SOCIAL FINANCE

Social Finance Limited
+44 (0) 20 7770 6836

87 Vauxhall Walk, London SE11 5HJ
socialfinance.org.uk

CS Regional Demand Model: Model Functionality Guide

How the model works

What is the model and what is it not?

The web application enables users to use their region's historic SSDA903 data on children's placements to produce a forecast of future placement activity.

The model is better thought of as a scenario modelling tool than a predictive model. In a predictive model, historic data is used to estimate relationships between input variables and outcomes, and then new input data is used to infer the future outcome. In children's placements, this could be a model that uses data on children's needs, the number of referrals to CSC etc. to predict the number of children entering care in a period. This is not what the demand modelling tool does.

Instead, the model allows the user to project trends from historic data into the future and then adjust those trends to model different scenarios for how the future might unfold.

The data that the application uses is:

1. SSDA903 Episodes file
2. SSDA903 Header file
3. SSDA903 UASC file

How is the care system represented in the model?

The model presents a simplified version of the care system, with children considered to be in states based on age and placement type.

For age, children can be: 0, 1-4, 5-10, 11-15 or 16+

For placements, children can be in: Fostering, Residential, Supported or Other. Note that internally and externally commissioned placements are grouped in the model (see Creating a budget forecast section later in this doc). PLACE codes in the SSDA903 are mapped to these categories as follows:

Model state	PLACE codes
Fostering	U1, U2, U3, U4, U5, U6
Residential	K2, P3, R1, S1
Supported	H5, K3, P2

Other	All other codes

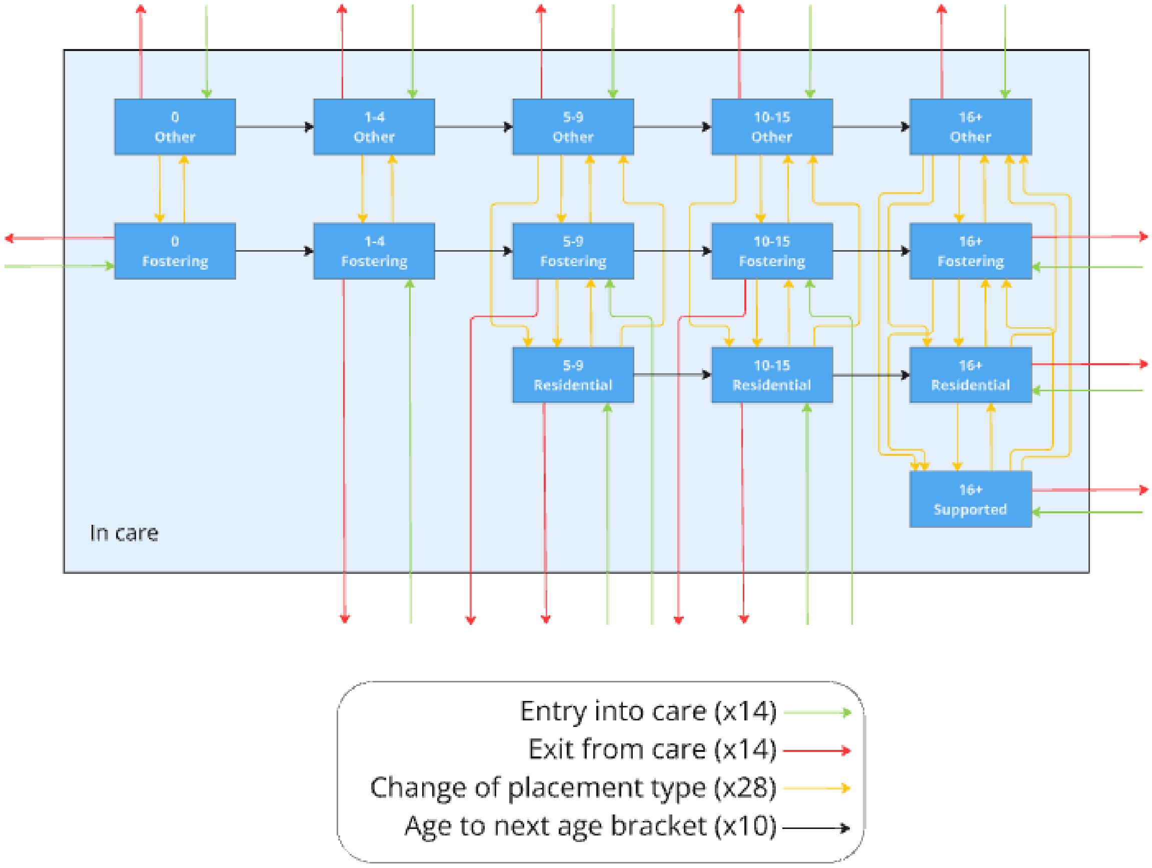
These states are currently fixed for users, but can be easily configured by maintainers of the application in the application code.

