SIMPLE VERSUS COMPLEX MODELS

Anne Morse, PhD
Estimates and Projections Area
Population Division

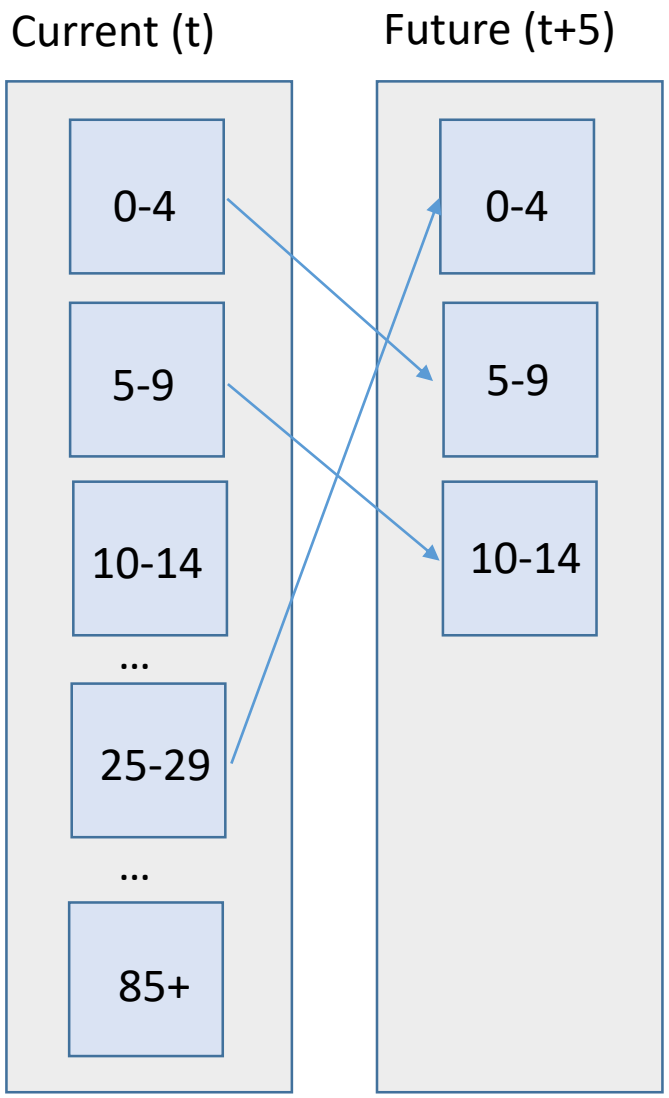
Outline

- Cohort-component model
- Considerations other than accuracy
- Definition of simplicity
- Theoretical reasons for the [possible] superiority of simplicity
  - The nature of the future/time-span
  - Decomposition/causal models
  - Overfitting/mistaking noise for signal

United States Census Bureau

The cohort-component model

- Projects cohorts forward in time by their components



United States Census Bureau

Considerations other than accuracy

- Internal consistency
- Use of most recent data
- Consideration of relevant variables
- Cost of development
- Ease of explanation
- Usefulness for policymaking

Source: Smith, Stanley K. "Further thoughts on simplicity and complexity in population projection models." *International Journal of Forecasting* 13.4 (1997): 557-565.

Considerations other than accuracy (cont.)

- Journals favor complexity
- Clients may be reassured by incomprehensibility

Sources:
Smith, Stanley K. "Further thoughts on simplicity and complexity in population projection models." *International journal of forecasting* 13.4 (1997): 557-565.
Green, Kesten C., and J. Scott Armstrong. "Simple versus complex forecasting: The evidence." *Journal of Business Research* 68.8 (2015): 1678-1685.

Considerations other than accuracy (cont.)

- Journals favor complexity
- Clients may be reassured by incomprehensibility

"The use of complex models and techniques may make the analyst appear to be intelligent and well informed….complex models and techniques also provide an imposing array of details behind which to hide when things go wrong

Sources:
Smith, Stanley K. "Further thoughts on simplicity and complexity in population projection models." *International journal of forecasting* 13.4 (1997): 557-565.
Green, Kesten C., and J. Scott Armstrong. "Simple versus complex forecasting: The evidence." *Journal of Business Research* 68.8 (2015): 1678-1685.

Considerations other than accuracy (cont.)

- Journals favor complexity
- Clients may be reassured by incomprehensibility

"The use of complex models and techniques may make the analyst appear to be intelligent and well informed….complex models and techniques also provide an imposing array of details behind which to hide when things go wrong, whereas simple models and techniques are transparent and may leave the analyst open to charges of overlooking important factors that could have improved the forecast.

