California County-Level Economic Forecast
Methodology Update

California County-Level Economic Forecast Methodology Update

October 2020

This publication was prepared for:

Transportation Economics Branch
Office of State Planning
California Department of Transportation
1120 "N" Street
P.O. Box 942874 (MS- 32)
Sacramento, CA 94274-0001
Fahimeh Bagheri
Economist
(916) 653–3087

This publication was prepared by:

The California Economic Forecast
Mark Schniepp, Director
5385 Hollister Ave Box 207
Santa Barbara, CA 93111
(805) 692–2498
www.californiaforecast.com

Methodology Update

This report updates the modeling approach and the data used to forecast the county level economic indicators for the annual California Department of Transportation County economic forecast.

The county models comprise an elaborate forecast system for projecting economic activity regionally in the state.

The modeling system is the only county level forecast in California where all county economies are forecast.

The modeling system has been continuously updated and improved since the year 2000 which represented the beginning year of this annually updated project.

The Econometric Model: A Brief Description

The county models are independent: each model consisting of a series of equations is autonomous from other county economies. While each county model is independent, it is nevertheless symmetrical with all other county models. The symmetry is important because all models produce forecasts for economic indicators common to all counties, using the same base years, the same inflation rate, the same units of measurement.

All models have the same outputs and the exogenous forecasts used in the equations for the county models are drawn from the same but always current pool of indicators that are generated by the California and U.S. forecasts updated routinely by the UCLA Anderson Forecast.

This is extremely important because all county forecasts can be summed to derive a statewide total. They can also be compared to one another (in an apples to apples comparison) to determine relative performance, such as per capita income, number of persons per household (average household size), retail spending per person, average annual salary per worker, workers per household, employment to population ratio, or people per vehicle.

The models are county-specific, and the specifications are built with the objective of considering unique attributes of each county economy.

Each county model is comprised of 6 blocks of equations: 40 to 50 stochastic behavioral relationships and 20 to 25 accounting identities.[1] The model is characterized by simultaneous interaction and determination of local employment, income, population, wages, retail spending, and the demand for housing.

---

[1] There are 58 counties and a minimum of 40 stochastic equations per county that need to be re-estimated and re-calibrated every year. That's 40*58= 2,320 equations that must be evaluated for plausibility, consistency, and stability. Some counties are now up to 64 stochastic equations.

The stochastic equations are estimated using the ordinary least squares regression method and the entire system is solved using the Gauss-Seidel algorithm.

The model is a "satellite model," requiring forecasts of various California and U.S. economic variables which are treated as exogenous to the local county area.

The county-level models are each moderately detailed. As we noted above, their equation systems are estimated as a model, independent of other counties. However, some interactions between counties have been accommodated where we have detected interdependence. For example, the visitor industry in Napa and Sonoma Counties, or transportation and warehousing in Riverside and San Bernardino Counties.

All of the stochastic equations in a county model are evaluated each time revised historical data or new data are introduced into the models. This is also true when a re-specification of an equation, or of the block structure occurs.

Outputs

The initial economic and demographic indicators that are forecast for each county are shown in the table below. Forecast values are prepared over a 30 year time period beginning with the year in which actual data are not yet available.

Forecasts are derived for each county independently.

Table 1
The principal economic indicators which are forecast by the California County econometric models

----
- Non-farm employment by principal two digit NAICS sector
- Farm employment
- Total wage and salary employment
- Unemployment rate
- Personal Income
- Per capita personal income
- Number of housing units permitted and total housing stock
- Taxable retail and total taxable sales
- Population (and births, deaths, net migration)
- Number of households
- Number of vehicle registrations
- Existing home sales
- Median housing values
- Total agricultural value of products
- The regional inflation rate
- The value of industrial production
- K-12 school enrollment

Methodology Update

MODEL STRUCTURE

General Characteristics

The county models are a macroeconomic structure consisting of interdependent equations. Each endogenous variable (determined by the model) is a function of other endogenous variables, exogenous variables (determined outside the model), and an error term. Implicitly, each equation may be represented as:

$$Y_{it} = f(Y_{jt}, X_{kt}, u_t)$$

where

Y_{it} = endogenous variable i in period t
Y_{jt} = endogenous variable j in period t
X_{kt} = exogenous variable k in period t
u_t = error term in period t

The determination of Y_{it} by a variable determined elsewhere in the model, is the essence of a simultaneous equation model. The endogenous variables interact within the model as they do in the real world.

