THE PENNSYLVANIA STATE UNIVERSITY
SCHREYER HONORS COLLEGE

DEPARTMENT OF ECONOMICS

DYNAMIC MODEL OF CLASSROOM PERFORMANCE

MONTANA MORRIS
SPRING 2018

A thesis
submitted in partial fulfillment
of the requirements
for baccalaureate degrees
in Physics and Economics
with honors in Economics

Reviewed and approved* by the following:

Joris Pinkse
Professor of Economics
Thesis Supervisor

Russell Paul Chuderewicz
Teaching Professor of Economics
Honors Adviser

* Signatures are on file in the Schreyer Honors College.

ABSTRACT

In primary education, a trend exists where the level of specialization of the teachers tends to increase as grade level increases. A transition from these generalists to specialists occurs as both the difficulty of the material increases and as the maturity of the students' increases. A dynamic performance model will be presented to illustrate this progression through optimal conditions. Not only does this presentation provide an explanation for the varying degrees of specialization, it does so within the confines of a disruption model, where misbehaving students negatively affect classroom productivity. Disruption models provide us with further insight into the choice of classroom size, such as the decrease in the number of classes when the wage market for teachers sees an increase, as Edward Lazear discussed in his disruption model from 2001. Thus, this model of administrative decision will provide a mathematical framework that explains both the change in class size and teacher specialization over the course of primary schooling for a given group of students.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to thank my parents for providing me with unconditional support throughout my academic career and for allowing me to pursue my interests. I would also like to thank my friends for making this journey an enjoyable one.

Finally, I would like to express my appreciation towards my thesis supervisor, Professor Joris Pinkse, for his promotion of curiosity in the classroom and for his patient guidance in this thesis process.

Chapter 1

Introduction

Economists have modeled educational production in a variety of ways to understand decision-making processes and their impacts in recent years. Education is so crucial to social welfare, as it enables children to grow into productive members of society over the course of their young lives. With such an undoubted economic impact, many academics have provided us with these models in order to provide insight into this progression in hopes of ultimately maximizing the systems in place. Many of these models aim to provide optimal choices subject to some set of classroom conditions. In this investigation, we will dive into a similar endeavor. We will begin with a general discussion of disruption models and additional concepts such as distance traveled that will be essential to the dynamic classroom performance model eventually presented.

This dynamic classroom performance model aims to explain optimal class sizes and degrees of subject specialization for a given group of students over the course of their academic careers. This model provides two clear modes where distance is traveled in education, either through social or intellectual adaptation. Over time, we will be able to see the interaction between these two performance parameters and the administrative choices made subject to them. We will explore key situations to ensure appropriate marginal returns over the course of this dynamic model, provide support as to why this model improves upon another highly regarded one, and then will present several shortcomings. Before concluding, this model will be used to qualitatively explain a dataset and provide a basis for further investigation and application.

Disruption Model of the Classroom

In 2001, Edward Lazear released an investigation into class size effects attempting to explain the inconsistent findings of past research. In certain situations, it appeared as larger class sizes resulted in greater productivity while in other circumstances a greater teacher to student ratio was preferred. A disruption model for a single group of students, holding factors such as educational level and ability constant, was presented to explain these intriguing empirical findings. With this disruption model, behavior drove these ambiguous class size effects. Essentially, it assumed that one bad apple would spoil the bunch for a given group of students' classroom experience, and Lazear deployed what he deemed "comparative statics" to understand the decision making of an educational administrator when considering how many teachers to employ at a given wage for a given group of uniform students.

For this model, some probability of good behavior, p, was used to describe the return to education. Considering each student's behavior as independent, although in reality some relationship likely exists, the probability of n students to create a healthy classroom environment is p^n. In the case where not all students behaved, $1 - p^n$, there simply was no return to education. This is a strong assumption but was initially adequate to gain intuition into the disruption model. Then, VZp^n was deemed the return to education, where Z was the total number of students and V was used to dollar value this return on the students. Considering this as an ultimate return to education, the economic cost of employing these teachers had to be considered as well where each teacher earned some wage, w, and the total number of classes was equal to total students divided by class size, Z/n. The following model and first order condition presented itself.

$$E(\pi) = VZp^n - \frac{Zw}{n}$$

$$First\ Order\ Condition:$$

$$VZln(p)p^n = -\frac{Zw}{n^2}$$

This optimality condition provides some crucial intuition. At higher wages, holding all else constant, a relatively larger class size, n, was the optimal decision for the educational institution. Furthermore, with a better behaving group of students, the resulting probability of a healthy classroom, p^n, was higher. This improved probability also resulted in a relatively larger class size at optimality (Lazear, 2001).

Distance Traveled in Education

While Lazear's unique model addressed differing class size effects and presented some strong intuition, it did not discuss any time evolution of this return to education. Much of the discussion in modern education surrounds maximizing the "distance traveled" by an institution's students. This class size decision simply does not exist in a single period, and this ultimate return to education in terms of a student's productivity is realized long after the education has been attained at graduation and beyond. The research surrounding distance traveled today primarily revolves around measuring this distance traveled in order to incentivize educators appropriately. Using the above model from Lazear, this implies that each of the individual teachers provides some different return to education, V. The following improvement on the above model would capture some of this type of analysis.

$$E(\pi) = \left[\sum_{i=1}^{\frac{Z}{n}} V_i \frac{Z}{n} p^n \right] - \frac{Zw}{n}$$

As is the case with the recent developments, the primary question has been how can we effectively measure this return on each classroom, V_i, in order to appropriately incentivize each individual teacher and differentiate wage, w, into wages, w_i, rather than taking the wage as a given from the labor market. Such attempts have been rather fruitless, with a primary focus on using standardized test scores and statistical analysis techniques. In “Teacher Effects, Value-Added Models, and Accountability,” Spyros Konstantopoulous presented some of the difficulties econometricians have faced when quantizing the differences in these distances traveled. He attributed much of the difficulty to an inability to adjust appropriately for specifications between students within a given group. In order to estimate teacher effects accurately, one must be able to adjust for any underlying differences between classrooms such as student ability. Another primary concern raised was the influence of choice on these forms of models; often statisticians assume the random assignment of students to classrooms, yet the effect of parental influence is undoubted. Parents typically choose which schools their children will attend and may even influence which teachers they may learn from within a given system (Konstantopoulous, 2014).

The above model in such a simplified form even explains these differences and difficulties in accepting value-added models. It is assumed that all teachers within a given group of students face the same probability of bad behavior and furthermore that all students have the same probability for bad behavior. This assumption is simply not the case, so accurately assessing the teacher’s effects on the return to education has become an extremely difficult task. We will leave the empirical difficulties to other discussions, but its connection to this discussion is extremely pertinent as the dynamic performance model presented aims to classify and understand this very ability of teachers to add value.

Specialization in Today's Classroom

Not only do factors such as ability and behavioral traits influence the amount of a value a teacher may add to his or her students, the actual course curriculum has a significant influence as well. For starters, this curriculum must align with that which is actually going to add economic value to the student in the long term. There has been a significant amount of investigation into signaling, the idea that graduation signals ability to employers and that the actual education is less relevant, or even irrelevant to productivity and placement, after its introduction as a theory in 1973 (Spence, 1973). Many economists have investigated and contributed to this job market signaling theory in the years that have passed, and while they have found merit at the secondary education level, near a boundary between education and workforce, very little research has applied signaling to the primary level, likely because it is simply inapplicable to students at a younger age. Technical skills such as mathematics develop a student's critical reasoning and logic while the language arts improve a student's ability to communicate which is of utmost importance as a productive member of society later in life. If we assume that education actually does provide value, then the question arises as to what methods are most effective at transmitting this information to the students.

At lower grade levels, we often see a single instructor teaching all of the subjects. As education level progresses, these instructors may become more specialized where one teacher handles language arts and social sciences while another handles mathematics and natural sciences. This specialization progresses until a single instructor is designated an individual topic at upper grade levels. With this very common structure in place, the pertinent researcher would inquire into its optimality and attempt to understand the phenomenon at work. According to research from Roland Fryer, a professor at Harvard University, specialization at the primary level simply does not boost standardized test results. Over the course of his research, he gathered data on 12,715 primary students over 50 school districts in Houston, Texas between 2013 and 2015. Not only did he not see an improvement with specialization, he saw an average impact on student achievement in specialized settings of -0.042 standard deviations in math and -0.034 standard deviations in reading comprehension per year (Fryer, 2016). Although test results may not be the best indicator of value added as discussed above, the data is remarkable nonetheless. If a student has the opportunity to learn mathematics from an established mathematician and linguistics from an established linguist, why would it still be optimal to have this individual study from a sole source?

