ECON 582: PIH and Time Series

Instructor: Dmytro Hryshko

Recall: utility function is quadratic, borrowing/saving using risk-free bond. The budget constraint is

$$a_{t+1} = (1+r)(a_t + y_t - c_t).$$

The lifetime budget constraint:

$$\sum_{j=0}^{\infty} \left(\frac{1}{1+r}\right)^j c_{t+j} = \sum_{j=0}^{\infty} \left(\frac{1}{1+r}\right)^j y_{t+j} + a_t.$$

Stochastic Euler equation:

$$c_t = E_t c_{t+1}, \quad \text{or}$$

$$E_t \Delta c_{t+1} = 0.$$

Excess sensitivity

Consumption is said to be a martingale. A natural way to test the martingale hypothesis is by running the following regression

$$\Delta c_{t+1} = \alpha + \beta X_t + \text{error},$$

and testing if $\beta = 0$. X_t is a variable known at t. For example, $X_t = \{y_t, \Delta y_t\}$—lagged income or lagged income changes.

If β is estimated to be statistically different from 0, consumption is said to be excessively sensitive to lagged information (income).

Such tests are called the excess sensitivity tests of consumption.

Since $E_t c_{t+1} = c_t$,

$$E_{t+1} c_{t+2} = c_{t+1}.$$

Apply the time-t conditional expectation operator to both sides of the equation to obtain

$$E_t E_{t+1} c_{t+2} = E_t c_{t+1}$$
$$E_t c_{t+2} = c_t.$$

Following similar steps, we can show that

$$E_t c_{t+k} = c_t, \text{for all } \ k \geq 1.$$

The lifetime budget constraint can be written as:

$$E_t \left[\sum_{j=0}^{\infty} \left(\frac{1}{1+r} \right)^j c_{t+j} \right] = E_t \left[\sum_{j=0}^{\infty} \left(\frac{1}{1+r} \right)^j y_{t+j} \right] + a_t ,$$

or

$$c_t \left[\sum_{j=0}^{\infty} \left(\frac{1}{1+r} \right)^j \right] = E_t \left[\sum_{j=0}^{\infty} \left(\frac{1}{1+r} \right)^j y_{t+j} \right] + a_t .$$

Thus,

$$c_t = \underbrace{\frac{r}{1+r} \left[a_t + E_t \sum_{j=0}^{\infty} \left(\frac{1}{1+r} \right)^j y_{t+j} \right]}_{\text{permanent income}} .$$

We can show (see the note) that

$$\Delta c_t = \underbrace{\frac{r}{1+r}\sum_{j=0}^{\infty}\left(\frac{1}{1+r}\right)^j (E_t - E_{t-1})y_{t+j}}_{\text{innovation in the permanent income}},$$

where $(E_t - E_{t-1})x_{t+k} = E_t x_{t+k} - E_{t-1}x_{t+k}$.

To understand more about the magnitude of consumption changes, we need to have some idea about the properties of a (stochastic) process of income $\{y_t\}$.

Time series

A time series is a collection of observations y_t, each recorded at time t.

We will talk of a time series as a collection of realizations of random variables Y_t.

The time of recording an observation belongs to some set T_0. If T_0 is a discrete set, a time series is called discrete (daily, monthly, annual time series are examples of discrete time series).

Sample path

In data, we observe a sample path of Y_t: e.g., $y_1, y_2, \ldots, y_T$.

We want to model the observed time series as a realization of a stochastic process $Y_t, t = 1, 2, \ldots, T$, keeping in mind that the process could have started before $t = 1$ and could run after $t = T$ (for example, Y_t can be recorded at $t = 0, \pm 1, \pm 2, \ldots$).

Auto-covariance function

We want to construct a mathematical/statistical model that would describe the data we observe.

For a single time series, the dependence between Y_t and Y_t, $Y_{t\pm1}$, $Y_{t\pm2}$, etc., is described by the auto-covariance function.

The *auto-covariance function* $\gamma(\cdot,\cdot)$ for a stochastic process $\{Y_t, t \in T\}$ is defined by

$$\gamma(i,j) = E\left[(Y_i - EY_i)(Y_j - EY_j)\right], \quad \forall i,j \in T.$$

$\gamma(i,i)$—the variance of Y at time i, $\gamma(i,i+1)$—the auto-covariance between Y's recorded at time i and $i+1$, etc. In general, those quantities can be time-dependent.

