Stochastic Modeling of the Dynamics of Multi-Classroom Education Using Teaching Games

Larry Liu
Stanford University
Stanford, USA
hrlarry@stanford.edu

Michael Frank
Stanford University
Stanford, USA
mcfrank@stanford.edu

ABSTRACT
How should a school administrator allocate a fixed budget towards increasing the number of classrooms and increasing student assessment to best increase student learning? This paper investigates how stochastic models accounting for the inherent uncertainty in student beliefs and teacher communication in teacher-student dyads can help school administrators figure out how to maximize student learning by optimally allocating resources. We replicate existing results from education literature about the effect of class sizes, homogenous ability classrooms, and assessment on student learning using computer simulation. We also unify these different determinants of student learning into a more holistic model and report the tradeoffs that committing budget to these design features presents. We demonstrated that we can create a pareto frontier for the number of teachers and assessments against a fixed budget, and find an optimal allocation of the budget for increasing student learning. We also demonstrate how the effectiveness of class sizes and assessment is contingent upon the learning concept being taught. We believe that the insights about student learning in multi-classroom settings that we've found and can continue to surface are difficult to surface in real-world studies with human subjects.

Author Keywords
Teaching; learning; reasoning; education; communication; pragmatics; Bayesian modeling; social cognition

INTRODUCTION
Formal education typically revolves around teachers trying to transmit knowledge of various forms to students. However, the interpersonal interactions involved in teaching has inherent ambiguity. Teachers don't have perfect knowledge of what students know or do not know about learning concepts, and there is noise in the communication of knowledge from teachers to students. This ambiguity lends itself well to the use of stochastic modeling to explore the dynamics in education.

While there exists previous research that uses stochastic modeling to represent teacher-student as well as single-classroom dynamics, to the author's knowledge there is no work that tries to model an entire school. At the school level, there are multiple classrooms and thus multiple teachers. Modeling education at the school-wide level presents the opportunity to explore tradeoffs that school administrators need to make when constrained by resources such as budgets, necessarily limiting the number of teachers that are available and the amount of assessment that students can take.

In this paper, we use stochastic modeling to optimize budget allocation for an administrative role by considering systems of multiple teachers and learners. Before describing the present study and the modeling strategy, we briefly discuss the existing literature.

BACKGROUND
Stochastic modeling of classroom dynamics literature
Some existing work has already been done to explore the teacher-learner dynamics with stochastic modeling. Previous research demonstrates the recursive structure through which teachers and learners make inferences about each other's intentions. In Shafto and Goodman (2008)'s model, the teacher has an initial belief about a student's understanding of a particular concept the teacher wishes to teach. Given this belief, the teacher chooses an example or set of examples to present to the student. The student then updates their belief about the learning concept with the new information presented by the teacher, reasoning that the teacher chose those specific examples to help calibrate the student's understanding towards the true concept. This recursive structure–wherein teachers select examples to maximize student learning while students consider that the teachers thoughtfully chose those examples for them–has demonstrably stronger inferences than models where the teacher and student agents don't reason about each other (Shafto, Goodman, and Frank, 2012).

Some existing work has deeply iterated on the recursive structure–that is, about the student's inferences about the teacher's inferences about the student and so on (Jager, 2010)–until asymptotic convergence. However, Frank and Goodman (2012) demonstrated that in modeling pragmatic communication, a two-step framework produces inferences about meaning beyond the literal. In the context of our teacher-student dyads, what this means is that even when only the teacher reasons about the student's understanding and not vice versa, the teacher will choose examples that are different from those that an inference-free teacher agent would choose to teach a learning concept. Existing economics research suggests that actual people usually employ somewhere between first- and second-order depth of reasoning (Nagel, 1995; Camerer, Ho, and Hong, 2004). These conceptual findings supports the use of two-step framework rather than deep recursion. Frank and Goodman showed that a model where the teacher makes decisions on examples to show with only first-order reasoning and similarly where students interpret teacher examples with only first-order reasoning is sufficiently strong. This is promising for the validity of simulated classrooms, because deep recursion is much more computationally demanding.

These teacher-learner dynamics become more interesting when modeling education at a classroom-wide and school-wide level. In classrooms, there are multiple students and thus multiple unique teacher-student dyads, each performing inferences as described above. As mentioned earlier, at the school level, there can be multiple teachers teaching multiple classrooms of students.

Educational policy literature.

Research in the field of education has produced various findings that are used to inform the design of school systems. For instance, researchers have suggested that larger classes produce worse learning outcomes for students (Glass and Smith, 1979; Slavin 1989). Somewhat orthogonally, some of posited that grouping students with similar ability into the same classroom improves learning outcomes (Slavin, 1987). Some argue that for the sake of differentiated learning, there is merit to more frequent or thorough assessments of student abilities so that teachers can better cater classroom materials and lessons to the students (Fuchs and Fuchs, 1986).

Crucially, these different features of school systems compete for the same common resource: money. It is expensive to hire more teachers. The creation and administration of assessments is also costly. Even if those individual features monotonically and non-asymptotically improved student learning outcomes, real-world budget constraints inevitably limit the school design possibilities for administrators.

Related Work

This paper draws inspiration from Michael Frank's modeling work. The models used for this paper builds off his previous models, and both can be found at `github.com/mcfrank/teaching`.

THE PRESENT STUDY

Existing work on modeling classroom dynamics has not placed real-world constraints on the parameters of the model. Most critically, theoretical models are not as actionable for education policymakers without realistic considerations of budget. As such, a lot of previous literature employs models that are potentially unrealistic.

Already, Frank has tried to address the student-teacher ratio in his work by modeling entire classrooms rather than individual teacher-student dyads (Frank, 2014). Because students usually vary in their understanding of learning concepts, a teacher instructing a group of students can, at best, only provide as much information gain per student as when a teacher instructs a single student.