In another investigation, the International Journal of STEM education surveyed primary educators for their insight into the young student's mind and classroom. For specialists at this level, planning time was the primary positive. They explained how the reduced amount of courses allowed them to more thoroughly plan interactive activities rather than old-school worksheets and better meet increasing standards. However, they all also noted the importance of getting to know the students' personalities and characteristics in order to understand the needs that the learning environment requires (Markworth, 2016). There must be some trade-off between the education and experience of a specialist versus a generalist, and some model must explain why the return to specialization was actually less than that of a generalist from the research carried out in Houston, Texas by Roland Fryer. In an attempt to depict these conditions with greater accuracy, we will time evolve Lazear's disruption model and make several adjustments as will be discussed in the next section.

Chapter 2

Theoretical Framework

Before diving deeper into the exploration of the educational decision-making process surrounding class specialization and class size, several mathematical objects will be defined that are pertinent to the later discussion. These mathematical objects simply create a solid framework for the passage of time and provide an ease of adjustment for future findings or research.

Payoff Vectors

To begin our definitions, we will introduce a payoff vector. This vector holds both the costs and returns to education at a given time. As these returns to education are not explicitly dollar-valued as revenue might be, allow this vector to be a slightly more abstract entity. At any given time, the expected profit will result from the collapse of the payoff vector. This occurs through a dot product between the payoff vector and a realization vector. At this time, the process likely seems rather trivial, but this will become more important as the discussion progresses. An example is provided below.

$$Payoff\ Vector, P = \begin{bmatrix} Return \\ Cost \end{bmatrix}$$

$$Realization\ Vector, R = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$E(\pi) = R^T P = Return - Cost$$

Educational Transformation

Defining the payoff vector allows for a simple representation of changes that may occur due to administrative decisions made over time. As these returns to education are more abstract than concrete revenue, consider this vector with a similar notion. Lazear's disruption model introduced a parameter to dollar-value the return per student, but in reality, the goal was to optimize the potential of a given set of students in the educational system. This dollar-valued productivity would not be realized until long into the future, so drawing this conclusion directly was made for simplicity. This payoff vector could also be defined as a decision vector, as it gives intuition to decisions made rather than profits earned. Now, with a payoff vector explicitly defined, we are able to evolve our returns over time through matrix transformations. Below shows the basic principle.

$$Transformation, T = \begin{bmatrix} Return\ Modification & 0 \\ 0 & Cost\ Modification \end{bmatrix}$$

$$Transformed\ Payoff\ Vector = TP$$

$$= \begin{bmatrix} Return\ Modification & 0 \\ 0 & Cost\ Modification \end{bmatrix} \begin{bmatrix} Return \\ Cost \end{bmatrix}$$

$$= \begin{bmatrix} (Return\ Modification)(Return) \\ (Cost\ Modification)(Cost) \end{bmatrix}$$

Very basic notation has taken up much time through this point and with purpose. To date, very little theoretical literature exists revolving around distance traveled in education. The focus has been solely on statistical methods and finding ways to quantify the progress of students.

They have worked through a variety of statistical modeling techniques, primarily centered on standardized test scores, in attempts to select out what teachers have been able to provide to students. Quantifying this distance traveled should be the ultimate goal. With a better understanding of that distance traveled, we can align teacher incentives appropriately, thus allowing for optimal decisions that improve the welfare of society as a whole, considering the utmost importance of education. While most time was spent in this empirical mindset, theoretical modeling was neglected, which would only improve those empirical applications in the long run.

If much of the rest of this paper is taken for granted, this foundation of transformations with regard to educational distance traveled should be appreciated and investigated further. It was used in a variety of other economic modeling scenarios and natural science applications, such as the case with quantum physics. In those quantum instances, objects undergo transformations where information is withheld from the experimenter. When measurements occur, the objects fundamentally change and only distributions of repeated experiments provide us with some picture of that withheld information. Such may be a very similar case with education. Certain characteristics, such as innate ability, which are so fundamental to productivity, are impossible to perceive. We can only force measurements to gain some form of intuition to the underlying attribute, as is the case with standardized testing and signaling theories.

Although in this case a payoff vector will evolve to understand administrative decision-making, the application to understand student progress may necessitate a different form. Potentially, this evolving vector is one of student characteristics, where some are constant under transformation, such as innate ability, while others are able to mold to societal requirements, like the case of probability of good behavior in this evolution of Lazear's disruption model. In providing a stronger theoretical framework, clues as to why it has been so difficult to quantify distance traveled may present themselves. Potentially, the data regarding standardized testing and distance traveled between different teachers has been so unclear because by design it should be. If these tests primarily measure something such as innate ability, where teachers hold no influence, insignificant results would make sense.

Chapter 3

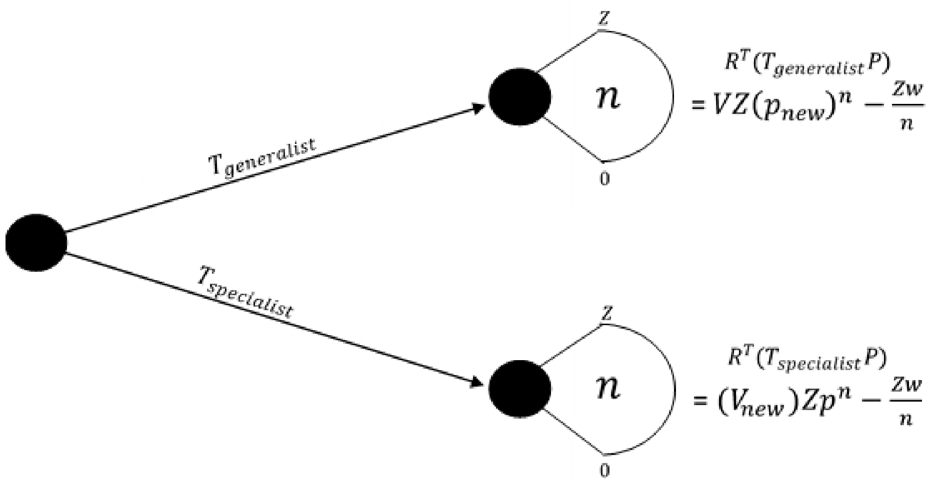
Specialization Decision

In this section, the formalisms presented in chapter 2 will be used to discuss the decision to specialize as well as its effect on optimal class sizes from an administrative perspective. The payoff vector and transformation notation will prove useful when expanding educational models into multiple periods. Before that discussion, I will introduce a model of administrative choice in a single time-period. Once that model is established, a dynamic disruption model will be introduced with several key flaws highlighted. I will then improve upon these weaknesses and build up towards the performance model.

Static Model

In one sole period, the administrative decision-maker must decide whether the return to the specialist or generalist model would be superior in only two steps. From a decision perspective, the process would unfold as such: The decision-maker would choose a class size, n, such that both situations optimize profits, where he or she choose either the generalist or specialist model. Then, the decision-maker chooses whether to specialize based on the greater of the two of those resulting profits. A summary of this situation is presented below.

Figure 1: Static Allocation Decision



$$R^T(T_{generalist}P_0) = VZ(p_{new})^n - \frac{Zw}{n} \; ;$$

$$choosing \; n \; s.t. \qquad VZln(p_{new}) * (p_{new})^n + \frac{Zw}{n^2} = 0$$

$$R^T(T_{specialist}P_0) = (V_{new})Z\, p^n - \frac{Zw}{n} \; ;$$

$$choosing \; n \; s.t \qquad V_{new}Zln(p)p^n + \frac{Zw}{n^2} = 0$$

In this basic situation, some interesting intuition presents itself. As Lazear already examined, in both cases, larger class sizes resulted optimally in situations where the probability of good behavior and teachers' wages were highest. However, in this short scenario, further insight is drawn into the interaction between these optimal class sizes and specialization. In situations where the school invested into the generalist model, the resulting class size ends up relatively larger than it would have been if specialization occurred given the improvement in behavior, as $p_{new} > p$. Of critical importance, however, is that it does not imply that generalist class sizes are typically larger than specialized class sizes. Rather, it implies that for a single institution, class sizes would grow over time when making the decision to avoid specialization.