The structure of the model is simultaneous, arranged in blocks of equations. Each block is comprised of a system of equations that define the block, or sector. All sectors are linked, meaning feedback exists between blocks. The equations within each block are either stochastic (that is, measured with error) or deterministic (i.e., are determined by an identity or formula having no measurable error).

The equations have been arranged in 6 blocks to aid in organizing the model.

Sector 1: Housing and New Building
Sector 2: Demographics
Sector 3: Income
Sector 4: Consumer Spending
Sector 5: Employment
Sector 6: all other equations including the Farm sector

For each sector, a particular set of endogenous variables are specified to meet the principal objectives of the county forecasting model, which are to generate forecasts for the indicators listed in Table 1. A number of other endogenous variables are needed as intermediate stages in the determination of the key variables that are typically be reported in the long term forecast tables.

Estimation Period

The database associated with each County was assembled from as far back in time that data have been recorded to the most current year for which actual information is available. Annual observations are used in the estimation and forecast.

Due to the varying availability of economic and demographic data at the sub-national level, each block in the system has its own number of observations associated with it. Consequently, the estimators calculated for the forecasting equations were derived from varying numbers of observations.

For the Employment block, all counties include NAICS data that began in 1990. For the larger counties, NAICs categorized employment data were backcast using the previous classification system for which the data was reported in, the Standard Industrial Classification system or SIC codes. The backcasted information begins in 1983 and ends in 1989. The additional 7 observations provide for more robust estimates and stability of the forecasts in the larger county models.

Income data for all Counties commences in 1969, and is available for all 58 Counties in California.

For the Housing Sector Block, the number of households and housing stock begin in 1980 for most counties. For some of the smaller counties, the data begin with the 1984 calendar year. The building permit data all begin in 1969 for all counties in California. Median home selling prices typically begin in 1990 for all counties. For some of the larger counties including those comprising Southern California and the Bay area, data as reported by the California Association of Realtors has been recorded since 1980.

The Consumer spending block which consists of retail sales and retail store permits begins for all Counties in 1969.

In the Demographic block, the observations begin in 1970 or 1980. Population in all cases begins in 1970. Net migration, births, deaths, and population by age also begin in 1970.

For the Farm sector, farming output or sales by county is available beginning in 1972 for most counties.

Equations withing the various blocks are estimated with as long a time series as possible, though that will vary depending on the county, the block, and the exogenous right-hand side indicators that form the equations. The length of time is limited by the indicator with the fewest historical data points.

Methodology Update

In general however, nearly all equations of the forecasting model for all counties rely on data starting in at least 1990. Consequently, through the 2019 forecast, estimated equations were generated from at least 30 years of historical data and often the time period is 40 years.

Other Indicators

The consumer price index (CPI) for the north and the southern regions of California, and the California composite CPI is available from 1920 to present. The statewide home mortgage rate begins in 1970.

Methodological Sequence for Developing the County Level Forecasts

1. Update County Level databases: 58 Counites x 66 variables or indicators per county. Ditto for the state. The update includes adding a new calendar year of actual data and revising the past history

2. Update County Level models with new and revised data. All county level equations are re-estimated to update and optimize the estimated coefficients

3. Solve each of the county forecasting models and produce a preliminary forecast

4. Evaluate the estimated equations and evaluate the forecast;
   a. perform validation criteria to evaluate the forecasts
   b. re-calibrate equations if necessary
   c. re-specify equations if necessary and produce a final forecast

5. Incorporate special considerations for the county forecasts, such as a recovery from natural disasters, the development of a new and significant industry, or new growth policies that have been established in the County. Adjust the forecast if necessary.

