



Lesson 6

Aligning Input Data with Parameters in Your Cost Model

Getting Real About the True Costs of Programs for Children and Youth

February 2024

Lesson Overview

6.1 Input Data in Cost Modeling.....	3
6.2 Factors That Influence Input Data in Cost Modeling.....	6
6.3 Common Input Data in Cost Modeling.....	7
6.4 Example from the Field	8
6.5 Equity Check	9
6.6 Decision Point: What Input Data Will You Include in Your Cost Model?	10
6.7 Key Takeaways	10
6.8 Going Deeper: Learn More About Input Data in Cost Models	11

Learning Objectives

At the end of this lesson, you will be able to do the following:

- Define what input data is in cost modeling and understand the difference between determined data and estimated data.
- Identify two or three factors that influence the accuracy and reliability of input data in cost modeling.
- Describe four or five common types of input data used in cost models.
- Understand how to engage key groups to check the accuracy and validity of input data used in cost models.

Lesson Resources

Example Cost Model ([contact us](#) for access)

6.1 Input Data in Cost Modeling

Input data is the actual data used in the cost model's calculations. Cost models include two main types of input data—determined data and estimated data.

Determined data (known data) is data that is known and fixed. It is not based on assumptions or estimates. Determined data is often used to set the boundaries or requirements for certain aspects of the cost model. Examples of determined data entered in a cost model could be child care licensing standards including required staffing ratios and group sizes; standardized wage scales; length of service provision each day and the number of days per year; and staffing qualifications, compensation, and access to professional supports.

Estimated data (unknown data) is data that is not known and must be estimated. It is based on

assumptions or guesses when certain information is unavailable or missing. Cost modelers use various techniques to fill in those blanks, for example they might do the following:

- Use national averages as proxies.
- Leverage data from similar programs or populations.
- Seek guidance from partners or subject-matter experts.
- Rely on historical data to estimate missing values.

These assumptions are entered into the cost model and are integral to the calculations and outputs. Examples of estimated data entered in a cost model could be the cost of materials, enrollment rates, or the demand for a new service.



What Is the Difference Between Parameters and Input Data in Cost Modeling?

Parameters are fixed and predetermined values that guide the cost modeling process, whereas input data is the actual data used in the cost model calculations. Parameters provide the framework and constraints for the cost model analysis while data inputs supply the specific information needed to perform the cost model calculations and generate meaningful outcomes/results. Input data represents both known (determined) and unknown (estimated) data. Here are some examples of these differences:

Parameters

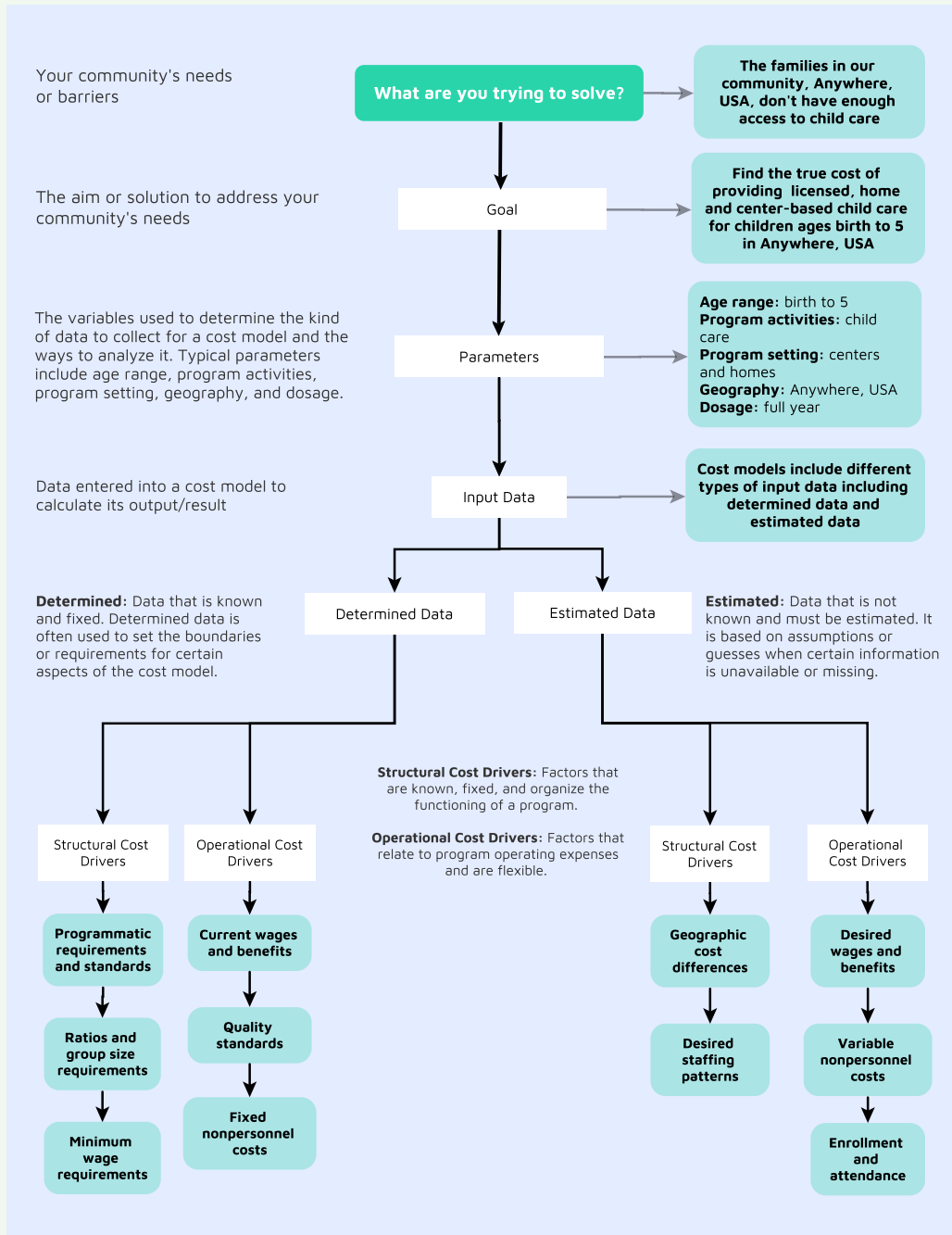
- Parameters are the variables you decide on to determine the kind of data you will collect. Two examples of parameters are the age range of the children and youth served by the program and different program settings such as center-based care or home-based care.

Input Data

- **Determined data:** Children-to-adults ratios can be researched through public online records, such as a state's child care licensing standards, and would be considered a determined data input. In Colorado, for example, the ratio of toddlers (ages 24 months to 36 months) per adults for child care centers is 1 staff member to 7 toddlers (1:7).
- **Estimated data:** Quality standards or guidelines often require you to make informed decisions when estimating input data. For example, if a quality standard specifies that programs use one of several approved types of new curriculum, you would want to reflect the cost of adhering to this guideline in nonpersonnel costs under a curriculum line item. The cost of this curriculum can be gleaned by doing some research on the costs of the most frequently chosen approved curricula and verifying whether a subscription is required to continue using the curriculum. Determining these costs also requires an educated guess on the number of hours needed to train staff on the new curriculum. Then, this cost might be included in additional substitute hours to cover the time that staff are out of the classroom for the curriculum training.

Figure 1 on page 4 further illustrates the role of input data and parameters in the context of cost model development.

Figure 1: Cost Model Flow Chart



The accuracy of a cost model depends on the quality of the input data. Though both determined data and estimated data are important, determined data should be used whenever possible because it can help improve the accuracy of the model's results. When estimated data is necessary, it is essential to use reliable sources and to make sure that the estimates are as accurate as possible. Experienced cost modelers are adept at combining both determined and estimated data into their cost models. By carefully

considering the type of data selected for each cost driver, cost modelers can ensure their models are accurate and reliable.