On the opposite end of the spectrum, class sizes would actually decrease over time in an environment where specialization continues to occur. The calculus is less revealing in this situation, but simulations allow us to arrive at this result. The case within post-secondary education, despite not being the primary application of this model, provides a clear illustration of the phenomenon. General education and entrance-to-major courses typically have a greater number of students than a major course near graduation.

Although all students are required to take math and language arts at the primary level while the post-secondary level provides more course flexibility, the situation should be no different, as the total student body, Z, has no effect on decisions. A student's choice to enter or not enter that population is irrelevant to the decision maker, thus this situation should mirror the conditions at the primary level, with a more pronounced affect due to a greater degree of specialization. Utilizing MATLAB with the parameters fit to create realistic class sizes, we can use optimization routines to recover these relationships. The following tables express this result.

Table 1: Returns at Varied Probabilities of Good Behavior

$Parameters: w = 10, Z = 100, V = 0.9$		
Optimal Class Size, n	Behavioral Probability, p	Resulting Return, $E(\pi)$
31.03	0.9750	8.80
31.36	0.9764	10.62
31.76	0.9777	12.52
32.24	0.9791	14.52
32.79	0.9805	16.62
33.45	0.9818	18.82
34.21	0.9832	21.15
35.11	0.9845	23.61

36.17	0.9859	26.22
37.43	0.9873	29.01
38.95	0.9886	31.99
40.82	0.9900	35.22

Table 2: Returns at Varied Returns per Student

$Parameters: w = 10, Z = 100, p = .975$		
Optimal Class Size, n	Classroom Return, V	Resulting Return, $E(\pi)$
31.03	0.9000	8.80
30.04	0.9364	10.48
29.14	0.9727	12.20
28.31	1.0091	13.95
27.55	1.0455	15.75
26.84	1.0818	17.57
26.18	1.1182	19.43
25.56	1.1545	21.32
24.99	1.1909	23.24
24.45	1.2273	25.19
23.94	1.2636	27.16
23.46	1.3000	29.15

The above tables can be viewed as snapshots over different parameters. The first selection holds all variables constant except for probability while the second situation varies only the return to a classroom given good behavior. As is evident from above, class sizes increase with increasing probability of good behavior. In the case of increasing returns from specialization, class sizes fall over time. However, both cases see improvements in the expected returns providing some crucial insight. Class sizes themselves are not the focus of the decision-maker; rather they react to the optimal choice. An educator would always choose a higher return to a well behaved classroom, $V > V_0$, and would always choose a higher probability of good behavior, $p > p_0$, disregarding any effect on class size.

Multi-Period Choice

When considering decisions made during only one period, the previous tables appear as mere snapshots providing some basic intuition. However, with a slight change in viewpoint, these tables become rather illustrative. Consider them instead as simulations over the course of 11 periods. Both cases began with a group of students at a probability of good behavior, p=.975, and a return to education, V=0.9, after say kindergarten. In this new description, a different transformation occurred to two initially identical groups, meaning both models began with the same selection of parameters. Then, the schools chose to be specialists or generalists indefinitely, facing the following transformations.

$$\left(T_{generalist}\right)^{12} = \begin{bmatrix} \left(\frac{p + 0.0014}{p}\right)^{n} & 0 \\ 0 & 1 \end{bmatrix}^{12} ; \left(T_{specialist}\right)^{12} = \begin{bmatrix} \frac{V + 0.0364}{V} & 0 \\ 0 & 1 \end{bmatrix}^{12}$$

Of note, these linear additions to probability towards good behavior and returns to education were chosen to illustrate realistic class sizes arbitrarily. Up until this point, the payoff vector notation did not seem particularly useful when a short-term decision occurred with returns quickly realized. Considering the fact that these educational decisions begin in kindergarten and end in twelfth grade, the application becomes more useful. Rather than considering a singular profit, we must consider a summation of profits over the course of all of those grades. This summation must account for the fundamental change that occurred at each grade level as well as the pure return from each grade level. I.e., there was some concrete return from that year of education, and some lasting effect from that year of education on the return of future education persisted, much in the way that content matter builds upon itself. We begin with a group of students, seemingly raw inputs. They undergo a series of transformation processes where each step adds value subject to some cost. These transformations result in a final product of substantially higher value than the return to each individual step, thus implying that not only are the additions important, the interaction between those additions are as well.

As each of these time periods progress, the administrator decides whether or not to specialize and then selects optimal class size subject to that decision of specialization. From an optimization perspective, the allocator would not only be selecting a transformation and class size based on this current grade level payoff; he or she would be selecting transformations based on future returns to education as well. The ultimate goal becomes optimizing a group of students' total return on their educational experience. The situation makes the most sense when we relate it back to the purpose of education. While the learning process may be enjoyable to an individual student, its true purpose, particularly in the public setting, is to improve the economic productivity of that group of students in the long term. The economy does not realize this productivity until after the educational experience is complete, so optimizing in a cumulative nature is more practical. The following illustrates this general idea.

$$E(\pi) = R^T(P_1 + P_2 + P_3 + \cdots + P_{13})$$

$$E(\pi) = R^T(T_1P_0 + T_2P_1 + \cdots + T_{13}P_{12})$$

$$E(\pi) = R^T(T_1P_0 + T_2T_1P_0 + \cdots + T_{13}T_{12}T_{11}T_{10}T_9T_8T_7T_6T_5T_4T_3T_2T_1P_0)$$

$$E(\pi) = R^T[\sum_{i=0}^{12}\{\prod_{j=0}^{i} T_jP_0\}]$$

Marginal Return Issues

With the resulting expected profit now well defined, consider the decisions made over the lifetime of an education. In an attempt to gain this clearer understanding, rectifying a fundamental flaw of the transformations previously introduced is necessary. In the simplistic case with indefinite generalists and specialists, the group of students evolved in a linear fashion. That is to say, each year they either gained a set amount of probability towards good behavior or improved their return to good behavior by a set amount. When considering the marginal return provided each year in both scenarios, an interesting situation presents itself. The two are experiencing an increasing marginal return. The discussion of specialization will be brief, as in this case, increasing marginal utility is plausible. When an individual continues to study, say mathematics, in an environment with very strong, specialized educators, the education may become more and more productive. This is the ultimate reasoning behind selecting a major in post-secondary education, but at the primary level, the situation is slightly different. For that reason, assuming a linear evolution in specialization is reasonable but necessitates the analysis discussed in the model presented later.

While the linear evolution in specialization does not come from a point of particularly flawed logic, the evolution of good behavior certainly does. Increasing marginal utility when improving behavior over time makes little sense. Within the confines of the model, this implies that moving an equal amount towards perfect behavior spurs greater and greater benefits. While this may be true in reality, a realization must then occur as to how difficult improvements become as a group nears perfect behavior. There must be some trade-off between generalists and specialists, as is evident in today’s school districts. This alludes to the functional form of this behavioral transformation, one closely tied to social development, and allows a postulation that the perfect classroom environment is only a theoretical environment with asymptotic movement towards it. The following formulation provides an improvement on this issue.

$$p = \frac{\varphi - 1}{\varphi} \; ; \varphi > 1$$

Now, the behavioral parameter, φ, can grow in a linear fashion, but the improvement in probability is unable to. One is unable to draw direct conclusions from the partial derivatives without an explicit formulation for optimal class size, as class size remains a choice variable following the change in the behavior parameter. Without this explicit substitution, imposing the optimality condition on class size prior to theorizing provides an adequate result.

$$choosing\ n = n^* \ s.t. \qquad VZln(p)(p)^n + \frac{Zw}{n^2} = 0$$

$$\frac{\partial\big(E(\pi)\big)}{\partial\varphi} = \frac{\partial\big(E(\pi)\big)}{\partial p} \times \frac{d(p)}{d\varphi} = n^* VZp^{n^*-1} \times \frac{1}{\varphi^2}$$

$$= \frac{n^* VZ(\varphi^2 - \varphi)^{n^*-1}}{\varphi^{2n^*}} > 0$$

A look to the first derivative with respect to the introduced behavioral parameter ensures the first important result that an increase in the behavioral parameter does in fact increase the expected profits. This is evident once the partial derivative is broken down. From the prior simulation, profits increased with the probability of good behavior, and through a complete derivative of the probability of good behavior, a strictly positive result is obtained. The product of these two positive results certainly leads to a positive result. This makes sense. The partial derivative of return with respect to our new parameter, φ, is positive when holding class size constant with all positive parameters. If class size is chosen to maximize profit, there should be no reason to select a class size that would negate this improvement of behavior. Now, a consideration of the second derivative to examine whether the improvement has fixed the issue of increasing marginal utility is warranted.