6. Produce the narrative explaining the forecast and update all forecast charts and tables

7. Using publication software, produce the forecast for each county as an independent chapter that will comprise the full document of forecasts for all California Counties, and the state.

Accommodating Special Circumstances

Every year, the rural counties are researched for significant changes in new development which will alter the forecast for new housing or non-residential building and therefore ultimate job creation. Special data pertaining to specific counties is gathered (or updated) to explain economic circumstances unique to that county. For example, because prisons can be the dominant driver of a small county's economy, prison inmate populations and employment are routinely updated for Lassen County, where three large prisons are located in Susanville. Forecasts of prison populations are obtained from the Department of Corrections.

In Yolo County, UC Davis dominates state employment. Consequently, enrollment for the campus is used to explain the variation in Yolo County state and local employment. Enrollment forecasts for the short term are typically provided by the UC system.

The methodology of accommodating special handing of particular sectors of a county's economy is updated over time from experience gained understanding the many nuances of the smaller county economies.

This is especially true for the 30 smallest counties in California.

Forecast Validation Criteria

County Validation

As part of the evaluation of the county point estimate forecasts over time, a number of ratios are constructed to validate the forecast. Ratios of the county forecast for indicator i to the same indicator for the state are calculated:

$X_{i,c}$ / $X_{i,California}$
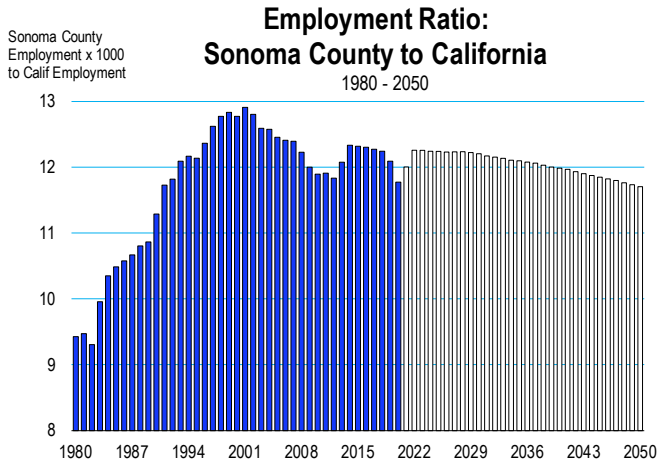
where $X_{i,c}$ is indicator i for county c, and $X_{i,California}$ is the same indicator for all of California.

The California indicator forecasts are produced independently by the UCLA Anderson Forecast.

This ratio is taken over the entire long term forecast period to evaluate the extent of the county forecast.

The forecast trajectory or path of $X_{i,c}$ / $X_{i,California}$

Methodology Update

**Employment Ratio:
Sonoma County to California**
1980 - 2050
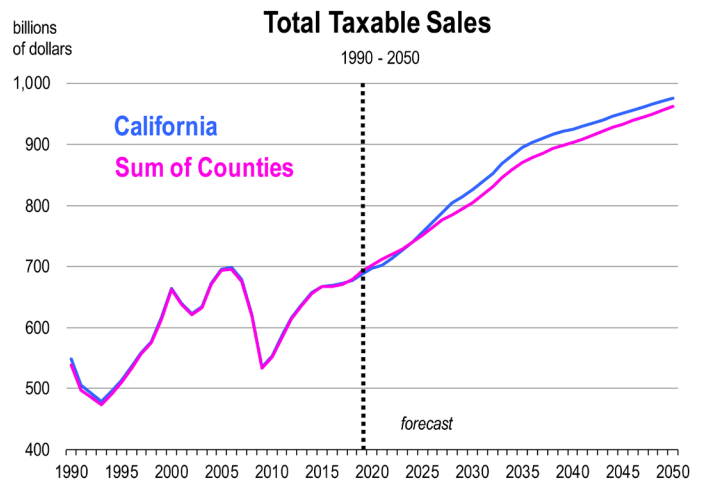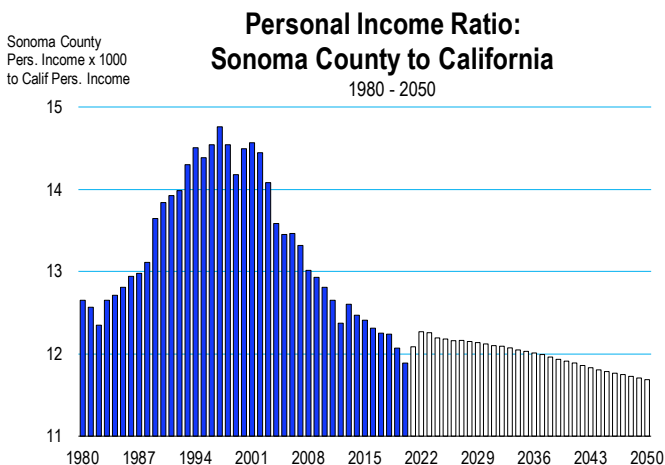
Sonoma County
Employment x 1000
to Calif Employment

