Project 3: Eliminating Child Care Deserts in New York State through Optimization

Tianlong Nan, Yaren Bilge Kaya, Daniel Bienstock, Zhixuan Ren, and Emily Cai

March 26, 2025

Child care, commonly referred to as day care, involves the supervision and care of children, typically ranging from two weeks to 12 years of age[1]. This service is particularly vital for working parents, offering a professional environment for children while their caregivers are engaged in employment[2].

An ongoing challenge in the United States (US) is the prevalence of "child care deserts"—regions where the demand for licensed child care far exceeds available slots. Over half of American children live in such areas[3]. For instance, due to the limited availability of child care services, many regions in New York State are designated as "child care deserts." This classification is determined by several factors, including:

- The total number of available child care slots;
- The population of children requiring care;
- The percentage of employed parents;
- The average income in the area.

Specifically:

1. In high-demand areas—defined as regions where at least 60% of parents are employed or the average income is $60,000 or less per year—an area is considered a child care desert if the number of available slots is less than or equal to half the population of children aged two weeks to 12 years.

2. In normal-demand areas, where employment and income levels do not meet the high-demand criteria, the threshold is lower: an area is classified as a child care desert if the available slots are less than or equal to one-third of the population of children within the same age range.

In this project we will focus on the issue of child care deserts in New York State (NYS). It is important to consider specific policies that apply to NYS, such as the increased capacity requirements for children aged 0-5. Babies and toddlers, in particular, tend to require more child care than older children, which is expected to be reflected in the state's policy. Beyond addressing overall demand, the NYS government ensures that children under the age of 5 have sufficient access to care. This means that the number of available slots for children in this age group must be at least two-thirds of the population of children aged 0-5.

[1] Bradley, R. H., & Vandell, D. L. (2007). Child care and the well-being of children. *Archives of pediatrics & adolescent medicine*, *161(7)*, 669-676.

[2] Adams, G., & Henly, J. R. (2020). Child care subsidies: Supporting work and child development for healthy families.

[3] `https://www.americanprogress.org/series/child-care-deserts/`

The Problem of Budgeting

New York State government aims to eliminate child care deserts across the state by increasing the number of child care slots in all areas, ensuring that no region is classified as a child care desert. To achieve this goal, the government plans to allocate funding for either building new child care facilities *or* expanding existing ones.

As a first step, NYS envisions an ideal scenario where new facilities can be built anywhere in the state.[4] There are three facility sizes available for construction, each with specific capacities and associated costs, as outlined in the provided data.

- When expanding existing facilities, the maximum increase in capacity to 1.2 times the current size (i.e. adding up to 20%), up to a maximum (prior capacity + new capacity) of 500 slots per facility.

Because expanding capacity at larger facilities is often easier and less costly on a per-slot basis compared to smaller facilities, the cost of expanding child care capacity should depend on the *scale of the expansion*[5]. In particular, the expansion cost is influenced by the facility's current capacity and the scale of the expansion:

- The expansion cost is piecewise-linear in the proportion of the increase in capacity;

Additionally, any newly created slots for children under the age of 5 require an extra cost of $100 per slot for specialized equipment.

Your task as consultants is to assist the NYS government in determining the **minimum amount of funding** (in total) needed to meet their target for each area, categorized by zip code.

The Problem of Realistic Capacity Expansion and Distance

After reviewing the idealistic scenario, NYS officials provided additional recommendations to better reflect the complexity of expanding child care facilities and choosing appropriate locations for new ones.

Recommendations on costs. Officials recognized that, due to space limitations, the marginal cost of expanding slots in existing facilities should depend on the scale of the expansion. In other words, the larger the expansion, the higher the marginal cost. As a result, instead of providing a fixed cost function, the officials propose a more realistic approach where the cost of adding slots follows a model where higher the expansion, higher the increase in the cost of adding more slots [6]. Specifically, where n is the number of slots that exist at a facility and m represents the number of slots to be expanded, the cost function is piecewise-linear in m/n, with

$$\text{slope} = \begin{cases} (20,000 + 200n) & 0 \leq m/n \leq 0.1, \\ (20,000 + 400n) & 0.10 \leq m/n \leq 0.15, \\ (20,000 + 1,000n) & 0.15 \leq m/n \leq 0.20. \end{cases}$$

[4] This means that any number of new facilities can be built in any area within NYS.

[5] Proportion of the increase in capacity.

[6] Think of it this way: when you only increase the number of slots by a small amount, it's relatively cheap. However, as you add more and more slots, the cost to add additional slots becomes higher, reflecting the increased difficulty of expansion.

Note that we will also have $n + m \leq 500$. This three-step cost curve is used in Problems B and C. For Problem A, choose one of the three slopes – that *single* slope will be valid for the entire range $0 \leq m/n \leq 0.20$.

