All U.S. States and Territories receive money from the federal government’s Child Care and Development Fund (CCDF) to operate child care subsidy programs, but each State and Territory has broad discretion in setting specific policies for their individual programs. Which families are eligible to receive subsidies, how much they must pay out-of-pocket, and how much providers will be paid are all policies that vary significantly by state and even by county and locality. This raises an important question for people interested in CCDF-funded programs: What do these policy differences mean for the parents and children served by subsidized child care and for the people providing the child care services?

This brief presents an example of how to answer these questions—by linking case-level information on subsidized families and children with detailed policy information from the federally-funded CCDF Policies Database. We illustrate this approach using a specific policy question—whether State/Territory policy choices appear related to the extent to which children with CCDF subsidies receive care from legally unregulated providers. However, our focus is not on the research results—which are intended to be illustrative rather than definitive—but rather on the process for this type of analysis. Linking administrative data with policy data has great potential for exploring the impacts of policy variations, and the CCDF Policies Database provides a rich source of policy data to use in these types of analyses.
Background

CCDF programs must operate within the overall policies established by federal law. Some of the key requirements are that families must have incomes under 85 percent of state median income (SMI), parents must generally be working or in school (or in some cases be looking for work), and the children being cared for must be either under age 13 or have special needs.\(^1\) However, many other policies—including the use of income limits lower than 85 percent of SMI, how much providers are reimbursed, how out of pocket costs are computed, and so on—are set by each place operating the program—each State, the District of Columbia, and each of five Territories.\(^2\)

When thinking about a single policy choice, how that choice will affect families or providers may seem obvious. For example, by definition, the lower the income limit for initial eligibility set by a State/Territory, the lower the incomes of families who are served by the program. However, the range of policies and their potential for interaction—with other policies and with characteristics of the population and the economy in each State/Territory—means that the results of policy choices are not always straightforward to observe. In some cases, researchers may want to conduct a more detailed analysis, looking at individual-level data of some sort and combining that with data on policies, across places and possibly also across time. In technical terms, the combination of policy data with individual-level data is a type of data linkage (also referred to as merging or matching).

A few examples of questions that might be addressed by linking individual-level data with CCDF policy data include:

- How are different copayment structures (the rules determining how much families have to pay out of pocket) related to the income levels of families receiving subsidies, when separated from other policies such as income eligibility limits?

- How much do differences in maximum reimbursement rates (the highest amounts that will be paid to providers) appear to be related to the characteristics of the providers who choose to participate?

- Do CCDF policy differences appear to be related to mothers’ employment rates?

Depending on the specific policy question, the individual-level data might be from national-level survey data or from program administrative data. For many different questions, policy information across time and across places could come from the CCDF Policies Database project, funded by the Administration for Children and Families, Office of Planning, Research and

---

\(^{1}\) Care can also be provided for children receiving child protective services.

\(^{2}\) “States/Territories” is used throughout the brief to refer to the 50 States, the District of Columbia, American Samoa, the Commonwealth of the Northern Mariana Islands, Guam, Puerto Rico, and the Virgin Islands. While not covered in this brief, the CCDF program also provides funding for the Tribes.
Evaluation (ACF/OPRE). Box 1 provides additional information about the CCDF Policies Database.

BOX 1

KEY FACTS ABOUT THE CCDF POLICIES DATABASE

- Over 1,800 policy variables
- Information starts in 2009
- New information released each year; 2014 information released winter 2015
- Information on every state, the District of Columbia, Puerto Rico, the Virgin Islands, American Samoa, Guam, and the Commonwealth of the Northern Mariana Islands
- Policy information is initially coded from caseworker manuals, then reviewed by State/Territory staff
- Captures policies that have been implemented (not planned changes)
- Picks up changes in policies and codes when they occur
- Captures some sub-state variation

This brief illustrates how individual-level data can be analyzed in concert with policy data from the CCDF Policies Database, focusing on the following policy question:

- Do State and Territory policy choices affect the use of legally unregulated care?

