

General, Age-Specific, and Total Fertility Rates in Montana, 2008-2017

Matthew Ringel, M.P.H.

Montana Department of Public Health and Human Services

ABSTRACT

Fertility rates have been steadily declining in the United States over the past several years, and this trend may have implications for the economy and for other elements of society. Some are concerned as declining fertility rates may be indicative of economic anxiety and could also cause future problems for the economy due to dwindling number of workers and consumers. Other see this as a positive change because it could indicate higher education and career achievement among women and it could lead to decreased consumption and pollution. Either way, this trend may be of interest to federal, state, and local governments. This paper shows readers how to use SAS to calculate general, age-specific, and total fertility rates and how to use linear regression to test for a significant trend. It uses Montana Birth Data and National Center for Health Statistics (NCHS) Population Data to illustrate the process. In Montana, between 2008 and 2017, there was a significant downward trend in both the general fertility rate and the total fertility rate. This trend was largely due to a substantial reduction in births among women between the ages of 15 and 24 and may be due to women delaying childbirth for educational purposes. Therefore, there is reason to believe that the current low total fertility rate may be temporary. This trend should be monitored into the future to see if it continues or reverses course.

INTRODUCTION

The total fertility rate in the United States has been steadily declining over the past several years, from 1.93 births per woman in 2010 to 1.77 births per woman in 2017 (Martin et al., 2018). Economists have taken note of this, as declining fertility rates may both contribute to and exacerbate economic downturns. People who are struggling financially may choose to have fewer children to compensate; and if enough families do this, the diminished workforce and number of consumers can lead to even further economic troubles (Sobotka, Skirbekk, and Philipov, 2009; Morgan, 2003). Declining rates of childbirth may also have some positive effects on society. Women may be choosing to have fewer children in order to focus on their educations and careers (Norville, Gomez, and Brown, 2003). In this case, the economic effects of a more well-educated and more experienced workforce could compensate for the smaller workforce. A smaller population may also produce less pollution and be better for the environment. For these reasons, this is a trend that federal, state, and local governments may also be interested in monitoring. This paper discusses three different types of fertility rates (general, age-specific, and total) and shows readers how to use both state health department birth data and NCHS population data to calculate fertility rates for their own states. It then shows how linear regression was performed on Montana data to test for a significant trend.

BIRTH DATA

State governments produce a birth certificate for each person born within the state, and data from birth records is collected and stored by the state for myriad purposes, including the production of public health statistics. States also send other states data regarding those states' residents. For example, if a Montana resident gives birth in Washington, Washington will share that data with Montana. This allows each state to produce statistics regarding its own residents. States have differing laws regarding the extent to which birth data is protected, but most states are willing to produce deidentified data or aggregate statistics to interested parties, often for a fee.

When calculating fertility rates, birth data is necessary to determine the numerators. This data should be limited to residents of the area of interest, whether that be a country, a state, or a county. This data should also be limited to births to women between the ages of 15 and 44, inclusive. This is because fertility rates only include women of reproductive age in the denominator and should only include them in the numerator for consistency. Since the vast majority of births occur to women between the ages of 15

and 44, this range is treated as the reproductive age. A where clause should be used to limit the birth data accordingly.

The following is an example of what that where clause might look like when the population of interest is a particular state, where state_of_residence is a variable indicating the mother's state of residence, X will likely be a number that corresponds with your state, and mother_age is the age of the mother:

```
where state_of_residence = X and mother_age >= 15 and mother_age <= 44;
```

Since the value of each numerator is the number of observations in each year, a good place to start is by adding a numeric variable, such as births, with the value of 1 for each observation. The structured query language procedure (sql) can then be used to sum the observations in each year.

The following is an example of code that can be used to produce numerators for each year of interest, where births is a variable with a value of 1 and birth_year is each baby's year of birth:

```
proc sql;
  create table births as
  select birth_year as year, sum (births) as all_births
  from birth
  group by year;
quit;
```

POPULATION DATA

Population estimates are typically obtained either directly or indirectly through the U.S. Census Bureau. NCHS population data is based on U.S. census data, but race is categorized differently. If race is not included in the analysis, the census estimates and the NCHS estimates should be the same, and NCHS population data are publicly available.

