamma(\theta) d\theta \qquad (26)
\end{aligned}$$

The second step uses Leibniz's rule for differentiation under the integral sign,

$$\frac{d}{d\alpha} \int_a^b f(x, \alpha) dx = \int_a^b \frac{\partial}{\partial \alpha} f(x, \alpha)\, dx + f(b, \alpha) \frac{db}{d\alpha} - f(a, \alpha) \frac{da}{d\alpha} \qquad (27)$$

and the third step makes use of the fact that $V = c(\underline{e_2}(e_1, \theta_c(e_1))$. Thus, if we set (26) equal to 0 we obtain an expression characterizing the critical points of $EU(\cdot)$,

$$\int_{\theta_c(e_1)}^{\infty} c'(\underline{e_2}(e_1, \theta)) \left(-\frac{d\underline{e_2}}{de_1}\right) \gamma(\theta) d\theta = c'(e_1) \qquad (28)$$

Now, we can further simplify this expression by assuming that cost is linear and by noting that $-\frac{d\underline{e_2}}{de_1} = \frac{f'(e_1)}{f'(\underline{e_2})}$ when $\alpha = \frac{1}{2}$. Thus (28) becomes,

$$\int_{\theta_c(e_1)}^{\infty} \left(\frac{f'(e_1)}{f'(\underline{e_2}(e_1, \theta))}\right) \gamma(\theta) d\theta = 1 \qquad (29)$$

Without specifying particular functional forms it is impossible to derive an explicit expression for e_1^*. However, whatever the solution happens to be, it is clear that the density function $\gamma(\theta)$ must have sufficient mass on the interval $(\theta_c(e_1^*), \infty)$, else (29) will not hold. Thus, whatever the optimal level of effort e_1 works out to be, the

Student must be sufficiently confident that they would be able to pass in the second Period with that amount of effort.[31] In fact, if we set $e_1 = 0$ we obtain the equation,

$$\int_{\theta_c(0)}^{\infty} \left(\frac{f'(0)}{f'(\underline{e_2}(0,\theta))} \right) \gamma(\theta) d\theta = 1 \tag{30}$$

Thus provided that the $f(\cdot)$ is sufficiently concave so that $\frac{f'(0)}{f'(\underline{e_2}(0,\theta))}$ is large (i.e. large diminishing returns to effort), and provided that the student is confident enough so that there exists significant probability mass on the interval $(\theta_c(e_1^*), \infty)$ the left hand side of (30) will almost surely be larger than the right hand side (i.e. $\frac{dEU}{de_1}|_{e_1=0} > 0$). This means that the optimal amount of effort in period 1 is almost always positive (as we would hope it would be). If, on the other hand, f is approximately linear so that effort is almost perfectly substitutable across time periods, and the student is not very confident in their ability, then it pays to not exert much effort in the first stage, learn your type, and then put in the lion's share of effort in the second stage, but only if that turns out to be optimal given what you've learned. This result is mainly due to the assumption that learning is perfect regardless of the amount of effort put in, so that the student can test the waters with only a minimum of effort, and bail out at little expense if they learn something unfavourable.

[31] More accurately, they would need to be sufficiently confident in the fact that they would be *willing* to pass in the second period given e_1. The distinction between 'willing' and 'able' is important because students are always able to pass, but they might not be willing to put in the effort unless it is optimal.