When the SSDA903 data is processed, children are assigned a state (e.g. “1-4 Fostering”) when their placement starts and they are considered to remain in that state until they either exit care or transition to another state, either by moving to another type of placement, or by ageing into the next age bracket.

Note that this means that if a child changes placement but does not change placement type, they do not transition in the model e.g. if a child in a fostering placement moves to another fostering placement, this is not considered a transition in the model. An impact of this is that the model is only representing the total number of a type of placement that are required at any one time to meet the demand of children in care, rather than the total number of placements that are being commissioned.

The ’Other’ group is primarily made up of children who are placed with parents, placed for adoption, in family centres and in secure children’s homes.

Figure 2 below shows a set of typical model states for a set of SSDA903 data and the transitions that can occur between them.



The rate at which these transitions occur is central to how the model works.

In brief, the model creates a forecast starting from a date the user selects by:

1. Calculating estimates for each of the transition rates using the historic SSDA903 data
2. Counting the number of children in each of the states (the state populations) at the beginning of the forecast
3. Stepping forward one day at a time, applying the relevant transition rates to each population

The following sections will explain these steps in more detail.

Transition rate calculation

The historic SSAD903 data loaded into the application by the admin user is used to calculate the 66 transition rates. The data is transformed into a daily view, and a count of the population of each state and the absolute number of each transition is made for each day. By dividing the number of transitions by the population count, a rate is derived for each day.

Next, an average rate is calculated for a ‘reference period’. The default reference period uses all of the years for which SSDA903 data is present e.g. if there are three returns for 2022, 2023 and 2024, the default reference period will be 01/04/2021 – 31/03/2024. The user can change the reference period in the application (see below).

For the entry rates, an average rate for the period is calculated by summing the population counts for each day and dividing by the number of days in the period.

For all other rates, a rate is calculated for each day in the reference period by dividing the number of transitions by the previous day’s population size. Then these daily rates are averaged to provide a rate for the period. Note that a daily rate is not calculated for days when the population is equal to 0.

The rates that are calculated this way are then used in the model forecast. As a result, it is important that the reference period and care population used for this calculation are as representative as possible of the future.

Whenever possible, it is better to derive rates from a large time period, as this will reduce the impact of anomalously low or high movements and produce a more representative average. The same is true for the care population. With a larger number of children in each model state, the rates calculated will be less prone to volatility due to random fluctuations.

Changing the inputs to rate calculations

The user can change the inputs to rate calculations in two ways:

1. Changing the reference period
2. Changing the care population

The reference period can be changed using the “Reference Start Date” and “Reference End Date” inputs on the “Model Base Forecast” page. In the first image below, the default reference period is shown. In the second image, the reference period has been changed by the user to a period when the care population was growing at a high rate. Consequently, the forecast populations to the right of the graph are also growing. This is due to the change in the transition rates calculated in the reference period.