$$choosing\ n = n^*\ s.t. \qquad VZln(p)(p)^n + \frac{Zw}{n^2} = 0$$

$$\frac{\partial^2\big(E(\pi)\big)}{\partial\varphi^2} = \frac{n^*(n^*-1)VZ(\varphi^2-\varphi)^{n^*-2}\,(2\varphi-1)}{\varphi^{2n^*}} + \frac{n^*VZ(\varphi^2-\varphi)^{n^*-1}\,(-2n^*)}{\varphi^{2n^*+1}}$$

$$= \frac{n^*VZ(\varphi^2-\varphi)^{n^*-2}}{\varphi^{2n^*+1}} \times [(n^*-1)(2\varphi-1)(\varphi) + (\varphi^2-\varphi)(-2n^*)]$$

$$= \frac{n^*VZ(\varphi^2-\varphi)^{n^*-2}}{\varphi^{2n^*+1}}(\varphi)(-2\varphi+n+1)$$

$$= \frac{(-2\varphi+n^*+1)n^*VZ(\varphi^2-\varphi)^{n^*-2}}{\varphi^{2n^*}} < 0$$

$$(-2\varphi+n^*+1) < 0$$

$$\varphi > \frac{n^*}{2} + \frac{1}{2}$$

At this point, the model requires an additional constraint to retain its application. The second derivative with respect to the behavioral parameter must be negative in order to satisfy the need for decreasing marginal returns to improvement in the behavioral parameter. All relevant parameters are positive, and therefore imposing the condition on the behavioral parameter relative to optimal class size selection is plausible when necessary. Often, models exist within some boundary, but this continued accumulation of boundaries and conditions results in a loss of simplicity and intuition. The primary goal is to understand the interaction between class size and specialization, not to explicitly evolve Lazear’s model of disruption, and although the refined definition for the probability of good behavior improved the model, it did so in a cumbersome manner.

Chapter 4

Classroom Performance Model

Despite the inadequate attempts to rectify issues such as increasing marginal returns, the concept of a behavioral parameter provided strong insight into the potential of some form of a disruption model. Lazear made an assumption that students' actions were independent and that this independence allowed a product over the set of students to express their collective probability of healthy behavior in a classroom. Furthermore, he assumed that each student held the same probability for misbehavior. In reality, neither of these assumptions hold true. Moreover, the probability of good behavior at the individual level follows some distribution, which hinges on outside factors such as the home life of students. His definition with regard to constant probability provided the necessary intuition in a single period model but proved inadequate when dynamics were involved.

Not only did the constant probabilities present theoretical issues, their independence did as well. If one student out of a group begins to deviate from acceptable behavior, others are more likely to follow. With this assumption violated, a clear dependence in events is evident. Furthermore, a reminder that the form intends to model the success of a classroom not necessarily the behavior of the students is warranted. With Lazear's disruption model, a single student's deviation from appropriate behavior results in no return from that educational environment. Once again, this result seems rather strict and inaccurate. Moving forward, an adjusted version of the model will be presented that is more adept at handling dynamic situations and overcomes several of these assumptions. Defining several new parameters and adjusting the viewpoint of several parameters already defined will start the discussion. Following this, manipulation of the model will occur to ensure its validity and consider its implications.

Parameter Definitions

To begin the discussion of this performance model, define a vector of student characteristics, $\vec{A}$. This vector includes any potential traits related to a set group of students, such as innate ability, socioeconomic status, etc. For this discussion, the vector will ultimately become trivial but provides a necessary basis for future exploration. Next, redefine the dollar-valued classroom return per student, V, to the dollar-valued intellectual return to education per student, which is a function of student characteristics, $\vec{A}$. This function is variable between periods, as was the dollar-valued classroom return in the dynamic form of Lazear's model. This implies that between time periods, a given set of students with a set of characteristics, $\vec{A}$, can receive different intellectual returns due to a phenomenon deemed intellectual adaptation.

Lastly, resolving the issues with the probability of a successful classroom as was previously defined by the independent product, p^n, must occur while presenting a second form of adaptation to compliment the intellectual adaptation discussed above. Introducing another function of student characteristics is necessary to do so. This function, call it $\alpha(\vec{A})$, will be a real-valued, positive object, which describes the social capacity of a group of students. This social capacity is as an abstract parameter that improves the learning environment as it increases. Thus, it relates inherently to the expected return on a classroom through another cumulative distribution related to class size. Furthermore, replacing the concept that there is some probability of receiving any return to a classroom with the concept that behavior, or social capacity, drives the fraction of maximal performance for a given classroom provides an improved relationship between the model and reality. Consider the following.

$$Fraction\ of\ Maximum\ Performance = \frac{\alpha(\vec{A}) - n}{\alpha(\vec{A})}\ ;\ \ \alpha(\vec{A}) > 0$$

$$\lim_{\alpha(\vec{A})\to\infty} \frac{\alpha(\vec{A}) - n}{\alpha(\vec{A})} = \frac{\infty}{\infty}$$

$$L'H\hat{o}pital's\ Rule \qquad \lim_{\alpha(\vec{A})\to\infty} \frac{\alpha(\vec{A}) - n}{\alpha(\vec{A})} = \frac{1}{1} = 1$$

A smooth definition for the success of a classroom was presented. When this parameter reaches some infinite value, as is described in the limit above, the perfect theoretical classroom environment is evident where the maximum intellectual return on education for a given group of students is achieved. Further, this function is forced to be real-valued and positive to ensure fractional returns are bounded between 0 and 1. One may question what would happen if class size, n, were chosen to be larger than our function, $\alpha(\vec{A})$, but as will be evident shortly, this case is not applicable. Furthermore, considering this success as some complex probability that any return occurs is no longer necessary; rather, there is some percentage of the maximum return, another positive simplification.

Before introducing the rest of the model, the vector of student characteristics previously discussed will be discarded. Introducing this vector when necessary will allow modeling with improved accuracy. For example, if a student has a strong mathematics teacher, he or she may improve his or her return based on his or her own innate logic abilities, or if a student has a strong language arts teacher, he or she may improve his or her return based on his or her own innate linguistic abilities. This implies logic and linguistic abilities are two separate characteristics with separate responses. Below shows the idea an arbitrary functional form.

$$\theta \equiv logic\ ability, \gamma\ \equiv linguistic\ ability, \vec{A} = \begin{bmatrix} \Theta \\ \Upsilon \end{bmatrix}$$

$$Prior\ to\ Exceptional\ Math\ Teacher, \qquad V(\vec{A}) = \ \Theta + \Upsilon$$

$$After\ Exceptional\ Math\ Teacher, \qquad V(\vec{A}) = \ 2\Theta + \Upsilon$$

The general idea is that these transformations can become infinitely complex and embedded into this model, allowing its utilization alongside student data, as was described in the framework of the first section. With this aside, I will return the social capacity function to a social capacity parameter and the intellectual return to education function to an intellectual return to education parameter.

Relational Logic

With the idea of student characteristics set aside, I complied our new parameters within the confines of the disruption model Lazear previously presented. This combination will provide the expected routine to a classroom in a single period, subject to the available constraints and variables. After its introduction below, I will manipulate the model to ensure that the underlying intuition remains sound.

$$E(\pi) = VZ(\frac{\alpha - n}{\alpha}) - \frac{Zw}{n}$$

$$First\ Order\ Condition:$$

$$\frac{\partial E(\pi)}{\partial n} = -\frac{VZn}{\alpha} + \frac{Zw}{n^2} = 0$$

$$n^* = +\sqrt{\frac{w\alpha}{V}}$$

Possessing an explicit expression for the optimal class size subject to the parameters of the model is a highly coveted result. Of note, I only consider the positive square root as a negative class size is not theoretically sound. This will provide a vast amount of improvement with regard to simplicity when considering movements in further directions spatially, such as when I attempted to take partial derivatives before. Now, I will check the validity of the logic. As the wage of the teachers, w, rises, optimal class size rises. This clearly coincides with Lazear's discussion. As the social capacity, α, of the students rises, optimal class size rises as well. In Lazear's model, higher probabilities of good behavior resulted in larger classes. In this case, the social capacity parameter is defined such that it directly relates to the probability of good behavior, which increases in optimal class size so the logic remains sound. Lastly, the inverse relationship between the dollar-valued return as in the previous model, the intellectual return as described in this case, and optimal class size is evident explicitly while only appearing through simulations previously.

