# 2022 Essex Summer School
# 3K: Dynamics and Heterogeneity

Robert W. Walker, Ph. D.
Associate Professor of Quantitative Methods
Atkinson Graduate School of Management
Willamette University
Salem, Oregon USA
rwalker@willamette.edu

August 10, 2022

2022 Essex $SS^2DA$: Dynamics-Heterogeneity

# Language of Time Series

Trend : pattern exists when there is a long-term increase or decrease in the data.

Seasonal : pattern exists when a series is influenced by seasonal factors (e.g., the quarter of the year, the month, or day of the week).

Cyclic : pattern exists when data exhibit rises and falls that are not of fixed period (duration usually of at least 2 years).

# Visuals: Seasonal plots

Data plotted against the individual "seasons" in which the data were observed. (In this case a "season" is a month.)

Something like a time plot except that the data from each season are overlapped.

Enables the underlying seasonal pattern to be seen more clearly, and also allows any substantial departures from the seasonal pattern to be easily identified.
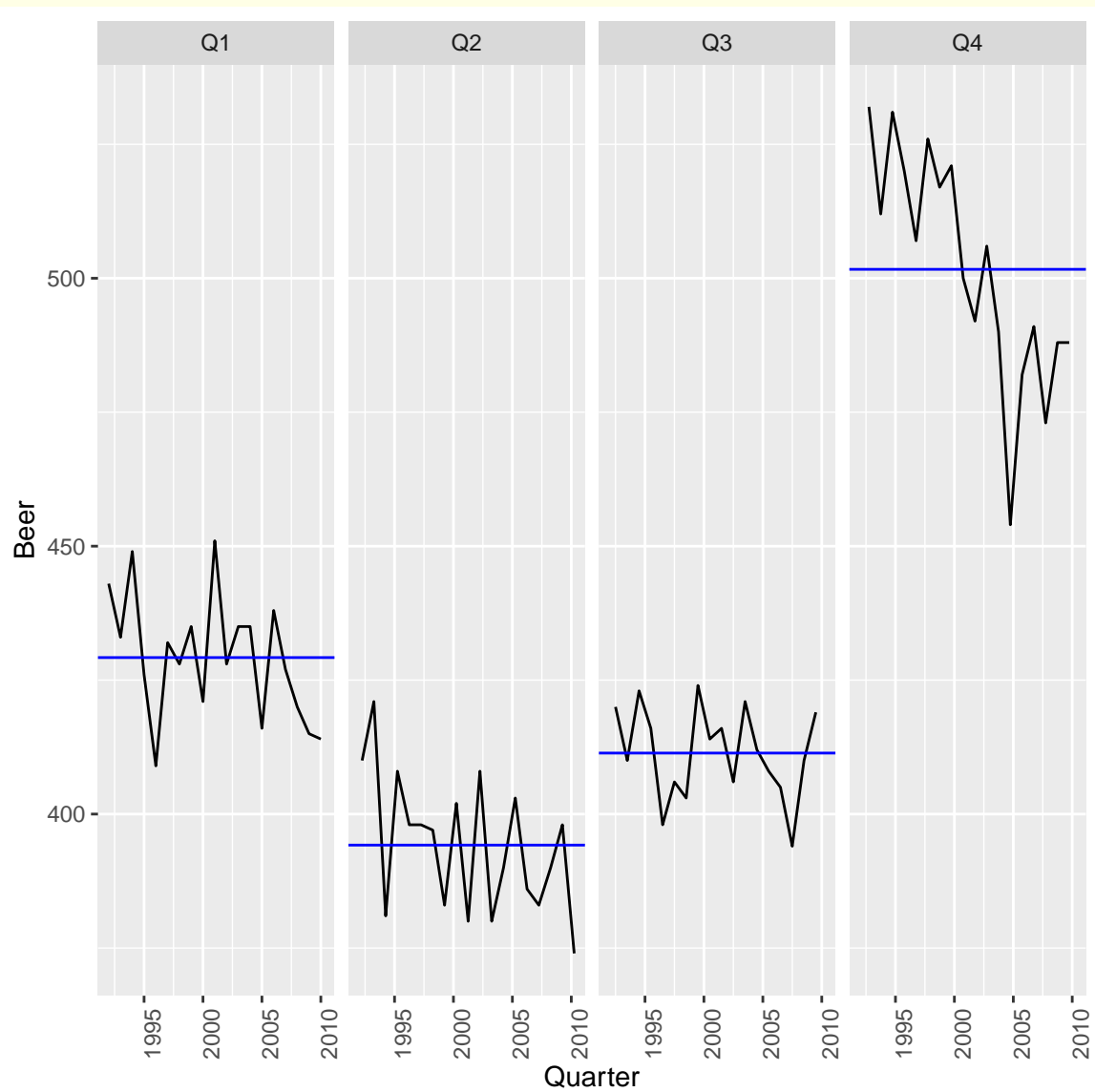
```
> library(fpp3); library(tidyverse)
> beer <- aus_production %>%
+    select(Quarter, Beer) %>%
+    filter(year(Quarter) >= 1992)
> beer %>% autoplot(Beer)
```
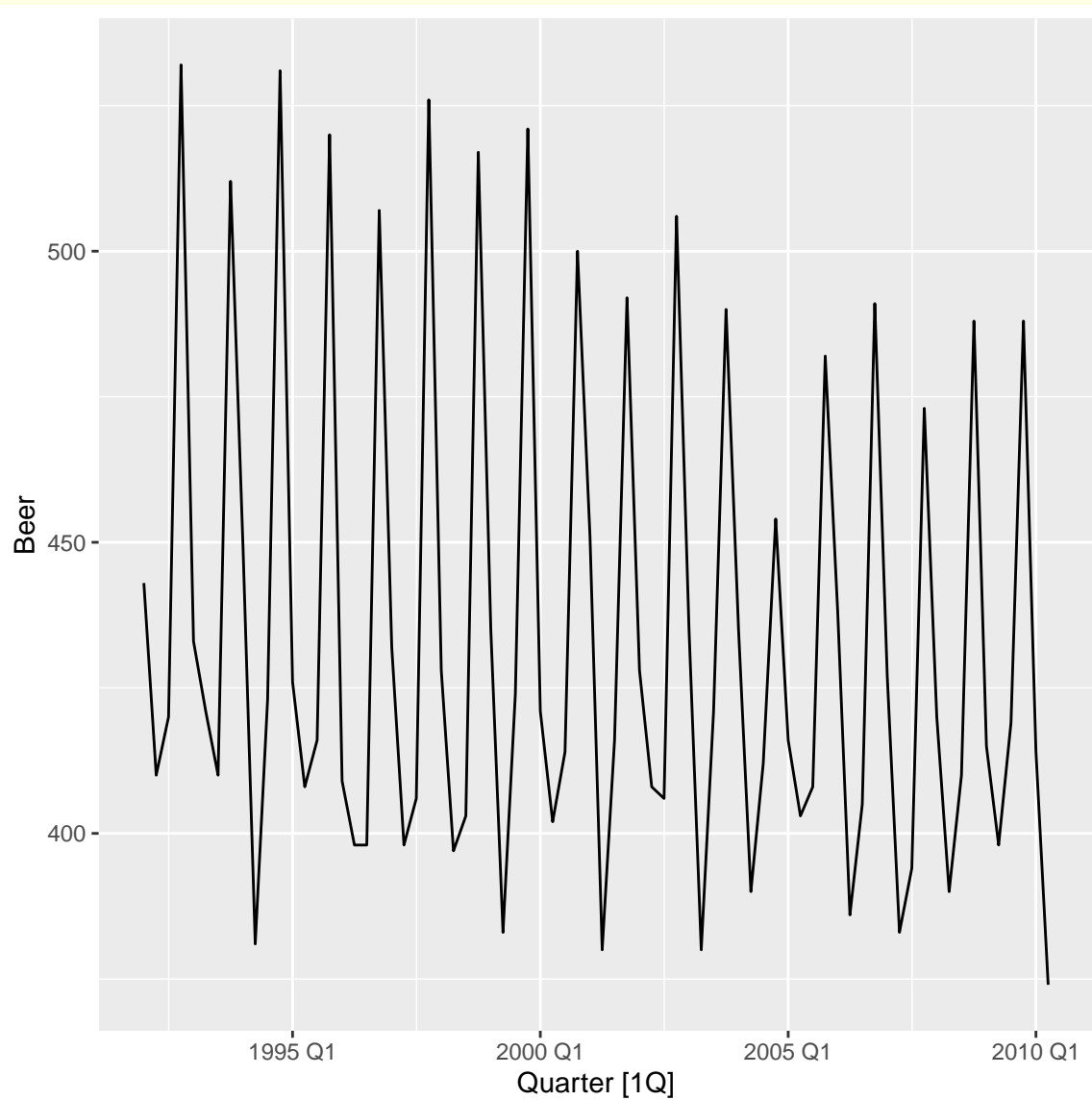
```
> beer %>% gg_season(Beer, labels="right")
```
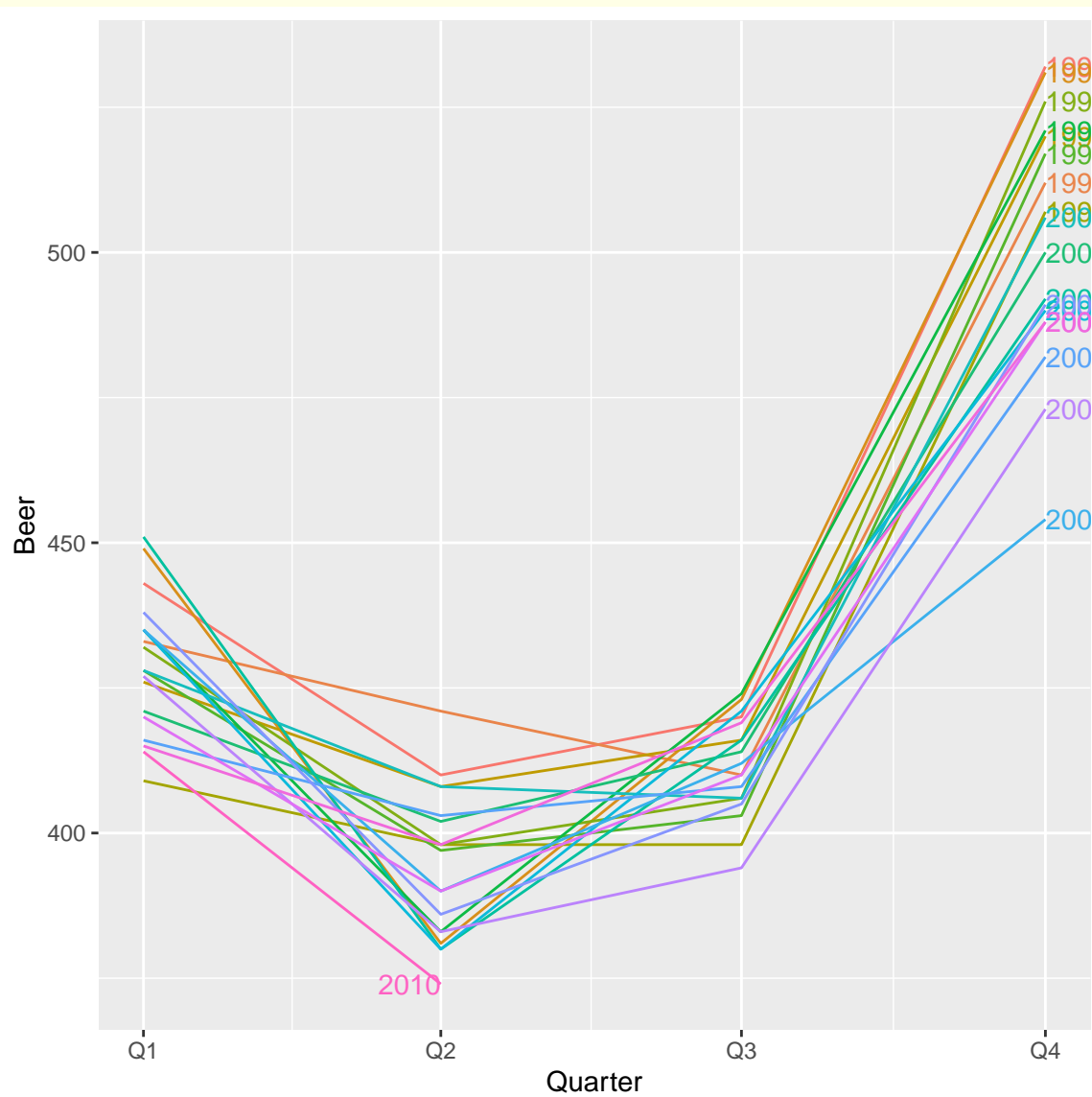
```
> beer %>% gg_subseries(Beer)
```

