

# The Economic Role of Paid Child Care in the U.S.

## A Report Series — Part 3: Economic Growth Modeling

### Glossary

May 2022



#### **Best-Fit Line**

A best fit line through a scatter plot of data best expresses the linear relationship between those points. The straight line provides the best approximation of the relationship between the data points. The slope of the line provides a quantitative estimate of both the direction and magnitude of the relationship. Best fit lines are also commonly referred to as trendlines or linear regression lines.

#### **Birth Rate**

Measures of the birth rate reflect the number of births in a population over time (typically one year). Commonly used measures of the birth rate include the crude birth rate and fertility rate.

#### **Capital Investment**

Capital investment is the purchase of tangible and intangible assets by firms, governments, and individuals for the purpose of pursuing their business and operating goals and objectives. Capital investment is measured in the child care report using net private fixed investment as defined by BEA. The estimated measure of capital is net of depreciation and includes the broad asset categories of equipment, structures, and intellectual property. Public sector assets are excluded from the analysis. State-level estimates are formed by partitioning national data on net private fixed assets at the industry level based on a region's share of national household earnings at the industry level.

#### **Child and Dependent Care Tax Credit (CDCTC)**

The Child and Dependent Care Tax Credit is a federal tax credit available to pay for the care of eligible children and adult dependents (qualifying persons) to enable taxpayers to work or look for work. To claim the credit, you (and your spouse if filing jointly) must have earned income during the year. Child and dependent care expenses must be work-related to qualify for the credit. There is a special rule for education. The spouse is treated as having earned income for any month that he or she is a full-time student, or physically or mentally not able to care for himself or herself. If filing a joint return, this rule also applies to either spouse. You can be treated as having earned income for any month you are a full-time student or not able to care for yourself. The credit is calculated based on earned income and covers a percentage of expenses incurred for the care of qualifying persons. For tax year 2021, the American Rescue Plan Act of 2021 extended the credit up to \$4,000 for one qualifying person and \$8,000 for two or more qualifying persons and made the credit temporarily refundable.

**Children of Child Care Age**

Children ages 14 and under are considered most likely to participate in formal or informal child care. The group of children of child care age is divided into two groups in the report: younger children ages 0 to 4 and older children ages 5 to 14. This definition follows the Current Population survey which measures paid child care usage for children ages 14 and under.

**Child Care Cost Burden**

The child care cost burden of a household reflects expenditures on paid child care as a percentage of household income. This measure reflects the notion that the cost burden of child care is best measured relative to ability to pay. The cost burden is calculated as child care expenditures divided by total household income. Burden can be measured on a per child basis or for all children in a household.

**Civilian Non-Institutional Population**

The civilian non-institutional population measures those persons ages 16 and older and their children not on active duty in the Armed Forces or residing in institutions (e.g., correctional institutions or long-term care facilities for the aged).

**Cointegration**

Cointegration is a statistical concept that refers to the long-run co-movement of two or more data series over time. If variables are found to be cointegrated, estimates can then be made of the long-run elasticity between two cointegrated factors over time. The concept of cointegration is closely tied to the notion of Granger causality but focuses on the long-run dimension of the relationship among a group of variables over time. A set of cointegrated variables maintain a long-run equilibrium relationship over time, with any short-run deviations from the long-run relationship corrected over time through an error-correction process. In fact, two cointegrated variables are expected to have short-run Granger causality present in at least one direction.

**Correlation**

Correlation is a statistical measure of the degree of linear dependence between two series over a specified period. Correlated series tend to move in coordination with one another over time. Positively correlated variables tend to move in the same direction; negatively correlated series tend to move in the opposite direction. If X and Y are correlated, the calculated correlation is the same for both series in each direction in the period.