Wage Condition

With confirmation that this model adequately describes all of the relationships embedded into Lazear's model, addressing an issue presented earlier becomes pertinent – the potential for a negative fraction of maximal performance. Ensuring this function's behavior is well defined and bounded between zero and one is necessary. By assuming decision makers allocate class sizes optimally, this holds true subject to a new condition on the wage of teachers.

$$\frac{\alpha - n^*}{\alpha} = \frac{\alpha - \sqrt{\frac{w\alpha}{V}}}{\alpha} = 1 - \sqrt{\frac{w}{V\alpha}} \in [0,1]$$

$$\sqrt{\frac{w}{V\alpha}} \in [0,1]$$

$$0 \leq \sqrt{\frac{w}{V\alpha}} \leq 1$$

$$0 \leq \frac{w}{V\alpha} \leq 1$$

$$\frac{w}{V} \leq \alpha$$

$$w \leq \alpha V$$

In the first steps, I broke down the percentage to maximal return and substituted the optimal class size. After substitution, I applied the necessary boundary conditions and proceeded with the necessary algebra. As only positive roots for class size are considered with relevant parameters real-valued and positive, the lower bound is guaranteed. The upper bound, however, does present some new intuition. The ratio of the wage paid to the teacher to the return received per student cannot exceed the social capacity parameter. The wage, w, is taken as is from the labor market. Therefore, with a group of students who are either extremely weak from a social capacity perspective, or extremely weak from an intellectual capacity perspective, it is simply suboptimal to educate them. In a different algebraic sense, the wage paid to a teacher must not exceed the product of that group of students' social and intellectual capabilities. The ultimate return to that classroom setting would be negative, and thus society is better off not participating. From a political perspective, this result is obviously unrealistic and immoral. Any citizen deserves a right to education, as it certainly relates to potential in life. However, from a social welfare perspective, it is rather revealing. Even Lazear began to mention that private schools were at an advantage for their ability to differentiate and decline students who misbehaved. This model provides formal evidence.

Chapter 5

Dynamic Classroom Choice

With the static model of classroom performance well defined, I will move the investigation to the dynamic decision process by beginning with a discussion of marginal returns to ensure appropriate behavior. I will then discuss the conditions on the model that result in certain optimal circumstances in an attempt to understand natural phenomena such as increasing specialization over the course of an education.

Marginal Return Improvements

If this model of classroom performance is to be used to understand dynamic educational decisions, ensuring the appropriate behavior of returns to intellectual and social adaptation is necessary. Both forms of adaptation should see increasing returns, but the marginal returns may vary and become ultimately more revealing. Fortunately, the process of understanding returns using partial derivatives is much easier now, as an explicit definition for optimal class size is available. I will begin this investigation with the social capacity parameter in hopes that the return to education will increase with an increasing parameter, α, and in further hopes that diminishing marginal returns are evident through a negative second derivative. This diminishing marginal return would provide a clear relationship with reality; a limit exists on how much an individuals' behavior can be improved.

$$Let\ E(\pi^*) = E(\pi)|_{n=n^*} = VZ\left(\frac{\alpha - \sqrt{\frac{w\alpha}{V}}}{\alpha}\right) - \frac{Zw}{\sqrt{\frac{w\alpha}{V}}} = VZ - 2Z\sqrt{\frac{wV}{\alpha}}$$

$$\frac{\partial E(\pi^*)}{\partial \alpha} = Z\sqrt{\frac{wV}{\alpha^3}} > 0$$

$$\frac{\partial^2 E(\pi^*)}{\partial \alpha^2} = -\frac{3Z}{2}\sqrt{\frac{wV}{\alpha^5}} < 0$$

The appropriate results were uncovered. All parameters involved are positive, and only positive square roots are considered for the optimal class size. Thus, increasing utility in our behavioral parameter exists alongside decreasing marginal utility. With these results well identified, I will investigate the relationship between the intellectual return parameter and the resulting profits through a similar, yet slightly more intriguing, exercise. In this instance, I hope to see increasing returns to an increase in our intellectual return parameter, and I will allow the marginal return to this increase to provide me with its own intuition.

$$\frac{\partial E(\pi^*)}{\partial V} = Z - Z\sqrt{\frac{w}{V\alpha}} > 0\ ;\ \sqrt{\frac{w}{V\alpha}} \in [0,1]$$

$$\frac{\partial^2 E(\pi^*)}{\partial V^2} = \frac{Z}{2}\sqrt{\frac{w}{\alpha V}} > 0$$

Increasing returns is evident as expected, and increasing marginal returns are presented as well. With further thought, this provides strong implications to reality. A school district rarely starts to specialize courses then return students to a generalist model. Increasing marginal returns to specialization implies a compounding nature, which would help explain why specialization tends to persist once it has started. Not only have does this model meet the necessary conditions for its validity, it has provided an interesting growth pattern to understand specialization persistence. Furthermore, this was done in an explicit manner from an algebraic standpoint, overcoming one primary difficulty with Lazear's model. Simulated data was necessary to visualize some concepts due to the complexity of the formulation. With a strong understanding of this model dynamics, I can begin to build the necessary transformations. In this case, both the intellectual and social parameters are capable of evolving in a linear fashion throughout time as summarized below.

$$T_{generalist} = \begin{bmatrix} \frac{(\alpha_{i+1} - n_{i+1})}{(\alpha_i - n_i)} \frac{\alpha_i}{\alpha_{i+1}} & 0 \\ 0 & \frac{n_{i+1}}{n_i} \end{bmatrix} = \begin{bmatrix} \frac{(\alpha_i + \Delta\alpha - n_{i+1})}{(\alpha_i - n_i)} \frac{\alpha_i}{\alpha_i + \Delta\alpha} & 0 \\ 0 & \frac{n_{i+1}}{n_i} \end{bmatrix}$$

$$T_{specialist} = \begin{bmatrix} \frac{V_{i+1}}{V_i} & 0 \\ 0 & \frac{n_{i+1}}{n_i} \end{bmatrix} = \begin{bmatrix} \frac{V_i + \Delta V}{V_i} & 0 \\ 0 & \frac{n_{i+1}}{n_i} \end{bmatrix}$$

Specialist Conditions

With the basic transformations well defined, an exploration into the education system today is warranted to understand the underlying conditions society is likely facing. Insight into why specialization persists once it begins and why generalization typically occurs at the onset of education can be drawn through the constraints of this model. In order to provide this illustration, I will consider a simple two-period education system. In only two periods, an administrator is able to choose between generalist and specialist two times and selects optimal class sizes subject to that choice in each period as previously described. In this situation, four possible expected returns to education present themselves as shown below.

$$E(\pi)_{GG} = VZ - 2Z\sqrt{\frac{wV}{\alpha + \Delta\alpha}} + VZ - 2Z\sqrt{\frac{wV}{\alpha + 2\Delta\alpha}}$$

$$E(\pi)_{GS} = VZ - 2Z\sqrt{\frac{wV}{\alpha + \Delta\alpha}} + (V + \Delta\mathrm{V})Z - 2Z\sqrt{\frac{w(V + \Delta\mathrm{V})}{\alpha + \Delta\alpha}}$$

$$E(\pi)_{SG} = (V + \Delta\mathrm{V})Z - 2Z\sqrt{\frac{w(V + \Delta\mathrm{V})}{\alpha}} + (V + \Delta\mathrm{V})Z - 2Z\sqrt{\frac{w(V + \Delta\mathrm{V})}{\alpha + \Delta\alpha}}$$

$$E(\pi)_{SS} = (V + \Delta\mathrm{V})Z - 2Z\sqrt{\frac{w(V + \Delta\mathrm{V})}{\alpha}} + (V + 2\Delta\mathrm{V})Z - 2Z\sqrt{\frac{w(V + 2\Delta\mathrm{V})}{\alpha}}$$

Prior to any discussion between an indefinite generalist and an indefinite specialist, I wanted to understand why the return to the generalist then specialist, $E(\pi)_{GS}$, tends to be strictly greater than the return to the specialist then generalist, $E(\pi)_{SG}$. With only two periods, the intuition may seem slightly more fluid, but consider the education process as two discrete blocks. In the first period, say elementary school, the focus is on improving social capability, while in the second period, the focus is on improving intellectual capability. As mentioned, these are two extremes while reality exists on some flexible scale. Despite this weakness, I proceed with the endeavor to see if any useful intuition presents itself.