Sources:
Smith, Stanley K. "Further thoughts on simplicity and complexity in population projection models." *International journal of forecasting* 13.4 (1997): 557-565.
Green, Kesten C., and J. Scott Armstrong. "Simple versus complex forecasting: The evidence." *Journal of Business Research* 68.8 (2015): 1678-1685.

Considerations other than accuracy (cont.)

- Journals favor complexity
- Clients may be reassured by incomprehensibility

"The use of complex models and techniques may make the analyst appear to be intelligent and well informed….complex models and techniques also provide an imposing array of details behind which to hide when things go wrong, whereas simple models and techniques are transparent and may leave the analyst open to charges of overlooking important factors that could have improved the forecast. Analysts who use relatively simple models or techniques not only must have the expertise to know when they are appropriate and how best to use them, but sometimes may need a thick skin as well."

Sources:
Smith, Stanley K. "Further thoughts on simplicity and complexity in population projection models." *International journal of forecasting* 13.4 (1997): 557-565.
Green, Kesten C., and J. Scott Armstrong. "Simple versus complex forecasting: The evidence." *Journal of Business Research* 68.8 (2015): 1678-1685.

United States Census Bureau

Definitions of simplicity and complexity

Simple	Complex
Depend only on past values	Include predictive covariates (Rogers, 1995; Smith and Sinich, 1992)
Linear and exponential extrapolations	Cohort-component, ARIMA time series, causal (Smith 1997)
Short computational time	Long computational time (Makridakis 2020)

Sources:
Green, Kesten C., and J. Scott Armstrong. "Simple versus complex forecasting: The evidence." *Journal of Business Research* 68.8 (2015): 1678-1685.
Makridakis, Spyros, Evangelos Spiliotis, and Vassilios Assimakopoulos. "The M4 Competition: 100,000 time series and 61 forecasting methods." *International Journal of Forecasting* 36.1 (2020): 54-74.
Rogers, Andrei. "Population forecasting: Do simple models outperform complex models?." *Mathematical Population Studies* 5.3 (1995): 187-202.
Smith, Stanley K., and Terry Sincich. "Evaluating the forecast accuracy and bias of alternative population projections for states." International Journal of Forecasting 8.3 (1992): 495-508.

United States® Census Bureau

Definitions of simplicity and complexity (cont.)

Simple	Complex
Depend only on past values	Include predictive covariates (Rogers, 1995; Smith and Sinich, 1992)
Linear and exponential extrapolations	Cohort-component, ARIMA time series, causal (Smith 1997)
Short computational time	Long computational time (Makridakis 2020)
Processes that are understandable to forecast users	Processes that are not understood (Green 2015)

Sources:
Green, Kesten C., and J. Scott Armstrong. "Simple versus complex forecasting: The evidence." *Journal of Business Research* 68.8 (2015): 1678-1685.
Makridakis, Spyros, Evangelos Spiliotis, and Vassilios Assimakopoulos. "The M4 Competition: 100,000 time series and 61 forecasting methods." *International Journal of Forecasting* 36.1 (2020): 54-74.
Rogers, Andrei. "Population forecasting: Do simple models outperform complex models?." *Mathematical Population Studies* 5.3 (1995): 187-202.
Smith, Stanley K., and Terry Sincich. "Evaluating the forecast accuracy and bias of alternative population projections for states." International Journal of Forecasting 8.3 (1992): 495-508.

United States® Census Bureau

1. The reality is simple

- Simple models:
  - No-change models
  - Linear extrapolation models

- No-change models may be hard to beat
  - Knowledge is insufficient
  - The situation is stable

Sources:
Green, Kesten C., and J. Scott Armstrong. "Simple versus complex forecasting: The evidence." *Journal of Business Research* 68.8 (2015): 1678-1685.

Pant, P. Narayan, and William H. Starbuck. "Innocents in the forest: Forecasting and research methods." *Journal of Management* 16.2 (1990): 433-460.
Rogers, Andrei. "Population forecasting: Do simple models outperform complex models?." *Mathematical Population Studies* 5.3 (1995): 187-202.

1. The reality is simple (cont.)

Why might no-change or linear extrapolation models produce more accurate results than complex models?

- The reality (especially in the short run) is often simple
- Time series are highly autocorrelated
- More recent data are typically more relevant
  - Use exponential smoothing

Sources:
Green, Kesten C., and J. Scott Armstrong. "Simple versus complex forecasting: The evidence." *Journal of Business Research* 68.8 (2015): 1678-1685.

Pant, P. Narayan, and William H. Starbuck. "Innocents in the forest: Forecasting and research methods." *Journal of Management* 16.2 (1990): 433-460.
Rogers, Andrei. "Population forecasting: Do simple models outperform complex models?." *Mathematical Population Studies* 5.3 (1995): 187-202.

