Financial Econometrics

- Week 6 -

Dynamic Regression models

VU Econometrics and Data Science

2024-2025

Paolo Gorgi

Vrije Universiteit Amsterdam

Extensions

Until now: we have focused essentially on models with time-varying conditional volatility.

Parameter-driven and **observation-driven** models can describe a wide range of other phenomena.

In general: time-varying parameters can be used to design models that capture changes in economies and financial markets.

Here: we discuss how to introduce time-varying coefficients in the linear regression model.

Application: Dynamic Capital Asset Pricing Model (CAPM).

Today's class

1. Observation-driven dynamic regression
   - Model specification
   - Maximum Likelihood estimation
   - Observation-driven regression with R
2. Parameter-driven dynamic regression
   - Model specification
   - Indirect inference estimation
3. Dynamic CAPM

Observation-driven dynamic regression

The linear regression model (i)

Linear Gaussian regression model:

$$y_t = \beta x_t + \varepsilon_t \ , \quad \varepsilon_t \sim NID(0, \sigma_\varepsilon^2),$$

where we assume that x_t and y_t have zero unconditional mean, and $\mathbb{E}(\varepsilon_t | x_t) = 0$.

Note: The population regression coefficient β describes the relation between y_t and x_t.

Note: The population regression coefficient β is given by

$$\beta = \frac{\mathbb{C}ov(x_t, y_t)}{\mathbb{V}ar(x_t)} = \frac{\sigma_{xy}}{\sigma_x^2}.$$

Note: The conditional expectation of y_t given x_t is $\mathbb{E}(y_t | x_t) = \beta x_t$.

The linear regression model (ii)

Recall from Inro Econometrics: β can be estimated by OLS $\hat{\beta}$.

The linear regression model is a powerful tool to obtain predictions and describe the relationship between variables.

Limitation: When dealing with time series data, the relationship between variables may be changing over time.

Solution: We can allow β to vary over time!

Application: We shall see how a time varying beta can be useful for the Capital Asset Pricing Model (CAPM).

Observation-driven regression (i)

Observation-driven regression model: specification based on GAS framework of Creal et al. (2013).

Observation equation:

$$y_t = \beta_t x_t + \varepsilon_t, \qquad \varepsilon_t \sim NID(0, \sigma^2).$$

Updating equation:

$$\beta_t = \omega + \phi\beta_{t-1} + \alpha(y_{t-1} - \beta_{t-1}x_{t-1})x_{t-1},$$

where the parameters ω, ϕ, α determine the dynamic properties of β_t.

- The time-varying regression coefficient β_t depends only on past values of y_t and x_t.

Observation-driven regression (ii)

Updating equation:

$$\beta_t = \omega + \phi\beta_{t-1} + \alpha(y_{t-1} - \beta_{t-1}x_{t-1})x_{t-1}.$$

Idea:

- The regression coefficient β_t measures the covariance structure between the variables.
- The regression coefficient β_t is updated using $(y_{t-1} - \beta_{t-1}x_{t-1})x_{t-1} = \varepsilon_{t-1}x_{t-1}$.
- $\varepsilon_{t-1}x_{t-1}$ is an estimate of the covariance between x_{t-1} and the error term of y_t.

Note: we do not need to specify a distribution for the regressor x_t.

Properties of the model

Properties of observation-driven regression

Note: the coefficient β_t is constant given past values of y_t, $Y^{t-1} = \{y_{t-1}, y_{t-2} \dots\}$, and past values of x_t, $X^{t-1} = \{x_{t-1}, x_{t-2} \dots\}$.

Therefore we can obtain obtain conditional moments and distribution of y_t given Y^{t-1} and X^t.

- **Conditional expectation:** $\mathbb{E}(y_t|Y^{t-1}, X^t) = \beta_t x_t$.
- **Conditional variance:** $\mathbb{V}ar(y_t|Y^{t-1}, X^t) = \sigma^2$.
- **Conditional distribution:** $y_t|(Y^{t-1}, X^t) \sim N(\beta_t x_t, \sigma^2)$.

Maximum Likelihood estimation (i)

Question: how can we estimate the parameter vector $\theta = (\omega, \phi, \alpha, \sigma^2)^\top$ of the observation-driven regression model?