$$E(\pi)_{GS} > E(\pi)_{SG}$$

$$-2\sqrt{\frac{wV}{\alpha + \Delta\alpha}} > \Delta\mathrm{V} - 2\sqrt{\frac{w(V + \Delta\mathrm{V})}{\alpha}}$$

$$\sqrt{wV}\left[\frac{1}{\sqrt{\alpha}} - \frac{1}{\sqrt{\alpha + \Delta\alpha}}\right] > \frac{\Delta\mathrm{V}}{2} - 2\sqrt{\frac{w\Delta\mathrm{V}}{\alpha}}$$

$$\sqrt{w} > \frac{\Delta\mathrm{V}}{2}\frac{\sqrt{\alpha(\alpha + \Delta\alpha)}}{(\sqrt{\alpha + \Delta\alpha} - \sqrt{\alpha} + \sqrt{\Delta\mathrm{V}(\alpha + \Delta\alpha)}}$$

$$1+\sqrt{\Delta V} > \frac{\sqrt{\alpha}}{\sqrt{\alpha+\Delta\alpha}} + \frac{\Delta V}{2}\sqrt{\frac{\alpha}{w}}$$

$$1 > \frac{\sqrt{\alpha}}{\sqrt{\alpha+\Delta\alpha}}\ ;\ \sqrt{\Delta V} > \frac{\Delta V}{2}\sqrt{\frac{\alpha}{w}}$$

$$\sqrt{w} > \sqrt{\frac{\Delta V\alpha}{4}}$$

$$w > \frac{\Delta V\alpha}{4}$$

With this derivation, I began with a fact routinely witnessed in society today. Specialization is more productive when done following generalization, assuming that improvement on social capacity above some certain amount allows optimal gain from an intellectual standpoint. From there, I was able to manipulate the inequality into two distinct parts, where I was able to uncover an approximate condition on wage. The wage, for a two period education system, must be greater than some threshold proportionate to the potential improvement due to specialization and the beginning social capacity for a given group of students.

At this point, drawing intuition does present some value. With a base threshold on wage, I accept that there must be some driving force to want to increase class sizes from a financial standpoint. If wages are not above a certain amount, the return to increasing the social parameter, which ultimately allows for increased class sizes, becomes less significant. Furthermore, this basis relates directly to the potential improvements in intellectual capacity and the starting social capacity parameter. A group of students, who begin with either an extraordinarily high starting social capacity or an extraordinarily high ability to gain from an intellectual standpoint, may result in a situation where it is be beneficial to specialize them initially then follow this specialization with generalization. Before accepting this result, I wanted to consider if in those situations, specializing in both periods is actually the optimal result, thus rendering the condition above sufficient to ensuring generalization always proceeds specialization.

$$E(\pi)_{SG} < E(\pi)_{SS}$$

$$-2\sqrt{\frac{w(V+\Delta\mathrm{V})}{\alpha+\Delta\alpha}} < \Delta\mathrm{V} - 2\sqrt{\frac{w(V+2\Delta\mathrm{V})}{\alpha}}$$

$$\sqrt{w} < \frac{\Delta\mathrm{V}}{2}\frac{\sqrt{\alpha(\alpha+\Delta\alpha)}}{\sqrt{(V+2\Delta\mathrm{V})(\alpha+\Delta\alpha)} - \sqrt{(V+\Delta\mathrm{V})\alpha})}$$

When considering the situation where dual specialization exceeds specialization followed by generalization, I uncovered a ceiling on the wage. By reconsidering the previous situation from the perspective where specialization followed by generalization exceeds the expected return from generalization followed by specialization, another ceiling on wage presents itself. In this case, I would prefer a strict condition on the ceiling resulting from dual specialization, such that any time specialization followed by generalization exceeds generalization followed by specialization, the administrator would optimally choose to specialize for the entire academic career for that group of students.

$$E(\pi)_{GS} < E(\pi)_{SG}\ ;\ \sqrt{w} < \frac{\Delta\mathrm{V}}{2}\frac{\sqrt{\alpha(\alpha+\Delta\alpha)}}{(\sqrt{\alpha+\Delta\alpha} - \sqrt{\alpha} + \sqrt{\Delta\mathrm{V}}\sqrt{\alpha+\Delta\alpha})}$$

$$E(\pi)_{SG} < E(\pi)_{SS}\ ;\ \sqrt{w} < \frac{\Delta\mathrm{V}}{2}\frac{\sqrt{\alpha(\alpha+\Delta\alpha)}}{\sqrt{(V+2\Delta\mathrm{V})(\alpha+\Delta\alpha)} - \sqrt{(V+\Delta\mathrm{V})\alpha})}$$

$$E(\pi)_{GS} < E(\pi)_{SG} < E(\pi)_{SS}$$

$$\frac{\Delta\mathrm{V}}{2}\frac{\sqrt{\alpha(\alpha+\Delta\alpha)}}{\left(\sqrt{\alpha+\Delta\alpha} - \sqrt{\alpha} + \sqrt{\Delta\mathrm{V}(\alpha+\Delta\alpha)}\right)} < \frac{\Delta\mathrm{V}}{2}\frac{\sqrt{\alpha(\alpha+\Delta\alpha)}}{\sqrt{(V+2\Delta\mathrm{V})(\alpha+\Delta\alpha)} - \sqrt{(V+\Delta\mathrm{V})\alpha})}$$

$$\sqrt{(V+2\Delta V)(\alpha+\Delta\alpha)}-\sqrt{\Delta V(\alpha+\Delta\alpha)}<\sqrt{\alpha+\Delta\alpha}-\sqrt{\alpha}+\sqrt{(V+\Delta V)\alpha}$$

Following this exploration, a new condition presents itself to ensure that a situation where specialization occurs prior to generalization is ultimately suboptimal. This inequality provides little intuition due to its complex nature. In hopes of gaining more understanding, I assumed the marginal increases in both intellectual and social capacity are small, a reasonable assumption as class sizes do not change drastically over grade levels, typically only several students over many periods, and follow with a first order Taylor approximation for the radicals presented above.

$$(\sqrt{V}+\frac{\Delta V}{\sqrt{V}})(\sqrt{\alpha}+\frac{\Delta\alpha}{2\sqrt{\alpha}})-(\sqrt{V}+\frac{\Delta V}{2\sqrt{V}})\sqrt{\alpha}>(\sqrt{\alpha}+\frac{\Delta\alpha}{2\sqrt{\alpha}})-\sqrt{\alpha}+\sqrt{\Delta V}(\sqrt{\alpha}+\frac{\Delta\alpha}{2\sqrt{\alpha}})$$

$$\sqrt{V\alpha}+\Delta V\sqrt{\frac{\alpha}{V}}+\sqrt{\alpha}+\frac{\Delta\alpha}{2\sqrt{\alpha}}-\sqrt{V\alpha}-\frac{\Delta V}{2}\sqrt{\frac{\alpha}{V}}>\sqrt{\alpha}+\frac{\Delta\alpha}{2\sqrt{\alpha}}-\sqrt{\alpha}+\sqrt{\Delta V\alpha}+\frac{\Delta\alpha}{2}\sqrt{\frac{\Delta V}{\alpha}}$$

$$\Delta V\sqrt{\frac{\alpha}{V}}-\frac{\Delta V}{2}\sqrt{\frac{\alpha}{V}}+\sqrt{\alpha}-\sqrt{\Delta V\alpha}-\frac{\Delta\alpha}{2}\sqrt{\frac{\Delta V}{\alpha}}>0$$

$$\sqrt{\frac{\Delta V^2V+4(1-\Delta V)^2}{V\Delta V}}>\sqrt{\frac{\Delta\alpha^2}{\alpha^3}}$$

$$\frac{\Delta V^2V+4-8\Delta V+4\Delta V^2}{V\Delta V}>\frac{\Delta\alpha^2}{\alpha^3}$$

$$\Delta V+\frac{4}{V\Delta V}+\frac{4\Delta V-8}{V}>\frac{\Delta\alpha^2}{\alpha^3}$$

Through the Taylor expansion, I was able to isolate an approximate condition that separated the social and intellectual parameters. In the case where it is more effective to follow specialization with generalization, the above condition is likely met such that indefinite specialization is the realized outcome. As previously discussed, high levels of starting social capacity would be one situation where it may be more effective to follow specialization with generalization than vice versa. However, from the above condition, high starting levels of social capacity, α, would likely allow indefinite specialization to improve on that outcome, particularly considering the assumption that the incremental increases are small which allowed for the Taylor expansion.