The present study employs similar techniques to test the effectiveness other aspects of school design grounded in education literature with stochastic modeling. The study seeks to provide quantitative theory to help evaluate the tradeoffs in some of the choices school administrators must make.

As in some of Frank's existing work (Frank, 2014), the fundamental unit of analysis in this paper is a teaching game. In each teacher-student dyad, a teacher tries to guide the student to discover a particular learning concept. For this paper, the learning concept we use is the weight of a biased coin (i.e. the parameter of a Bernoulli distribution). As Frank discusses, while this is an extremely simplistic case of a learning concept, many interesting dynamics of the classroom become clear due to the simplicity. We are able to observe many dynamics of optimal teaching where arbitrarily complex learning concepts may obscure the discernable patterns in student learning.

MODEL

In our model, there are three unique types of agents: students, teachers, and administrators. Conceptually, a school consists of an administrator, at least one teacher, and at least one student. The number of teachers never exceeds the number of students (i.e. teachers cannot have empty classrooms).

The objective of the model is for the administrator to maximize the aggregate information gain of all students in the school.

The administrator has a full roster of all students who will be attending the school and can choose between various designs or features of school operations, each with its own costs. The cost constraints are as follows:

- The administrator is constrained by a particular budget B.
- Hiring each teacher has a fixed cost C_T. Each teacher is able to lead their own subset of students from the roster, representing a classroom. For the sake of simulation, all additional costs in maintaining an additional classroom (property, utilities, capital equipment, school supplies, etc.) are considered part of cost C_T.
- There is a fixed cost of assessment of students C_A. Assessments reduce the noise (improves the precision) of a teacher's beliefs about a student's initial level of understanding of the learning concept. For the sake of simulation, all additional costs of developing, validating, and administering assessments are considered part of C_A.

The agents' functions are described below.

Student

Each student has a prior belief about the bias of a Bernoulli variable whose true bias the teacher is trying to teach. This prior belief is represented as a Beta distribution whose α and β parameters are integers drawn from a Uniform distribution between 1 and 10 inclusive. Each student's prior distribution is generated independently from every other student's prior distribution. That is, the student's prior belief is represented as a $Beta(\alpha, \beta)$.

During the teacher inference phase, a student is able to use a set of examples from the teacher with x heads and y tails to update their prior distribution to create a new Beta distribution that represents the posterior distribution of their belief about the bias of the coin. This posterior distribution is created using the closed-form solution of a Beta distribution. That is, the student's posterior belief is represented as $Beta(\alpha + x, \beta + y)$.

Teacher

Each teacher is assigned a classroom roster of students on which to perform inference. This roster contains information about the students' prior beliefs about the learning concept, represented as the α and β parameters in a Beta distribution. These parameters can have noise that simulates how teachers do not have perfect knowledge of their students' beliefs of the learning concept. We test the effect of imperfect teacher knowledge about student beliefs on learning rate in Simulation 3.

The teacher is also given a target concept to teach to the students. This target concept is the bias of a Bernoulli variable. That is, the teacher wants to teach students what p in a distribution $Bernoulli(p)$ using only outcomes $\{0, 1\}$ from that distribution.

The target bias is encoded as the mean a Beta distribution to introduce noise in the learning concept. The parameters of this target Beta distribution α_T and β_T are selected prior to any presentation of examples to students. This target is shared across all teachers in the school system.

During the teacher inference phase, the teacher randomly generates sets of examples to reason about teaching to students in their roster. Using the roster consisting of student prior beliefs, the teacher calculates the amount of student learning they believe the students would receive using the examples, and weights that set of examples proportionally. See the Information Gain subsection for details on information gain.

Finally, the teacher reports to the administrator the probability distribution of their roster of students receiving a particular amount of information gain.

Administrator

The administrator gets a roster of students in the entire school and also determines the target concept to teach the students.

The administrator can sort all of the students by their prior beliefs about the learning concept. The real-world analog of this process is the conducting of placement tests prior to a school term. This is done with the closed-form solution of Beta distributions, where the mean value of the distribution is equivalent to

$$\mu = \frac{\alpha}{\alpha + \beta}$$

. We test the effectiveness of sorting students in Simulation 1.

The administrator inference phase involves randomly simulating school systems. Simulation involves four parameters:

- The administrator can choose how to distribute the students into subsets to assign to teachers as their classroom rosters. Primarily, this involves deciding whether or not to sort students by prior beliefs about the target learning concept. (Simulation 1)
- The administrator can choose the number of classrooms they want to have. This is represented by the number of teachers that they hire, and is inversely proportional to class size. (Simulation 2)
- The administrator can choose how many assessments they want to conduct on the students. Increasing the number of assessments reduces the noisiness of the teachers' imperfect beliefs about student priors. (Simulation 3)
- The administrator can choose the target learning concept to be taught. In our simulation, it is the parameters of a target Beta distribution, which in turn effects a point target of the bias of a coin. This random generation of the target is symmetrical between [0,1] across 0.5. This is treated as moderating variable to the other parameters' impact on information gain. (All simulations).

The administrator weights the various policies that they simulate based on the aggregate information gain of the students in the entire school (compared to just each classroom for each teacher's inference), and is able to infer the most effective school design within budget constraints.

Controls

In our simulation the administrator controls for these parameters. Changes to these parameters may still have some impact on the effect sizes since they are used in some of the generative processes that create our synthetic data, but we hold them constant across all simulations. They are arbitrarily selected, but in an order magnitude consistent with Frank (2014)'s prior work.