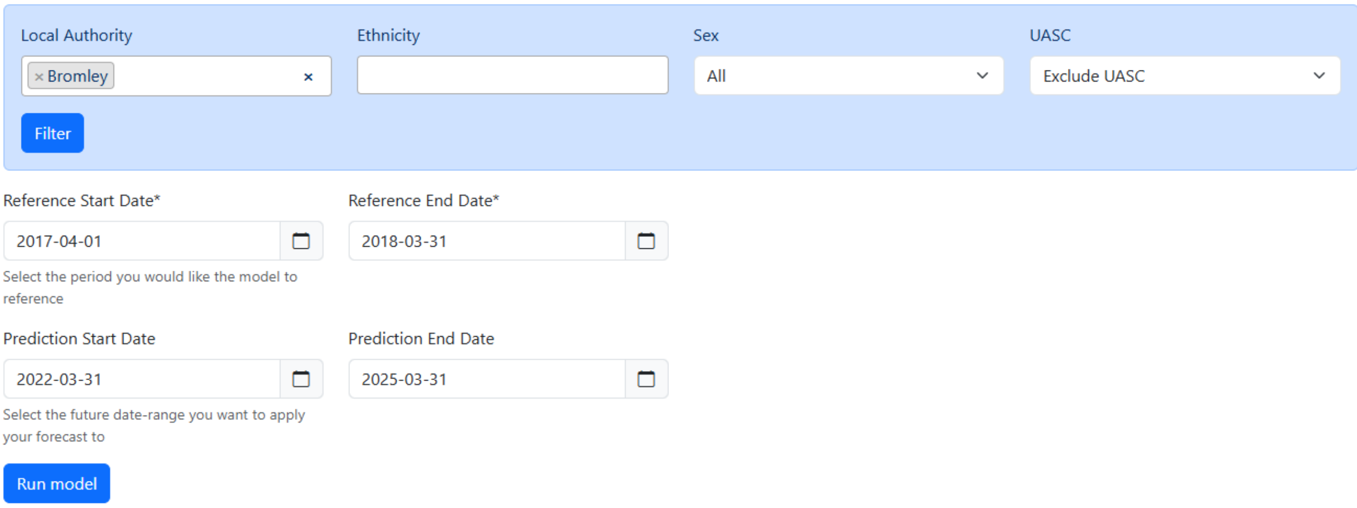
Figure 3a & b: Two different reference periods influencing the model forecast; in the first image, the reference period inputs are highlighted in red

a. default reference period

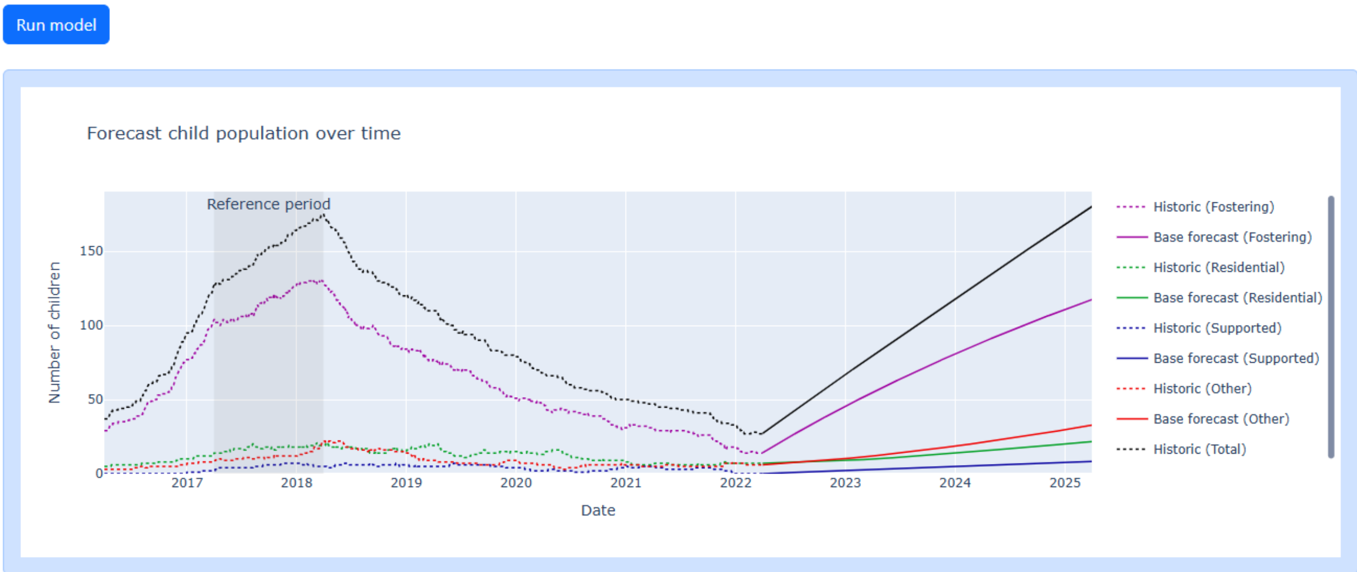
b. user-selected reference period

The transition rates can also be influenced by the care population selected. This is managed by using the filters at the top of the Model Base Forecast page. In Figures 3c and 3d below, the same reference period as in Figure 3b above has been created using a different population, filtered as shown to include only one of the LAs in the region and to exclude children with a UASC status.

c. user-filtered care population



d. model using user-filtered care population and user-selected reference period



A note on ageing rates

The rates at which children age from one age bracket to the next are calculated like any other, by using the real rates at which children transitioned between these brackets in the reference period. The factor that will influence this rate is the distribution of the ages of children in each age bracket during the reference period e.g. if most of the children in the 10-15 age bracket were 14 or 15, the rate of ageing out of that age bracket will be higher than if most of the children were 10 or 11. It is this distribution that the rate really represents in the model.