Results presented later in the brief reveal that several policies are related to the extent to which subsidized children are in legally unregulated care. For example, there appeared to be more use of legally unregulated care when providers were allowed to charge families the difference between the provider’s usual rate and the maximum reimbursement rate paid to the provider by the subsidy program. No final policy conclusions should be drawn from this analysis, as some potentially important factors were not included. However, even this illustrative analysis suggests policies are related to outcomes, and combining individual data and policy data can help uncover those connections.
Legally Unregulated Care

Providers are sometimes referred to as “legally unregulated” when they provide care outside the scope of a State’s/Territory’s broader child care regulations. However, they must still meet certain standards in order to be paid through the CCDF system. Some of the “legally unregulated” care subsidized through the CCDF program is care in the child’s home; however, some family child care homes, group homes, or child care centers may be “legally unregulated”, depending on a particular State’s or Territory’s licensing rules. The extent of legally unregulated care is of great interest to policymakers and child care researchers due to potential impacts on child care quality and stability.

The portion of CCDF-subsidized children cared for by legally unregulated providers has fallen substantially since the early years of the CCDF program; 26 percent of children with CCDF-subsidized care in federal fiscal year (FFY) 2000 were in legally unregulated care, compared with 13 percent in the preliminary program data for FFY 2014. However, there is wide variation in the prevalence of legally unregulated care for CCDF-subsidized children across the States/Territories. In FFY 2014, the portion of CCDF-subsidized children in legally unregulated care was under 5 percent in 20 States/Territories, but it was 20 percent or higher in 12 States/Territories (figure 1).

Differences in levels of legally unregulated child care used by CCDF recipients across States/Territories could be due to numerous factors, including differences in child care markets, overall economic circumstances, characteristics and preferences of subsidized families, and State/Territory rules regarding who may operate as a legally unregulated child care provider. States’/Territories’ CCDF policies related to copayments, provider reimbursements, and provider requirements could also play a role by affecting the characteristics of providers who are participating in the CCDF system and/or by affecting the choices made by parents.

---

3 The statutory language for the Child Care and Development Block Grant Act of 2014 (the most recent reauthorization of the block grant), refers to legally unregulated providers as license-exempt providers. Under the new law, all States/Territories will have to require CCDF providers, including legally unregulated providers, to meet health and safety standards that include background checks, annual inspections, and pre-service/orientation and ongoing trainings. States/Territories will have the option to exempt relative providers from the requirements. Additional information on CCDBG reauthorization is available from the Office of Child Care at [http://www.acf.hhs.gov/programs/occ/ccdf-reauthorization](http://www.acf.hhs.gov/programs/occ/ccdf-reauthorization).

4 Child care centers are usually non-residential child care facilities that typically care for a larger number of children at one time than residentially-based facilities. Family child care homes are residential child care programs. Group child care homes are similar to family child care homes but are usually allowed to care for more children at the same time than a family child care home.

To try to understand the role of subsidy policies, one could look at specific policies one-by-one, comparing State/Territory choices on a particular policy to the State/Territory data on the percent of subsidized children in legally unregulated care. However, the number of policies and the complex interactions among policies, family characteristics, and provider characteristics would make it difficult to draw out relationships using this approach.

A more comprehensive approach is to combine individual-level data on CCDF-subsidized children and their child care providers from the program’s administrative data with data on a range of policy options, and to analyze those data with a statistical method that can tease out the individual impact of just one characteristic or policy, conceptually holding the others constant. The rest of this brief walks through a simple approach to that type of analysis. The
steps that are involved are: selecting the administrative data, selecting the policy data, combining the files, and conducting the statistical analysis.