To help you better distinguish between parameters and input data—because they *do* differ from each other—see page 3 and read *What Is the Difference Between Parameters and Input Data in Cost Modeling?*

What Causes Variations in Cost Model Results?

Cost model results—the final price tags or cost outputs—will vary based on the following general factors:

- Different parameters you choose to focus on that guide and constrain your cost model
- Decisions you make about the specific data inputs to include in your model
- Different scenarios you explore to compute true costs for various programs and services

Consider the example in figure 2 that shows how three different scenarios for teacher salaries led to three different cost model results (shown in the Monthly Operating Costs column) for a preschool program. Two key parameters guiding the focus of this cost model are levels of quality and teacher compensation. Actual values (determined input data) for current teacher salaries (\$12/hour) are used as input data in the first option (Cost Model A). These input values are then adjusted to explore two additional scenarios where teacher salaries are increased to \$15/hour (Cost Model B) and \$22/hour (Cost Model C).

Seeing the results of these three options side-by-side reflects the power of a cost model because it visualizes the benefits and drawbacks of different decisions regarding the allocation of resources and their respective opportunity and equity costs. Let’s analyze this further, as follows:

- **Cost Model A** represents the status quo where

teachers are paid \$12/hour—an unlivable wage that is so low, teachers qualify for public support. Cost Model A is built using determined data (current wages) and reflects the inequitable reality that the early childhood workforce lives in poverty and subsidizes the true cost of early childhood education through substandard compensation rates.

- **Cost Models B and C** model the conditions that *should* exist if early childhood educators are compensated fairly, with a living wage that reflects their professional skills and knowledge. Cost Models B and C more accurately represent the true costs of educating young children in the cost outputs for a program’s operating expenses.

This example illustrates several reasons why it is important to carefully consider and discuss the different types of input data you plan to use in your cost model, as follows:

- **Your choice of input data can significantly impact the accuracy and reliability of the cost model results.** For example, consider staffing patterns. Once you begin to dig into the licensing requirements for staffing and ratios, you’ll start to imagine a typical classroom, such as one staff person per four infants in a class of nine children maximum. You may think, “That would mean three staff”—but at what job level and with what full-time equivalent? Three full-time teaching staff will have different results than one full-time lead teacher, one full-time certified teacher, and one part-time assistant teacher per classroom. Once

Figure 2: Sample Scenarios for Preschool Teacher Salaries

	Number of Full-Time Equivalent Teachers	Teacher Salary by Hourly Rate	Rent and Utilities	General and Administrative Costs	Monthly Operating Expenses
Cost Model A	10	\$12	\$1,000	\$800	\$21,000
Cost Model B	10	\$15	\$1,000	\$800	\$25,800
Cost Model C	10	\$22	\$1,000	\$800	\$37,000

you decide on your staffing pattern scenarios, you can present them to the program operators to check their validity and create space for changes in the data inputs/data values early on in the cost model development.

- **Discussing a cost model's input data decisions up front helps key contributors understand how a cost model works.** Being transparent about the input data entered in a cost model can increase trust and credibility with program providers and the broader community, because they can see how the cost estimates were derived and assess the accuracy of the data (determined data, estimated data, and scenarios that explore various options using different data values).
- **Individuals involved in the cost model process develop a greater understanding of the limitations and potential biases in cost models.** Your communication with the individuals supporting your cost model, as well as general community members, should be continuous throughout the cost modeling process to check your biases and uncover limitations. An example of this would be assuming that all youth participating in out-of-school time programming attend programs provided through their school or a large center and, as a result, limiting input data to only these programs. While school-based programs are an important part of out-of-school time learning experiences, they may be inaccessible to some young people because of fees or transportation. Community members, local leaders, and program providers will have knowledge about other programs that their community has, such as libraries, recreational centers, or civic engagement clubs where young people can learn, play, and grow.
- **Talking through which input data to include in a cost model can lead to cost model outputs that more likely reflect the true costs and realities of a program or service.** For example, you may include a certain staffing support position, such as a substitute teacher, in your cost model but a program may not have a pool of substitutes or prospects to use. Talking through this can illuminate current operational challenges, like staffing shortages, and how they are reflected in the cost model as well as what the ideal program with full staffing would cost.
- **The process of identifying input data can facilitate discussions and decision-making among the team supporting the cost model.** This process helps everyone involved consider the

potential impact of different assumptions on the cost model outputs and the program or service being modeled.