Weak stationarity

A stochastic process $\{Y_t, t = 0, \pm 1, \pm 2, \dots\}$ is *weakly stationary* if

$$
\begin{aligned}
EY_t^2 < \infty, \quad \forall t \\
EY_t = \mu, \quad \forall t \\
\gamma(i,j) = \gamma(i+t, j+t) = \gamma(i-j) = \gamma(j-i), \quad \forall i, j, t = 0, \pm 1, \pm 2, \dots .
\end{aligned}
$$

If, e.g. $i - j = 1$, for a weakly-stationary process $\gamma(1)$ can be calculated as

$$E\left[(Y_t - EY_t)(Y_{t-1} - EY_{t-1})\right] = E\left[(Y_t - EY_t)(Y_{t+1} - EY_{t+1})\right], \forall t.$$

Similarly, $\gamma(2)$ is

$$\gamma(2) = E\left[(Y_t - EY_t)(Y_{t-2} - EY_{t-2})\right] = E\left[(Y_t - EY_t)(Y_{t+2} - EY_{t+2})\right],$$

etc.

If you have a model for Y_t, the mean, variances and auto-covariances can be estimated by simulating the model S times for $t = 0, \pm 1, \pm 2, \ldots$ and taking the average across S simulations.

For example,

$$EY_t = \frac{1}{S} \sum_{s=1}^{S} y_t^s,$$

$$EY_t^2 = \frac{1}{S} \sum_{s=1}^{S} (y_t^s)^2,$$

where y_t^s is the value assumed by Y at time t in simulation s. In real data, we do not have the luxury of observing the process repeatedly but we can infer the mean, variances, and auto-covariances of the process by calculating sample analogs of population moments.

Random walk

Most of observed time series are not stationary. An example is a random walk.

Random walk can be described by a process $Y_t = Y_{t-1} + u_t$, where $u_t \sim iid(0, \sigma^2)$, $t = 1, \ldots, T$, and $Y_0 = 0$.

Note that $Y_t = \sum_{j=1}^{t} u_j$ and $EY_t^2 = t\sigma^2$.

Random walk is *not* covariance-stationary since it violates one of the conditions of weak stationarity (in this case, finite variance). $\Delta Y_t (= Y_t - Y_{t-1})$, however, is covariance-stationary.

For $y_1, y_2, \cdots, y_T$, a sample path of a stationary process Y_t, we can estimate the sample auto-covariance function.

The *sample auto-covariance function* is defined by:

$$\hat{\gamma}(k) = \frac{1}{T} \sum_{j=1}^{T-k} (y_{j+k} - \overline{y})(y_j - \overline{y}), \quad 0 \leq k < T,$$

where $\hat{\gamma}(k) = \hat{\gamma}(-k)$, $-T < k \leq 0$, and $\overline{y} = \frac{1}{T} \sum_{t=1}^{T} y_t$.

ARMA processes

Stationary ARMA processes specify Y_t as a function of *current and past* realizations of white noise.

A stochastic process, U_t, is called *white noise* (WN) with mean zero and variance σ^2 if

$$\begin{aligned} \gamma(0) &= \sigma^2 \\ \gamma(k) &= 0, \quad k \neq 0. \end{aligned}$$

An ARMA(p,q) process is described, for each $t = 0, \pm 1, \pm 2, \dots,$ by the following equation:

$$Y_t - \phi_1 Y_{t-1} - \phi_2 Y_{t-2} - \cdots - \phi_p Y_{t-p} = U_t + \theta_1 U_{t-1} + \theta_2 U_{t-2} + \cdots + \theta_q U_{t-q},$$

where $U_t \sim WN(0, \sigma^2)$.

