LFD Medicaid Model

Julia Pattin

General Approach

Data driven model

- Based on observations from prior years
- Based on observed expenditure trends
- Goal is to optimize accuracy and precision

The Data

The model relies entirely on Medicaid paid claims data

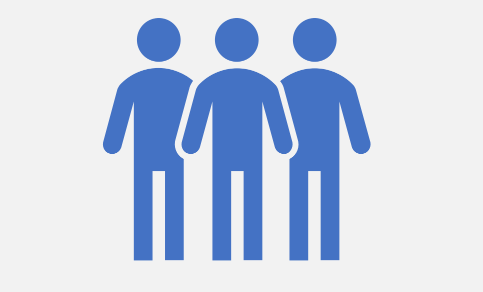
- Data spans from 2004 – now & is reported monthly
- Claims may be paid up to 3 years after the date of service, so the claims data for FY 2018 – FY 2021 is incomplete
  - The values used for FY 2019, FY 2020, and FY 2021 are forecast a rate of payment completion for each service in each month, so these values are informed by actual payments but still have a small amount to be paid and are forecasts, not actual values
- Projected individually for approximately 95 service categories
- A few categories are filled in with estimates from the DPHHS model

The Data

- Some service categories are waiver-based and are available according to the number of spaces authorized by the federal government and funding allocated by the legislature
- When these waivers are expanded, there is an additional step of modeling take-up of these waiver slots, the rate of which is forecast using historic take-up rates according to waiver enrollment

- Use independent economic variables – IHS econometrics as used throughout LFD forecasts, incorporated only when model fit is improved
- Presented as an annual estimate
- Model does not include the HELP Act (Medicaid Expansion) population

The Data - Enrollment



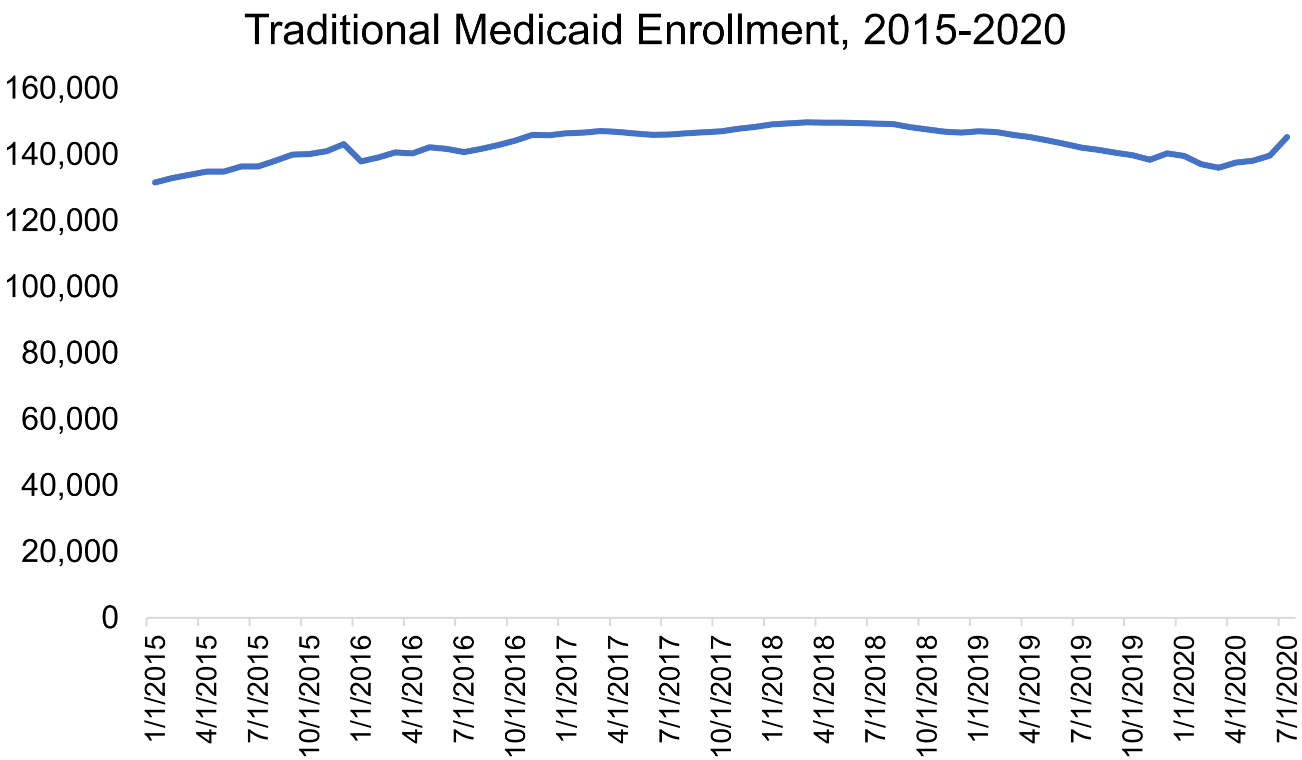
- Enrollment is not directly factored into the model, but it has an indirect effect on predictions