- We simulate every configuration of classrooms on 100 unique random sets of students.
- The number of students. We use use 100 students.
- The number of examples shown by the teacher. We use 5 examples.
- The agnostic beliefs about the best example set. Our teachers have no agnostic beliefs about the number of heads or tails that should be shown, uniformly choosing between 0 and 5 heads out of 5 total examples to show.
- The strength of teaching concepts. These are embedded in the magnitude of the target α and β parameters. All target concepts have a total psuedocount (sum of prior α and β) of 10.
- The strength of student beliefs. These are embedded in the magnitude of the student prior α and β parameters. All students have a total pseudocount (sum of prior α and β) of 11. This is slight offset from the strength of the teaching concept to avoid edge cases where a student belief distribution perfectly matches the teaching concept.

- The variation in student beliefs. We randomly generate student prior parameters by uniformly sampling from the (α,β) tuples [(1,10), (2,9), (3,8), (4,7), (5,6), (6,5), (7,4), (8,3), (9,2), (10,1)].

Information Gain

Frank (2014) proposed a information gain formula using Kullback-Leibler divergences between student distributions and the target distributions to evaluate the similarity of the distributions (Cover and Thomas, 2012). Two KL divergences are calculated for each student, one for the distance between the target distribution B_T and the student prior distribution B_S, and one for the distance between the target distribution B_T and the student posterior distribution $B_S + e$ after updating their beliefs upon seeing example e. The difference between the posterior KL divergence and the prior KL divergence is the information gain. That is,

$$IG(e) = D_{KL}(B_T||B_S) - D_{KL}(B_T||B_{S+e}) \tag{1}$$

However, since Frank's proposed model only accounts for a single example, we derived a modified version of the single-evidence information gain formula that generalizes to any number of examples in an example set E. Derivation details can be found in the appendix. The final form for h heads and t tails is reproduced here:

$$\begin{aligned} IG(E) = &\sum_{k=0}^{h+t-1} \log(\alpha_S + \beta_S + k) \\ &- \sum_{i=0}^{h-1} \log(\alpha_S + i) - \sum_{j=0}^{t-1} \log(\beta_S + j) \\ &+ (h)\psi(\alpha_T) + (t)\psi(\beta_T) \\ &+ (h+t)\psi(\alpha_T + \beta_T). \end{aligned} \tag{2}$$

where ψ represents the digamma function, and α_S, β_S, α_T, and β_T are student prior alpha, student prior beta, target alpha, and target beta parameters, respectively.

RESULTS

In each simulation here, we report the aggregate (sum of) information gain above baseline for 100 students given any particular school design. The baseline structure differs in each simulation, are discussed in the following sections. Information gain at a per-student level would simply be one-hundredth of each of these results.

Critically, this representation of results does not visually capture the final inferential step performed by the administrator. The administrator utilizes information gain to infer which design to use. Because our results in each simulation demonstrate near-strict dominance of some parameterization of design choices over others, the administrator's inference is somewhat uninteresting. We discuss these outcomes in the following sections on each simulation.

Simulation 1: Sorting students by prior beliefs

We hypothesized that sorting students into classrooms by prior beliefs will increase information gain compared to randomly assigning students to classrooms. By sorting students by prior beliefs, each classroom will have more homogenous prior beliefs, so the students will find similar sets of examples that a teacher could provide to their classroom useful.

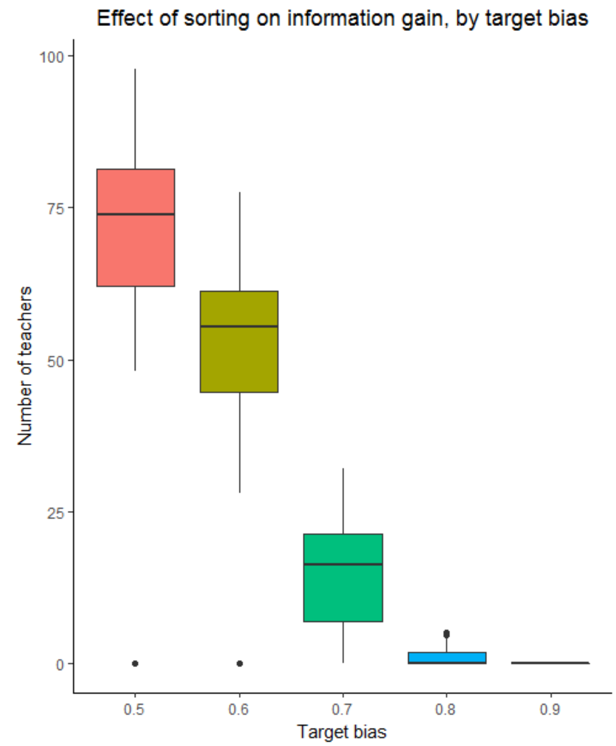

Figure 1. Students' amount of information gain when sorted into classrooms based on prior beliefs over unsorted baseline

Our baseline for this simulation was the unsorted information gain under the same parameters. For instance, the sorted aggregate IG of a school with 5 teachers and a target bias of .6 is compared to the unsorted aggregate IG of a school with 5 teachers and a target bias of .6 as a baseline.

Our simulated findings support our hypothesis and are consistent with real-world education research on tracking student abilities (Slavin, 1987). We find that sorted students show greater information gain than if the same set of students are distributed into unsorted classrooms, as shown in Figure 1, F(1,4998) = 1422, p<.001. The administrator consistently opted to sort students into classrooms over random assignment.

There is also reasonable support for an additive moderation model, F(1,4997) = 228.3, p < .001, because the strength of the effect of sorting seemed to vary depending on target belief as well as the number of teachers, as seen in Figure 2.