Forecast calculation

With the transition rates calculated, the model needs just one more input to be able to generate a forecast. This is the dates on which the forecast will start and end.

Forecast start and end date

The forecast start date determines the state population sizes at the beginning of the forecast. The default for this input is the last day in the historic data, as this is the last date for which those population sizes are known. The user can set this date to be any date for which the population sizes are known.

The default for the forecast end date is two years after the forecast start date. The user can set this date to any date after the forecast start date. However, we advise not creating very long forecasts as the uncertainty around the forecast is expected to grow as the time from the forecast start date increases.

Forecasting changes in population size

With the transition rates and the starting population, the model can produce a forecast. It does this by stepping forward one day at a time and applying the transition rates to the populations. For the entry rates, this is a simple as adding the average rate to the population each day e.g. a rate of 0.5 means that 1 child is added every 2 days. For the other rates, the population in the state is multiplied by the average rate e.g. if the exit rate for the 16-18 Fostering state was 0.25, a quarter of the children in that state will exit care each day.

To understand why this is the method chosen, the following section examine different ways we can create a forecast using the transition rates derived from the historic data.

For children that already have a state in the model, for each day the event that happens (exiting care, moving to another state, or making no change) can be thought of as a result drawn from a multinomial distribution, where the probabilities of each result are defined by the transition rates.

For children that don’t have a state in the model, the children who are entering care, their probability of entering into a state can be thought of as a result drawn from a poisson distribution, with the probability defined by the entry rate for that state.

One way to get a forecast would be to treat each child in the model as an individual entity and each day select from the transitions available (including remaining in the same state) based on their probabilities of occurring, essentially drawing from the multinomial distribution for their state. If we had 10 children in a state, we would draw from the distribution 10 times and assign an event drawn (exit care, move to another state etc.) to each child. Then we would also draw from the poisson distribution to see if any children entered that state. If we did this for each child and state every day for two years, we would produce a simulated forecast based on the probabilities of transitions occurring that we had observed in the historic data.

We would want to run that simulation many times to reduce the effect of outlier scenarios that might occur with only a single run, such as the unlikely possibility that every child in care exited on day one. By running the simulation many times, we expect the average of the results to be similar to the effect of applying the average of the distributions each day, instead of sampling from the distribution randomly.

In this application, instead of running simulations, we use a mathematical approach to derive the result of averaging over a very large number of simulations. Instead of treating each child individually, we take the population of each state on each day and apply the mean of each transition rate to that state to determine the number of children who make each type of transition.

For the multinomial distributions (exits and transitions between states, including ageing) in practice this is achieved through calculating the product of two matrices, one with the populations in each state, and another with the average transition rates for all transitions. This is because the mean of the multinomial distribution for any given result (any given transition) is:

$$np_i$$

where n is the number of events (in this case the state population) and p_i is the probability of the result (the transition rate).

For the poisson distributions (entries), it is achieved by simply adding the average entry rate to the population each day. This is because the mean of the Poisson distribution for an event is λ, where λ is the number of events expected in a given time interval i.e. the average rate.

This creates a very smooth trajectory in the population sizes that lacks any of the stochastic variation that would be seen in a simulation. But the end result should be representative of what the population sizes would be under the observed rates of transition.

The method chosen also allows us to explicitly calculate the variance of the outcomes of the transitions, as these are known properties of each probability distribution.

In the case of the multinomial distributions, this is calculated for each outcome as:

$$np_i(1 - p_i)$$

while for the binomial distribution, it is λ as before.

We calculate the variance across the transitions for each day and cumulatively sum them as time goes on to reflect the increasing uncertainty in events as we take further steps away from the start of the forecast.

We represent the variance on the forecast graphs as a confidence interval of two standard deviations above and below the mean, calculated as:

$$2 * \sqrt{variance}$$

If we were to run our simulation many times, the area within the confidence intervals shows how far away from the mean result ~95% of the simulated forecasts would be expected to fall.