If enrollment increases…

Higher enrollment *generally* results in higher expenditures

Those higher expenditures are then used to predict future expenditures

…then future predictions usually increase too



Time Series Models

What is a time series?

➢ A sequence of values taken at successive, equally-spaced points in time

A time series forecast can be used to predict the likely outcome of a time series, given knowledge of the most recent outcomes

- More recent observations are weighted more heavily
- Implicitly accounts for inflationary factors
- Must satisfy certain statistical assumptions in order to be valid
- Can’t intrinsically adjust for policy changes

Components of a Time Series Model

Trend Component	Cyclical Component	Seasonal Component	Residual Variation
➢ The general tendency of an increase or decrease over time	➢ Medium-term cycles, generally over 2 or more years	➢ Short-term cycles, generally fluctuations within a year	➢ Unpredictable, random influences outside of regular patterns
			

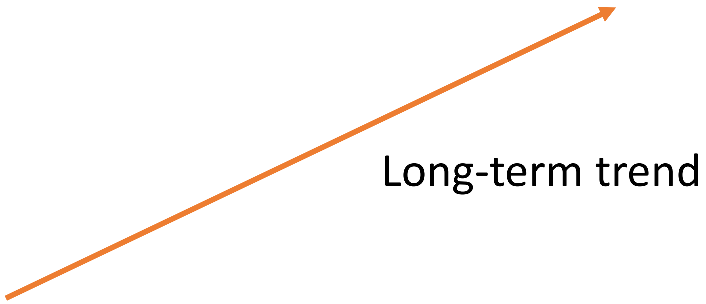
Components of a Time Series Model

Trend

Long-term trend

Components of a Time Series Model

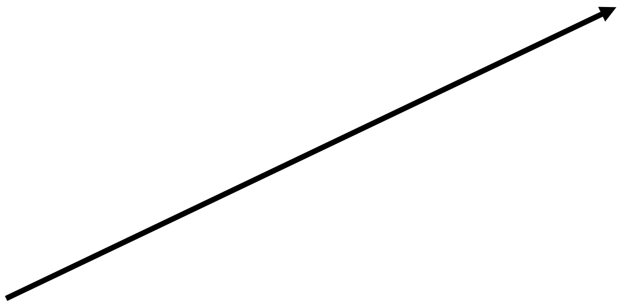
Trend



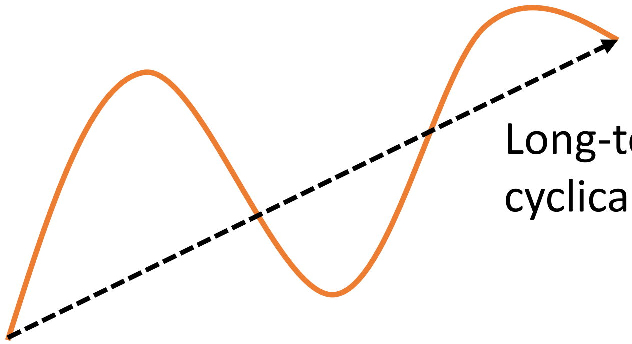
- This **trend** is essentially where general medical inflation is captured