Take a moment to focus on the content below in Data Input Decisions for Early Childhood Education Cost Modeling for more information about the decisions cost modelers typically make.



Data Input Decisions for Early Childhood Education Cost Modeling

When developing cost models for early childhood education programs and services, you will typically need to identify specific input data about the following areas:

- Program delivery models (part-day, school day, or full-day care—aka number of hours of care provided)
- Number of days of care per year
- Average number of hours worked per staff member
- Staff turnover or vacancy rates
- Rate at which staff enroll in benefits
- Child enrollment rates

6.2 Factors That Influence Input Data in Cost Modeling

Various factors will influence the range and types of input data you enter into a cost model. Understanding and acknowledging the effects of these factors will help you develop more accurate and reliable cost models. These factors include the following:

- **Historical data:** The data used to build a cost model is often based on past expenditures or experiences, which can shape assumptions about future costs. For example, when collecting data from child care providers, we ask them for financial information from a few recent fiscal years, such as their payroll, tax returns, and annual operating budget.
- **Geographic location:** The cost of labor, materials, and other resources can vary by geographic location, so input values about costs may vary based on where the program or service is implemented. You can see this, for example,

when exploring a new salary scale to include in the cost model. Cost of living data values will vary based on the cost of housing, insurance, food, transportation, and other basic necessities in a particular location.

- **Community input:** Input from program staff, administrators, policymakers, and community members can influence decisions made about which data inputs are entered (or not entered) into a cost model. For example, in the area of family child care, many times, providers in homes may not view certain activities as expenses, such as having a family member assist with child care or paying a neighbor to do maintenance work. By gathering information directly from the providers, you can include these support costs and people as data inputs entered into the cost model, making the cost model outputs more reflective of the true cost of care in family child care homes.
- **External factors:** External factors such as changes in laws, regulations, or economic and political conditions can also influence the data inputs used in a cost model. During the COVID-19 pandemic, for instance, staff-to-child ratios had to be reduced because of emergency health regulations. These reduced ratios resulted in a need for additional staff to care for the same number of children, leading to a higher cost per child.

These factors are dynamic and can change over time—often as a result of changes in the program, the community, and the economy. That means you will need to revisit the data inputs entered into a cost model, and potentially revise them periodically, to ensure that the cost model continues to reflect the true costs and realities of the program or service being modeled.

6.3 Common Input Data in Cost Modeling

Examples of the most common input data entered into cost models include the following:

Staffing data: This data includes specific adult-child ratios, as well as the salaries, level of education, and experience required for different staff positions.

- Example: There is one teacher for every four infants and toddlers or 12 preschool children.
- Example: Lead teachers must have a bachelor's degree.

Program data: This includes factors such as the length of the program day, the age range of children served, and the type and intensity of the education and/or interventions provided.

- Example: A child care program operates for eight hours per day, serves children 3–5 years, provides meals, and offers mental health services to children with identified mental health needs.



Reviewing Your Input Data with Community Members

To ensure the accuracy and validity of the input data used in your cost model, it is important to establish mechanisms for soliciting feedback from state and community members. This can be done in various ways, including holding focus groups, conducting surveys, and convening advisory committees.

In Illinois, for example, state leaders built a cost model to determine the cost of providing high-quality early childhood education to all children in the state. To ensure the input data they used was accurate, they solicited feedback, both verbally and via an online form, from parents, educators, community leaders, and policymakers.