In essence, the more slots you try to add, especially beyond certain thresholds, the more expensive it becomes to add each additional slot. The cost grows in stages, so recognize that the larger the expansion, the harder and more costly it becomes. The restriction on the scale of expansion is equivalent to say: it costs arbitrarily large to expand a facility to more than 20% of its current capacity, thus no one wants do that.

Recommendations on locations. At the same time, NYS officials have been evaluating potential locations for new facilities. To avoid an over-concentration of child care centers in specific regions, they recommend imposing a distance limitation between facilities, within each area. This means that, within each area, no two facilities, whether new or existing, can be located too close to one another. The minimum distance is 0.06 mile.

Your task is to help the NYS government determine the **minimum amount of funding** needed to achieve their goals, considering both the more complex cost structure and the distance limitation for new and existing facilities.

The Problem of Fairness

To address the issue of child care deserts, governments have introduced significant funding to help child care providers maintain and expand their services. In New York State (NYS), the government has received a total of $1 billion to improve the availability of child care services statewide[7] . However, beyond simply eliminating child care deserts, the government also wants to ensure that access to child care is distributed fairly across all regions, while maximizing overall social coverage.

One key fairness measure officials recommend is to minimize the gap in child care availability between different areas. Specifically, they want to ensure that the difference in the ratio of available child care slots to the total population of children between any two areas does not exceed 0.1. This means that no region should be significantly better or worse off than another in terms of child care access.

Under this fairness constraint, the government's goal is to maximize a social coverage index. This index is based on the weighted sum of child care coverage for two groups: children under 5 and all children, with a 2:1 weighting in favor of younger children (under 5). The higher weight reflects the greater importance of child care coverage for younger children in promoting overall social well-being.

Your task is to help the NYS government determine the best strategy to maximize this social coverage index while ensuring that no area is classified as a child care desert and staying within the $100 million budget. If it is impossible to satisfy the fairness requirement under the given budget, report this issue.

Available Data

- `child_care_regulated.csv`: This dataset contains the information of (existing) child care facilities in New York State.
- `population.csv`: This dataset contains the population of children in different ranges of ages (e.g., 0-5, 5-10, 10-14, etc.) in each zipcode region in New York State.

[7] `https://ocfs.ny.gov/programs/childcare/deserts/#overview`

- `avg_individual_income.csv`: This dataset contains the average individual income in each zipcode region in New York State.
- `employment_rate.csv`: This dataset contains the employment rate in each zipcode region in New York State.
- `potential_locations.csv`: This dataset contains the potential locations available for building a facility in New York State.
- We estimate the cost of building a new facility based on this article. To build a new facility, you can assume the government needs to spend a fixed (one-time) amount of money based on the capacity of the facility:

Facility Size	# of Slots (Ages 0-5)	Cost of New Facility ($)
100 slots (Small)	50 slots	65,000
200 slots (Medium)	100 slots	95,000
400 slots (Large)	200 slots	115,000

Table 1: Construction cost estimates for different sizes of child care facilities

Problem A: Budgeting

Sets

- A_h: the set of areas (zip codes) in NYS with high demand for child care services.
- A_n: the set of areas in NYS with demand for child care services that is not high.
- A: the set of all areas in NYS); $A = A_h \cup A_n$. Note that $A_h \cap A_n = \emptyset$.
- J_a: the set of all current facilities in area a, where $j \in J_a$.

Parameters

Remark. *For those parameters which are described as simple scalars in the project statement, i.e., $r_a, e_a, f^{e,b}, f^{e,c}$ and f_i^n, c_i^t, c_i^u for $i = 1, 2, 3$, it is optional to define specific parameters for them. Doing so may help you understand the behavior of a solution.*

- Demographic information
  - r_a: the rate of employment of area a;
  - e_a: the average income level of area a;
  - p_a^t: the total number of children ages 2 weeks-12 years in area a;
  - p_a^u: the total number of children under 5 (babies and toddlers) in area a;
- Cost information
  - f_i^n: the cost of building a new facility of type i; (here, and below, $i = 1, 2, 3$ represents the facility type). This data is as in Table 1 provided above.
  - From the piecewise-linear cost function (for expanding a facility) choose one of the three slopes;
  - $f^{e,u} = 100$: the additional cost of extending an existing facility by one slot for children under 5.
- Operational information
  - c_i^t: the total capacity of a newly-built facility of type i for children (in all age ranges);
  - c_i^u: the capacity of a newly-built facility of type i for children under 5;
  - $n_{a,j}^t$: the total number of existing slots in facility j in area a for all age ranges;
  - $n_{a,j}^u$: the number of existing slots in facility j in area a for child under 5.