Step 1: Selecting the Administrative Data

For cross-state analysis of CCDF-subsidized children and families, the best data source is the CCDF “801” data. For each fiscal year, States/Territories submit individual-level data on the families and children served by their CCDF programs, with all the States/Territories required to submit the same items of information. Many places submit information on all of their cases, whereas others submit a large sample. The data include information about the families, the children, and the types and amounts of child care provided under CCDF. Each observation in the file includes information on one family and their children, for a particular month. The full documentation of the 801 data, including a description of each variable and how it is coded, is available on the website of the Child Care and Early Education Research Connections project.6

For this brief’s policy question, the two most important pieces of information to know about each subsidized child are the child’s State/Territory of residence and whether the care the child received was legally unregulated. However, several other characteristics of families and children could also be important. For the example analysis, we also included the following information from the 801 data:

- One-parent or two-parent family
- Poverty status
- Whether subsidized care is being provided for employment or for school/training
- Age of the child
- Race/ethnicity of the child
- Gender of the child

Any of these characteristics could affect a family’s likelihood of choosing legally unregulated care instead of other types of care. For example, analysis of the FFY 2011 801 data shows that children ages 3 and 4 are least likely to be in legally unregulated care, and 12-year-olds are the most likely. Therefore, even without differences in policies, areas that serve a higher portion of older children might show more use of legally unregulated care.

6 Data files (in multiple formats, including SAS, Stata, R, and ASCII) and codebooks for the CCDF Administrative Data (801 data) are available from Research Connections at http://www.researchconnections.org/childcare/series/215/studies/35293?q=801&type=Data+Sets &paging.startRow=1
Step 2: Selecting the Policy Data

For this illustrative analysis, we focused only on policies available in the CCDF Policies Database (box 2 provides information about the policy areas covered by the CCDF Policies Database). Four policies available in the Database were identified as having a plausible connection to the incidence of legally unregulated CCDF-funded care.

- **Whether a subsidy can be paid to a provider living in the child’s home:** As of October 2014, 31 States/Territories allowed CCDF subsidy payments to a provider who lives in the child’s home. Since that type of provider is generally legally unregulated, families in these States/Territories might be more likely to be using care that is legally unregulated. (This policy variable is not intended to capture whether in-home care (care provided in the child’s home) is allowed; it only captures whether the State/Territory will pay a provider who lives in the home with the child.)

- **Levels of reimbursement rates relative to providers’ regular charges:** States/Territories are required to conduct periodic market-rate surveys to inform the setting of reimbursement rates for CCDF providers. However, in many places, the reimbursement rates are lower than what is usually charged by many providers. This could lead to more families choosing lower-cost care, and could be related to a higher portion of children being in legally unregulated care.

- **If the copayment amounts depend on the choice of child care arrangement:** In most States/Territories, a family’s required copayment (what the family must pay out-of-pocket), is not affected by the family’s choice of child care arrangement. Copayments are usually based on a family’s size and income level, and sometimes based on the number of children receiving care or the amount of time spent in care, but do not usually vary by provider type or cost. In other words, if there are two families with identical characteristics, but one chooses a legally unregulated provider and the other chooses a regulated child care center (which likely charges a higher price for care), in most States/Territories those families will owe the same copayment. However, in seven States/Territories (as of October 2014), the copayment was higher when a family chose higher-cost care. For example, for families with income in a certain range, the copayment may be computed as 10 percent of the maximum reimbursement rate for the selected type of care. Since reimbursement rates vary by type of care, the copayment would be lower if the family selected a lower-cost, in-home provider (who might be legally unregulated) instead of a higher-cost, child care center. That scenario could induce more families to choose legally unregulated care.

- **If providers are allowed to charge families the difference between their regular price and the CCDF program’s reimbursement rate:** In some cases, the care that a family wants to use may cost more than the State’s/Territory’s maximum reimbursement rates. In that scenario, many States/Territories—39 as of October 2014—allowed the
provider to charge the family the difference. The rationale for this policy is that it allows families greater choice. Other States/Territories, in order to limit families’ out-of-pocket expenses, do not allow providers to make those charges. In those places, the providers may choose to care for the children at the rate that will be paid by the CCDF program or they may decline to provide care for children with CCDF subsidies. Policies that allow providers to apply charges above the copayment schedule may be associated with families being less likely to use more-expensive care, and more likely to be in legally unregulated care.