Components of a Time Series Model

Trend

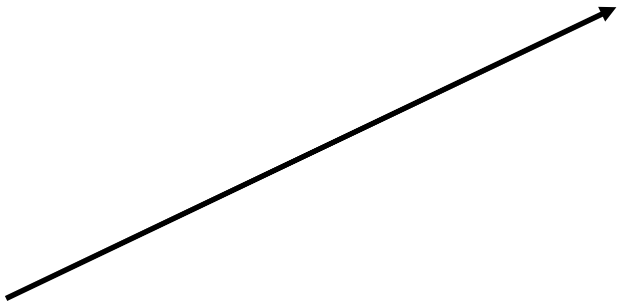


Cyclical

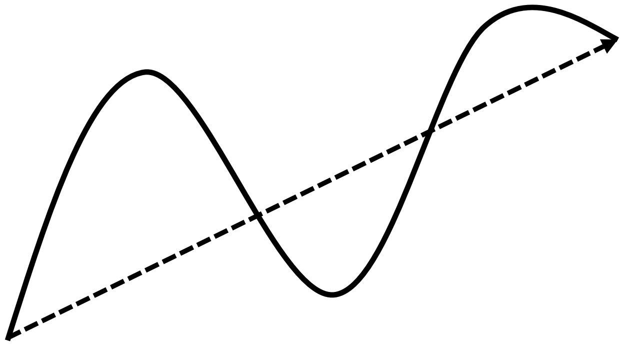


Components of a Time Series Model

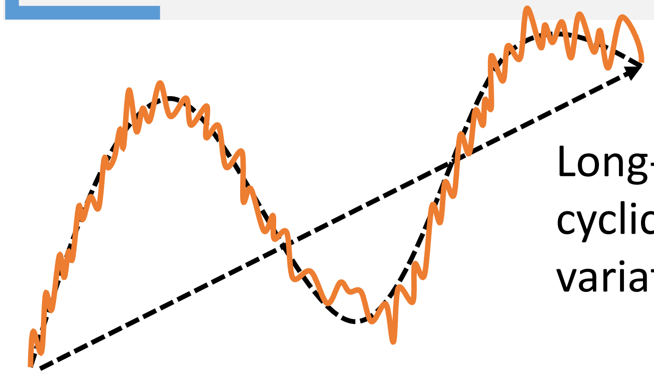
Trend



Cyclical



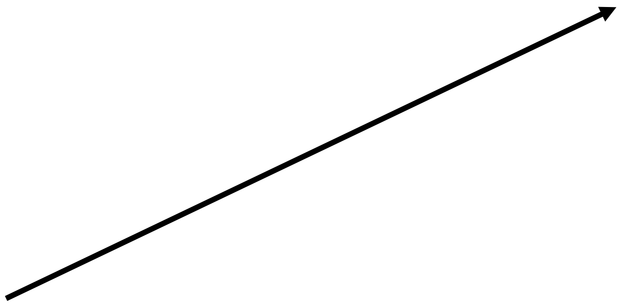
Seasonal



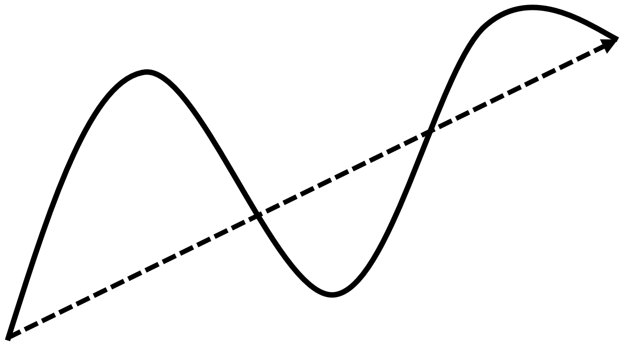
Long-term trend with
cyclical & seasonal
variations