was more impacted in the county during the pandemic than in California, due largely because tourism is more concentrated in Sonoma relative to the state. The forecast shows employment relative to California returning to the same path it was on for much of the previous 20 year history.

Another example is the ratio of county personal income to California personal income. History indicates that personal income in Sonoma County has been declining as a share of total statewide personal income since 2001. Following the pandemic recession, Sonoma personal income recovers sharply but the declining share of the state is reinstated over the forecast. The forecast then demonstrates that Sonoma personal income is forecast to grow at a rate that is consistent with its history.
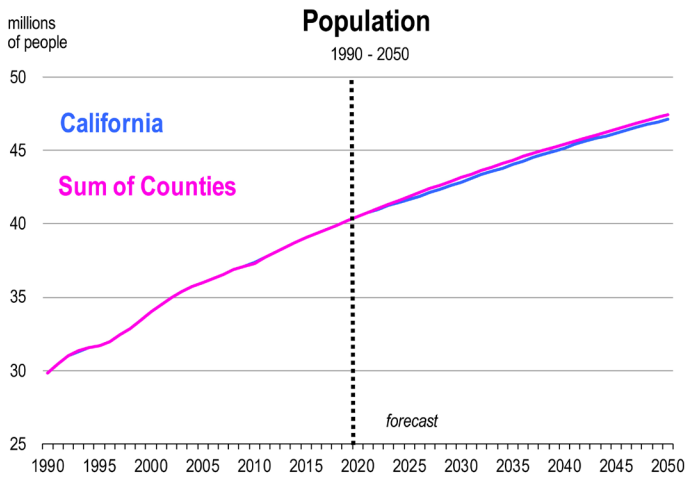
Aggregate Validation

Aggregate validation occurs when the sum (or average) of indicator i is then compared with indicator i for the state. The aggregate indicator should generally move in tandem with the state indicator through the forecast period. Typically the acceptable error range for most indicators is 5 percent. However for volatile series such as net in-migrating population which itself is measured with error, higher forecast errors are considered acceptable.

should be relatively constant or trending according to its historical movement. Then the county's forecast for indicator i is typically deemed reasonable. If the forecast deviates from the historical path, there may be issues with the forecast, or such a deviation may be explained by circumstances known to characterize X_i for county c over time.

Either way, the calculation of county forecast ratios with the state provide us with information necessary to either validate the forecast, dismiss it, or accept it when either known or expected circumstances warrant it.

For example, the ratio of employment for Sonoma County and California demonstrates that employment relative to California

For example, the sum of taxable sales for all counties, over the 1990 to 2050 period, is compared to the statewide forecast prepared by the UCLA Anderson Forecast. The county sum deviates from the statewide total by less than 4 percent over time. For total employment the deviation is less than 3 percent over the entire forecast period.

**Personal Income Ratio:
Sonoma County to California**
1980 - 2050

Sonoma County
Pers. Income x 1000
to Calif Pers. Income

Total Taxable Sales
1990 - 2050

billions
of dollars

California

Sum of Counties

forecast

Methodology Update

Population
1990 - 2050

millions
of people

California

Sum of Counties

forecast

For population, the deviation between the county sum and the statewide total is less than 1 percent.