2. Decomposition and causal models

- Simple models: directly project the variable of interest
- Complex models:
  - Decompose the variable into parts

    or

  - Project using a casual variable

2. Decomposition and causal models (cont.)

Why might decomposing or using a causal model produce less accurate results?

- Basic causal factors may act more directly on the total than on the parts

Source:
O'neill, Brian C., et al. "A guide to global population projections." *Demographic Research* 4 (2001): 203-288.

2. Decomposition and causal models (cont.)

Why might decomposing or using a causal model produce less accurate results?

- The data may be worse
  - Of the causal variable
  - Of the decomposed components

- The data may be
  - Short
  - Noisy
  - Non-representative

Source:
O'neill, Brian C., et al. "A guide to global population projections." *Demographic Research* 4 (2001): 203-288.

2. Decomposition and causal models (cont.)

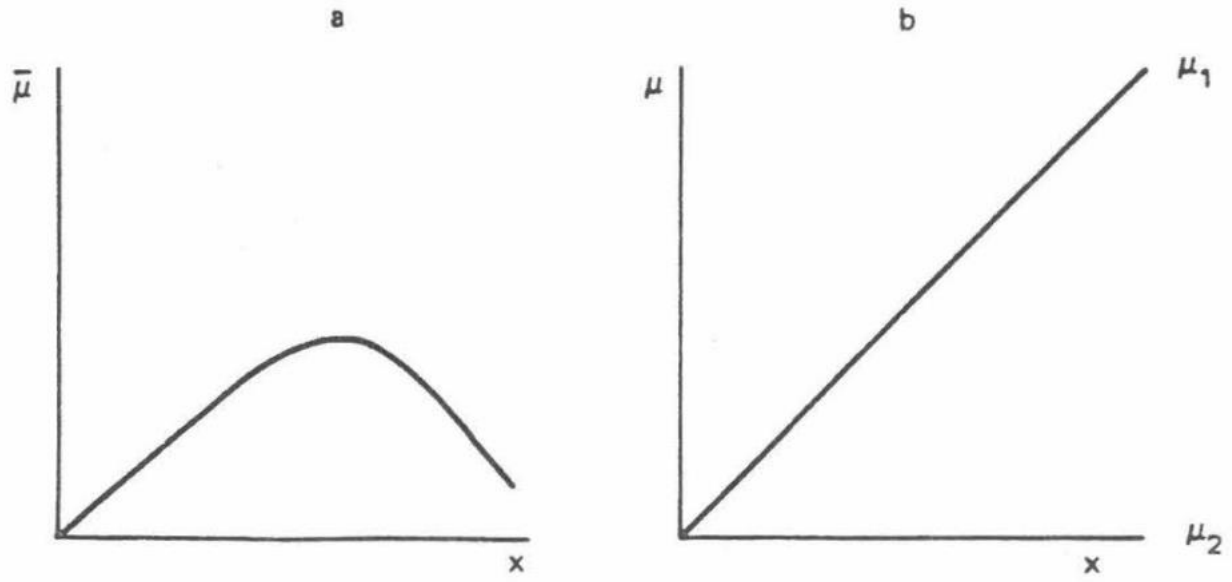


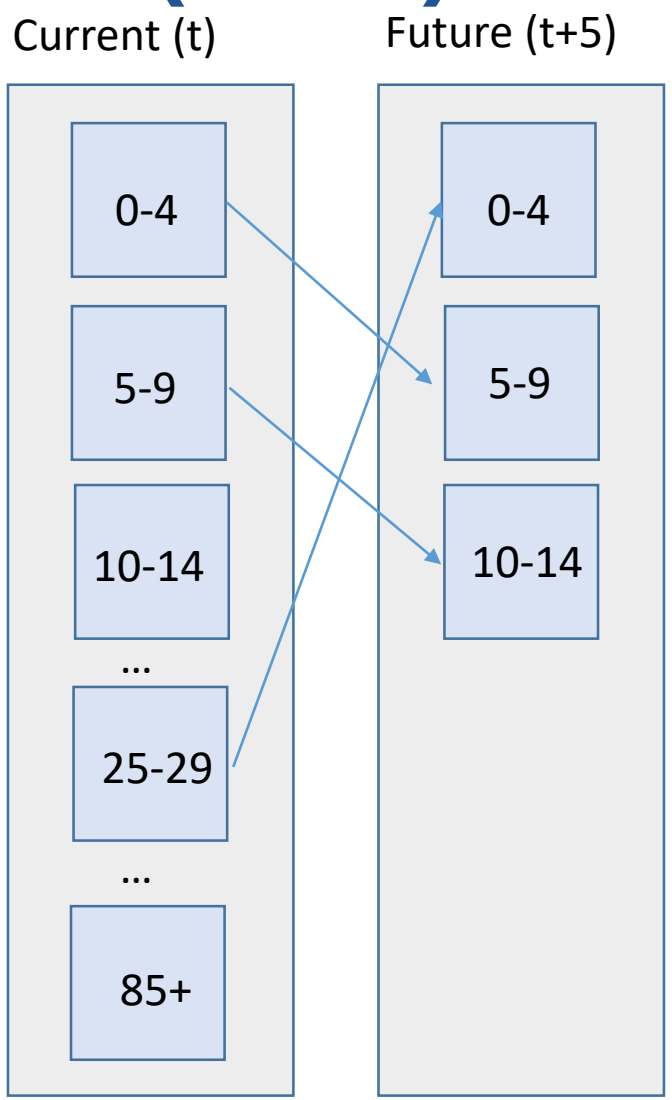
Figure 5. The observed hazard rate may increase and then decline if the hazard rate for one subcohort is increasing and the other subcohort is immune. The curve for μ was calculated from (2), (3), and (4) using $\mu_1(x) = .002x$, $\mu_2(x) = 0$, and $\pi(0) = .95$. The curves are shown for values of x from 0 to 75.