**Cost-of-Living**

Cost of living reflects differing prices across geographic areas for a range of typical living expenses including housing, food, energy, and other items. Measures of the cost of living are often used to compare how costly it is to live in one geographic area versus another. Cost of living adjustments are made in the report using state-level regional price parity (RPP) indexes produced by the Bureau of Economic Analysis (BEA) along with the national implicit price deflator to adjust for national price changes over time.

**Crude Birth Rate**

The crude birth rate is the number of births per 1,000 population in a geographic area.

**Current Population Survey (CPS)**

The Current Population Survey, also commonly referred to as the household survey, is a sample-based monthly survey of about 60,000 eligible households. It provides a comprehensive body of data on the U.S. labor force by demographic and labor force characteristics.

**Current Population Survey (CPS) Annual Social and Economic Supplement (ASEC)**

A widely used supplement to the Current Population Survey is the Annual Social and Economic Supplement (ASEC) conducted by the Census Bureau every February, March, and April. The supplement collects data on health insurance coverage, work experience, income from all sources, receipt of noncash benefits, poverty, migration, geographic mobility, and other special topics. The CPS ASEC also collects data on the number of children in paid child care and the expenditures of households and families using paid care. Use of the ASEC requires a tradeoff from monthly to annual data but provides a broader sample and larger universe than the basic CPS.

**Educational Attainment**

Educational attainment refers to the highest level of education that an individual has completed. Attainment is often measured using the number of years of education completed, especially when used to describe the average attainment across the population of a geographic region. Attainment is distinct from the level of schooling that an individual is attending currently.

**Elasticity**

Elasticity is an economic concept used to measure the percentage change of one economic variable in response to a change in another. The response is deemed elastic (or highly responsive) if the resulting change in a variable is more than proportional to the initial change and inelastic (or not highly responsive) if less than proportional.

**Employment-Population Ratio**

The employment-population ratio (or employment ratio) is a measure of labor force attachment that measures the share of the population activity employed. The ratio is calculated as the number of employed workers divided by the civilian noninstitutional population. The employment ratio does not consider unemployed workers as attached to the labor force. As a result, the employment ratio is far more volatile than the labor force participation rate across the economic cycle.

**Family**

A family is defined in the Current Population Survey (CPS) as a group of related individuals who are all members of the same household. Multiple families can be domiciled within the same household.

**Female Labor Force Participation Rate**

The female labor force participation rate measures the rate of participation of women in the labor force.

**Fertility Rate**

The fertility rate is the number of births per 1,000 women ages 15 to 44 in a geographic area.

**Goods-Producing**

Goods-producing sectors of the economy are those that produce products rather than services. These typically include NAICS sectors covering farming; forestry, fishing, and related activities; mining; and manufacturing.

**Granger Causality**

Granger causality is a statistical test of the usefulness of one variable in forecasting future values of another. Granger causality is present between two variables if future forecasts of variable X are improved by using variable Y in its prediction, above the level present when using only information about the history of X. Granger causality can be present in a single direction from either X to Y or Y to X, in both directions (bi-directional), or may not be present at all. If there is no Granger causal relation found from Y to X, Y is deemed strictly exogenous to X in providing useful forecasting information. Granger causality also differs greatly from measuring the correlation between two time series. Correlation simply measures the linear dependence between two series over a specified period. If X and Y are correlated, the calculated correlation is the same for both series in each direction in the period. Granger causality, however, measures statistical predictability in both directions and in the time dimension.

**Great Recession**

The Great Recession refers to the steep decline in economic activity associated with the U.S. recession lasting from December 2007 to June 2009, as well as downturns in national economies globally. It is the longest recession in the post-World War II period and generally considered the most significant economic downturn since the Great Depression.

**Household**

Survey data from the Current Population Survey (CPS) Annual Social and Economic Supplement (ASEC) are organized using samples of households or dwellings. A household is defined as all persons who occupy a single dwelling unit. A dwelling unit is a room or group of rooms intended for occupation as separate living quarters and having either a separate entrance or complete cooking facilities for the exclusive use of the occupants. In a small percentage of cases, multiple family units occupy a household.