When calculating fertility rates, population data are necessary to produce the denominators. These data should be limited to women between the ages of 15 and 44, inclusive.

The following is an example of what a where clause to limit this dataset might look like, where X is a number corresponding to the female sex:

```
where sex = X and age >= 15 and age <= 44;
```

The population dataset is unlike the birth dataset in that it is not one line per person, but rather one line per combination of certain demographic characteristics, such as age, race, and sex. It also includes a variable for the population estimate for each of those specific demographic combinations. Since the value of the denominator is the sum of the values of the population variable for that given year, sql can be used to sum the population values for each year.

The following is an example of code that can be used to produce denominators for each year of interest, where popcount is the population estimate for each combination of demographic characteristics, and year is the year:

```
proc sql;
  create table femalepopulation as
  select year as year, sum (popcount) as female_population
  from population
  group by year;
quit;
```

GENERAL FERTILITY RATE

The general fertility rate is the number of live births per thousand women between the ages of 15 and 44. Once you have numerators and denominators in different datasets, you can merge the datasets and calculate the fertility rate.

The following is an example of code that can be used to merge the numerators and denominators and calculate a general fertility rate, where `general_fertility` is the general fertility rate in births per thousand women of reproductive age:

```
data fertility;
  merge births femalepopulation;
  by year;
  general_fertility = all_births / female_population * 1000;
run;
```

AGE-SPECIFIC FERTILITY RATES

A general fertility rate is usual for understanding how a population is growing or shrinking in the short term, but it can be skewed due to the age of the population. Age-specific fertility rates give a more detailed picture of the situation and are necessary for calculating a total fertility rate and projecting long-term population change. Age-specific fertility rates are similar to general fertility rates, except both the numerator and denominator are limited to women of a smaller age range, typically five-year increments. A good first step for performing this calculation is to flag the births according to the age of the mother.

The following is an example of code that can be used within a data set to flag births to mothers between the ages of fifteen and nineteen, where `mother_age` is the age of the mother in years:

```
if mother_age >= 15 and mother_age <= 19 then fifteen = 1;
```

The following is an example of code that can be used to produce numerators for age-specific fertility rates in addition to the general fertility rate, where `births` is a variable with a value of 1 and `birth_year` is each baby's year of birth:

```
proc sql;
  create table births as
  select birth_year as year, sum (births) as all_births, sum
  (fifteen) as fifteen_births, sum (twenty) as twenty_births, sum
  (twentyfive) as twentyfive_births, sum (thirty) as thirty_births,
  sum (thirtyfive) as thirtyfive_births, sum (forty) as
  forty_births
  from birth
  group by year;
quit;
```

This procedure can also be performed on the population dataset in order to create denominators for the age-specific fertility rates.

TOTAL FERTILITY RATE

The total fertility rate is the estimated number of children a woman will have during her lifetime assuming age-specific fertility rates remain unchanged. The most precise way to calculate it is by adding each of the single-year fertility rates but five-year fertility rates are an acceptable proxy. If five-year rates are used instead of single-year rates, the result must be multiplied by five.

Below is an example of a line of code that can be added to the data step shown above to calculate total fertility rates, where `total_fertility` is the total fertility rate and the other variables are the age-specific fertility rates:

```
total_fertility = (fifteen_fertility + twenty_fertility +
  twentyfive_fertility + thirty_fertility + thirtyfive_fertility +
  forty_fertility) * 5 / 1000;
```

The result was divided by 1000 because age-specific fertility rates are expressed in births per thousand women whereas a total fertility rate is the lifetime estimate for a single woman. Therefore, the result must be divided by 1000 after it is multiplied by five.

TESTING THE TREND

After calculating fertility rates for each of the years, you can use linear regression to test the trend for a significant increase or decrease. Linear regression is used to test for a correlation between any number of independent variables and a dependent variable. Independent variables are not influenced by the other variables but may influence the independent variable. For example, a person's age and sex (independent variables) play a role in determining a person's height (dependent variable), but not the other way around. When determining if there is a significant trend (either upward or downward) in the fertility rate, you are testing for a correlation between a dependent variable (fertility rate) and an independent variable (year). You can determine if a trend is significant by calculating the two-tailed p-value for the association. The p-value is the probability of a having trendline that steep or steeper due to chance alone. A value of less than 0.05 is generally accepted as a significant trend. A two-tailed p-value is necessary, as opposed to a one-tailed p-value, because you want to consider the possibility of either an upward or a downward trend, not just one or the other.