Sources:
Vaupel, James W., and Anatoli I. Yashin. "Heterogeneity's ruses: some surprising effects of selection on population dynamics." The American Statistician 39.3 (1985): 176-185.

2. Decomposition and causal models (cont.)

Why might decomposing or using a causal model produce less accurate results?

- Input variables may be harder to project than the output variable



Source:
O'neill, Brian C., et al. "A guide to global population projections." *Demographic Research* 4 (2001): 203-288.

2. Decomposition and causal models (cont.)

Why might we decompose even if the input variable is harder to project than the output variable?

- Stakeholder needs
- Increased accuracy
  - Increased accuracy from utilizing that “people age one year at a time” may outweigh loss of accuracy from unreliable inputs

Source:
O'neill, Brian C., et al. "A guide to global population projections." *Demographic Research* 4 (2001): 203-288.

2. Decomposition and causal models (cont.)

- We expect the cohort-component method to be more accurate than a simple method when the trends of the input variables are relatively stable

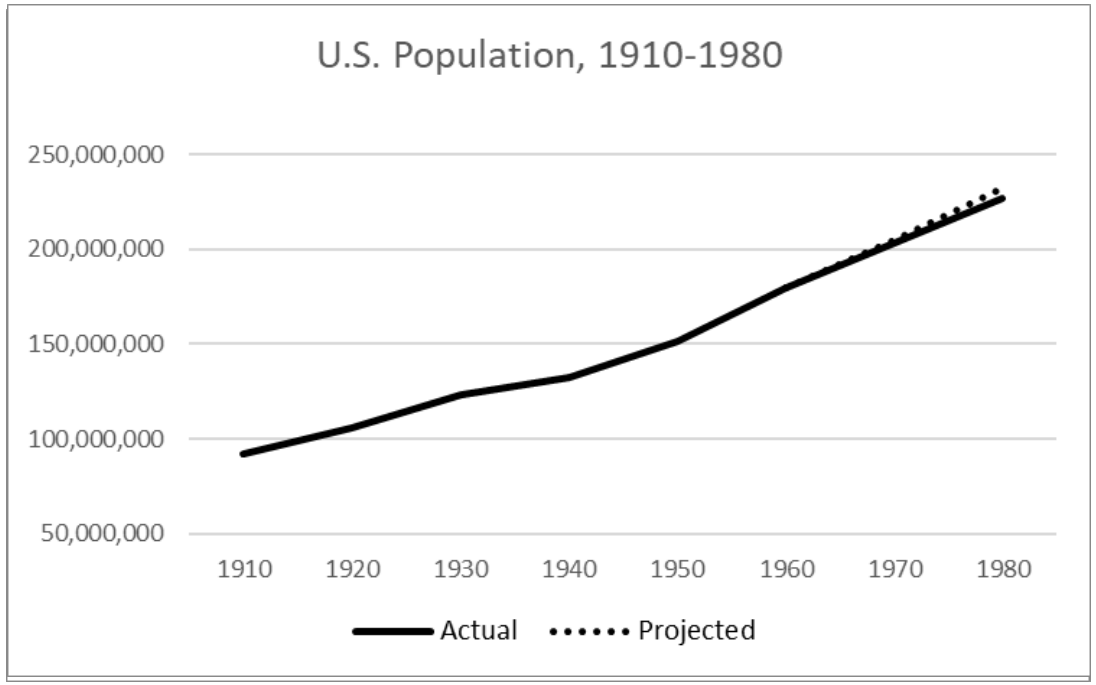
- We expect the cohort-component method to be less accurate than a simple method when there is a major turning point in the input trend
  - This makes the piece more difficult to predict than the whole

Source:
O'neill, Brian C., et al. "A guide to global population projections." *Demographic Research* 4 (2001): 203-288.