Through the online form, for example, state leaders received important feedback that challenged them to revise some aspects of their cost model to address children with disabilities, dual language learners, and family child care programs. After receiving this feedback, they met with the people and agencies that had submitted comments to clarify and improve their cost estimates. For example, staff from the Latino Policy Forum provided invaluable assistance and information about accurate costs to support young dual language learners. This ensured that the cost model captured the appropriate additional costs, and the number of children were identified for these special populations.

Source: *Illinois Cost Model for Early Childhood Education and Care Services. Updated February 2021*

Facilities data: This includes the cost of facilities including rent, utilities, insurance, and maintenance.

- Example: Rent for an infant-toddler program is \$22.50 per square foot, utilities cost \$450 per month, and maintenance costs are 2% of the facility's total value.

Cost of living data: This means the cost of living in the geographic area where the program is located and includes the cost of housing, transportation, and other basic needs in a specific community or state.

- Example: A cost model might recognize that the cost of living in certain communities or counties is 25% higher than the state median and that factor impacts the salary and benefits offered to program staff.

Find out why it is valuable to discuss your input data with the individuals who will be affected by the cost model's outcomes—and potentially revise your data as a result—in *Reviewing Your Input Data with Community Members* on page 7.

6.4 Example from the Field

The *Modeling the Cost of Child Care in the District of Columbia (2016)* report was written by the Office of the State Superintendent of Instruction to analyze the cost of early childhood services in the District of Columbia based on costing out each level of the Quality Rating and Improvement System for center-based and home-based programs. The cost model aimed to provide a more comprehensive understanding of the cost of high-quality early childhood education in the District of Columbia, and to inform policymaking and resource allocation decisions. Specifically, the goals motivating this cost model included the following:

- Identify the fiscal impact of the District of Columbia's Quality Rating and Improvement System standards on centers and homes.
- Identify key cost drivers.
- Explore the differential costs between programs that serve primarily (or exclusively) infants and toddlers and those that serve primarily (or exclusively) 3- and 4-year-old children.
- Create greater transparency about how the child care market operates.
- Provide information that would allow key community members and policymakers in the District of Columbia to test a range of alternative rate-setting and subsidy policy recommendations using the information gleaned from the cost model to understand the clear fiscal impact of each decision.

The District of Columbia tool modeled costs to address each of these goals by analyzing different scenarios using various data input values for the cost of delivering services, the assumed sources of revenue, and different types of programs of varying sizes. Several decisions about the cost model parameters influenced the input data built into the model, as follows:

Personnel Data for Center Programs

- Adult-child ratios were based on child care licensing regulations.
- Lead teacher wages increased as the programs' quality levels increased.
- Classroom staffing included increased coverage during opening and closing (based on a 10-hour day), daily breaks, and annual leave.
- As quality levels increased, additional staffing time was assumed to cover time for planning, reflective supervision, child assessment, and attending professional development activities.
- The model included one full-time director for each center. As with teacher compensation, director salaries increased as program quality increased.

Nonpersonnel Data for Center Programs

- All providers are participants in the Child and Adult Care Food Program.
- Among enrolled children, 80% are subsidy eligible and 20% are private pay.
- To allow the programs to break even, or realize positive net income, 98% enrollment is needed. This is higher than the 85% industry standard and would require that the District of Columbia contract with programs where the city either purchased slots or based the subsidy payments on enrollment versus daily attendance, which is current standard practice.
- Of the revenue generated from tuition, fees, and copayments, 3% would become bad debt.

Different scenarios were analyzed based on a range of options of program size, age mix of children served, and percent of children with disabilities/special needs. The total cost, total revenue, and net profit/loss ratios were computed for each scenario.



6.5 Equity Check

You will need to account for several equity considerations when aligning input data to parameters in your cost model. We encourage you to reflect on and discuss the following questions:

Are You Using Determined Data (Versus Estimated Data) Whenever Possible to Improve the Accuracy of Your Cost Model Outcomes?

- The accuracy of a cost model depends on the quality of the input data. Are you investing the time and effort to locate determined data for as many of your cost model inputs as possible?
- Are you identifying and using the most recent determined data in your model to improve the accuracy of your results?
- Are you ensuring that your input data reflects and includes traditionally marginalized communities?