Several technical issues arise in selecting and using the data for this type of analysis. One technical challenge is that in most cases, some manipulation or calculation is needed to convert the information in one or more variables from the CCDF Policies Database into a variable that can be used for quantitative analysis. That was the case with each of the four policies above. For three of the policies, a minimal amount of work was required. For example, to develop a variable telling if the State/Territory ever allows people living in the child’s home to provide subsidized care, it was necessary to combine information from four detailed Database variables that each focus on a different type of in-home resident—relative or non-relative as well as included or not included in the CCDF assistance unit. Specifically, if any of those four variables was coded to say that the person could provide care under CCDF, then the variable for the analysis was coded as “yes”.

The policy concept that required the most work in order to create an analysis variable from the Database information was the level of reimbursement rates relative to unsubsidized costs. States/Territories have dozens of different reimbursement rates, varying by types of providers, ages of the children being cared for, amount of care (full-time or part-time), when the care is being provided, and often by sub-area of the State. Also, what is important for this analysis is not the specific dollar amounts of the maximum reimbursement rates in a specific State/Territory, but where they fall relative to the costs of child care in that State/Territory. To create a variable that could be used within the analysis, we focused on 4-year olds receiving full-time care in child care centers. We then compared the highest reimbursement rate in the State/Territory for that type of care to the average cost of care for an urban 4-year-old according to the National Association of Child Care Resource and Referral Agencies.\(^7\) Across the States/Territories, the median ratio between the maximum reimbursement rate and the average cost of care was 93 percent; the range in the values was 60 percent to 128 percent. (Of course, this single simple measure does not fully capture State/Territory variations in the relationship of provider reimbursements to unsubsidized child care costs; those relationships might vary by age of child, by provider type, and by sub-state area.)

Policy Information Included in the CCDF Policies Database

- Basic eligibility criteria—age requirements for children, parent/guardian activity requirements (including job-search eligibility policies)
- Income limits—initial and continuing limits for each family size
- How family units are defined for determining eligibility
- How income is defined for determining eligibility and copayments
- How copayments are computed—general procedures and detailed income categories and copayment levels
- Variations in copayments for special cases—more than one child, part-time care, groups exempted from copayments, etc.
- Copayment administration—who collects copayments, policies for absences, etc.
- Application and verification procedures
- Other administrative procedures—appeals, terms of authorization, etc.
- Priority policies and waiting list policies
- Certification periods and procedures for reauthorization
- Requirements for reporting changes during a certification period
- Provider reimbursement rates—including variations by provider type and other variations (e.g., quality levels)
- Requirements for legally unregulated providers, including health and safety requirements
- Other provider policies—who can provide care, on-site visits, termination, etc.

Different kinds of manipulations or calculations would be required for different kinds of policy variables. For example, to use information on level of copayments in a state in a quantitative analysis, the researcher would have to make choices to simplify the detail of copayment policies into one or more specific variables. One possibility would be to use one of
the columns of information about copayments in the CCDF Policies Database’s Book of Tables—which shows monthly copayments for various prototypical family situations and income levels.

Another kind of manipulation is needed for cases in which States/Territories take a wide range of options to a particular policy question—such as how to define the assistance unit in complex household situations—and the options are neither numeric nor ordered in any way. In those types of cases, researchers must review the options (including nuances coded in “notes” fields in the Database and in footnotes in the project’s published tables) to determine how best to adapt the information for use in quantitative analysis.

A second technical issue is raised by the fact that policies change over time. For this analysis, we used the policy data for all the States as of October 1, 2010, since we were using FFY 2011 801 data, and October 2010 is the first month of FFY 2011. A more complex analysis could have used policies from different months of the year; the CCDF Policies Database includes information on policy changes and when they occurred.

Finally, as with any data file, the CCDF Policies Database includes some missing data, due to the fact that some information could not be located in policy materials or obtained from State/Territory contacts. When there is missing data for a policy variable of interest, researchers must choose to either drop the variable, drop the State/Territory with the missing data, or locate a value for the missing policy data through other sources.

Step 3: Combining the Data

To allow sophisticated quantitative analysis, it is necessary to augment the information about each child that is available in the 801 data (State/Territory of residence, type of child care, and the other demographic information mentioned earlier in the brief) with State/Territory-level policy information. In this analysis, for every child in a particular State/Territory, the same policy information from the CCDF Policies Database was appended.