Decision variables

- $y_{a,i}$ integral: the number of new facilities of type i to build in area a ($i = 1, 2, 3$);
- $m_{a,j}^u$ integral: the number of extended slots for children ages 0-5 in facility j in area a;
- $m_{a,j}^t$ integral: the total number of extended slots for children in facility j in area a;

Objective

The objective should account for the cost of the extensions (the m variables) as well as the new facilities (the y variables).

Constraints

1. Child care coverage. You should have a constraint for every area a with high need (in A_h) and one constraint for every area with normal need (A_n) (one-half and one-third coverage requirement, respectively). The constraint should use the p_a^t values, appropriately, to represent coverage requirement, to be fullfilled using the $m_{a,j}^t$ and the $y_{a,i}$.
2. There is enough capacity for children under 5: Here there is a constraint involving the p_a^u (appropriately, there is a two-thirds rule) as requirements, to be met using the m^u and the y variables.
3. The number of new slots for children under 5 cannot exceed the total number of new slots:
$$m_{a,j}^u \leq m_{a,j}^t, \quad \forall\, j \in J_a,\ a \in A.$$
4. Constraints on the maximum number of slots and expansions. These constraints involve the m variables.

Problem B: The Problem of Realistic Capacity Expansion and Distance

Parameters

- All parameters except $f^{e,c}$ in Problem A are available in this problem.
- $f^{e,c,(1)}$: the existing-capacity-based cost of extending an existing facility within the segment of $[0, 0.10]$;
- $f^{e,c,(2)}$: the existing-capacity-based cost of extending an existing facility within the segment of $[0.10, 0.15]$;
- $f^{e,c,(3)}$: the existing-capacity-based cost of extending an existing facility within the segment of $[0.15, 0.20]$;
- $B(a)$: the (finite) set of all potential locations to build new facilities upon.

Decision variables

- $z_{a,l} \in \{0, 1\}$: whether or not to build a new facility at the potential location l in area a, for each of ;
- $z_{a,l,i} \in \{0, 1\}$: whether or not to build a new facility of type i at the potential location l in area a;
- $x^{(1)}_{a,j} \in \mathbb{R}_{\geq 0}$: the ratio of extended slots over the current slots in facility j in area a (in the range of $[0, 0.1]$);
- $x^{(2)}_{a,j} \in \mathbb{R}_{\geq 0}$: the ratio of extended slots over the current slots in facility j in area a (in the range of $(0.1, 0.15]$);
- $x^{(3)}_{a,j} \in \mathbb{R}_{\geq 0}$: the ratio of extended slots over the current slots in facility j in area a (in the range of $(0.15, 0.2]$);
- $m^u_{a,j} \in \mathbb{Z}^+$: the number of extended slots for children ages 0-5 in facility j in area a;
- $m^t_{a,j} \in \mathbb{Z}^+$: the total number of extended slots for children in facility j in area a;
- $b_{a,j} \in \{0, 1\}$: whether the facility j in area a is expanded ($= 1$ if expanded and 0 otherwise).

Objective

Minimize expansion cost.

Constraints

- All child care deserts are eliminated:
- There is enough capacity for children under 5:
- The expansion rate constraints:
- The number of new slots for children under 5 should not exceed the total new slots:

- The constraints on the maximum number of the slots we can extend:
- One location can only be used to build one facility:

$$\sum_{i \in \{1,2,3\}} z_{a,l,i} \leq z_{a,l}.$$

- The minimum distance between any two facility locations within an area is not too small.

Problem C

Parameters

- All the parameters of Problem B are available in this problem.

Decision variables

- All the decision variables in Problem B are decision variables in this problem as well.
- $t \in \mathbb{R}_{\geq 0}$: an auxiliary decision variable. It represents the minimum weighted coverage ratio.

Objective

$$\max \quad t$$

Constraints

- All constraints in Problem B should be preserved.
- The fairness requirement:
- t is the minimum weighted coverage ratio:

$$t \leq \frac{2}{p_{a,b}} \left(\sum_{j \in J_a} (\bar{n}_{a,j,b} + m_{a,j}) + \sum_{i \in I} c_i^5 \sum_{l \in B(a)} z_{a,l,i} \right) + \frac{1}{p_{a,t}} \left(\sum_{j \in J_a} \left(1 + x_{a,j}^{(1)} + x_{a,j}^{(2)} + x_{a,j}^{(3)}\right) n_{a,j} + \sum_{i \in I} c_i^{18} \sum_{l \in B(a)} z_{a,l,i} \right), \quad \forall\, a \in A.$$

- Budget constraints.

Wed.Mar.26.210138.2025@freestyle