2. Decomposition and causal models (cont.)

- We expect the cohort-component method to be less accurate than a simple method when there is a major turning point in the input trend

Projection with a simple (average rate) model



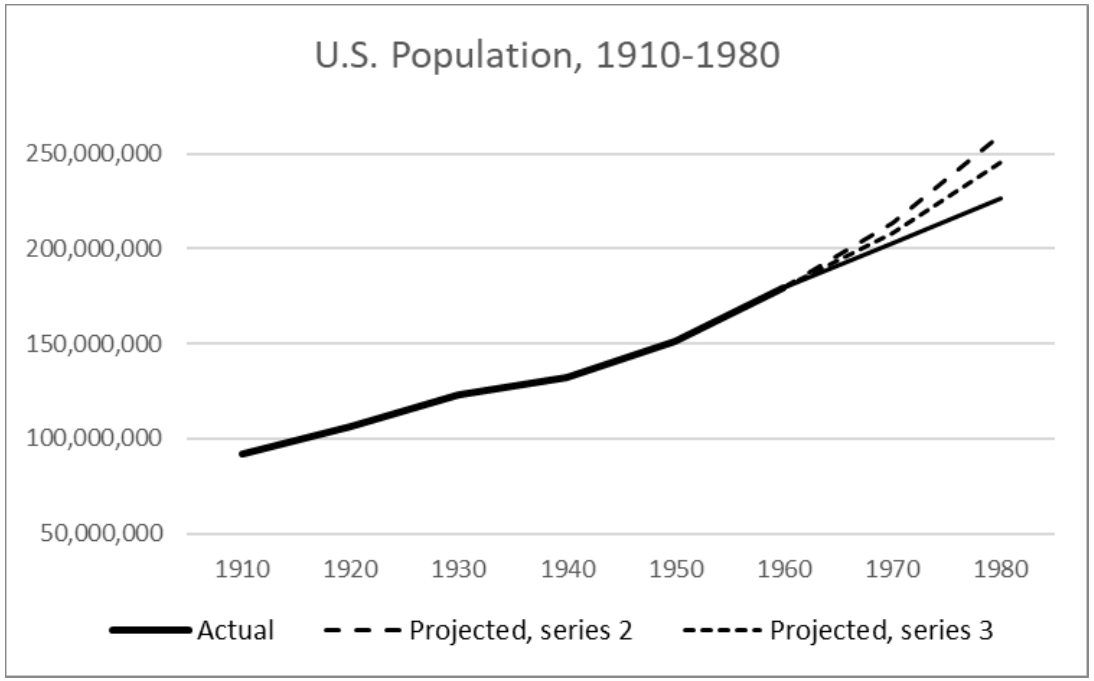
Sources:
Historical Population Change Data (1910-2020), U.S. Census Bureau

Pant, P. Narayan, and William H. Starbuck. "Innocents in the forest: Forecasting and research methods." *Journal of Management* 16.2 (1990): 433-460.
Rogers, Andrei. "Population forecasting: Do simple models outperform complex models?." *Mathematical Population Studies* 5.3 (1995): 187-202.

2. Decomposition and causal models (cont.)

- We expect the cohort-component method to be less accurate than a simple method when there is a major turning point in the input trend