Answer: we can easily estimate θ by Maximum Likelihood.

Important: the likelihood function is available in closed form!

Note: the conditional distribution of y_t given X^t and Y^{t-1} is $y_t|(Y^{t-1}, X^t) \sim N(\beta_t x_t, \sigma^2)$.

Therefore the conditional density function of y_t is

$$p(y_t|Y^{t-1}, X^t) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_t - \beta_t x_t)^2}{2\sigma^2}\right).$$

Maximum Likelihood estimation (ii)

Recall: we can write the log-likelihood function as the sum of the conditional log-densities.

The **log-likelihood** function is

$$L_T(\theta) = \sum_{i=2}^{T} -\frac{1}{2}\left(\log\sigma^2 + \frac{(y_t - \beta_t x_t)^2}{\sigma^2}\right),$$

- β_t is obtained using the updating equation

$$\beta_t = \omega + \phi\beta_{t-1} + \alpha(y_{t-1} - \beta_{t-1}x_{t-1})x_{t-1}.$$

- The initialization β_1 can be set equal to the unconditional expectation of β_t, i.e. $\beta_1 = \omega/(1-\phi)$.

Maximum Likelihood estimation (iii)

ML estimator: the ML estimator $\hat{\theta}_T$ is given by

$$\hat{\theta}_T = \arg\max_{\theta \in \Theta} L_T(\theta).$$

Note:

- The likelihood is maximized using numerical methods as we have seen for GARCH models.
- Standard asymptotic properties apply: the ML estimator is consistent and asymptotically normal.

ML estimation with R (i)

ML with R: `llik_OD_regression.R`, `estimate_OD_reg.R`.

Write the likelihood function: `llik_OD_regression()`

Input: observed data `y`, `x` and parameter vector `par`

Output: average log-likelihood `llik`

First: The function `llik_OD_regression()` is defined

```
llik_OD_regression <- function(y,x,par){
```

Second: parameter values are defined using the input `par`

```
n <- length(x)
omega <- par[1]
phi <- exp(par[2])/(1+exp(par[2]))
alpha <- exp(par[3])
s2 <- exp(par[4])
```

ML estimation with R (ii)

Third: β_t is initialized and a *for loop* with the updating equation is used to obtain β_t

```
beta <- rep(0,n)
beta[1] <- omega/(1-phi)
for(t in 2:n){
  beta[t] <- omega+phi*beta[t-1]+alpha*(y[t-1]-beta[t-1]*x[t-1])*x[t-1]
}
```

Finally: the average log-likelihood is computed

```
l <- -(1/2)*log(s2)-(1/2)*(y-beta*x)^2/s2
llik <- mean(l)
return(llik)
```

Optimize the likelihood: optimization is the same as we have seen for GARCH models (see `llik_OD_regression.R`).

Estimating β_t with R (i)

Estimating β_t with R

Once we have the ML estimate $\hat{\theta}_T$, we can obtain the filtered coefficient β_t by running the updating equation evaluated at $\hat{\theta}_T$.

Set initialization:

```
n <- length(xt)
beta <- rep(0,n)
beta[1] <- omega_hat/(1-phi_hat)
```

Run for loop:

```
for(t in 2:n){
   beta[t] <- omega_hat+phi_hat*beta[t-1]+
                    alpha_hat*(yt[t-1]-beta[t-1]*xt[t-1])*xt[t-1];
}
```

Estimating β_t with R (ii)

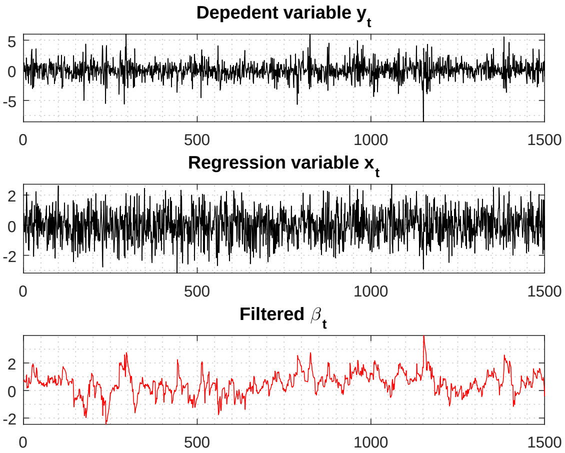

Figure: Time series of y_t and x_t and estimated β_t.