**Household Income**

Household income represents all combined forms of gross income, both earned and unearned, for all members of a household ages 15 and over.

**IPUMS-CPS**

IPUMS provides census and survey data from around the world integrated across time and space. IPUMS integration and documentation makes it easy to study change, conduct comparative research, merge information across data types, and analyze individuals within family and community context. Data and services available free of charge. IPUMS CPS harmonizes microdata from the monthly U.S. labor force survey, the Current Population Survey (CPS), covering the period 1962 to the present. Data include demographic information, rich employment data, program participation and supplemental data on topics such as fertility, tobacco use, volunteer activities, voter registration, computer and internet use, food security, and more. IPUMS-CPS, University of Minnesota, [www.ipums.org](http://www.ipums.org).

**Labor Force**

The labor force includes all persons in the civilian noninstitutional population classified as either employed or unemployed. The labor force does not change as individuals move from employment to unemployment, and vice versa. The labor force changes only when new entrants enter the labor force or existing participants exit.

**Labor Force Status**

Labor force status measures the degree of labor force attachment for persons ages 15 years and older. Persons are generally classified as either in the labor force or not in the labor force. Those in the labor force are further classified as either employed or unemployed. Many persons are not in the labor force due to school, retirement, health, personal choice, and other factors. Members of the Armed Forces are excluded from most measures of work status.

**Labor Force Attachment**

Labor force attachment is a general economic term referring to a person's status as a participant in the labor force. Persons attached to the labor force include those either employed or unemployed. Those who are unattached do not participate in the labor force. The two most widely used measures of the degree of labor force attachment for the population of a geographic area are the labor force participation rate and the employment-population ratio.

**Labor Force Participation Rate**

The labor force participation rate is the most widely cited measure of labor force attachment and is calculated as persons in the labor force (either employed or unemployed) divided by population (civilian noninstitutional) ages 16 and over. In other words, it captures the percentage of the population of a geographic area that is either employed or unemployed and looking for work. The inclusion of the unemployed is the key characteristic of the participation rate versus the employment ratio, another popular measure of labor force attachment.

**Maternal Labor Force Participation**

Maternal labor force participation refers to the labor force participation of women with children.

**Mean Reversion**

Mean reversion refers to the tendency of observations in a data series to gradually move towards the long-term mean of the series. Large deviations away from the mean are expected to be followed by a return toward the mean in a mean reverting series.

**Median Household Income**

For households, the median income represents the level of household income where half the households in a geographic region (including those with no income) earn more and half earn less. Median household income is also referred to as the midpoint of the income distribution or the 50th percentile of household income.

**Outlier**

An outlier is an observation or data point that differs significantly from others in the same sample. Outliers can be due to measurement error or may simply reflect unusual and unexpected behavior among the observations in the sample. Outliers are sometimes excluded from the data set to gauge the sensitivity of any statistical findings to the presence of the outlier(s).

**Paid Child Care**

Paid child care is defined in the report as any form of child care arrangement for a child ages 0 to 14 for which a parent makes a direct expenditure on care to enable them to work. This follows the definition of paid child care used in the Current Population Survey (CPS). Paid options can include both formal and informal care arrangements such as neighbors or friends, which may or may not be regulated by states.

**Panel Data**

Panel data refers to data observations categorized for a given entity or data measure observed across time. Panel data is also known as longitudinal or cross-sectional time series data. Panel data used throughout the child care report is defined for multiple states (cross sections) and multiple time periods.

**Panel Model**

Panel model techniques are statistical tools and methods that simultaneously utilize the information contained in the economic behavior of an entity or data measure across time. Unit root, Granger causality, and cointegration tests used throughout the child care report are all panel modeling techniques that use a 50-state panel dataset. The use of a panel of states rather than national data can provide for more robust estimates of the fundamental factors driving paid care usage.