Below is an example of code that can be used to test the trend for the general fertility rate, each age-specific fertility rate, and the total fertility rate:

```
proc reg data = fertility;
    model general_fertility = year;
    model fifteen_fertility = year;
    model twenty_fertility = year;
    model twentyfive_fertility = year;
    model thirty_fertility = year;
    model thirtyfive_fertility = year;
    model forty_fertility = year;
    model total_fertility = year;
run;
```

Running this code may result in the following message within the log:

```
WARNING: The range of variable YEAR is so small relative to its mean that there may be
loss of accuracy in the computations. You may need to rescale the variable to have a
larger value of RANGE/abs(MEAN), for example, by using PROC STANDARD M=0;
```

This can be resolved in the data step by subtracting all values of year by the lowest value. This would result in the earliest year being year zero. A format can then be applied to restore the original values of year in the output.

Figure 1 shows the trend in Montana and can serve as an example of what output for a linear regression model might look like in general. It also shows an important statistic, R-Square. R^2 is the percent of the variation within the dependent variable that can be explained by the model. In this case, the model has only one independent variable, year. Therefore, 73% of the change in Montana's total fertility rate can be attributed to the year.

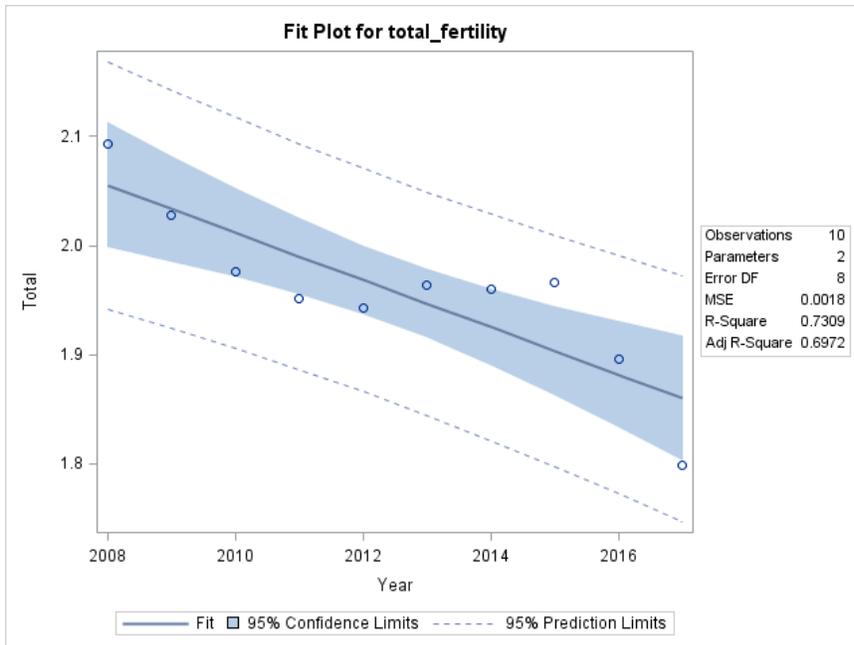


Figure 1. Fit Plot for Total Fertility in Montana

Figure 2 shows the table of parameters. This can be used to estimate the value of the dependent variable given certain values of the independent variables. It also shows you two-tailed p-values for each variable.

Parameter Estimates						
Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	Intercept	1	2.05517	0.02487	82.62	<.0001
year	Year	1	-0.02172	0.00466	-4.66	0.0016

Figure 2. Parameter Estimates for Total Fertility in Montana

MONTANA RESULTS

Figure 3 shows the general fertility rate, each of the age-specific fertility rates, and the total fertility rate for the years 2008-2017.