For net migration, the county sum generally follows the statewide forecast path, but the average annual deviation is 13 percent.

Particular county models for indicator i will be re-specified and re-forecast if the aggregate validation criteria produces a large error for a series that is generally smooth.

Net Migration
1990 - 2050

migrants

California

Sum of Counties
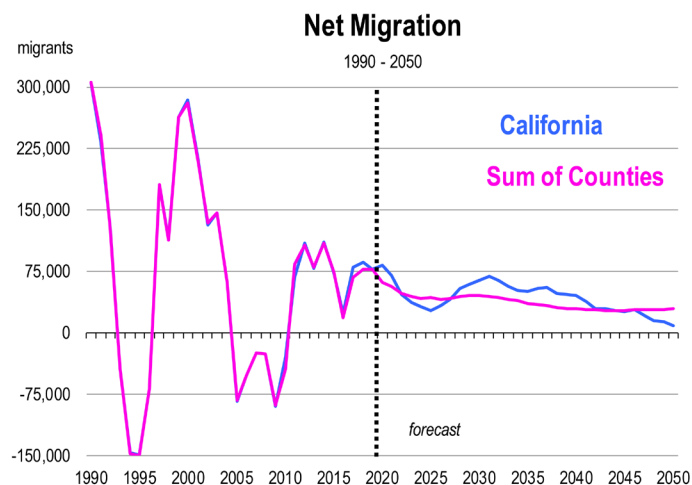
forecast

ENDOGENOUS FACTORS
(ECONOMIC INDICATORS THAT ARE FORECAST)

These variables are left hand side variables that are modeled using a behavioral relationship specification comprised of both exogenous factors and other endogenous variables.

There are more endogenous economic indicators forecast as part of the modeling system than we present in the county forecast presentations. This is because many more endogenous variables need to be forecast because they are used as exogenous factors used to determine the core economic indicators.

Sectors of the Model

The model is arranged into 6 sectoral blocks of equations. However the blocks are not recursive, that is, they are not estimated independently and determined (or solved for) sequentially. The models are characterized by simultaneous interaction and determination of local employment, income, population, wages, and housing demand.

Housing and New Building

Stochastic equations

Number of households (HH)
single family units (SFU)
Multiple family units (MFU)
Residential building value permitted, constant dollars
Non-residential building value permitted, constant dollars
Average building value for new residential units, constant dollars
Median home selling price, constant dollars
Number of existing home re-sales

Identities

Housing stock: HS = HS(t-1) + UNITS(t-1)

Single and multifamily housing permits: SFU + MFU = UNITS

Ratio of single family units to total residential units permitted: SFU / UNITS

Total building value permitted, the sum of residential and non-residential value

The ratio of the county median price to the national median selling price:

Methodology Update

Demographics

Stochastic equations

Births (calendar year series)
Deaths (July series)
Deaths (calendar year series)
Net in-migrating population (July-June series)
Number of registered vehicles
Unemployment rate
Employed labor force
Civilian labor force
Number of registered passenger cars

Identities

Population on July 1: (POPJUL) = POPJUL(t-1) + births(t) − deaths(t) + net in-migration(t)

Population growth: (POPJUL(t)-POPJUL(t-1))/ POPJUL(t-1)

Change in population: POPJUL(t) − POPJUL(t-1)

Persons per vehicle: POPJUL / number of vehicles

Average household size: POPJUL / HH

Income

Stochastic equations

Transfer payment income, constant dollars
Property (or asset) income, constant dollars
proprietor income, constant dollars
Residence adjustment income, constant dollars
Average earnings per worker, constant dollars (RYEPW)

Identities

Total wage and salary earnings, constant dollars = total employment*RYEPW

Total personal income, constant dollars = Total wage and salary earnings, constant dollars + transfer, property, proprietor, and residence income, constant dollars