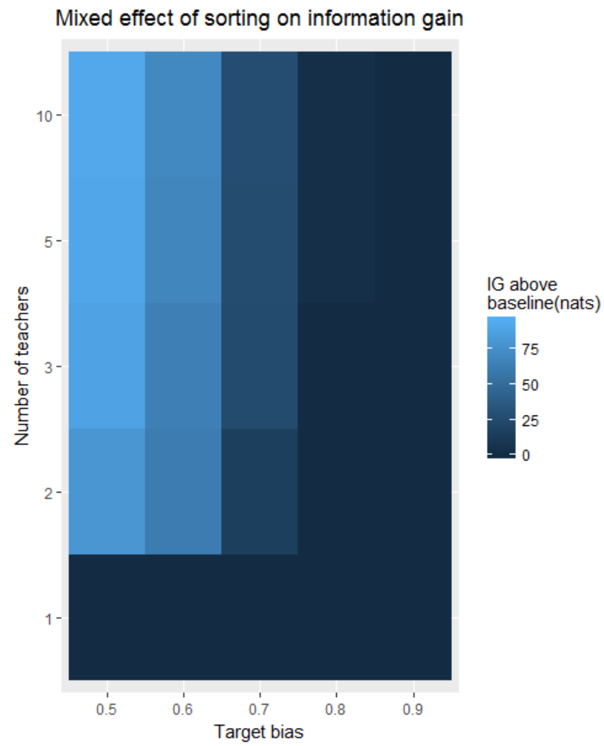

Figure 2. Information gain based on the composite target bias and number of teachers.

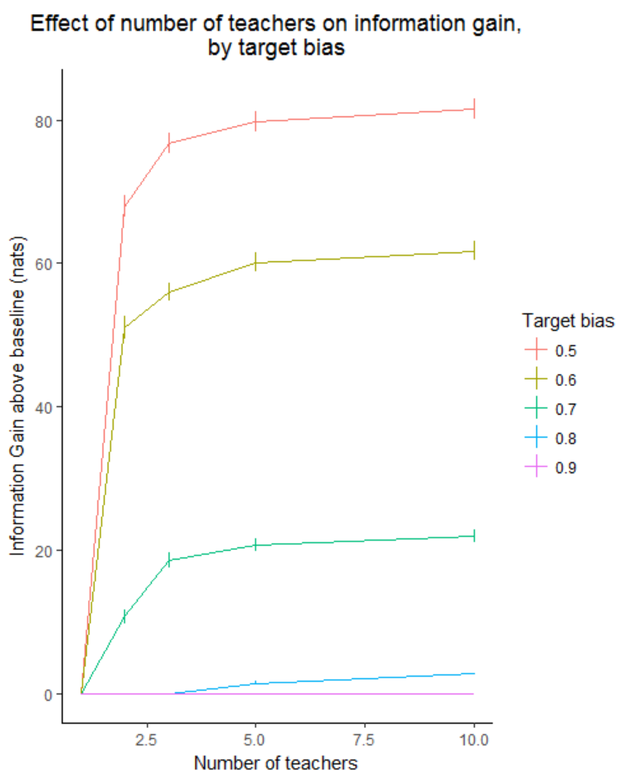

Figure 3. Students' amount of information gain when sorted into classrooms based on prior beliefs over unsorted baseline

Simulation 2: Number of teachers and class size

We hypothesized that increasing the number of teachers will strictly improve student learning rate. We further guessed that the improvements to learning rate will have diminishing returns, even though they will still increase learning. By having more teachers, students can get better individualized sets of examples from the teacher that appropriately helps them calibrate their prior beliefs towards the target concept.

Our baseline for this simulation was the sorted information gain under the same target bias parameters. Our simulated findings support our hypothesis and are consistent with real-world education research on class size (Glass and Smith, 1979; Slavin, 1989). We find that increasing the number of teachers, which decreases the number of students per class, results in greater information gain, as shown in Figure 3, $F(1,2498) = 181.2$, $p < .001$. The administrator consistently opted to maximize the number of teachers used.

Simulation 3: Noisy students and assessments

We hypothesized that as the number of assessments conducted to measure student ability increases, so will learning rate because teachers can update their beliefs about student prior beliefs. This allows them to more accurately select useful examples to present as they refine their model of student beliefs. As with the number of classrooms, we hypothesized that there would be diminishing returns of the number of assessments on the improvements to learning rate.

Our baseline for this simulation was the information gain under the same target bias parameters while using the true student beliefs. In other words, we measured the dropoff in information gain when examples are picked for the imperfect student priors compared to the information gain when examples are picked for true (omniscient) student priors. Throughout the simulation, we controlled for the number of classrooms by setting the number of classrooms to 5.

Our simulated findings support our hypothesis and are consistent with real-world education research on assessments (Fuchs and Fuchs, 1986). We find that increasing the number of assessments to reduce the noise in teacher knowledge of student prior beliefs results in a smaller dropoff in information gain, as shown in Figure 4, $F(1,2498) = 1004$, $p < .001$. The administrator consistently opted to maximize the number of assessments performed.

Worth noting is that student noise introduces uncertainty in two stages of the model. First, the administrator can sort students into classrooms less effectively. The classrooms distributed based on noisy beliefs about student priors will still be more homogenous than random classrooms, but will be less homogenous than classrooms based on students' true prior beliefs, which were used in Simulations 1 and 2. Secondly, what the teachers believe are the most optimal example sets to show their particular class of students may not actually produce optimal learning; the student agent still learns and updates their beliefs after example exposure based on their true prior beliefs. Further work could try to divorce effects on information gain introduced by the uncertainty introduced in these two areas.