Adjusting transition rates

Using a reference period to define transition rates will only get us so far in being able to craft a flexible set of scenarios to represent the future. The application also allows users to change most of the rates in the model to fine-tune expectations or to explicitly compare scenarios based on different sets of assumptions.

This can be done in the Model Adjusted Forecast page. On this page, the rates that can be adjusted are split into three tables: Children entering care, Children leaving care and Rate of change between placement types.

To make adjustments, you can view the table by:

Children entering care | Children leaving care | Rate of change between placement types

Once you have adjusted the table you will be able to see the base forecast alongside your adjusted forecast to compare. You can continue to edit the table across each of the views.

Rates in these tables include any previous adjustments you might have made.

Rates of ageing between age brackets are currently unavailable for adjustment.

The rates shown are the average rates used in the forecast calculations, represented as decimals. All of the rates will be shown, even when they are equal to 0. When a rate is equal to 0, this means that no such transition occurred in the reference period chosen. This could be due to one of two things:

- the reference period is too short and did not capture any transitions
- the population in the relevant state is too small, or completely absent

E.g. if the rate of 5-10 Fostering -> 5-10 Residential is 0, it may be because there were no (or very few) 5-10 year olds in fostering placements in your reference data.

When you edit rates, you are given the option to multiply or add to the existing rate.

Adjust entry rates

You may want to edit the number of children per year entering care to give a more accurate picture of future forecasting or simply to explore what changes you may expect to see.

Rates are applied to the daily population of children. For example, a rate of 0.5 to 10-16 Fostering would mean that every two days a child would enter this placement.

Rate multiplication multiplies the initial base rate produced by the model parameters you selected.

Age group	Placement	Base entry rate	Multiply Rate	Add to Rate
5 to 10	**Fostering**	0.0918	Multiply Rate	Add to Rate
	Other	0.011	Multiply Rate	Add to Rate
	Residential	0.0123	Multiply Rate	Add to Rate
10 to 16	**Fostering**	0.1623	Multiply Rate	Add to Rate
	Other	0.0144	Multiply Rate	Add to Rate
	Residential	0.0274	Multiply Rate	Add to Rate

Users can multiply any non-zero rate by any positive number. Users can add to entry rates by any positive or negative number, but the entry rates have a floor of 0. Users can add to exit and placement transition rates by any positive or negative number but the rates have a floor of 0 and a ceiling of 1.

When a user adjusts a rate, it will replace the rate derived from the reference period and will be used in the forecast calculation, producing an 'adjusted forecast' alongside the 'base forecast' (which continues to use the rates derived from the reference period).

However, adjustments to exit and placement transition rates will only be applied as specified if they do not cause the sum of all the rates for that model state to sum to more than 1. In most cases, if an adjustment changes the sum of all rates in this way, the rate that will be adjusted 'behind the scenes' is the rate at which a child makes no transition and remains in the same state.

Example: for the 5-10 fostering state, we have the following rates:

5-10 fostering -> exit care = 0.1

5-10 fostering -> 5-10 residential = 0.1

5-10 fostering -> 5-10 other = 0.15

5-10 fostering -> 10-16 fostering = 0.05

The sum of these rates is 0.4, which means that the rate at which children will remain in this state is equal to 0.6. If the exit rate is increased to 0.3, this will be accepted and the effect will be a reduction in the rate at which children remain in the state to 0.4, with the other rates remaining unchanged.

However, if the exit rate is increased to 0.8, the sum of rates is now > 1 even if we assume a rate of remaining in the state of 0. In this scenario, the adjusted rate of 0.8 will be accepted, but the three other rates associated with the state (the transitions to residential and other and the ageing rate) will be scaled down to sum to 0.2 in order to accommodate the adjustment.

If multiple adjustments are made that cause the sum of rates to equal > 1 and it is not possible to scale back down to 1 by altering only the rates that have not been adjusted, the adjusted rates themselves will also be scaled down.

In our previous example, let's says adjustments were made as following:

5-10 fostering -> exit care = ↑ **0.6**

5-10 fostering -> 5-10 residential = ↑ **0.6**

5-10 fostering -> 5-10 other = ↑ **0.6**

5-10 fostering -> 10-16 fostering = 0.05

The adjustments cannot be accepted, as they sum to more than 1. In this instance, the rates that were not adjusted (ageing and remaining in state) will be reduced to 0, and the three rates that were adjusted will each be scaled down to 0.3333 to maintain the overall sum of rates for the model state at 1.