Per capita personal income, constant dollars: Personal Income / Population

Wage ratio = County average salary / California average salary

Consumer Spending (retail sales)

Stochastic equations

Retail sales (taxable retail sales), constant dollars
Number of retail outlets or stores
Total taxable sales, constant dollars

Identities

Retail sales per store = retail sales / retail outlets

Ratio of retail sales to personal income = retail sales / personal income

Employment (non-farm sector)

Stochastic equations

employment in mining
employment in construction
employment in manufacturing
employment in durable manufacturing
employment in transportation, communications, and utilities
employment in retail trade
employment in wholesale trade
employment in information
employment in financial activities
employment in professional and business services
employment in education and healthcare services
employment in leisure, accommodation, and recreation services
employment in other services
employment in state and local government (ESLG)
employment in federal government (EFG)
Number of proprietors (self-employed workers)

Identities

Employment in government = ESLG+EFG

Total wage & salary employment (ETWS) = sum of all non-farm employment sectors plus the farm sector

Change in total employment: ETWS(t) − ETWS(t-1)

Employment to population ratio: ETWS / Population

Growth rate of employment: (ETWS(t)-ETWS(t-1))/ ETWS(t-1)

Methodology Update

Farm Sector and Misc. Equations

Stochastic equations

Wage and salary employment in farming
Total agricultural crop value, constant dollars
Southern and Northern California inflation rate (I)
K-12 school enrollment

Identities

Consumer Price Index, Southern (Northern) California (CPI) = CPI (t-1)*(1+[I(t)/100])

EXOGENOUS VARIABLES

There are approximately 100 to 120 exogenous variables that we had selected in the initial development of the model. Most of these factors have remained relevant for use in the models over time. However, as the economy changes, new exogenous factors may be added to models to explain the variation in county level economic indicators.

Not all of these exogenous variables are used. However, these variables have been found to be important in the original specification tests based on goodness of fit criteria together with their theoretical propriety. The exogenous variables are updated annually and made available for updating the equations in the model and/or enhancing the specifications as needed.

Currently, all blocks in the model are driven by exogenous factors, as well as endogenous factors that are determined in other blocks of the general model.

The exogenous variables include the following:

(1) California economic and demographic variables

(2) National economic variables

(3) Local county demographic variables: These factors are age specific population counts from the Department of Finance. The model uses 10 of these to drive various equations in the Employment and Demographic blocks of the model.

(4) Housing variables: the California median home price, California re-sales, mortgage rates, and notices of default and foreclosures in California

(5) Special circumstance exogenous factors as needed. This would include forecasts by the Department of Corrections of prison populations, forecasts of UC enrollment for particular campuses, or forecasts of State Budget revenues and/or expenditures by the Legislative Analyst's Office

Most of the exogenous variables used to drive the county level forecasts come from the UCLA Anderson forecast for the Nation and the State. These forecasts are updated four times per year. We use the most recent update, typically the June forecast of each year to drive the county level forecasts which are routinely completed in September or October of the same year. This part of the modeling infrastructure that we have developed over the years is entirely in place. Therefore, new exogenous forecasts from UCLA can be incorporated into our County models within a day or two of their release.

The local county demographic variables include age specific populations that are estimated by the Department of Finance, Demographic Research Unit every 2 years. They produce forecasts for these age specific population indicators through the year 2060.

Housing sale and price indicators are developed in an independent housing model for California and the Nation. See below.

HOUSING MODEL

The purpose of the housing model is to forecast home prices and existing home sales for California, because these forecasts are not part of the UCLA Anderson Forecast for the State and Nation.

The housing model uses exogenous inputs from the UCLA California and National forecasts.

Mortgage rates and economic variables such as employment, income, and building are used to predict the future direction of the housing market in terms of sales and housing values. The model does not attempt to forecast future housing cycles, but rather provides reasonable trend forecasts for what can be expected given the future demand for housing, plausible income estimates, and availability or constraints on supply represented by new home production.

The national median home value for new housing and for existing housing are interrelated. An exogenous forecast for one will provide us with an attendant forecast for the other.