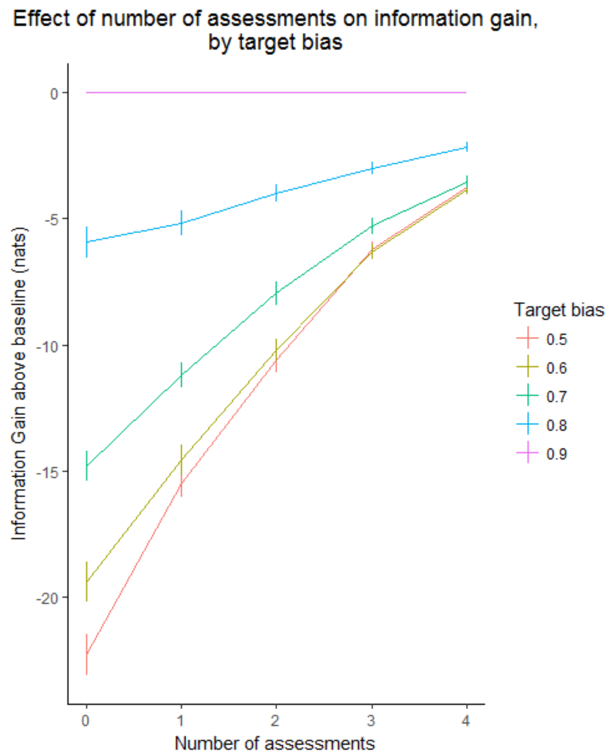

Figure 4. Students' amount of information gain by amount of assessment, where each assessment reduces the noise in teacher beliefs about the student priors

Simulation 4: Pareto frontier with imposed budget constraints

We hypothesized that for a particular budget constraint and fixed costs of teachers and assessment, there will be a pareto frontier of possible allocations of that budget towards teachers and assessments on which there will be an optimal allocation that maximizes student learning. Furthermore, we hypothesized that the individual results about number of teachers number of assessments would hold within levels of the other. This is the unification of all previous simulations into one model that allows us to introduce a real-world constraint of budget.

In this simulation, we test all integer combinations of number of teachers between 1 and 10 with number of assessments between 0 and 5. We assigned the cost of each teacher to be $C_T = \$10$ and the cost of each assessment $C_A = \$20$.

Our baselines were the worst-performing within-target setup. This always happened to be the 1-teacher 0-assessment setup, which is consistent with existing education literature, since a single teacher cannot significantly differentiate learning, and still has a lot of uncertainty about student beliefs because no assessments are performed.

Our simulated findings support our hypothesis. There is considerable support for an interaction model between teachers, assessments, and the target bias, F(7, 112) = 48.61, p<.001. To visualize the pareto frontier, we draw a heatmap of scores above baseline along teacher and assessment axes. We continue to see that the effects continue to be weaker for more extreme learning concepts, t(294) = -11.09, p<.001. This result can be seen in Figure 5.

Since real-world learning concepts are essentially non-negotiable, the effectiveness of a particular allocation of budget should be judged within-target, i.e. an optimal setup within the target bias = 0.8 level should not be penalized for being ineffective in comparison to the optimal setup for teaching a target bias of 0.5. We normalized the information gain above baseline within each target bias level. The results can be seen in 6

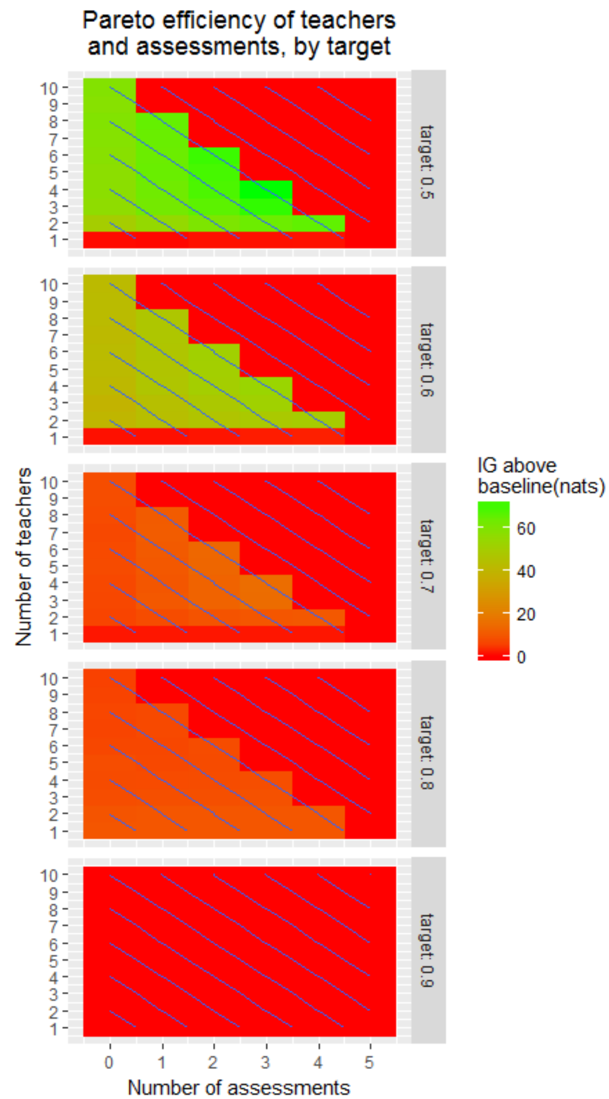

Figure 5. Pareto frontiers *as absolute measures of information gain above within-target baseline*. Diagonal lines show different budget levels, while the green line demarcates specifically the $B = 100$ budget. Costs of each teacher C_T is \$10, and cost of assessment $C_A$is \$20.

Consistent with Simulation 3, increasing the number of assessments on average improves learning rate up until you cannot afford any more assessments, across all levels of teachers and target biases. Consistent with Simulation 2, increasing the number of assessments improves learning rate up until you cannot afford any more assessments, across all levels of assessments and target biases of 0.5, 0.6, and 0.7.