Although this is intriguing, it is of relatively little value due to its inherent ambiguity and provides a strong basis for improving the transformations of this model. A holistic model for education should explicitly show that this generalization must come prior to specialization at optimal outcomes without imposing additional constraints. Even with no wage floor, it would seem that specialization is best suited to follow generalization, but within the confines of this model, a new constraint must be applied, a less than desirable result. The basic, linear form for the dollar-valued intellectual capacity parameter is likely one inadequacy. More interaction between the intellectual and social parameters is another plausible solution.

Despite these clear deficiencies, this dynamic performance theory still presents significant improvement in transitioning Lazear's static disruption model into multiple time periods by introducing social and intellectual capacity parameters. I was able to improve upon the clear issues regarding marginal returns and was able to explicitly formulate optimal class sizes. With further theoretical exploration, I should be able to maintain these concepts within the contexts of transformations while gaining a more clear understanding as to why the school system retains its formal structure. These improvements will be left to future discussions.

Chapter 6

Application

With the dynamic performance model well defined and the underlying intuition relatively sound, I wanted to look to actual datasets in order to see its application. Data should drive the model-building process, and then the model should provide the intuition as to why this structure presents itself. If the data and intuition do not mesh appropriately, the cycle towards improvement should continue. In this step, I allow my model to provide intuition to data, looking towards a sole dataset over one single period for simplicity to do so, and consider whether that underlying intuition is valid.

Although the model's intention is to follow a singular group of students over their educational career, a cross section is adequate for an initial investigation, as only high-level, averaged data values are considered anyways. These permissions imply that this analysis expresses only the qualitative potential of the model rather than claiming any concrete relationships as fact.

Schools and Staffing Survey Dataset and Methodology

Between 1987 and 2011, the National Center of Education Statistics (NCES) conducted the Schools and Staffing Survey (SASS) designed to investigate a variety of variables in school districts throughout the United States. The NCES conducted questionnaires seven times throughout that time and integrated data across both public and private school districts spanning kindergarten through twelfth grade. These surveys addressed a wide range of topics such as teacher demand, teacher compensation, principal perceptions, and general school conditions. For the brunt of this investigation, we will examine a small cross section from the final dataset drawn in 2011.

The survey from 2011 consisted of five types of questionnaires with slight modifications depending on the application in either private or public institutions. These five questionnaires consisted of a principal questionnaire, a teacher questionnaire, a school district questionnaire, a school questionnaire, and school library media questionnaire. The sampling frame for public schools resulted from an adjusted version of the 2009-2010 Common Core of Data (CCD), which dropped unique cases such as those institutions who taught only kindergarten or those funded by the Bureau of Indian Education while they drew the sampling frame for public schools from the Private School Universe Survey (PSS). Each sampling set surveyed all principals, districts, and media centers while the teachers were further sampled from a frame collected via a list provided by the school district.

All of these data points were collected through a combination of mail and internet reporting techniques with telephone and in person follow-up as necessary. The U.S. Census Bureau then conducted the data processing where a computer program first screened a completed questionnaire for quality control. Following this automatic process, a manual investigation of errors and flagged interviews proceeded on a case-by-case basis. This process recoded each survey as to whether or not it would be included in the ultimate dataset. They then defined the response rate per scoped area as the number of accepted responding questionnaires divided by the total sampled cases. They pooled these results together by using the inverse probability of selection to weight each individual scope (*Schools & Staffing Survey)*.

Public School Teacher Data File

One specific selection, the “Public School Teacher Data File” will be applied to this model of classroom choice, in an attempt to understand both specialization levels and classroom sizes within the United States public system. The public section of the data was selected for analysis over the private or combined section for several key reasons. For starters, choosing between public and private was essential to the discussion. Although both sectors likely fit the model provided based on Lazear’s research, the two face slightly different situations. For example, Lazear commented on the ease of ability for private schools to select out students who are not behaving appropriately. This allowed them to increase the probability of good behavior and thus profitability of educational outcomes. Although in principle, either private or public would be reasonable fits for this model under different parameters, we aim to hold as many constants as possible thus forcing us to choose one over the other.

With this choice between the private and public sectors’ datasets, the public school districts provided a greater number of data points. Not only was its sample set larger, it received a relatively higher response rate than the private school sample, at 77.7% versus a 69.9% as reported by the US Department of Education. With a higher response rate, a more representative dataset is likely, thus less time must be spent examining potential nonresponse bias. Furthermore, the incentives of the public institution likely align more closely with the incentives of a government entity, such as the NCES. They likely want to see improvements via educational research on a broader scale, so aligning these efforts seems natural. Secondly, they derive most of their funding at the state or local level. Therefore, the individual responses will have little interaction with their funding which diminishes the return to inflating or falsifying responses.

While these aspects were important in making this choice, the difference in stratification of the school levels was by far the most crucial. Specific grade-level evidence is unavailable in both cases, but SASS did provide educational level on a smaller spectrum, discrete basis. In the private sector school districts, the NCES used a three level classification – elementary, secondary, and combined. With the public sector, they presented four classifications – elementary, middle, secondary, and combined. This additional segment provides a significant improvement when considering the evolution of students and choices over time. Unfortunately, without more concise data points, the combined school districts will not be considered. These are likely much smaller institutions, which may face different incentives anyways, so their lack of consideration seems reasonable.

While this short explanation gave some sense of the methodology the NCES utilized and provided insight into the assumptions and the specific choices of data, it does so in very brief terms with minimal consideration for the accuracy of the resulting figures. The goal of this paper is to understand the underlying economics, which may have produced the results provided rather than confirm the accuracy of the sampling techniques. With an entity as founded as the NCES, assume the statistics provided are accurate and representative moving forward (*Schools & Staffing Survey).*

Class Sizes across Grade Levels

In the following exploration, I utilize the dynamic classroom performance model to understand average class sizes across elementary, middle, and high school for public schools as presented in the following table from a qualitative perspective, dialing into class sizes for two key reasons. For starters, only class size and wage are measurable variables to input into the model, and data expressing both together was difficult to obtain. Without both as an option, public teacher wage was easily assumed to be constant despite varying drastically by geographic region. Although this assumption is invalid, it is adequate for preliminary analysis leaving only class size behind to investigate. Secondly, by assuming that administrators make decisions optimally, these select class sizes provide insight into the underlying social and intellectual parameters at play.

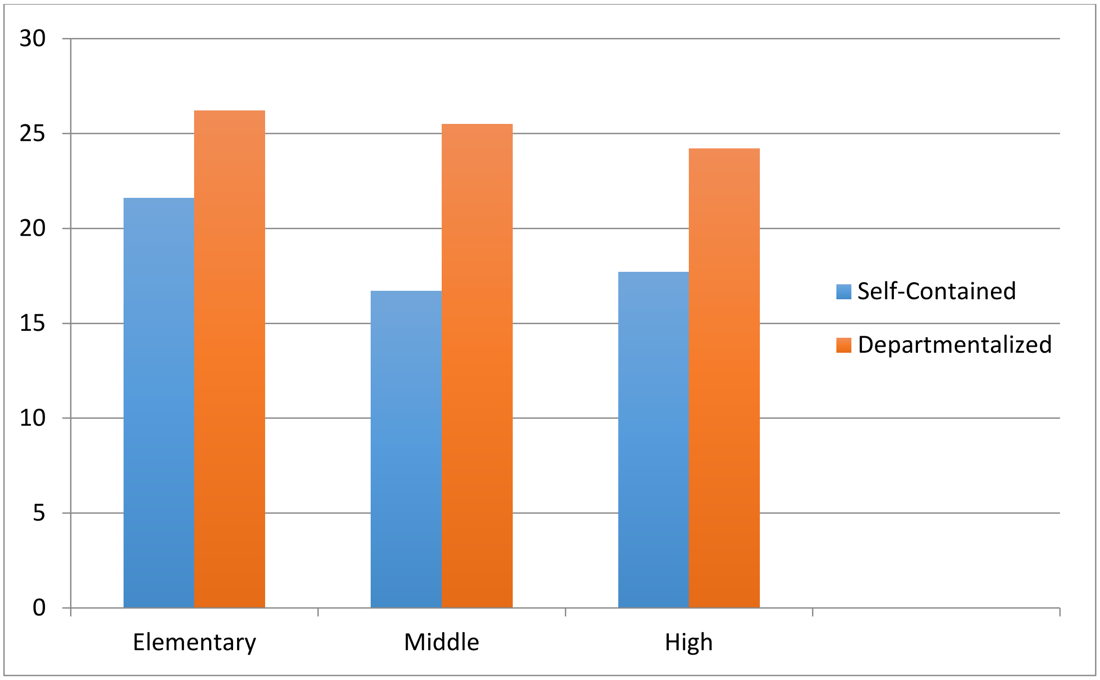
Once again, I accepted that the model was not perfectly aligned with these large scale, average values, and with that lack of alignment, multiple, reasonable explanations may present themselves for the variation in class sizes through outside variables. Regardless, I chose to utilize the averages, compiled through the NCES database, displayed in the following table.