Methodology Update

Median Home Price
1980 - 2050
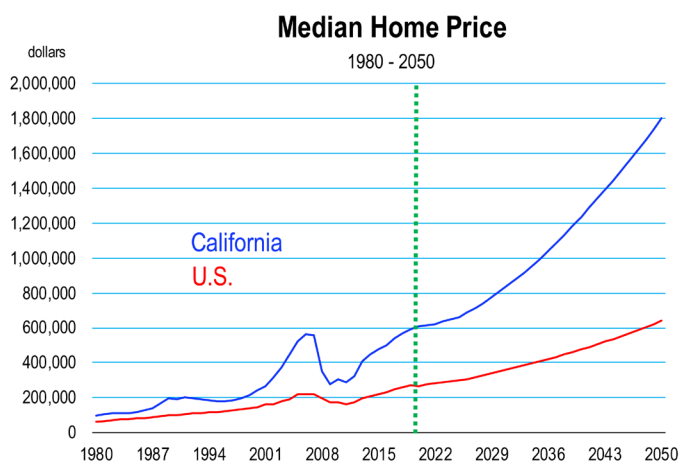
Furthermore, the movement in national home values is correlated with movements in statewide housing prices. The U.S. median home selling value for existing homes has a strong correlation with median home selling values in California.

Home sales in California are correlated with the rate of home sales in most counties (because after all, home sales in the counties are the component parts of statewide home sales). And the variation in California homes sales can be explained by statewide forces such as job creation or demographic trends.

However, the variation in county home sales in further influenced by local job creation, population growth, and homeowner distress that might be specific to a particular county, such as a natural disaster or the departure (or arrival) of a large and significant employer.

The forecast for California home values is driven by national home price movements, mortgage rates, and economic factors indicative of the business cycle. When the California home price forecast is used to drive the county level home price, all of the factors that produced the California forecast are embodied in the county level home price forecast. And much of the variation is explained by statewide housing price movements.

But other local influences such as housing production, job creation, or population growth may also account for specific within-county variation in home prices, as they do with home sales.

Because of the critical importance that homeowner distress had on the housing market during the Great Recession, notices of default and foreclosures were added to the housing model. They have been much less important over the last 10 years as selling values have steadily risen in California and home foreclosures have largely been irrelevant over this time period. Instead, the demand for housing has been relatively strong in tandem with job and income creation. Hence, home prices have been nearly runaway over the last decade in California.

Note of the direction of the housing market in 2020

The actual direction of the housing market during calendar 2020 indicates relatively strong demand resulting in an accelerated pace of housing price growth. Because the forecast for new housing supply remains constrained in many areas of California for the foreseeable future, the modest pace of demand growth will push general housing values higher in California, and in most if not all of the regional housing markets in the state.

DATABASE, DATA SOURCES

The database is an extensive collection of County-level economic and demographic variables from a myriad of sources in California. The database spans the period: 1947 to 2019 (though for most indicators, the data series begins in the 1980s).

Indicators in all County Models (Primary Data Source)

Taxable Retail Sales (Department of Tax and Fee Administration)
Retail Store Outlets (Department of Tax and Fee Administration)
Total Taxable Sales (Department of Tax and Fee Administration)
Personal Income (Bureau of Economic Analysis)
Components of Personal Income (Bureau of Economic Analysis
Total Employment (Employment Development Department)
Employment by Sector (Employment Development Department)
Unemployment Rate (Employment Development Department)
Vehicle Registrations (Department of Motor Vehicles)
Births, Deaths (Centers for Disease Control and Prevention)
Population, Net Migration (Department of Finance)
Population by Age Group (Department of Finance)
Residential building permits (CIRB)
Non-residential bldg. Permits (CIRB)
Median Home Selling Price (Corelogic)
Home Sales (Corelogic)
Agricultural Production (County Agricultural Commissioners)
Households (Department of Finance)

Methodology Update

Indicators in all County Models (Primary Data Source) *continued*

Housing Stock (Department of Finance)
Industrial Production (Bureau of Economic Analysis)
Registered Vehicles (Department of Motor Vehicles)
Los Angeles Area Consumer Price Index (Bureau of Labor Statistics)
Bay Area Consumer Price Index (Bureau of Labor Statistics)
California Consumer Price Index (Bureau of Labor Statistics)
Public School Enrollment (Department of Education)

All county-level dollar variables are deflated using the local consumer price deflator or the statewide price deflator. The base year is the most recent calendar year just completed.