In more complex analyses, different information might be appended for children of different ages, using different types of child care, and so on. For policies that vary by sub-state area, the policy that is applicable to a child’s or family’s specific location could be appended—but only if the child- or family-level data includes detailed geographic information. As mentioned above, policy changes during a particular year could be included; specifically, the policies in place in the exact month of a particular 801-data observation could be merged together with that child’s information. Alternatively, the policies appended to the child’s record could be the policies in place in the month that the child care assistance began (which is also recorded in the 801 data).

8 The public-use 801 data do not identify a family’s specific location within a state, although future revisions to the 801 data could potentially include this information.
From a technical standpoint, the merging can be conducted using whatever statistical software package is being used for the analysis. Each package has some type of command to merge two data files together according to one or more matching variables. In this example analysis, the only matching variable needed was the State/Territory variable (figure 2).

FIGURE 2
Example of Computer Code Merging Child-Level 801 Data and State-Level Policy Data Using the SAS Software System⁹

Step 4: Conducting the Analysis

Before conducting the quantitative analysis, the researcher may want to limit the group of children or families in some way. For this analysis, we chose to exclude children living in the Territories, since child care markets in the Territories may differ substantially from those in the 50 States and DC. We also excluded children age 13 and older, since the process for choosing a provider may be quite different for the parents of teenagers with special needs.

Focusing on the remaining children, we used statistical software to estimate a “probit” equation. A probit is a type of regression that can be used when trying to explain an outcome that has only two possible values—in this case, whether the child is in regulated care or legally-unrelated care. The analysis looks at that outcome variable relative to the other variables that the researcher thinks may be related to the outcome in some way. The results of the probit can indicate the extent to which a particular variable is associated with a higher or lower incidence of legally unregulated care, holding constant all the other variables in the analysis—including the child and family characteristics and the other policy variables. Specifically, the results of the probit model provide an estimate of the probability of a “yes” outcome for any combination of characteristics.

Two caveats are important to note. First, the model does not indicate whether a particular policy variable causes the outcome, only that it is related to the outcome. Second, there are certainly other factors that could be related to the use of legally unregulated care that are not

---

⁹ In this example, the researchers had previously created a SAS data file named “CCDFdata” with the desired 801 data with children as the unit of analysis, and had created a SAS data file named “DBData” with one record per State, including the desired policy variables.
included in this example analysis, such as economic circumstances, the characteristics of the overall child care market in the area where the child lives, and the State’s/Territory’s rules for who may operate as a legally unregulated provider. Those factors could vary not only across States/Territories but also within States/Territories; for example, regulated care might be less available in more rural areas compared with urban areas.

With those caveats in mind, the analysis suggests that each of the four policy variables is related to the likelihood that CCDF recipients will use legally unregulated care, and that all of those relationships are statistically significant (not just due to chance caused by the particular samples drawn in the States/Territories that do not provide data on the full universe of subsidized children). The most intuitive way to understand the results is to calculate the model’s estimates of the probability of being in legally-unrelated care for a child with a certain set of characteristics, starting with one set of policies and then varying the policies. For this example, we look at the probabilities for a white boy in the age range of 8 to 10, in a family with an unmarried mother who is employed, and with family income below the poverty guideline.

We start by assuming that the following policies are in place for the base case: individuals living in the home with the child are not allowed to provide care, the maximum reimbursement rate equals 90 percent of the average cost of center-based care, the copayment does not vary with the family’s child care choice, and providers cannot charge families an amount beyond the copayment when their full price exceeds the reimbursement rate. With those policies in place, the probit equation says there is an 11.6 percent chance that the child will be in legally unregulated care (table 1).

Changing any one of the policies results in a statistically significant change to the probability that the child is in legally-unrelated care, as follows:

- If individuals living in the home with the child are allowed to provide care, but all the other policies stay as in the base case, the probability that the child is in legally unregulated care increases from 11.6 to 18.0 percent.

- If the maximum reimbursement rate equals 65 percent of the average cost of unsubsidized care, and all the other policies stay as in the base case, the probability increases from 11.6 to 12.7 percent.