When You Use Estimated Data, How Are You Ensuring the Reliability and Accuracy of the Assumptions Influencing Your Calculations?

- Can you triangulate, or use multiple sources of data, to get a more accurate understanding of a specific cost? For example, combining the use of national averages as proxies with input from community members and experts who can

provide estimates of local costs. Using varied sources—combining quantitative data with qualitative lived experience data—will allow for a fuller understanding.

- How will you consider context sensitivity—meaning, in what ways have you accounted for the limitations and effects of using broad population data versus community- or program-specific data?

Are You Ensuring That Community Members and Local Leaders Understand How a Cost Model Works and How Decisions Are Being Made to Inform the Local Cost Model?

- How will you keep community members informed and invite their participation throughout the entire cost model process, from planning to implementation to updating the cost model over time?
- How will you use community input to build, sense-check, and adjust the cost model and its output? (It's important that key individuals are involved from the beginning and not just responding to a cost model that's already been built).
- How will you ensure that you follow through on any commitments made to community members to use their feedback to inform the cost model?

- How will you document and communicate the sources and limitations of the input data used in the cost model?

How Will You Ensure That Your Cost Model Responds to Variation in the Cost of Labor, Materials, and Other Resources Based on Geographic Location, Cost of Living, or Other Factors?

- If you are asked to calculate true costs for a program or service, yet you know that significant variation exists in costs across neighborhoods and/or communities, how will you recognize these differences in your cost model?

6.6 Decision Point: What Input Data Will You Include in Your Cost Model?

Now take some time to focus on what input data you will include in your cost model by asking yourself these questions:

- **What is the service delivery model?** Are services provided in one setting or a range of learning settings, such as community-based organizations, public schools, and family child care homes?
- **What is the funding framework?** What funding streams provide revenue to support service delivery?
- **Will you need to address compensation parity?** How will you do so? Salary and benefit schedules for child and youth educators in nonschool settings often are built to be commensurate with local school district educators and vary as a factor of educator qualifications, role, and tenure.
- **Will you include licensing requirements?** These will vary but likely be major cost drivers (e.g., group sizes, ratios by child and youth age, staff qualification minimums, staffing requirements, space requirements).
- **How will you account for high-quality requirements?** These may include quality rating improvement standards, state early learning standards, and/or accreditation standards, such as those for the National Association for the Education of Young Children or National Association for Family Child Care.
- **Will you include universal or targeted income eligibility thresholds?**
- **What are the facilities requirements?**

6.7 Key Takeaways

Takeaway 1: Input data is the actual data used in the cost model's calculations. Cost models include two main types of input data—determined data and estimated data. **Determined data** (known data) is data that is known and fixed. It is not based on assumptions or estimates. **Estimated data** (unknown data) is data that is not known and must be estimated. It is based on assumptions or guesses when certain information is unavailable or missing.

Takeaway 2: Various factors influence the accuracy and reliability of input data in cost modeling:

- **Historical data:** The data used to build a cost model is often based on past expenditures or experiences, which can shape assumptions about future costs.
- **Geographic location:** The cost of labor, materials, and other resources can vary by geographic location, so input values about costs may vary based on where the program or service is implemented.
- **Contributor input:** Input from program staff, administrators, policymakers, and community members can influence decisions made about which data inputs are entered (or not entered) into a cost model.
- **External factors:** External factors such as changes in laws, regulations, or economic and political conditions can also influence the data inputs used in a cost model.

Takeaway 3: Examples of the most common input data entered into cost models include the following:

- **Staffing data:** This data includes specific adult-child ratios, as well as the salaries, level of education, and experience required for different staff positions.
- **Program data:** This includes factors such as the length of the program day, the age range of children and youth served, and the type and intensity of the education and/or interventions provided.
- **Facilities data:** This includes the cost of facilities including rent, utilities, insurance, and maintenance.
- **Cost of living data:** This means the cost of living in the geographic area where the program is located and includes the cost of housing, transportation, and other basic needs in a specific community or state.