Components of a Time Series Model

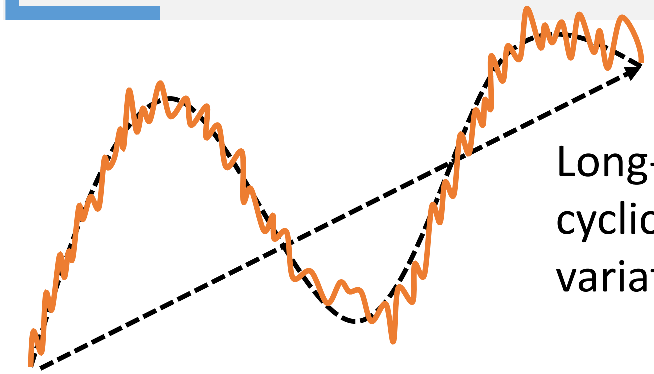
Trend



Cyclical



Seasonal

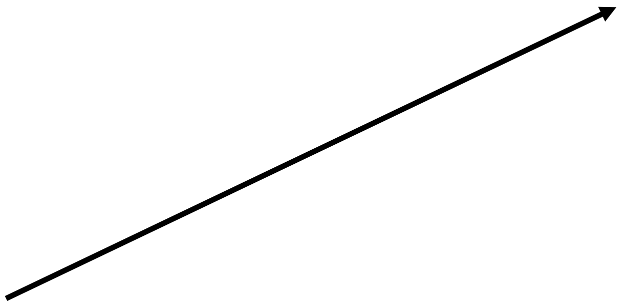


Long-term trend with cyclical & seasonal variations

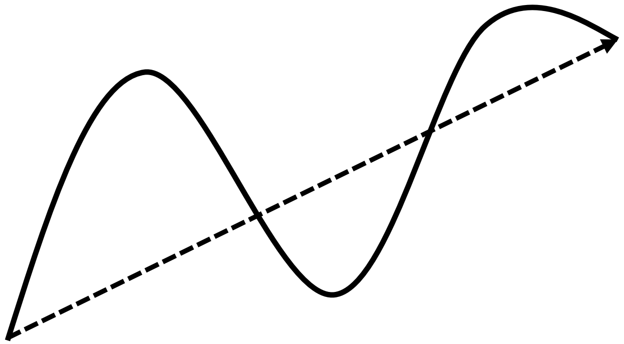
- If a specific service has **seasonal** variation borne out in the data, it can be captured in by the model
  - For example, there may be increased hospital expenditures in the winter due to accidents on icy roads

Components of a Time Series Model

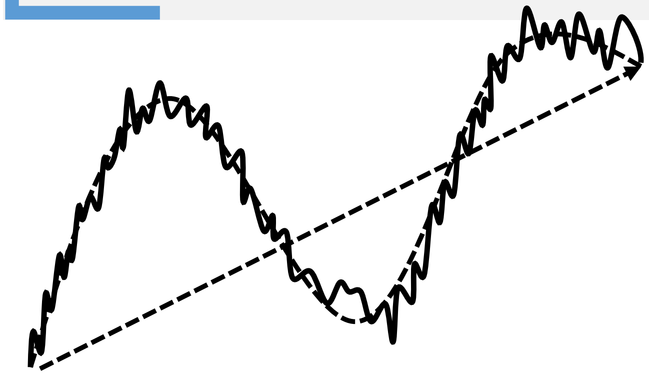
Trend



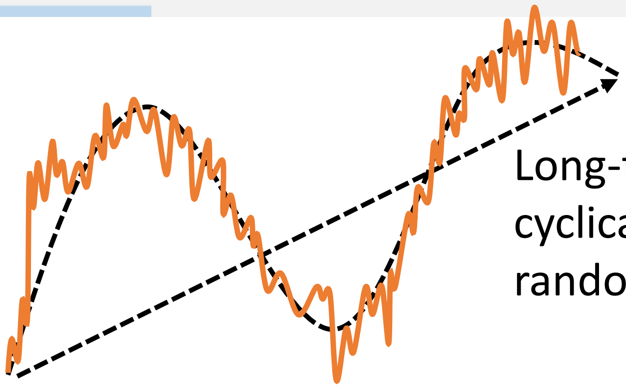
Cyclical



Seasonal



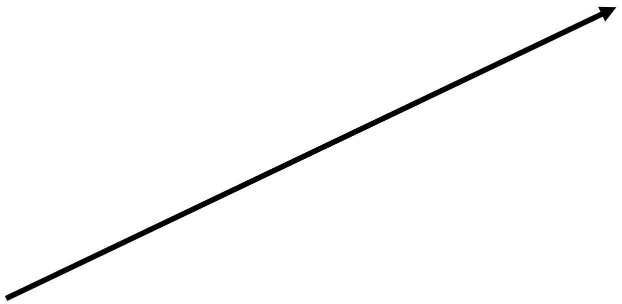
Residual Variation



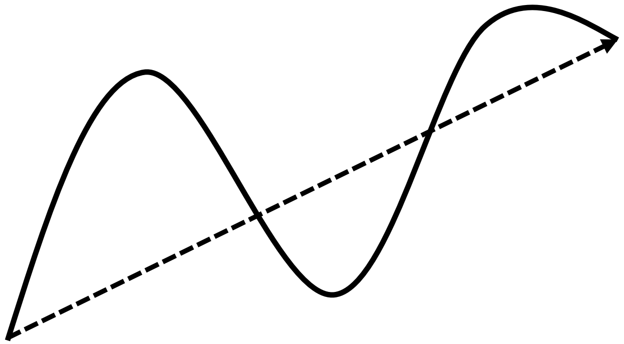
Long-term trend with cyclical, seasonal, & random variations