Projection with a complex (cohort-component) model

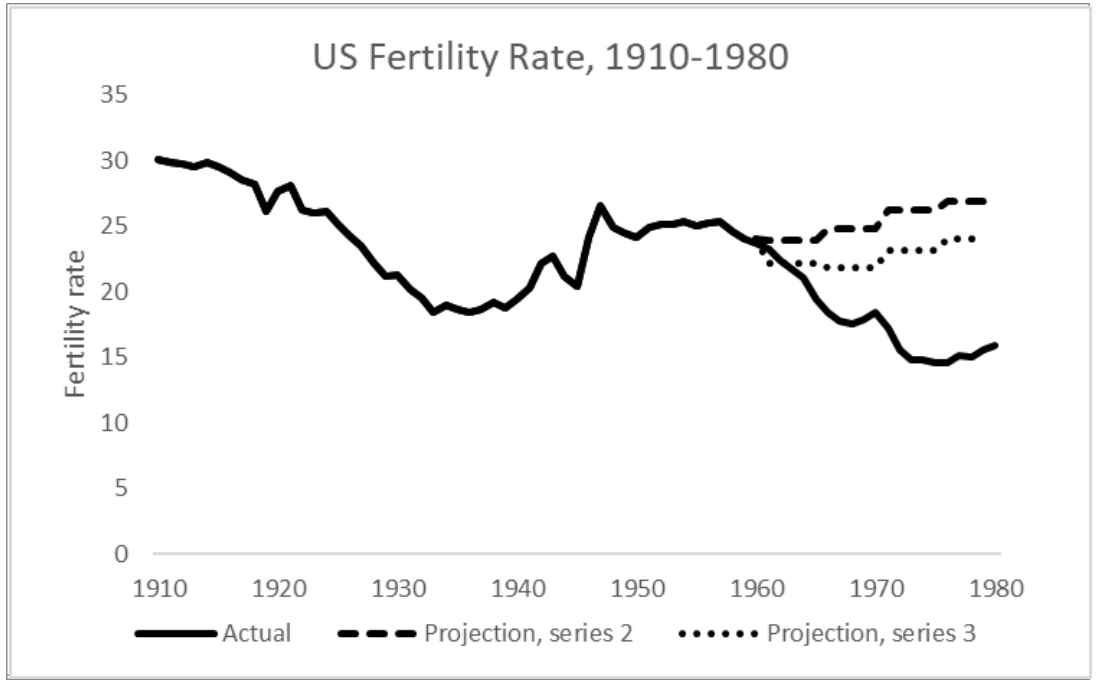


Sources:
Historical Population Change Data (1910-2020), U.S. Census Bureau
Zitter, Meyer and Siegel Jacob. Illustrative Projections of the Population of the United States by Age and Sex 1960 to 1980. *Current Population Reports.* 25.187 (1958)

2. Decomposition and causal models (cont.)

- We expect the cohort-component method to be less accurate than a simple method when there is a major turning point in the input trend