In a more compact notation, a mean zero ARMA(p,q) process is defined by

$$\phi(L)Y_t = \theta(L)U_t,$$

where $\phi(L) = \phi_0 L^0 - \phi_1 L - \phi_2 L^2 - \ldots - \phi_p L^p$,

$\theta(L) = \theta_0 L^0 + \theta_1 L + \theta_2 L^2 + \ldots + \theta_q L^q$,

$\theta_0 \equiv 1$, $\phi_0 \equiv 1$, and L is the lag operator so that $L^k x_t = x_{t-k}$, $\forall k = 0, \pm 1, \pm 2, \ldots$.

An MA(q) process is obtained by setting $\phi(L) \equiv 1$:

$$Y_t = \theta(L)U_t,$$

where $\theta(L) = 1 + \theta_1 L + \theta_2 L^2 + \ldots + \theta_q L^q$.

Similarly, an AR(p) process is obtained by setting $\theta(L) \equiv 1$:

$$\phi(L)Y_t = U_t.$$

An ARMA(p,q) process is said to be *stationary* if the roots of the AR polynomial,

$$1 - \phi_1 z - \phi_2 z^2 - \ldots - \phi_p z^p = 0,$$

are greater than 1 in modulus, i.e., lie outside the unit circle.

E.g., an AR(1) process: $(1 - \phi L) Y_t = U_t$. It is stationary if the root of $1 - \phi z = 0$ is greater than 1 in absolute value. It happens if $|z| = |\phi^{-1}| > 1$, or if $|\phi| < 1$. We already know that if ϕ is equal to 1, the process is not covariance-stationary; the same applies to all AR(1) processes with $|\phi| > 1$.

We can express an AR(1) process as

$$Y_t = \frac{U_t}{(1 - \phi L)} = (1 + \phi L + \phi^2 L^2 + \phi^3 L^3 + \dots) U_t$$
$$= U_t + \phi U_{t-1} + \phi^2 U_{t-2} + \phi^3 U_{t-3} + \dots .$$

An AR(1) process with $|\phi| < 1$ is an example of a *causal* process—the process that assumes some value at time t, which is independent of all future shocks $\{U_s\}, s > t$.

Thus, an AR(1) process can be represented by an MA process of infinite order with particular restrictions on the moving average coefficients.

An ARMA(p,q) process is said to be *invertible* if the roots of the MA polynomial,

$$1 + \theta_1 z + \theta_2 z^2 + \ldots + \theta_q z^q = 0$$

are greater than 1 in modulus, i.e., lie outside the unit circle.

For an MA(1) process, this means that $|\theta| < 1$.

Note also that an MA(1) process

$$Y_t = (1 + \theta L)U_t = (1 - (-\theta)L)U_t$$

can be expressed as $\frac{Y_t}{(1-(-\theta)L)} = U_t$, or

$$(1 + (-\theta)L + (-\theta)^2 L^2 + (-\theta)^3 L^3 + \ldots)Y_t = U_t,$$

or

$$Y_t = \theta Y_{t-1} - \theta^2 Y_{t-2} + \theta^3 Y_{t-3} - \ldots + U_t = -\sum_{j=1}^{\infty}(-\theta)^j Y_{t-j} + U_t.$$

That is, an invertible MA(1) process can be represented by an AR process of infinite order.

The Auto-Covariance Function

If the process is causal (stationary), and expressed as

$$Y_t = \psi(L)U_t = \sum_{j=0}^{\infty} \psi_j U_{t-j}, \text{ with } U_t \sim WN(0, \sigma^2), \text{ we can}$$

calculate the auto-covariance function as:

$$\gamma(0) = \sigma^2(\psi_0^2 + \psi_1^2 + \psi_2^2 + \psi_3^3 + \psi_4^2 + \dots) = \sigma^2 \sum_{j=0}^{\infty} \psi_j^2$$

$$\gamma(1) = \sigma^2(\psi_0\psi_1 + \psi_1\psi_2 + \psi_2\psi_3 + \psi_3\psi_4 + \dots)$$

$$\gamma(2) = \sigma^2(\psi_0\psi_2 + \psi_1\psi_3 + \psi_2\psi_4 + \psi_3\psi_5 + \dots)$$

$$\gamma(3) = \sigma^2(\psi_0\psi_3 + \psi_1\psi_4 + \psi_2\psi_5 + \psi_3\psi_6 + \dots)$$

$$\vdots$$

More succinctly, $\gamma(k) = \sigma^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+|k|}$.