In practice, we don't recommend setting such drastically high rates for these transitions if the reference data does not suggest that they are realistic. Setting rates like in the example above would mean that on day one of the forecast, all of the children currently in 5-10 fostering would leave fostering and that any subsequent entrants to this state would all also leave the following day. It's worth bearing in mind that these are daily rates and that even a rate of 0.1 means that, on average, a placement for that state will only last for 10 days.

These scaling adjustments will be reflected if you return to the tables on the Model Adjusted Forecast page, where you can always see the actual rate the model is using for the adjusted forecast after any user adjustments and automatic scaling have been applied.

Creating a budget forecast

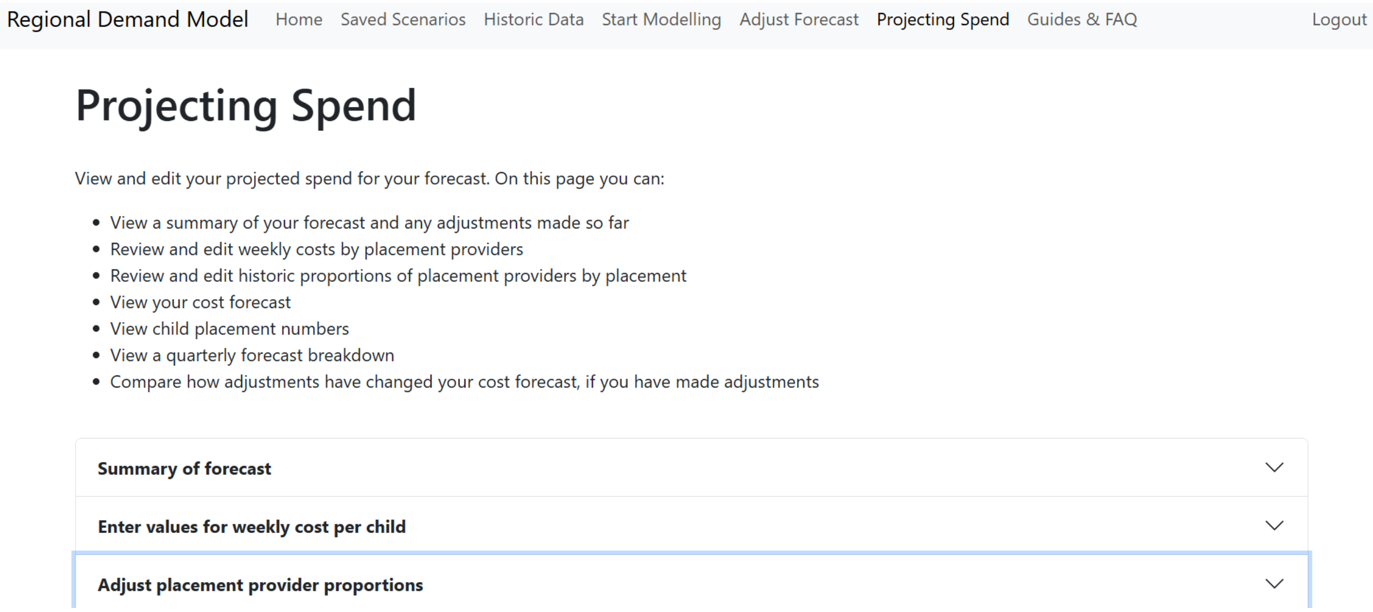
With a forecast, the final element of the application enables a user to transform the model forecast into a more detailed view of future placements and translate this into a cost forecast.

There are two key pieces of functionality here:

- providing more detail to placement types in forecast
- converting into a cost

More detailed placement types

At this point, the application enables users to provide a more detailed view of how different placement types will be commissioned. This functionality is in the "Adjust placement provider proportions" drop-down tab on the "Project spend" page.