**Personal Income**

Personal income includes all forms of income that persons receive in return for their provision of labor, land, and capital used in current production and the net current transfer payments that they receive from business and from government.

**Per Capita Income**

Per capita income measures the amount of income earned per person in a geographic region. Per capita income is commonly used as a measure of standard of living of the population in a region.

**Prime Working Age Women (ages 25-54)**

Women of prime working age are those ages 25 to 54 who actively participate in the labor force. These women are more likely to participate in the labor force than younger and older women and have likely completed pre-career education and training.

**Probability Value (p-value)**

A probability value, or p-value, is a statistical parameter used within hypothesis testing that determines the probability of obtaining the observed results assuming a given probability distribution of the test statistic and that the null hypothesis is true. In other words, it is the predetermined level of probability at which statistical significance is found. A p-value of 0.05 (5 percent) or lower is typically considered the threshold of statistical significance.

**Public Preschool Education**

Public preschool includes a range of publicly funded early childhood education programs accessed by children before they begin compulsory education at the primary school level. Public pre-kindergarten (or Pre-K) programs are commonly available to children ages 4 to 5 in many states (i.e., 5-year-old children not yet enrolled in public kindergarten). In some states, public preschool also serves 3-year-old children. Publicly funded preschool could be located in a school or in a mixed delivery setting such as child care centers and family child care homes depending upon state or local school district decisions.

**Quartile**

A quartile is a statistical tool used for summarizing data by dividing the observations into four groups that are more-or-less of equal size. Data is often ranked along some measure of the value of the underlying data and then assigned to quartiles. As with other forms of quantiles (e.g., terciles, quintiles, deciles, etc.), quartiles provide a convenient means of comparing data across grouped intervals.

**Real Personal Income**

Personal income calculated at its nominal, or current, value and then adjusted for the effects of inflation over time is deemed real personal income. At the state level, an additional adjustment is made to nominal personal income to reflect state-level differences in cost-of-living when calculating real personal income. The cost-of-living adjustments are made using Regional Price Parity (RPP) indexes developed by the Bureau of Economic Analysis.

**Sample Size**

Sample size refers to the number of individual observations in a sample of data.

**Service-Providing**

The service-providing sectors of the economy produce intangible services instead of goods. A range of services are produced by both private and public sector entities under the NAICS classification system.

**Stationarity**

A stationary data series will have a mean, variance, and autocorrelation structure that is stable over time. Visually, stationary series tend to be mean reverting and do not trend strongly upward or downward. They also do not have periodic patterns such as seasonality. A non-stationary data series, or one with a unit root, may have to be differenced one or more times to achieve stationarity. The level of integration, denoted as  $I(i)$ , is used to describe the number of times ( $i$ ) a data series must be differenced to achieve stationarity.

**Statistical Causality**

The statistical notion of causality tests for the increased predictability of the future path of one variable,  $X$ , using another variable,  $Y$ . While not addressing the issue of economic causality in the traditional sense, tests of statistical causality provide an empirical measure of the historical responses and timing embodied in the relationships among data series. Granger causality is a common approach to testing for statistical causality. Because economic causality also operates in the time dimension, economic causal relations are often informed using Granger-type methods, particularly in forecasting applications.

**Statistical Significance**

In statistical hypothesis testing, a result is statistically significant if it is deemed unlikely to have occurred due to chance given the stated hypothesis tested. Statistical significance is usually determined by rejection of the null hypothesis.

**Subsidies and Cost Offsets**

Several federal and state subsidies, tax credits, and other forms of cost offsets are available to assist families in meeting the cost of paid child care. Subsidies and offsets examined in the report include those provided through the Child Care and Development Fund (CCDF), the Temporary Assistance for Needy Families (TANF) block grant, and the Child and Dependent Care Tax Credit (CDCTC).

**Time Series Analysis**

Time series analysis describes a group of statistical techniques and methods for analyzing time series data to extract meaningful characteristics of the data. These techniques are used most often to examine relationships present between variables over different points in time. The Granger causality and cointegration tests used in the reports are widely used methods of time series analysis.