Indicators in Select County Profiles (Primary Data Source) and Relevant Counties

Oil Prices (Department of Energy)
• Kern
Airline Passengers (SFO, LAX, San Diego Intl. Airport)
• San Mateo
• San Diego
• Los Angeles
Wine Grape Production (US Department of Agriculture)
• Napa
• Sonoma
California General Fund Balance (Legislative Analyst's Office
• Sacramento
University Enrollment (UC and CSU Systems)
• Yolo
• Butte
Prison Staffing and Population (Department of Corrections)
• Del Norte
• Lassen
• Amador
Cannabis Permits (Dept. of Food and Agriculture)
• Humboldt
• Trinity
• Mendocino

2019 Data

Every county model begins with collection of the most readily available data. All data updates include revisions to the historical data by the issuing agency. For all counties, identical sources and vintages of data are used, organized as follows:

• Agriculture
• Building (Units permitted, residential and commercial values)
• Vehicle registrations
• Housing
• Population
• Income
• Taxable sales and permits
• Employment by industry

Agriculture

Agriculture includes the total value of all crop production in the county, excluding timber. The crop reports for every county have been released through 2018 by each county's agricultural commissioner. To estimate 2019 values, trends from the previous 10 years of data were used.

Building

All building data comes from the Construction Industry Research Board. The data includes the number of new single and multi-family unit building permits issued, and their corresponding values, the values of all new non-residential permits (broken down into commercial, industrial, and other), and the value of all renovation activity on residential and non-residential structures. All building data is updated through the calendar 2019 year.

Vehicle Registrations

Vehicle registration data comes from the California Department of Motor Vehicles. The county data includes the number of automobiles, trucks, trailers, and motorcycles. The county total represents the sum of these categories. The most current annual vehicle registration data is available through 2019.

Housing

The housing section is divided into two parts. First, home sales and median selling price data come from Corelogic. Second, housing stock data, including total units, single-family units, multi-family units, and the number of households are acquired from the California Department of Finance, Demographic Research Unit.

Methodology Update

Population

Births and deaths are obtained from the Centers for Disease Control and Prevention (available through 2018 and estimated for 2019). Net migration is acquired from the California Department of Finance (available through 2019). Population by age group is acquired from the California Department of Finance (available through 2019, with forecasts through 2050).

Income

Components of personal income are acquired from the Bureau of Economic Analysis. The latest information is through 2018; consequently, the 2019 components of income must be estimated. These data are estimated using regression models that have been built for each county in California, where county income is regressed against State income (and other indicators if appropriate).

Taxable Retail Sales and Total Taxable Sales

The California Department of Tax and Fee Administration is the source for taxable retail sales and total taxable sales by county. Data is available for the 2019 calendar year.

Employment By Industry

The Labor Market Information Division of the Employment Development Department provides estimates of employment by county.

The Current Employment Statistics (CES) Program issues monthly estimates of employment by 2 and 3 digit NAICS. We use 2 digit NAICS employment in each county-level forecast model. Currently, the data is updated through calendar 2019 for use in the current year (2020) model assignment. However, because preliminary CES data is published monthly, we are able to evaluate how the first year employment forecast is tracking the actual information for the partial (current) year in which the forecast is being produced. We are able to make adjustments to the county models so that the first year forecast moves in alignment with the cumulative CES data for the current year.

In some county profiles, tables are included that show industry employment at 3 digit, 4 digit, 5 digit, or 6 digit NAICS sectors. The source for these data is the Quarterly Census of Employment and Wages. The most recent year for these data is the 2019 calendar year.