Parameter-driven dynamic regression

Regression with time-varying parameters (i)

Parameter-driven dynamic regression model

Observation equation:

$$y_t = \beta_t x_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma_\varepsilon^2).$$

Transition equation:

$$\beta_t = \alpha_0 + \alpha_1 \beta_{t-1} + \eta_t, \quad \eta_t \sim N(0, \sigma_\eta^2).$$

- β_t evolves exogenously (not a function of past observations).
- Relationship between y_t and x_t changes at each point in time.
- Changes can be smooth or abrupt depending on parameters α_0, α_1 and σ_η^2.

Regression with time-varying parameters (ii)

Note:

- No need to specify a distribution for x_t: exogenous regressor.
- The coefficient β_t is not a constant even if we condition on past y_t and x_t.

Question: How can we estimate the parameter vector $\theta = (\alpha_0, \alpha_1, \sigma^2_\eta, \sigma^2_\varepsilon)^\top$ of the *parameter-driven* regression model?

Answer: Simple, we can use the Indirect Inference estimator $\hat{\theta}_{TH}$

$$\hat{\theta}_{TH} = \arg\min_{\theta \in \Theta} d\big(\hat{B}_T, \tilde{B}_H(\theta)\big).$$

Note: The likelihood function of this specific model can be written in closed form using the Kalman Filter. You will study the Kalman filter in the Econometrics Master.

Estimation by indirect inference (i)

Indirect inference estimation

Problem: we cannot generate values for x_t (no model for x_t).

Question: how can we obtain the auxiliary statistics $\tilde{B}_H(\theta)$ from simulated data?

Answer:

1. Generate M series of length T for y_t by using observed x_t in the simulations.
2. Obtain an auxiliary statistics from each of the generated series $\tilde{B}_i(\theta)$, $i = 1, \ldots, M$.
3. Average the *auxiliary statistics* $\tilde{B}_H(\theta) = \frac{1}{M}\sum_{i=1}^{M} \tilde{B}_i(\theta)$ (this way we have $H = M \times T$).

Estimation by indirect inference (ii)

The **auxiliary statistics** must be simple to compute and capture the time-varying parameter dynamics.

The parameters α_0, α_1, σ^2_η and σ^2_ε determine:

1. Average covariance between y_t and x_t.
2. Average variance of regression residuals.
3. Dependence structure in the covariance between y_t and x_t.

Important: if we run simple regression of y_t on x_t we describe average dependences between the two variables!

Estimation by indirect inference (iii)

Idea: Regress y_t on x_t, save residuals $\hat{e}_t = y_t - \hat{\beta} x_t$, and proceed as follows:

- Use OLS estimate $\hat{\beta}$ to capture (1).
- Use variance of residuals to capture (2).
- Use autocovariance in $\{\hat{e}_t x_t\}_{t=1}^T$ to capture (3).

The *auxiliary statistic* from the observed data $\hat{B}_T$ is

$$\hat{B}_T = \begin{bmatrix} \hat{\beta} \\ \hat{\gamma}_0 \\ \vdots \\ \hat{\gamma}_p \end{bmatrix},$$

where $\hat{\gamma}_l = T^{-1} \sum_{t=1}^T \hat{e}_t x_t \hat{e}_{t-l} x_{t-l}$ for $l = 0, 1, \ldots, p$.

Estimation by indirect inference (iv)

Similarly: we obtain $\tilde{B}_H(\theta)$ from the simulated data

In particular, we obtain M estimates $\tilde{B}_i(\theta)$ from M simulated series $\{\tilde{y}_{i,t}(\theta)\}_{t=1}^T$ for $i = 1, \ldots, M$:

- Run M regressions of $\tilde{y}_{i,t}(\theta)$ on x_t and get OLS estimates $\tilde{\beta}_i(\theta)$.
- Obtain M residual sequences $\tilde{e}_{i,t}(\theta) = \tilde{y}_{i,t} - \tilde{\beta}_i(\theta)x_t$.