Projection with a complex (cohort-component) model

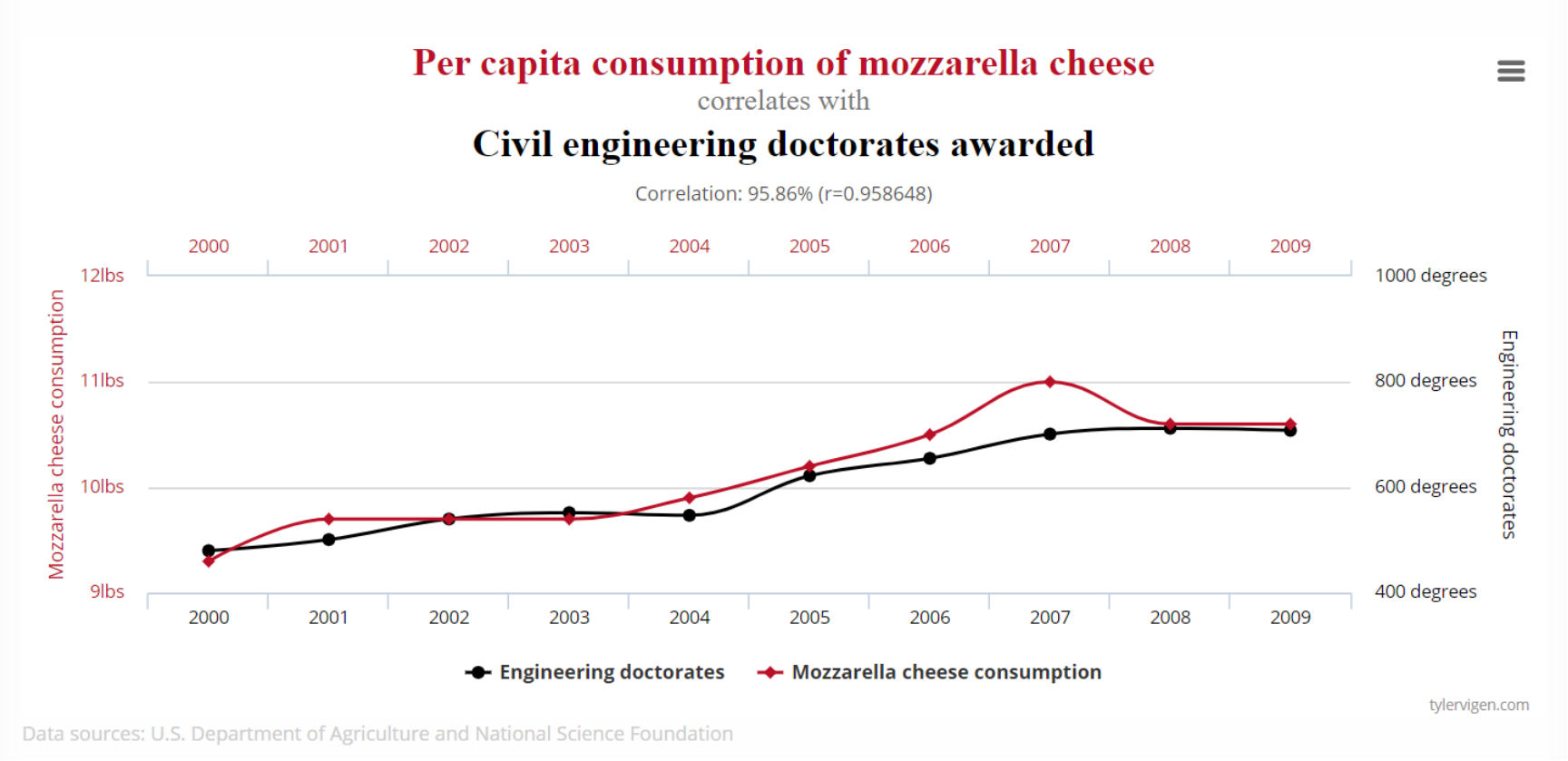


Sources:
Hamilton BE, Lu L, Chong Y, et al. Natality trends in the United States, 1909–2018. National Center for Health Statistics. 2020.
Zitter, Meyer and Siegel Jacob. Illustrative Projections of the Population of the United States by Age and Sex 1960 to 1980. *Current Population Reports.* 25.187 (1958)

2. Decomposition and causal models (cont.)

Why might using a causal model produce less accurate results?

- A weak causal relationship



Source:
Vigen, Tyler. Spurious correlations. Hachette UK, 2015. https://tylervigen.com/spurious-correlations

2. Decomposition and causal models (cont.)

Decomposition is most useful when:

- There is valid and reliable information about each element
- Elements are subject to different causal forces
- Elements are easier to predict than the whole
- The pieces are additive
  - Multiplicative might lead to compounding errors

Sources:

Green, Kesten C., and J. Scott Armstrong. "Simple versus complex forecasting: The evidence." *Journal of Business Research* 68.8 (2015): 1678-1685.

3. Overfitting

- Complex forecasting methods mistake random noise for information
- The key is in determining whether the patterns represent noise or signal
- Need for theory!
- Predictions may be *more prone* to failure in the era of Big data

United States Census Bureau

Sources:
Silver, Nate. The signal and the noise: Why so many predictions fail-but some don't. Penguin, 2012.

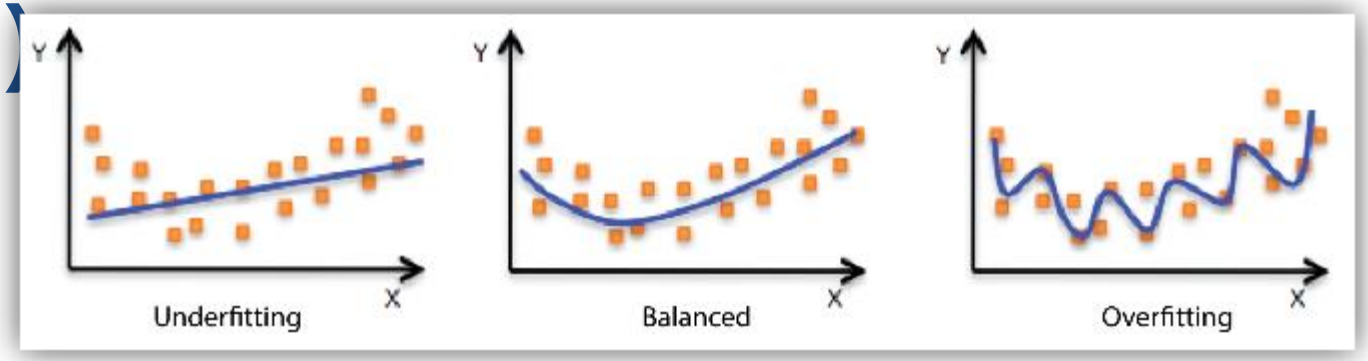
3. Overfitting (cont.)

"More complex methods might promise to extract more information from data, but such methods tend to mistake noise for information. As a result, more complex methods make more serious errors, and they rarely yield the gains they promised."

Sources:

Pant, P. Narayan, and William H. Starbuck. "Innocents in the forest: Forecasting and research methods." *Journal of Management* 16.2 (1990): 433-460.