Takeaway 4: To ensure the accuracy and validity of the input data used in your cost model, it is important to establish mechanisms for soliciting feedback from state and community members. This can be done in various ways, including holding focus groups, conducting surveys, and convening advisory committees.

6.8 Going Deeper: Learn More About Input Data in Cost Models

Your data inputs drive the outcomes for your cost model. Whether they are determined by an external organization or are estimated by your informed decisions, the data will breathe life into the tool. Some inputs hold more weight than others, impacting the overall costs to different degrees. The following sections explore a few inputs to consider as you choose what will work best for your community.

Livable Wage

You might wonder how to ensure that you are not maintaining the status quo of paying providers a nonlivable wage. Your goal is to make sure that you include not only the current wage scales, but also one (and sometimes more) that will reflect the value that providers bring to children and youth.

A reliable starting resource for developing a new salary and wage scale is the [MIT Living Wage Scale](#). Searchable by state and county, this scale provides the hourly rate workers must make to support themselves or themselves and their family. The hourly rates

increase based on a family’s composition, including adults and children. Here is where your estimated data input makes its entrance!

See figure 3 for an example of what the hourly rate by family composition looks like in Wayne County, MI.

As you can see, the living wage for one adult with no children (\$16.57) is very different from a family with two adults (one of whom works) and two children (\$37.15), which is more than double! If you’re trying to make an informed decision about what hourly rate you will need to start at, you are going to need to estimate what an average family composition is in Wayne County, MI. That will be your starting hourly rate for the most entry-level direct service provider and increases in the salary scale will be based on upward mobility, experience, and credentialing.

Full-Time Versus Part-Time Care

As the needs of families change, the demand for calculating costs outside of full-time care grows. How would you account for students who receive care full time versus part time? If this is a need you see in your community, you can include variables in your model to produce those costs as estimated input data. One way to do this is to add a variable that converts the days of care into percentages, with full-time care equating to 100%. To develop a part-time percentage, you will need to estimate how many total days of care would be considered part-time. Figure 4 on page 12 shows an example of what this would look like.

Figure 3: Living Wage in Wayne County, MI

	1 Adult				2 Adults (1 Working)			
	0 Children	1 Child	2 Children	3 Children	0 Children	1 Child	2 Children	3 Children
Living Wage	\$16.57	\$38.92	\$54.64	\$73.93	\$26.54	\$32.30	\$37.15	\$41.05
Poverty Wage	\$6.53	\$8.80	\$11.07	\$13.34	\$8.80	\$11.07	\$13.34	\$15.61
Minimum Wage	\$10.10	\$10.10	\$10.10	\$10.10	\$10.10	\$10.10	\$10.10	\$10.10

Figure 4: Full-Time Versus Part-Time Care

	Enter for Average Center (Required)
Number of days care is provided annually (full-time care)	261
Percent of days care is provided annually (full-time care)	100%
Number of days care is provided annually (part-time care)	130.5
Percent of days care is provided annually (part-time care)	50%

Quality Levels

More often than not, high-quality programs are a priority for communities. When you cost out programs at different quality levels, you must recognize the additional supports needed by sites and staff to attain high-quality levels of programming. These additional supports are where you come in with your estimated input data. For example, to go from the first quality level to the second, a provider would need to reduce the ratios of adults per child or youth ratios, so you would need to create a scenario that reflects the smaller ratio.

Priority Populations

To work toward more equitable services for children and youth, you may want to include young people who are farthest away from resources in your cost model to focus investments toward them. How would this affect your cost model?

Priority populations are any groups you or another community deems as needing additional services to work toward equity. You can tailor the additional costs to invest in the priority population, depending on what needs and barriers the population experiences.



Children’s Funding Project is a nonprofit social impact organization that helps communities, states, and Native nations expand equitable opportunities for children and youth through strategic public financing. [childrensfundingproject.org](https://www.childrensfundingproject.org)