The *auxiliary statistic* from the simulated data $\tilde{B}_H(\theta)$ is

$$\tilde{B}_H(\theta) = \frac{1}{M}\begin{bmatrix} \sum_{i=1}^M \tilde{\beta}_i(\theta) \\ \sum_{i=1}^M \tilde{\gamma}_{0,i}(\theta) \\ \vdots \\ \sum_{i=1}^M \tilde{\gamma}_{p,i}(\theta) \end{bmatrix},$$

where $\hat{\gamma}_{i,l} = T^{-1}\sum_{t=1}^T \tilde{e}_{i,t}(\theta)x_t\tilde{e}_{i,t-l}(\theta)x_{t-l}$ for $l = 0, 1, \ldots, p$

Indirect inference with R (i)

Estimation with R: `sim_m_REG.R`, `estimate_REG_II.R`.

Function to obtain $\tilde{B}_H(\theta)$: `sim_m_REG()`

```
sim_m_REG <- function(e,x,par){
```

Inputs: parameter vector `theta`, the observed regressor `x`, and the simulated errors `e` of both the observation equation and the transition equation.

Note: matrix `e` contains two vectors of length $H = M \times T$ of $N(0,1)$ random variables.

Output: the auxiliary statistics $\tilde{B}_H(\theta)$

Indirect inference with R (ii)

First: define parameters and T, H from the inputs

```
a0 <- par[1]
a1 <- exp(par[2])/(1+exp(par[2]))
s_eta <- exp(par[3])
s_eps <- exp(par[4])

output <- 0

n <- length(x)
H <- length(e[,1])
M <- H/n
```

Indirect inference with R (iii)

Second: obtains innovation with correct variance

```
eta <- sqrt(s_eta)*e[,1]
eps <- sqrt(s_eps)*e[,2]
```

Third: simulate data M paths of length T from the model

```
for(m in 1:M){
   b <- rep(0,n)
   b[1] <- a0/(1-a1)
   for(t in 2:n){
      b[t] <- a0+a1*b[t-1]+eta[(m-1)*n+t]
   }
   y <- b*x+eps[((m-1)*n+1):(m*n)]
```

(FOR LOOP continues...)

Indirect inference with R (iv)

Finally: obtain auxiliary statistic $\tilde{B}_H(\theta)$

```
   hb <- cov(y,x)/var(x)
   yr <- y-hb*x
   xy <- yr*x
   acvfxy <- acf(xy, lag.max=15, type ="covariance", plot=F)$acf[-1]
   output <- output+c(var(yr),hb,acvfxy)/M
   }
return(output)
}
```

Note: we used an autocovariance function with 15 lags.

Note: output delivers the average statistic!

Indirect inference with R (v)

We are now in a position to obtain II parameter estimates!

Optimize $d\big(\hat{B}_T, \tilde{B}_H(\theta)\big)$**:** `estimate_REG_II.R`

First: obtain $\hat{B}_T$

```
hb <- cov(yt,xt)/var(xt)
yr <- yt-hb*xt
xy <- yr*xt
acvfxy <- acf(xy, lag.max=15, type ="covariance", plot=F)$acf[-1]
sample_m <- c(var(yr),hb,acvfxy)
```

Second: generate error terms `e`

```
n <- length(xt)
M <- 20
H <- M*n
e <- cbind(rnorm(H),rnorm(H))
```

Indirect inference with R (vi)

Third: set initialization for numerical optimization

```
a1 <- 0.95
a0 <- cov(xt,yt)/var(xt)*(1-a1)
s_eta <- 0.2
s_eps <- var(yr)
par_ini <- c(a0, log(a1/(1-a1)), log(s_eta), log(s_eps))
```

Finally: optimize the II criterion and obtain `theta_hat`

```
est <- optim(par=par_ini,
       fn=function(par) mean((sim_m_REG(e,xt,par)-sample_m)^2),
        method = "BFGS")
```

Dynamic CAPM

Capital Asset Pricing Model (i)

The Capital asset pricing model (CAPM) describes how financial markets price assets.