Year	General	Age 15 - 19	Age 20 - 24	Age 25 - 29	Age 30 - 34	Age 35 - 39	Age 40 - 44	Total
2008	69.6	38.9	105.4	131.1	96.9	38.9	7.4	2.09
2009	68.3	38.4	102.5	123.9	94.1	38.9	7.6	2.03
2010	67.1	35.2	96.9	116.8	96.5	42.2	7.6	1.98
2011	66.6	29.3	96.3	120.2	95.9	40.0	8.6	1.95
2012	66.2	28.7	89.8	121.4	96.1	44.8	7.7	1.94
2013	67.3	27.8	91.1	126.4	95.6	44.1	7.7	1.96
2014	67.3	26.7	83.6	126.8	101.4	43.8	9.6	1.96
2015	67.7	25.6	82.2	124.5	105.3	46.1	9.4	1.97
2016	65.5	23.8	78.2	119.3	101.0	46.5	10.4	1.90
2017	62.1	21.3	73.5	108.5	101.1	45.6	9.6	1.80

Figure 3. Fertility rates in Montana

The general fertility rate has significantly trended downward

(Parameter Estimate = -0.49, $R^2 = 0.55$, $p = 0.014$).

The fertility rate among women age 15-19 has significantly trended downward

(Parameter Estimate = -1.93, $R^2 = 0.93$, $p < 0.001$).

The fertility rate among women age 20-24 has significantly trended downward

(Parameter Estimate = -3.44, $R^2 = 0.98$, $p < 0.001$).

The fertility rate among women age 25-29 has had no significant linear trend.

(Parameter Estimate = -1.05, $R^2 = 0.26$, $p = 0.136$).

The fertility rate among women age 30-34 has significantly trended upward

(Parameter Estimate = 0.89, $R^2 = 0.57$, $p = 0.012$).

The fertility rate among women age 35-39 has significantly trended upward

(Parameter Estimate = 0.87, $R^2 = 0.81$, $p < 0.001$).

The fertility rate among women age 40-44 has significantly trended upward

(Parameter Estimate = 0.31, $R^2 = 0.72$, $p = 0.002$).

The total fertility rate has significantly trended downward

(Parameter Estimate = -0.02, $R^2 = 0.73$, $p = 0.002$).

CONCLUSION

Both the general and the total fertility rate have been declining in Montana over the past decade. This is due primarily to a substantial decrease in births to women between the ages of 15 and 24. On the contrary, there has been a smaller yet significant upward trend in births to women between the ages of 30 and 34. Since ages 15 to 24 roughly correspond to the high school and college years, a plausible explanation for Montana's declining fertility rate is greater educational achievement among Montana women. It is possible that the decline in the total fertility rate is due to delayed childbirth instead of forgone childbirth. Such a phenomenon would produce a downward spike in the total fertility rate that

would eventually rebound (Kohler and Ortega, 2002). This was observed in the 1970's when the U.S. total fertility rate reached a low of 1.76 (Martin et al., 2017). Future research could investigate this possibility by exploring how educational attainment among women has changed during the past decade. Future research should also explore ways to enable women who would like to have more children be able to accomplish this. Such research could investigate the effectiveness of interventions such as workplace daycare, tax credits, and paid family leave.

REFERENCES

- Kohler, HP; Ortega, JA. 2002. "Tempo-Adjusted Period Parity Progression Measures, Fertility Postponement and Completed Cohort Fertility." *Demographic Research*, 6(6):91–144.
- Martin, JA et al. 2017. "Births: Final Data for 2015" *National Vital Statistics Reports*, 66(1). National Center for Health Statistics.
- Martin, JA et al. 2018. "Births: Final Data for 2017" *National Vital Statistics Reports*, 67(8). National Center for Health Statistics.
- Morgan, SP. 2003. "Is Low Fertility a Twenty-First-Century Demographic Crisis." *Demography*, 40(4):589–603.
- Norville, C; Gomez, R; Brown, RL. 2003. "Some Causes of Fertility Rates Movements." University of Waterloo.
- Sobotka, T; Skirbekk, V; Philipov, Dimiter. 2009. "Economic Recession and Fertility in the Developed World." *Population and Development Review*, 37(2):267–306.