# Day 3: Dynamic Models

The ARIMA approach is fundamentally inductive. The workflow involves the use of empirical values of ACF's and PACF's to engage in model selection. Dynamic models engage theory/structure to impose more stringent assumptions for producing estimates.

# Time Series Linear Models/Dynamic Models

First, a result.

# Aitken's Theorem?

In a now-classic paper, Aitken generalized the Gauss-Markov theorem to the class of Generalized Least Squares estimators. It is important to note that these are GLS and not FGLS estimators. What is the difference? The two GLS estimators considered by Stimson are not strictly speaking GLS.

Definition:

$$\hat{\beta}_{GLS} \;=\; (\mathbf{X}'\Omega^{-1}\mathbf{X})^{-1}\mathbf{X}'\Omega^{-1}\mathbf{y} \tag{1}$$

Properties: (1) GLS is unbiased. (2) Consistent. (3) Asymptotically normal. (4) MV(L)UE

# A Quick Example

The variance/covariance matrix of the errors for a first-order autoregressive process is easy to derive. The matrix is banded; observations separated by one point in time are correlated $\rho$. Period two is $\rho^2$; the corners are $\rho^{T-1}$. The diagonal is one. What I have actually described is the correlation; the relevant autocovariances are actually defined by $\frac{\sigma^2 \rho^s}{1-\rho^2}$ where $s$ denotes the time period separation. It is also straightforward to prove (tediously through induction) that this is invertible; it is square and the determinant is non-zero having assumed that $|\rho < 1|$. The $2x2$ determinant is $\frac{1}{1-\rho^2}$. The $3x3$ is $1 * (1-\rho^2) - \rho(\rho - \rho^3) + \rho^2(\rho^2 - \rho^2)$. The first term is positive and the second term is non-zero so long as $\rho \neq 0$. But even if $\rho = 0$, we would have an identity matrix which is invertible.

$$\Phi = \sigma^2 \Psi = \sigma_e^2 \begin{pmatrix} 1 & \rho^1 & \rho^2 & \cdots & \rho^{T-1} \\ \rho^1 & 1 & \rho^1 & \cdots & \rho^{T-2} \\ \rho^2 & \rho^1 & 1 & \cdots & \rho^{T-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{T-1} & \rho^{T-2} & \rho^{T-3} & \cdots & 1 \end{pmatrix}$$

given that $e_t = \rho e_{t-1} + \nu_t$. A Toeplitz form....