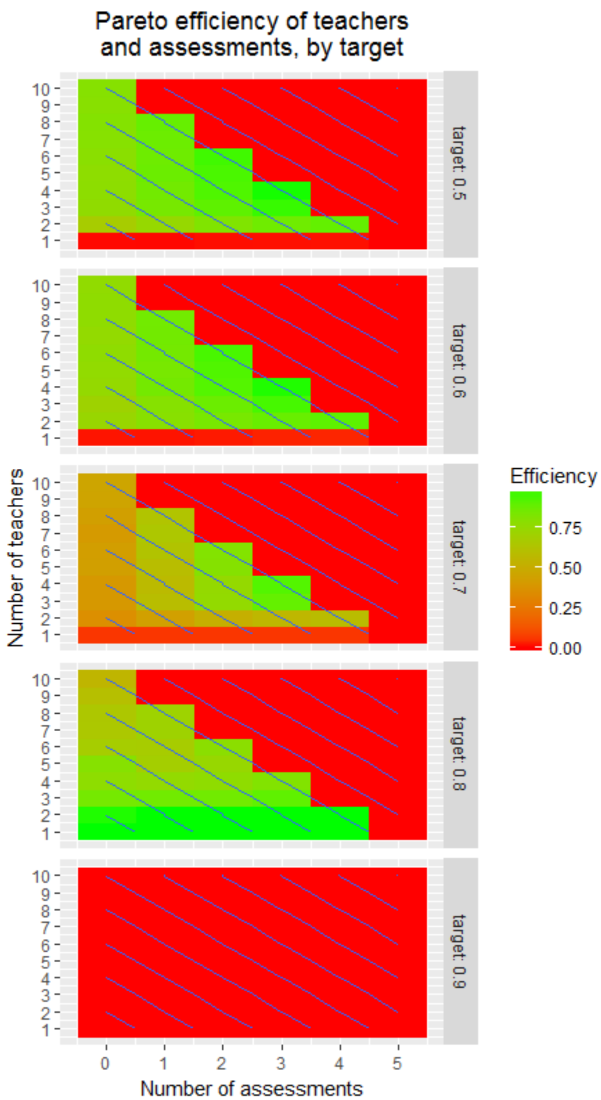

Figure 6. Pareto frontiers *as efficiency against best performing within-target setup*. Diagonal lines show different budget levels, while the green line demarcates specifically the $B = 100$ budget. Costs of each teacher C_T is \$10, and cost of assessment C_A is \$20.

An interesting result, however, is that having noisy students diminishes the effectiveness of increasing the number of teachers at certain a target bias of 0.8. While we aren't sure of the specific mechanism by which this happens, we suspect that having more teachers introduces additional risk of wrongly sorting students' prior beliefs, which would increase the heterogeneity of classrooms. Since a target belief of 0.8 is a possible target belief to show precisely with 5 examples (the basis of our simulations) but also extreme enough that most students's prior belief will fall on one side of that target, increasing the number of teachers might actually reduce the effectiveness of the teacher's choice of examples.

A novel finding that emerged not described in existing education literature is that there are different patterns of optimality along the pareto frontiers. As the extremity of the target bias moves from 0.5 to 0.7, we see that the IG-optimal allocation of budget shifts from about 5-6 teachers and 2 assessments towards more assessments at the expense of the number of teachers.

CONCLUSION

Overall, we were able to replicate many real-world findings from education literature. We found the sorting students into ability level classrooms, increasing assessment, and decreasing class size all increase information gain.

One of the more interesting findings was the influence of the target bias on the effects we were finding. Based on our results, target bias significantly moderates each of the effects we found. When teaching an extremely difficult (or easy, given symmetry) learning concept such that all student prior beliefs are on one side of the concept, the payoffs to information gain that sorting, class sizes, and assessment have diminish because regardless of the level of noise, teachers have relative certainty about the students' relative level of understanding. Similarly, target biases that are very close to central (around 0.5) have a lot to gain from diminishing noise in the system.

Finally, we found that the relative effectiveness of number of teachers and and the number of assessments varies depending on the target learning concept. We saw that the most optimal distribution of budget on the pareto frontier shifted towards reducing student noise when the target bias was more extreme–this seemed consistent with intuition, since having more teachers doesn't allow them to better select examples if the administrator produces greater sorting errors, resulting in heterogeneous classrooms where each teacher has less certainty about whether their examples will be useful for any particular student or not.

DISCUSSION

Implications

While a lot of aspects of school learning (peer effects, social dynamics, etc.) are not built into our model and would be difficult to capture algorithmically, we believe that there is a lot to be gained from simulating classroom dynamics. Using stochastic modeling is particularly useful when exploring the effectiveness of education policy measures for a handful of reasons. Fundamentally, the merits of stochastic modeling arise from the use of synthetic data. Since the studies are computer-simulated, there are no human subjects in the studies, and all of the data is synthetic. This allows researchers to take greater liberties in experimental design because there are no ethical concerns. For instance, we can execute a simulation that will intentionally misclassify synthetic students into classrooms that are not appropriate for their achievement level to understand the negative impact doing so has on both the misclassified student and his peers. Such a study would not be ethically permissible in an actual school.

Furthermore, no two schools are the same, and the use of modeling and synthetic data enables rapid iteration over different simulated schools with different simulated students. We can quickly test whether the purported benefit to student learning of various education interventions holds with, for example, a limited budget, or highly varied student achievement levels, or immensely large class sizes. There is also no risk of student subject attrition that may jeopardize real-world school studies; synthetic students do not exhibit unpredictable absenteeism, transfer schools, take sick days, etc. As such, we can customize our interventions based on school parameters, validate the robustness of our results, and better infer generalized findings about education policy without incurring astronomical experimental monetary costs nor require lengthy longitudinal study. With a complete stochastic model, we can quickly scale the number of hypotheses we test and the variety of schools we test them on.