Note: In financial markets dominated by risk-averse investors, higher-risk securities are priced to yield higher expected returns than lower-risk securities.

Important: The CAPM assumes that there are no sources of risk except the systematic market risk as the idiosyncratic risk of a specific asset can be eliminated through diversification.

Therefore: The expected return of an asset depends on the exposition to the market risk: higher risk, higher returns.

Capital Asset Pricing Model (ii)

CAPM model:

$$R_i = r^f + \beta_i (R^m - r^f),$$

where

- R_i is the expected return of stock i;
- r^f is the risk-free rate;
- R^m is the expected market return;
- β_i is called the beta of stock i;

Capital Asset Pricing Model (iii)

Important: β_i provides a measure of exposition to the systematic risk for stock i

- If $\beta_i = 1$, then the stock is as risky as the market and therefore it has the same expected return;
- If $\beta_i < 1$, then the stock is less risky than the market and therefore it has a lower expected return;
- If $\beta_i > 1$, then stock is more risky than the market and therefore it has a higher expected return.

Capital Asset Pricing Model (iv)

The CAPM can be expressed as a regression model

$$r_{i,t} = r^f + \beta_i (r_t^m - r^f) + \varepsilon_{i,r},$$

where

- $r_{i,t}$ is the return of asset i at time t;
- r_t^m is the market return at time t;
- $\varepsilon_{i,r}$ is an idiosyncratic error term (specific to stock i).

Capital Asset Pricing Model (v)

Varaince decomposition CAPM:

$$\mathbb{V}ar(r_{i,t}) = \beta_i^2 \mathbb{V}ar(r_t^m) + \mathbb{V}ar(\varepsilon_{i,r}),$$

where

- $\mathbb{V}ar(r_t^m)$ is the systematic risk (market risk);
- $\mathbb{V}ar(\varepsilon_{i,r})$ is the idiosyncratic risk (diversifiable risk).

Note: Only exposition to the systematic risk gives higher returns as idiosyncratic risk can be eliminated through diversification.

Capital Asset Pricing Model (vi)

Estimation of the CAPM can be implemented though a simple OLS regression.

In practice:

- The market returns r_t^m are measured using some broad market indexes such as the S&P500;
- The risk free asset r^f is measured using short term treasury bills. The risk free can be time varying r_t^f.

Dynamic CAPM (i)

The CAPM model assumes that the β_i of a stock does not change over time.

Problem: This assumption is somewhat restrictive and often rejected by the data.

Solution: Extend the CAPM to have a time varying $\beta_{i,t}$ using dynamic regression

$$r_{i,t} = r^f + \beta_{i,t}(r_t^m - r^f) + \varepsilon_{i,t}$$

$$\beta_{i,t} = \alpha_0 + \alpha_1 \beta_{i,t-1} + \eta_t,$$

where $\{\varepsilon_t\}_{t\in\mathbb{Z}} \sim NID(0, \sigma_\varepsilon^2)$ and $\{\eta_t\}_{t\in\mathbb{Z}} \sim NID(0, \sigma_\eta^2)$.

Dynamic CAPM (ii)

Alternative specification of time-varying β: A time varying beta can also be specified indirectly using GARCH models.

Idea: In the standard CAPM, the beta is given by

$$\beta_i = \frac{\mathbb{C}ov(r_t^m, r_{i,t})}{\mathbb{V}ar(r_t^m)} = \frac{\sigma_{m,i}}{\sigma_m^2}$$

Therefore: Time varying covariance $\sigma_{m,i,t}$ and variance $\sigma_{m,t}^2$ can be obtained using a bivariate GARCH model for r_t^m and $r_{i,t}$. Then, the time-varying beta is defined as

$$\beta_{i,t} = \frac{\sigma_{m,i,t}}{\sigma_{m,t}^2}.$$

Dynamic CAPM (iii)

Concluding remarks:

1. The time-varying CAPM model can be specified as a *parameter-driven* regression or an *observation-driven* regression.
2. Alternatively, bivariate GARCH models can be used to obtain time-varying betas.
3. Extension of the CAPM is given by the multiple factor model that includes other factors besides the market risk to explain expected returns of assets.