Components of a Time Series Model

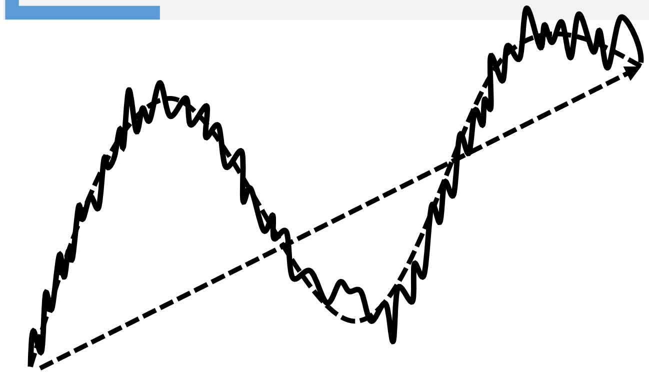
Trend



Cyclical

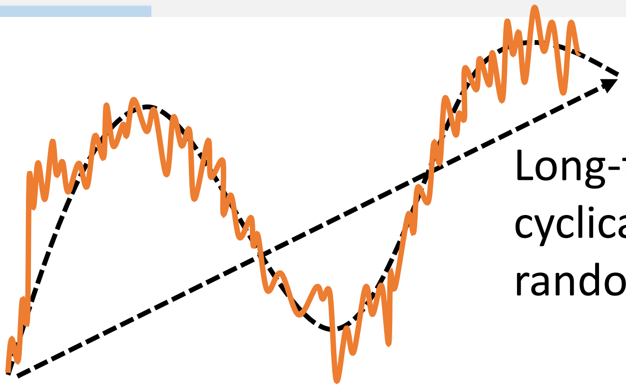


Seasonal



- **Residual variation** (also called random noise or random variation) has no pattern & cannot be predicted
- Cannot be replicated by repeating an experiment again

Residual Variation



Long-term trend with cyclical, seasonal, & random variations

Seasonal ARIMA Time Series

Seasonal

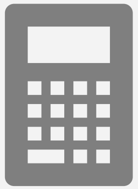
- Incorporates any **cyclical** trends in the data (any long-term patterns of ups and downs seen over time)

Autoregressive (AR)

- Incorporates any **cyclical** trends in the data (any long-term patterns of ups and downs seen over time)

Integrated (I)

- Software performs a mathematical transformation, makes a prediction, & un-transforms result



Moving Average (MA)

- Incorporates the overall **trend**, the changing of the average value over time

Seasonal ARIMA Time Series Model

(ARIMA - Autoregressive Integrated Moving Average)

AR: Autoregressive

Accounts for the use of previous observations as inputs into a regression equation which predicts the future observations

I: Integrated

Most statistical forecasting methods are based on the assumption that time series can be rendered approximately **stationary** through mathematical transformations