Our series of simulations are designed to replicate real-world results from education literature in a quantitative, simulated fashion. The first few simulations looked at a lot of existing educational theory in isolation, while our final simulation attempts to unify these different design aspects of a school system into a more holistic view. We found that increasing the number of teachers (and thus decreasing class size) and increasing the amount of assessment to improve precision about student beliefs increases student learning, as predicted in real-world school studies.

Importantly, we were able to generate pareto frontiers allocating a fixed shared resource of budget into the dimensions of teachers and sorting. The ability to make sense of the tradeoffs that school administrators and policymakers need to make when designing their schools can help improve education. This would be particularly impactful in communities of low socioeconomic status where academic achievement tends to be lower because budgets are particularly constrained and attracting teaching staff is difficult.

Limitations and Future work

Some of the limitations of the present work is in the assumptions that we make to build our model. Most obvious is the simplicity of the learning task we described during the introduction. Most real-world learning tasks are more complicated and may have multiple objectives. Furthermore learning in the real-world is affected by a lot of more abstract facets: peer learning and social environment, parenting styles, stereotype threat, among other things. These are not captured in the model, and would be difficult to capture in any model.

The agnostic prior beliefs we 1) use for teachers to generate examples from; 2) generate student prior beliefs from; and 3) create target bias distributions with are all uniform distributions. These are not necessarily the case. Respectively, 1) teachers may have prior beliefs about which examples are more effective, akin to pedagogical knowledge or pedagogical content knowledge (Cochran, 1991); 2) students may tend towards certain common prior beliefs/misconceptions about a particular learning concept; 3) more extreme learning concepts may be more rare, and thus the admin should weight more heavily the pareto frontier of moderate target bias values.

Nevertheless, our results are promising and justify further work in this domain. Of particular interest are simulations of a lot of real-world scenarios in education. We list a few of them below:

- Misclassification of a student: what is the extent of the negative impact that happens when an admin wrongly sorts a student when distributing them into classrooms sorted by student prior beliefs?
- Absenteeism: what happens to learning rate when a student misses certain examples presented by the teacher?
- Stubbornness/Difficulty to teach: how do variations on the strength of student prior beliefs about learning concepts affect learning rate of all students in classroom settings?

ACKNOWLEDGEMENTS

Many thanks to Michael Frank, Michael Henry Tessler, and Noah Goodman for valuable discussion.

APPENDIX

Derivation of Generalized Multi-Sample Information Gain Formula

As described in the text, we are interested in computing

$$IG(e) = D_{KL}(B_T||B_S) - D_{KL}(B_T||B_{S+e}) \tag{3}$$

where the divergence measure is computed in closed form for e.g., B_T and B_S, as

$$\begin{aligned} D_{KL}(B_T||B_S) = & \log(\frac{B(\alpha_S,\beta_S)}{B(\alpha_T,\beta_T)}) \\ & + (\alpha_T-\alpha_S)\psi(\alpha_T) + (\beta_T-\beta_S)\psi(\beta_T) \\ & + (\alpha_T-\alpha_S+\beta_T-\beta_S)\psi(\alpha_T+\beta_T). \end{aligned} \tag{4}$$

where ψ denotes the digamma function and $B(a,b)$ denotes the beta function. We can substitute Equation 4 into Equation 3 to get

$$\begin{aligned} IG(e) = & \log(\frac{B(\alpha_S,\beta_S)}{B(\alpha_T,\beta_T)}) \\ & + (\alpha_T-\alpha_S)\psi(\alpha_T) + (\beta_T-\beta_S)\psi(\beta_T) \\ & + (\alpha_T-\alpha_S+\beta_T-\beta_S)\psi(\alpha_T+\beta_T) \\ & - \log(\frac{B(\alpha_{S+e},\beta_{S+e})}{B(\alpha_T,\beta_T)}) \\ & - (\alpha_T-\alpha_{S+e})\psi(\alpha_T) - (\beta_T-\beta_{S+e})\psi(\beta_T) \\ & - (\alpha_T-\alpha_{S+e}+\beta_T-\beta_{S+e})\psi(\alpha_T+\beta_T) \end{aligned} \tag{5}$$

Consider the case where e is a series of h 1's (heads) and t 0's (tails). Then $\alpha_{S+e} = \alpha_S + h$ and $\beta_{S+e} = \beta_S + t$, so we can simplify Equation 5 to

$$
\begin{aligned}
IG(e) =& \log(\frac{B(\alpha_S,\beta_S)}{B(\alpha_T,\beta_T)}) - \log(\frac{B(\alpha_S+h,\beta_S+t)}{B(\alpha_T,\beta_T)}) \\
&+ (\alpha_T-\alpha_S)\psi(\alpha_T) + (\beta_T-\beta_S)\psi(\beta_T) \\
&+ (\alpha_T-\alpha_S+\beta_T-\beta_S)\psi(\alpha_T+\beta_T) \\
&- (\alpha_T-\alpha_S-h)\psi(\alpha_T) - (\beta_T-\beta_S-t)\psi(\beta_T) \\
&- (\alpha_T-\alpha_S-h+\beta_T-\beta_S-t)\psi(\alpha_T+\beta_T) \\
=& \log(\frac{B(\alpha_S,\beta_S)}{B(\alpha_T,\beta_T)} \cdot \frac{B(\alpha_T,\beta_T)}{B(\alpha_S+h,\beta_S+t)}) \\
&+ (h)\psi(\alpha_T) + (t)\psi(\beta_T) + (h+t)\psi(\alpha_T+\beta_T). \\
=& \log(\frac{B(\alpha_S,\beta_S)}{B(\alpha_S+h,\beta_S+t)}) \\
&+ (h)\psi(\alpha_T) + (t)\psi(\beta_T) + (h+t)\psi(\alpha_T+\beta_T).
\end{aligned}
\tag{6}
$$