If the variance is stationary, we can rewrite,

$$\sigma_e^2 = \frac{\sigma_\nu^2}{1 - \rho^2}$$

A comment on characteristic roots....

# Cochrane-Orcutt

We have the two key elements to implement this except that we do not know $\rho$; we will have to estimate it and estimates have uncertainty. But it is important to note this imposes exactly an AR(1). If the process is incorrectly specified, then the optimal properties do not follow. Indeed, the optimal properties also depend on an additional important feature.

# What does the feasible do?

We need to estimate things to replace unknown covariance structures and coverage will depend on properties of the estimators of these covariances. Consistent estimators will work but there is euphemistically "considerable variation" in the class of consistent estimators. Contrasting the Beck and Katz/White approach with the GLS approach is a valid difference in philosophies.[1]

---

[1]We will return to this when we look at Hausman because this is the essential issue.

# Incremental Models

$$y_t = a_1 y_{t-1} + \epsilon_t$$

is the simplest dynamic model but it cannot be estimated consistently, in general terms, in the presence of serial correlation. Why? The key condition for unbiasedness is violated because $\mathbb{E}(y_{t-1}\epsilon_t) \neq 0$. OLS will not generally work.

# Incremental Models with Covariates

$$y_t = a_1 y_{t-1} + \beta X_t + \epsilon_t$$

The problem is fitting and the key issue is white noise residuals post-estimation. But we have to assume a structure and implement it.

# Distributed Lag Models

$$y_t = \alpha + \beta_0 X_t + \beta_1 x_{t-1} + \ldots + \epsilon_t$$

The impact of $x$ occurs over multiple periods. It relies on theory, or perhaps analysis using information criteria/F [owing to quasi-nesting and missing data]. OLS is a fine solution to this problem but the search space of models is often large.

In response to this problem, we have structured distributed lag models; there are many such schemes.

- Koyck/Geometric decay:
  short run and long-run effects are parametrically identified

$$y_t = \alpha + \beta(1 - \lambda) \sum_{j=0}^{\infty} \lambda^j X_{t-j} + \epsilon$$

- Almon (more arbitrary decay):

$$y_{it} = \sum_{t_A=0}^{T_F} \rho_{t_A} x_{t-t_A} + \epsilon_t$$

with coefficients that are ordinates of some general polynomial of degree $T_F >> q$. The $\rho_{t_A} = \sum_{k=0}^{T_F} \gamma_k t^k$.

# Autoregressive Distributed Lag Models

$$y_t = \alpha + \gamma_1 y_{t-1} + \beta_0 X_t + \beta_1 x_{t-1} + \ldots + \epsilon_t$$

OLS is often used if iid; $\epsilon_t$ is unrelated to $y_{t-1}$. If not iid: GLS is needed. The authors argue that the lagged dependent variable often yields white noise for free. As they also note, there is a deBoef and Keele paper showing the relationship between these models and a form of error correction models. More on that tomorrow.

# Structural vs. Non-structural

Data analysis can quite yield models comparisons among competing dynamic structures. The key issue is that the analyst need divine the process; what is the relevant error process and what is the structure and timing of effects alongside the potential question of incremental adjustment. We need good theory for that.

Given such theory, we can take an equations as analysis approach, measure the variables, and derive reduced forms, and then recover parameter estimates deploying simultaneous equations methods. Very large such systems were a core part of early empirical macroeconomics. The failures of such systems led to the proposal of alternatives.

Chris Sims suggested a more flexible approach: the VAR.

# Vector AutoRegression

Choose a relevant set of lag lengths and write each variable in the system as a function of lags of itself and other variables to the chosen lengths. The key insight is that this VAR is the reduced form to some more complicated as yet unspecified structural form. But if the goal is to specify how variables related to one another and to use data to discover Granger causality and responses to impulse injected in the system.