Since $\sum_{j=0}^{\infty} |\psi_j|$ (termed "absolute summability") implies $\sum_{j=0}^{\infty} \psi_j^2$ (termed "square summability"), assuming the former guarantees that the variance of Y_t, $\gamma(0)$, is finite, one of the assumptions of weak stationarity.

Auto-covariance function for an AR(1) process

If $(1-\phi L)Y_t = U_t$, then $\psi_j = \phi^j$, $j = 0, \ldots, \infty$.

$$\gamma(0) = \sigma^2(1+\phi^2+\phi^4+\phi^6+\ldots) = \frac{\sigma^2}{1-\phi^2}$$
$$\gamma_1 = \sigma^2\phi(1+\phi^2+\phi^4+\phi^6+\ldots) = \phi\frac{\sigma^2}{1-\phi^2}$$
$$\gamma(k) = \phi\gamma(k-1), k \geq 1.$$

Auto-covariance function for an MA(1) process

If $Y_t = (1 + \theta L)U_t$, $\psi_0 = 1$, $\psi_1 = \theta$, $\psi_j = 0$, $j > 1$.

$$\gamma(0) = \sigma^2(1 + \theta^2)$$
$$\gamma(1) = \theta\sigma^2$$
$$\gamma(k) = 0, \quad k > 1.$$

If the process is an MA(q), the auto-covariance function is zero for $k > q$.

Back to the PIH

$$\Delta c_t = \underbrace{\frac{r}{1+r} \sum_{j=0}^{\infty} \left(\frac{1}{1+r} \right)^j (E_t - E_{t-1}) y_{t+j}}_{\text{innovation in the permanent income}} .$$

We need to know $(E_t - E_{t-1})y_t$, $(E_t - E_{t-1})y_{t+1}$, $(E_t - E_{t-1})y_{t+2}$, $\ldots$.

Assume that the process is a covariance-stationary MA(∞) process:

$$Y_t = \theta(L)U_t = U_t + \theta_1 U_{t-1} + \theta_2 U_{t-2} + \theta_3 U_{t-3} + \ldots .$$

Then,

$$\begin{aligned}
(E_t - E_{t-1})y_t &= u_t \\
(E_t - E_{t-1})y_{t+1} &= \theta_1 u_t \\
(E_t - E_{t-1})y_{t+2} &= \theta_2 u_t \\
(E_t - E_{t-1})y_{t+3} &= \theta_3 u_t \\
&\vdots
\end{aligned}$$

For this income process,

$$\begin{aligned}\Delta c_t &= \frac{r}{1+r}\left[u_t + \frac{\theta_1}{1+r}u_t + \frac{\theta_2}{(1+r)^2}u_t + \frac{\theta_3}{(1+r)^3}u_t + \dots\right]\\ &= \frac{r}{1+r}u_t\left[1 + \frac{\theta_1}{1+r} + \frac{\theta_2}{(1+r)^2} + \frac{\theta_3}{(1+r)^3} + \dots\right]\end{aligned}$$

Note that $1 + \frac{\theta_1}{1+r} + \frac{\theta_2}{(1+r)^2} + \frac{\theta_3}{(1+r)^3} + \dots = \theta(L)|_{L\equiv\frac{1}{1+r}}$.
Thus,

$$\Delta c_t = \underbrace{\frac{r}{1+r}\left[\theta\left(\frac{1}{1+r}\right)\right]}_{\text{MPC out of the shock}} u_t.$$

If Y_t is some general ARMA(p,q) process, $\phi(L)Y_t = \theta(L)U_t$ so that $Y_t = \frac{\theta(L)}{\phi(L)}U_t$, it can be shown that

$$\Delta c_t = \frac{r}{1+r}\frac{\theta\left(\frac{1}{1+r}\right)}{\phi\left(\frac{1}{1+r}\right)}u_t.$$

Note that the polynomial $\phi(L)$ may have a unit root.