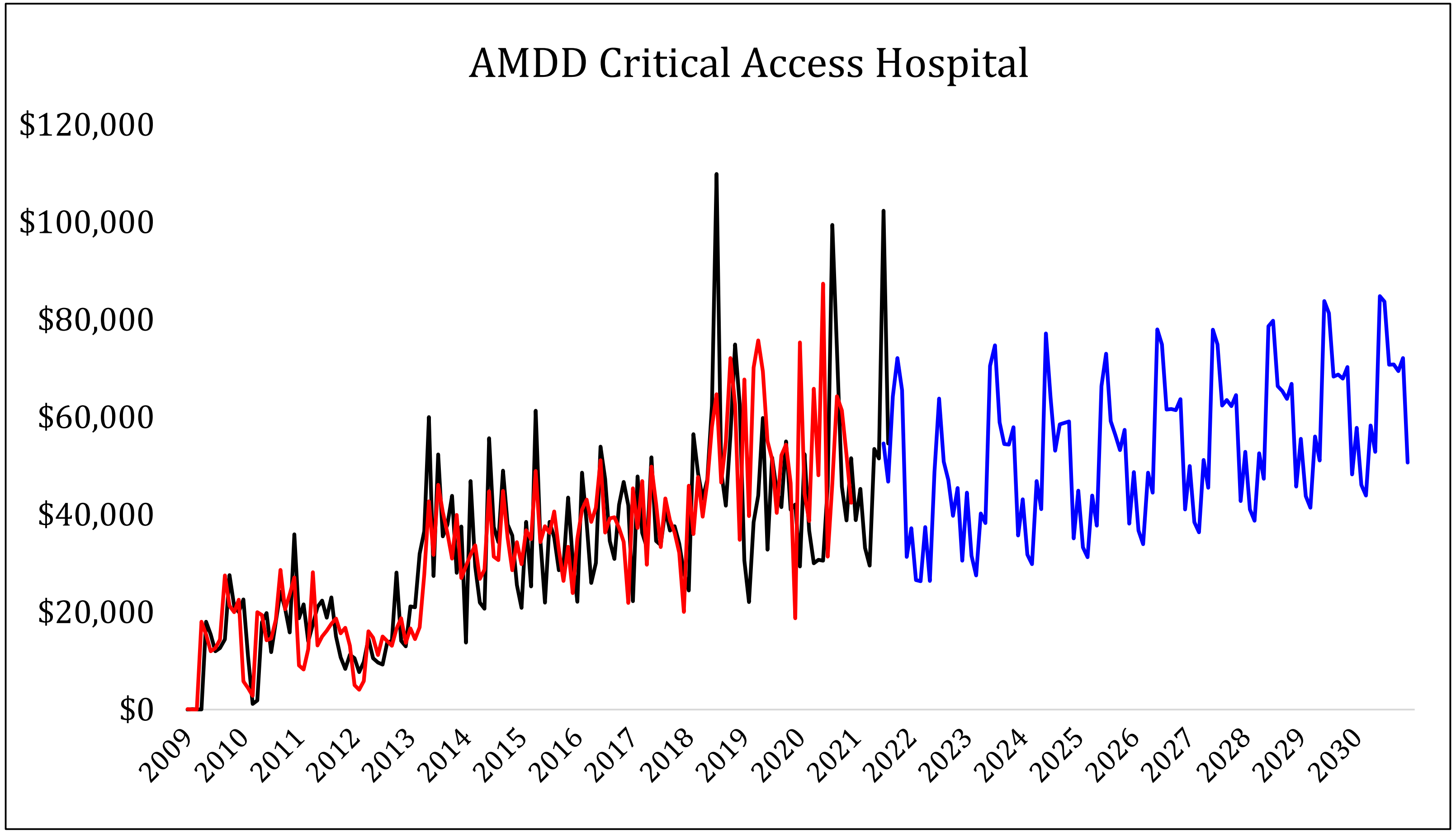
MA: Moving Average

Accounts for the error of the model as a combination of previous error terms; accounts for medical inflation, etc.

Seasonal

Allows for seasonal cycles and moving averages

Time Series Example



Wrap up

Summary

Data driven statistical time series model, relying on 17 years of data and making predictions based on prior information and observed trends

Modeled at a monthly level, but presented as an annual estimate

Questions?

LFD

MONTANA LEGISLATIVE FISCAL DIVISION

Some
Statistical Methodology

Stationarity

Most statistical forecasting methods are based on the assumption that time series can be rendered approximately stationary through mathematical transformations

- **Stationarity** – statistical properties such as mean, variance, autocorrelation, etc. are all constant over time

- A stationary series is relatively easy to predict: you simply predict that its statistical properties will be the same in the future as they have been in the past
- Predictions for the stationary series can then be "untransformed," by reversing the mathematical transformations, to obtain predictions for the original series – done internally within software
- Stationarize a time series through differencing $(Y_t - Y_{t-1})$

Stationarity

- Obtain meaningful sample statistics such as means, variances, and correlations with other variables – useful as descriptors of future behavior *only* if the series is stationary
  - If the series is consistently increasing over time, the sample mean and variance will grow with the size of the sample, and they will always underestimate the mean and variance in future periods
  - If the mean and variance of a series are not well-defined, then neither are its correlations with other variables
  - Accurate sample statistics are important
    - The **mean** (average) is our estimate
    - Though variance is not presented as a part of the estimate, it is still important
    - Time series models are based largely upon correlations, so accurate correlation values are necessary