In this drop-down, the user can break their model forecast into more detailed placement categories as follows:

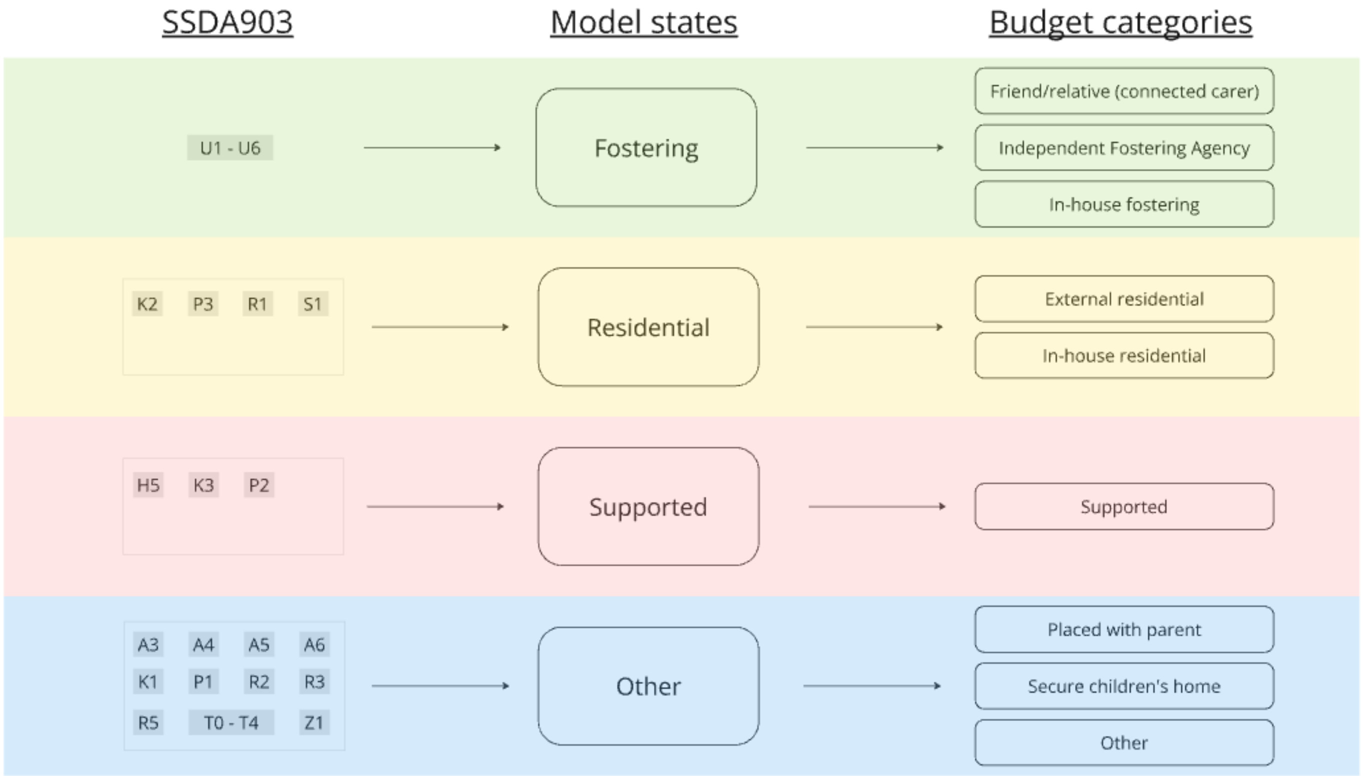
- fostering -> fostering (friend/relative), fostering (IFA), fostering (in-house)

- residential -> residential (external), residential (in-house)
- other -> placed with family, secure home, other

Supported offers no additional sub-types, as there are no sub-types for this within the SSDA903. A full diagram of how this relates to SSDA903 placement types is shown in Figure 4 overleaf.

The model forecast is split into these sub-types based on proportions for each sub-type e.g. if the proportions for the residential sub-types are each 0.5, half of the model forecast population for residential placements will be assigned to external residential placements and half to in-house residential placements.

Figure 4: How placements are categorised across the model



The default proportions used are derived from the user-selected reference period of the SSDA903 data. Over the period selected, the application calculates the proportion of all residential placements that were in-house and external and carries these proportions through to the prediction. If the reference period is changed, these proportions may also change.

In the drop-down tab, the user has the option to over-write these proportions with values of their choice. Similarly to editing transition rates, if the user edits proportions such that they do not equal 1 overall for the model placement category, the proportions will be scaled to ensure that the totals equal 1.

These proportions (either default or user-defined) are then used in the subsequent graphs and tables showing a more detailed breakdown of the model forecast by placement type.

Converting into cost

The drop-down option “Enter values for weekly cost per child” is hopefully self-explanatory. This allows the user to enter a cost for each placement sub-type which will be used to convert the model forecast (in number of children) into a forecast for daily cost.

An option can be used to add inflation to this cost, which will increase the cost of placements by the value selected with each successive year in the forecast.