- If the copayment computation changes so that the copay is higher when the parent chooses higher-cost care, but all other policies stay as in the base case, the probability increases from 11.6 to 20.0 percent.

- If providers are allowed to charge parents for the portion of their regular fee that exceeds the maximum reimbursement rate, but all other policies stay as in the base case, the probability of being in legally unregulated care increases from 11.6 to 25.1 percent – more than double the probability in the base case.
**FIGURE 3**
Estimated Probability of Being in Legally Unregulated Care (Illustrative Analysis)

<table>
<thead>
<tr>
<th>Are in-home providers ever allowed?</th>
<th>Base case</th>
<th>Policy Test 1</th>
<th>Policy Test 2</th>
<th>Policy Test 3</th>
<th>Policy Test 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Reimbursement rate as a percent of the average cost of care?</td>
<td>90%</td>
<td>90%</td>
<td>65%</td>
<td>90%</td>
<td>90%</td>
</tr>
<tr>
<td>Does the copay vary with the choice of care?</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Can provider charge the family an amount beyond the copay if the full price exceeds the reimbursement rate?</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

**Probability that the child is in legally-unregulated care:**

<table>
<thead>
<tr>
<th></th>
<th>0%</th>
<th>5%</th>
<th>10%</th>
<th>15%</th>
<th>20%</th>
<th>25%</th>
<th>30%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base case</td>
<td>11.6%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Policy Test 1</td>
<td>18.0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Policy Test 2</td>
<td>12.7%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Policy Test 3</td>
<td>20.0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Policy Test 4</td>
<td>25.1%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure Note: Example child: Boy aged 8-10, white, unmarried mother who is working, family income below poverty Base case policies: In-home providers not allowed; maximum reimbursement rate is 90% of average cost of care; copay does not vary with choice of care; providers are not allowed to charge families the difference between their full price and the maximum reimbursement rate.

**Implications**

The analysis of merged administrative data and policy data has strong potential for CCDF program research. The 801 data is a readily-available source of information on the individual characteristics of a very large number of actual CCDF subsidy recipients; and the CCDF Policies Database is a readily-available source for detailed information on CCDF policies over time and across places. Merging the two data sources allows analyses that could help uncover relationships between State/Territory policy choices and program outcomes.
Looking Forward/Resources

The CCDF policies shown here are taken from the CCDF Policies Database. For additional detail regarding the policies available in the Database, see the following resources:

- Full data files: [http://www.researchconnections.org/childcare/resources/17048](http://www.researchconnections.org/childcare/resources/17048)

If you have questions or comments, please contact us at CCDFDatabase@urban.org
About the Authors

**Linda Giannarelli** is a senior fellow in the Income and Benefits Policy Center at the Urban Institute, focusing on analysis of government programs that support lower-income families. She serves as project director for the CCDF Policies Database. She also directs work involving the TRIM3 microsimulation model, and has used TRIM3 to assess the potential anti-poverty impacts of child care subsidies.

**Sarah Minton** is a research associate in the Income and Benefits Policy Center at the Urban Institute. Her research focuses on child care subsidy policies and other government programs that serve low-income families. She serves as co-project director for the CCDF Policies Database.

**Christin Durham** is a research associate in the Income and Benefits Policy Center at the Urban Institute. She conducts research and evaluation for projects related to workforce development as well as poverty and the safety net.

Acknowledgments

The authors thank Silke Taylor for conducting the statistical programming and Owen Haaga for his assistance in reviewing the results. We are indebted to Kathleen Dwyer and Meryl Barofsky of OPRE, Joe Gagnier and Andrew Williams of OCC, and Gina Adams and Monica Rohacek of the Urban Institute for very helpful reviews of earlier drafts of this brief. We also thank Lorraine Blatt for expert assistance with graphics.

The views expressed in this publication do not necessarily reflect the views or policies of the Office of Planning, Research and Evaluation, the Administration for Children and Families, the U.S. Department of Health and Human Services, the Urban Institute, or the Urban Institute’s trustees or funders.