**Time Series Data**

Time series data is a collection or sequence of data observations collected over time intervals. Time series data is commonly collected on an hourly, daily, weekly, monthly, quarterly, or annual time interval and indexed in time order.

**Todo-Yamamoto Method (TY Method)**

All Granger causality tests in the child care report are implemented within a VAR model framework using the method of Todo and Yamamoto (1995). The TY method allows for causality testing among a group of data series within a system framework. The system includes an equation for each data series with the series as the dependent variable and the remaining variables as explanatory (independent or right-hand side or) variables. The TY method is noteworthy in that it is robust to the presence of unit roots, or the order of integration of the time series. The base VAR used in the tests is augmented, or overfit, by including an additional lag of the level of each variable as an additional exogenous variable in each equation of the VAR.

**Traded Activity (or Openness)**

Traded activity is defined as production for trade outside a region, or a region's degree of openness. The concept traces its origins to the notion of enhancing the 'basic' industries located within a region. Basic industries produce goods and services that are exported for sale outside the local market. This includes trade with other states as well as internationally. States with large manufacturing, mining, and Federal government sectors (including military) tend to have the most traded activity with outside regions. Traded activity captures spending from outside the region which in turn helps support the development of the region's 'non-basic' sectors. Non-basic industries are believed to merely recirculate existing purchasing power, which exerts less influence on overall regional growth than an equivalent injection of spending from outside the region.

**Unit Root Test**

Unit root tests are used to test the stationarity of a data series and establish its degree of integration. I(0) variables are stationary in levels (no unit root) and require no differencing, while I(1) variables have a unit root and must be differenced once to achieve stationarity. An I(2) series is one that must be differenced twice to achieve stationarity. Most nonstationary series are I(1) and become stationary after differencing once. Few data series require differencing twice (or more) to achieve stationarity.

**Unpaid Child Care**

Some families may use unpaid child care, which reflects time children spend out-of-the-home. However, for purposes of this report series, only the use of paid child care was reviewed. The series compares the average income of families with children age 14 and younger that use paid child care compared to families with children of the same age that do not use paid child care. The same analysis is also included for families with children under age 5 that use paid care compared to families with children under age 5 that do not use paid care.

**Vector Autoregressive (VAR) Model**

Vector autoregressive models are a time series technique used to investigate the relationships among a group of time series variables. The estimated model includes an equation for each data series with the series as the dependent variable and the remaining variables as explanatory (independent or right-hand side or) variables. Each equation in the VAR model includes only each variable's lagged (or past) values, lagged values of the other variables in the model, and an error term. The VAR model imposes no structural assumptions on the data but instead treats all data in the model as endogenous to the system.

**Women of Working and Childbearing Age (ages 18-54)**

The population of women ages 18 to 54 are of both working age and childbearing age. These women are the most likely to use paid child care services for children ages 0 to 14. This measure captures a broader group of women than prime working age women (ages 25-54) by including younger women ages 18-24 who are typically of childbearing age.

## Abbreviations and Acronyms

BLS	Bureau of Labor Statistics
BEA	Bureau of Economic Analysis
CCDF	Child Care and Development Fund
CDCTC	Child and Dependent Care Tax Credit
CED	Committee for Economic Development of The Conference Board
CPS	Current Population Survey
CPS ASEC	Current Population Survey, Annual Social and Economic Supplement
ECPP	Education Early Childhood Program Participation
HHS	US. Department of Health and Human Services
IPUMS-CPS	IPUMS-Current Population Survey
IRS	Internal Revenue Service
NAICS	North American Industrial Classification System
NIEER	National Institute for Early Education Research
RPP	Regional Price Parity
SIPP	Survey on Income and Program Participation
SPM	Supplemental Poverty Measure
TANF	Temporary Assistance for Needy Families
VAR	Vector Autoregressive Model