And, since

$$B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}, \tag{7}$$

we can rewrite the first term and reduce:

$$
\begin{aligned}
IG(e) =& \log(\frac{\frac{\Gamma(\alpha_S)\Gamma(\beta_S)}{\Gamma(\alpha_S+\beta_S)}}{\frac{\Gamma(\alpha_S+h)\Gamma(\beta_S+t)}{\Gamma(\alpha_S+\beta_S+h+t)}}) \\
&+ (h)\psi(\alpha_T) + (t)\psi(\beta_T) + (h+t)\psi(\alpha_T+\beta_T). \\
=& \log(\frac{\Gamma(\alpha_S)\Gamma(\beta_S)\Gamma(\alpha_S+\beta_S+h+t)}{\Gamma(\alpha_S+h)\Gamma(\beta_S+t)\Gamma(\alpha_S+\beta_S)}) \\
&+ (h)\psi(\alpha_T) + (t)\psi(\beta_T) + (h+t)\psi(\alpha_T+\beta_T). \\
=& \log(\frac{\Gamma(\alpha_S)}{\Gamma(\alpha_S+h)}) + \log(\frac{\Gamma(\beta_S)}{\Gamma(\beta_S+t)}) \\
&+ \log(\frac{\Gamma(\alpha_S+\beta_S+h+t)}{\Gamma(\alpha_S+\beta_S)}) \\
&+ (h)\psi(\alpha_T) + (t)\psi(\beta_T) + (h+t)\psi(\alpha_T+\beta_T). \\
=& -\log(\frac{\Gamma(\alpha_S+h)}{\Gamma(\alpha_S)}) - \log(\frac{\Gamma(\beta_S+t)}{\Gamma(\beta_S)}) \\
&+ \log(\frac{\Gamma(\alpha_S+\beta_S+h+t)}{\Gamma(\alpha_S+\beta_S)}) \\
&+ (h)\psi(\alpha_T) + (t)\psi(\beta_T) + (h+t)\psi(\alpha_T+\beta_T).
\end{aligned}
\tag{8}
$$

When $h = 0$, the first log term reduces to 0. When $t = 0$, the second log term reduces to 0. Otherwise, since for positive integer n,

$$\frac{\Gamma(x+n)}{\Gamma(x)} = \prod_{k=0}^{n-1}(x+k), \tag{9}$$

we can reduce the previous formulation a bit further when $h \neq 0$ and $t \neq 0$, to

$$
\begin{aligned}
IG(e) =& -\log(\prod_{i=0}^{h-1}(\alpha_S+i)) - \log(\prod_{j=0}^{t-1}(\beta_S+j)) \\
&+ \log(\prod_{k=0}^{h+t-1}(\alpha_S+\beta_S+k)) \\
&+ (h)\psi(\alpha_T) + (t)\psi(\beta_T) \\
&+ (h+t)\psi(\alpha_T+\beta_T).
\end{aligned}
\tag{10}
$$

For computation, to prevent overflow we calculate the series equivalently as follows:

$$
\begin{aligned}
IG(e) =& \sum_{k=0}^{h+t-1}\log(\alpha_S+\beta_S+k) \\
&- \sum_{i=0}^{h-1}\log(\alpha_S+i) - \sum_{j=0}^{t-1}\log(\beta_S+j) \\
&+ (h)\psi(\alpha_T) + (t)\psi(\beta_T) \\
&+ (h+t)\psi(\alpha_T+\beta_T).
\end{aligned}
\tag{11}
$$

which is Equation (2) in the text.

REFERENCES

Camerer, C. F., Ho, T. H., & Chong, J. K. (2004). A cognitive hierarchy model of games. The Quarterly Journal of Economics, 861-898.

Cochran, K. F. (1991). Pedagogical Content Knowledge: A Tentative Model for Teacher Preparation.

Frank, M. C. (2014). Modeling the dynamics of classroom education using teaching games.

Frank, M. C., & Goodman, N. D. (2012). Predicting pragmatic reasoning in language games. Science, 336(6084), 998-998.

Fuchs, L. S., & Fuchs, D. (1986). Effects of systematic formative evaluation: A meta-analysis. Exceptional children, 53(3), 199-208.

Glass, G. V., & Smith, M. L. (1979). Meta-analysis of research on class size and achievement. Educational evaluation and policy analysis, 1(1), 2-16.

Jager, G. (2010). 9 Game-Theoretical Pragmatics. Handbook of logic and language, 467.

Nagel, R. (1995). Unraveling in guessing games: An experimental study. The American Economic Review, 85(5), 1313-1326.

Shafto, P., & Goodman, N. (2008). Teaching games: Statistical sampling assumptions for learning in pedagogical situations. In Proceedings of the 30th annual conference of the Cognitive Science Society (pp. 1632-1637). Austin, TX: Cognitive Science Society.

Shafto, P., Goodman, N. D., & Frank, M. C. (2012). Learning from others the consequences of psychological reasoning for human learning. Perspectives on Psychological Science, 7(4), 341-351.

Slavin, R. E. (1987). Ability grouping and student achievement in elementary schools: A best-evidence synthesis. Review of educational research, 57(3), 293-336.

Slavin, R. E. (1989). Class size and student achievement: Small effects of small classes. Educational Psychologist, 24(1), 99-110.