Table 3: NCES Public Class Size Statistics from 2011

Elementary School		Middle School		High School	
Self-Contained	Departmentalized	Self-Contained	Departmentalized	Self-Contained	Departmentalized
21.6	26.2	16.7	25.5	17.7	24.2

In the case of this dataset, self-contained classes are representative of the generalist model. With a self-contained classroom, a sole teacher handles an entire group of students for the whole term for all subjects. For departmentalized coursework, a specialized model ensues where each teacher handles a given subject over several different select groups of students. Take a look into some general trends evident in this dataset through the following graphic.

Figure 2: NCES Public Class Size Statistics from 2011



From this graphic, it is evident that average class size tended to drop as grade level progressed for those following a departmentalized program. For those in the self-contained setting, the trend is relatively less clear, with a significant drop in class size from elementary school to middle school, followed by a slight increase in size from middle school until high school. An immediate insight one may draw is that self-contained class size did not consistently rise with grade level as was predicted through our optimality conditions. Focusing on improving behavior does allow class sizes to grow; however, recall that at some point, it often becomes optimal to switch from the generalist setting into the specialist setting.

Therefore, the drop in class size between elementary school and middle school for generalist situations might imply that most individuals developed the necessary social skills to move into more specialized coursework between elementary school and middle school. The students with still relatively low social capacity were then left in the generalized setting resulting in the drop of class size for that grade level.

Following through with this explanation, students then experienced little cross over between the different models from middle school to high school. Those with exceptionally low social capabilities stay in the self-contained classes throughout the remainder of their academic careers. This allows for the slight increase in class size between middle school and high school due to the improvements in behavior in the self-contained setting. Meanwhile, class sizes dropped slightly between middle school and high school in the specialized setting as was once again predicted by an improvement in intellectual capacity.

This brief, qualitative overview simply provides insight into the usefulness of the model in understanding the forces that drive classroom decisions. With a weak alignment between the subset of data and the actual model, further fitting of the parameters is unnecessary and would provide little intuition. However, there are many other avenues where this type of modeling may become useful. For example, data is available regarding the amount of school psychologists present alongside classroom size data. Potentially, one may draw a positive correlation between the number of psychologists and the size of the classes. This may imply that forces, outside of teachers, are able to improve social capacity and this improvement in social capacity may allow for greater class sizes subject to the cost of these outside forces. Regardless, the general idea is that this or similar models may provide useful explanations to variations typically seen in data.

Chapter 7

Conclusion

After diving into Lazear's disruption model of education and considering the importance of distance traveled, I recognized the necessity of a dynamic decision model, which incorporates class size and the degree of teacher specialization, to understand administrative decisions made over time. I began building this dynamic model by introducing an underlying transformation framework to easily capture the essential changes that occur as education progresses. The implications of this foundation when modeling distance traveled may be far reaching with further applications.

Following the theoretical framework, a careful exploration of Lazear's model presented clear inadequacies when utilized in multiple periods, where slight modifications in his probability of a successful classroom provided improvement. Despite these modifications, I was unable to see appropriate marginal returns in a clean, consistent manner. A social capacity parameter, directly correlated to the probability of good behavior, handled this discrepancy. This parameter became directly related to the fraction of maximal performance in a given classroom through a new distribution, dependent on class size. This distribution accurately depicted the conditions Lazear had uncovered in a single period but was ultimately more equipped to handle a dynamic return to education. Not only did this new presentation allow for decreasing marginal return towards increasing the social capacity parameter, it allowed for an explicit formulation of optimal class size while providing a new, rather intriguing condition on wage. If wage is greater than the product of a group's social and intellectual capacity, it is suboptimal to educate them.

Finally, I examined if the model explicitly proved why specialized teachers proceed generalized teachers in the typical setting. Unfortunately, the conditions placed onto the model in this situation became relatively complex, rendering further improvement necessary. With improvements aside, I was still able to consider a basic dataset to understand class size movements and specialization choices occurring across United States public schools today from a qualitative perspective. Moving forward, the model should be adjusted to gain further intuition as to why specialization occurs following generalization in education. With an appropriate model, insight should present itself as to specific situations where it is optimal to send a group of students into some specialized form of education.

BIBLIOGRAPHY

Fryer, Roland. (2016). The Pupil Factory: Specialization and the Production of Human Capital in Schools. *National Bureau of Economic Research*.

Konstantopoulos, Spyros. (2014). Teacher Effects, Value-Added Models, and Accountability. *Teacher's College Record*, 116(1).

Lazear, Edward. (2001). Educational Production. *The Quarterly Journal of Economics*, 116(3), 777-803.

Markworth, Kimberly, Brobst, Joseph, Ohana, Chris, & Parker, Ruth. (2016). Elementary Content Specialization: Models, Affordances, and Constraints. *International Journal of STEM Education*, *3*(1).

Schools & Staffing Survey. US Department of Education: National Center for Education Statistics. Retrieved from https://nces.ed.gov/surveys/sass/index.asp.

Spence, Michael. (1973). Job Market Signaling. *The Quarterly Journal of Economics*, 87(3), 355-37.

ACADEMIC VITA

Montana D. Morris

montanamorrisshc@gmail.com

Education

The Pennsylvania State University | Schreyer Honors College University Park, PA

Majors: Physics, B.S. | Economics, B.S.
Minor: Mathematics
Graduation Date: May 2018

Work Experience

KCF Technologies, Inc. | Sentry Engineering Intern State College, PA

Vibrational Signal Analysis | March 2017 – May 2017

Analyzed engineering data for major manufacturing companies in the automotive, paper and pulp, and oil industries
Communicated technical reports with maintenance and management of clients weekly

Software Development | May 2017 – August 2017

Wrote a prototype machine-learning algorithm which modeled the response of key performance indicators to operational changes
Created a self-calibrating model to predict broken-down operational costs from input KPI's and plausible financial data
Developed an interface between the financial and operational algorithms with ability to produce system optimization routines

Technical Consulting | July 2017 – August 2017

Engaged with customer engineering teams to identify system weaknesses and provided actionable proposals to management
Uncovered a fluid condition for a paper and pulp customer when investigating a quality issue
Identified two major cost factors for a hydraulic fracturing client

McLanahan's Penn State Room State College, PA

Laborer | April 2016 – September 2016

Unloaded truck shipments and stocked assorted food, apparel, and household products
Helped reorganize the shelving units to improve sales and customer flow

Oaktree Country Club West Middlesex, PA

Laborer | May 2015 – August 2015

Worked within the maintenance department and typically shoveled, edge, or weeded various parts of the golf course
Operated equipment such as dump trucks and large tractors during occasional excavating projects

Research / Other Involvements

Penn State Economics Department University Park, PA

Undergraduate Research Assistant | January 2017 – May 2017

Examined the accuracy of a published trade model via empirical methods with the guidance of a faculty professor
Utilized STATA to manipulate and build a model for millions of trade data points
Validated the relationship of import exposure on unemployment in local economies with a relatively high manufacturing output

Management 296 – PSU Independent Studies

Student Consultant | September 2015 – May 2016

Worked alongside five other Penn State Students for AccuWeather to better utilize their mobile applications in foreign markets
Planned further exploration for German markets through future students who study abroad
Proposed a simplified approach for Indian markets due to a lack of technology

Phoenix THON Penn State Organization

Founding Member | September 2015 – May 2018

Created an organization with 25 peers to take ownership for our own campaign against cancer
Raised over $5000 across the organization for Penn State's Dance MaraTHON

Academic Achievements: John and Elizabeth Holmes Teas Fellow, Elsbach Honors Physics Scholarship, Sigma Pi Sigma Physics Honor Society, Philips Healthcare Scholarship, Pentz Memorial Scholarship, Academic Excellence Scholarship | Computer Languages: Python, R, C++, STATA