3. Overfitting (cont.)



Overfitting occurs when the model is fit so closely to the training data that it loses its generalizability

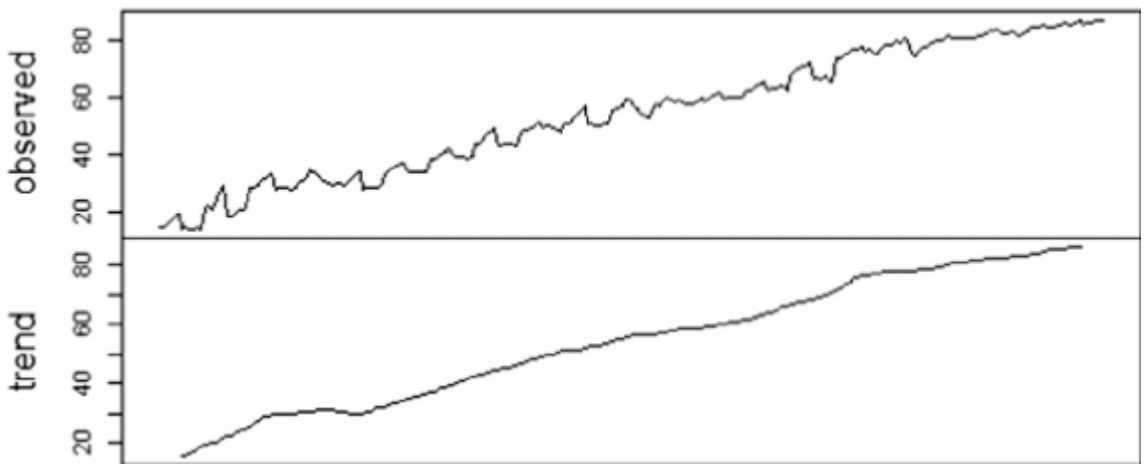
More likely to overfit when:

- Data are noisy
- Poor understanding of fundamental relationships

United States Census Bureau

Sources:
Silver, Nate. The signal and the noise: Why so many predictions fail-but some don't. Penguin, 2012.

3. Overfitting (cont.)



More likely to overfit when:

Data are noisy ➔

Make data less noisy ➔

Smooth and de-seasonalize!

Sources:
Makridakis, Spyros, Evangelos Spiliotis, and Vassilios Assimakopoulos. "Statistical and Machine Learning forecasting methods: Concerns and ways forward." *PloS one* 13.3 (2018):
Pant, P. Narayan, and William H. Starbuck. "Innocents in the forest: Forecasting and research methods." *Journal of Management* 16.2 (1990): 433-460.
Mirzavand, Mohammad, and Reza Ghazavi. "A stochastic modelling technique for groundwater level forecasting in an arid environment using time series methods." *Water Resources Management* 29.4 (2015): 1315-1328.

3. Overfitting (cont.)

Overfitting and measures of fit

"Goodness of fit was not necessarily related to post-sample accuracy. In fact, often the opposite is true."

Squared errors ☹
AIC ☺
Confidence intervals ☺

Sources:
Makridakis, Spyros, et al. "Confidence intervals: An empirical investigation of the series in the M-competition." *International Journal of Forecasting* 3.3-4 (1987): 489-508.
Makridakis, Spyros, Evangelos Spiliotis, and Vassilios Assimakopoulos. "Statistical and Machine Learning forecasting methods: Concerns and ways forward." *PloS one* 13.3 (2018):
Makridakis, Spyros, Evangelos Spiliotis, and Vassilios Assimakopoulos. "The M4 Competition: 100,000 time series and 61 forecasting methods." *International Journal of Forecasting* 36.1 (2020): 54-74.
Pant, P. Narayan, and William H. Starbuck. "Innocents in the forest: Forecasting and research methods." *Journal of Management* 16.2 (1990): 433-460.

United States® Census Bureau

3. Overfitting (cont.)

"There is no reason for such complex methods to be less accurate than simple statistical benchmarks, at least once their shortcoming of over-fitting is corrected."

Sources:
Makridakis, Spyros, Evangelos Spiliotis, and Vassilios Assimakopoulos. "The M4 Competition: 100,000 time series and 61 forecasting methods." *International Journal of Forecasting* 36.1 (2020): 54-74.

3. Overfitting (cont.)

"Men may construe things after their fashion/ Clean from the purpose of the things themselves."

—Shakespeare's *Cicero*

United States Census Bureau

Sources:
Makridakis, Spyros, Evangelos Spiliotis, and Vassilios Assimakopoulos. "The M4 Competition: 100,000 time series and 61 forecasting methods." *International Journal of Forecasting* 36.1 (2020): 54-74.

Conclusion

- Consciously articulate all considerations that influence your choices
- Carefully consider when to include a causal model or when to decompose
  - Is the causal theory strong?
  - Are the data usable?
  - Is the element easy to predict?

United States Census Bureau

Conclusion

- Consciously articulate all considerations that influence your choices
- Carefully consider when to include a causal model or when to decompose
- Get rid of unnecessary noise in your data:
  - Smooth and deseasonalized!
- Don’t overfit your data
  - Use a good measure of fit method
    - Don’t place too much weight on outliers
    - Account for the number of parameters used

United States Census Bureau

Thank you!

Anne Morse

anne.morse@census.gov

Estimates and Projections Area

Population Division

U.S. Census Bureau

United States Census Bureau