Excess smoothness

Aggregate income in macro data is well fit by the following model

$$\begin{aligned}
\Delta y_t &= \mu + \alpha \Delta y_{t-1} + u_t \\
(1-L)y_t &= \mu + \alpha(1-L)Ly_t + u_t \\
(1-L)(1-\alpha L)y_t &= \mu + u_t \\
y_t &= \frac{\mu}{(1-L)(1-\alpha L)} + \frac{u_t}{(1-L)(1-\alpha L)} \\
y_t &= \tilde{\mu} + \frac{u_t}{(1-L)(1-\alpha L)}.
\end{aligned}$$

Utilizing our formula for consumption changes in accordance with the PIH

$$\Delta c_t = \frac{r}{1+r}\frac{1}{(1-\frac{1}{1+r})(1-\frac{\alpha}{1+r})}u_t = \frac{1+r}{1+r-\alpha}u_t.$$

Excess smoothness, contd.

Note that α is estimated to be positive in the aggregate data. Thus, the MPC out of the shock to aggregate income should be *greater* than 1. Implications for the variances are such that

$$var^{PIH}(\Delta c_t) = \left(\frac{1+r}{1+r-\alpha}\right)^2 \sigma_u^2 > \sigma_u^2.$$

The variance of the innovation to consumption should be larger than the variance in the innovation to income. In the data, the reverse is true.

This result is known as the excess smoothness of consumption.

PIH and structural income processes

Let's assume that the income process consists of two components—the permanent component and the transitory component.

$$
\begin{aligned}
y_t &= \tau_t + w_t \\
\tau_t &= \mu + \tau_{t-1} + u_t^P \\
w_t &= \theta(L) u_t^T,
\end{aligned}
$$

where τ_t is the permanent (in macro: "trend") component, u_t^P is the permanent shock;

w_t is the mean-reverting, stationary component (in macro: "cycle"), u_t^T is the transitory shock.

$u_t^P \sim iid(0, \sigma^2_{u^P})$, $u_t^T \sim iid(0, \sigma^2_{u^T})$, and u_t^T, u_t^P are uncorrelated at all leads and lags.

We assume that a consumer is able to differentiate between the permanent shocks (e.g., due to such events as promotion/demotion, permanent disability, etc.) and transitory shocks (e.g., those emanating from temporary sickness, short spells of unemployment, bonuses, overtime, etc.) to his income.

We want to predict the consumer's reaction to these distinct shocks assuming the PIH is true.

We can express income in terms of current and past shocks as:

$$(1-L)y_t = \Delta\tau_t + \Delta w_t = \mu + u_t^P + (1-L)\theta(L)u_t^T, \quad \text{or}$$
$$y_t = \tilde{\mu} + (1-L)^{-1}u_t^P + \theta(L)u_t^T,$$

where $\tilde{\mu} = (1-L)^{-1}\mu$.

The PIH implies:

$$\Delta c_t = \frac{r}{1+r}\left[\left(1-\frac{1}{1+r}\right)^{-1}u_t^P + \theta\left(\frac{1}{1+r}\right)u_t^T\right],$$

where $\theta(L) = 1 + \theta_1 L + \theta_2 L^2 + \dots$, and
$\theta\left(\frac{1}{1+r}\right) = 1 + \frac{\theta_1}{1+r} + \frac{\theta_2}{(1+r)^2} + \dots$.

In micro data, $\Delta y_t \sim MA(2)$ (e.g., Abowd and Card 1989). Thus, $w_t \sim MA(1)$ and $\theta(L) = 1 + \theta L$. It follows that

$$\Delta c_t = u_t^P + \frac{r}{1+r}\left(1 + \frac{\theta}{1+r}\right) u_t^T = u_t^P + \frac{r}{1+r}\left(\frac{1+r+\theta}{1+r}\right) u_t^T.$$

Meghir and Pistaferri (2004): $\hat{\theta} \in [0.17, 0.5101]$. Assume $r = 0.02$.

For $\theta = 0.17$, $\frac{r}{1+r}\left(\frac{1+r+\theta}{1+r}\right) \approx 0.023$; for $\theta = 0.5101$, $\frac{r}{1+r}\left(\frac{1+r+\theta}{1+r}\right) \approx 0.029$—a very tiny fraction.

Thus, the MPC out of the permanent shock is 1 and the MPC out the transitory shock is at most 0.03 (3 cents per $1.)