

12 – La nuova macroeconomia classica

Cominciamo lo studio della macroeconomia trattando il caso dei mercati perfetti.

La **nuova macroeconomia classica (NMC)** studia per l'appunto i **mercati perfetti**, ossia mercati dove valgono tre ipotesi fondamentali:

- Gli individui sono illimitatamente razionali: ossia sono in grado di prendere decisioni ottimali.
- L'informazione è liberamente disponibile a tutti.
- I mercati sono di concorrenza perfetta (le imprese sono *price-taker*).

Nella prima parte del capitolo cercheremo di determinare il livello del PIL, dell'occupazione, del salario reale e del tasso di interesse reale in un'**economia stazionaria**, in cui cioè:

- Le famiglie consumano interamente il reddito.
- Le imprese non effettuano investimenti aggiuntivi.
- Lo stock di capitale e lavoro sono costanti e dunque anche il PIL.

Il modello sarà poi successivamente arricchito in modo da poter ignorare queste ipotesi e poter spiegare anche l'andamento fluttuante del PIL, la formazione di risparmio e investimento e l'utilizzo della moneta.

12.1 Il metodo neo-classico

Analizziamo ora i principi della metodologia della NMC.

Il primo principio della NMC è una struttura logica tipica dell'**equilibrio generale macroeconomico**, ossia:

- Identificazione e caratterizzazione dei soggetti economici rilevanti (famiglie e imprese).
- Identificazione e caratterizzazione dei mercati in cui tali soggetti interagiscono (mercato del lavoro, del capitale, del prodotto finale).
- Determinazione di prezzi e quantità scambiate in equilibrio su tutti i mercati.

Il secondo principio della NMC è il **principio dell'agente rappresentativo** (microfondazioni). Ossia, poiché l'economia è composta da un numero molto elevato di consumatori, imprese e prodotti, dobbiamo trovare il modo di aggregarli. Questo modo consiste nel considerare una famiglia rappresentativa, un'impresa rappresentativa e un unico bene omogeneo.

Il terzo principio della NMC, nonché un aspetto molto importante delle decisioni razionali riguarda il futuro, le decisioni *forward looking* o intertemporali: in questo campo la NMC ha introdotto un'innovazione metodologica, ossia le **aspettative razionali**. Tuttavia le aspettative razionali richiedono strumenti di analisi troppo sofisticati, assumeremo dunque un caso limite di queste, ovvero che gli agenti economici abbiano una conoscenza certa di tutte le variabili, presenti e future.

12.2 Prodotto e occupazione

Studiamo dunque, nel caso di un'economia stazionaria, i problemi di ottimo dell'impresa rappresentativa, della famiglia rappresentativa e di quali risultati dobbiamo attenderci dalla loro interazione sui vari mercati.

12.2.1 L'impresa rappresentativa

Ricordiamo che la nostra è un'economia capitalistica in cui le imprese sono unità produttive private, il cui capitale è posseduto dalle famiglie e la cui ragione sociale è la massimizzazione del reddito del capitale (profitto). Supponiamo tuttavia, come illustrato dalla NMC, che vi sia un'unica impresa rappresentativa.

DOMANDA DI FATTORI

Partiamo dalla descrizione della tecnologia produttiva, rappresentata dalla funzione di produzione:

$$Y_t = Z_t Y(K_t, N_t)$$

ove Y_t è la produzione, Z_t è una variabile non ancora ben definita, K_t è il capitale e N_t è il lavoro.

Supponiamo che essa sia crescente al crescere di ogni fattore (capitale e lavoro), con produttività marginale (PM) decrescente per ogni fattore e con rendimenti di scala costanti (ossia la produzione aumenta o diminuisce nella stessa proporzione in cui aumentano o diminuiscono entrambi i fattori). L'ultima proprietà ci garantisce che $Z_t Y(K_t, N_t) = Y(Z_t K_t, Z_t N_t)$, cioè Z_t rappresenta variazioni comuni della produttività dei fattori che hanno un effetto proporzionale sul prodotto totale. Questa variabile è importante per catturare gli effetti prodotti da innovazioni tecniche e organizzative ed è dunque legata alla/coincide con la **produttività totale dei fattori (PTF)**.

La nostra impresa rappresentativa è dotata di una quota fissa di capitale K_t che deve sfruttare al massimo al fine di massimizzarne il rendimento $K_t R_t$, ove R_t è il tasso di interesse annuo lordo pagato per unità di capitale. R_t comprende il tasso di interesse netto r_t dovuto alla proprietà e comprende anche l'usura del capitale nel processo produttivo ($0 \leq d \leq 1$) che assumeremo pari al 100%, dunque $R_t = 1 + r_t$. Assumeremo anche che la proprietà ogni anno rifornisce l'impresa dell'intero capitale usurato, cioè fa un piano di ammortamento del 100%, $A_t = K_t$.

Bisogna inoltre considerare l'impiego del fattore lavoro N_t dato il saggio di salario reale (per ora lavorata) W_t , il quale è stabilito dal mercato. Non essendoci settore pubblico o estero, il PIL coincide con il RNL, ossia l'intero valore della produzione Y_t viene distribuito alle famiglie, ossia

$$Y_t = N_t W_t + K_t R_t$$

Si noti che $N_t W_t$ è la quota di prodotto nazionale che va al fattore lavoro, mentre $K_t R_t$ è quella che va al fattore capitale.

Il problema di massimo è dunque il seguente

$$\begin{array}{ll} \max & K_t R_t = Y_t - N_t W_t \\ \text{t.c.} & Y_t = Z_t Y(K_t, N_t) \end{array}$$

L'impresa domanda lavoro N_t^d fino al punto in cui la produttività marginale del lavoro (PML) uguaglia il salario, ossia

$$PML := Z_t Y'(N_t^d) = W_t$$

1. Fissati il capitale e la PTF, la domanda di lavoro è decrescente rispetto al salario.
2. La domanda di lavoro è crescente rispetto al capitale e alla PTF.

Il massimo tasso di interesse (lordo) che l'impresa può pagare sul capitale deve soddisfare la condizione:

$$PMK := Z_t Y'(K_t^d) = R_t$$

1. Il massimo tasso di interesse sul capitale è decrescente rispetto alla dotazione di capitale.
2. Il massimo tasso di interesse sul capitale è crescente rispetto al livello del fattore lavoro e della PTF.

Per ora ci occupiamo di un'economia nella quale le imprese hanno una dotazione di capitale data, e quindi pagano il tasso d'interesse lordo R corrispondente secondo la funzione K^d .

L'OFFERTA DI PRODOTTO

Ora l'impresa deve chiedersi che livello di produzione attuare. Conoscendo le quantità di lavoro e capitale di massimo profitto, che denoteremo rispettivamente con N_t^* e K_t^* , il **livello di produzione ottimale** è dato dalla quantità dei fattori produttivi di massimo profitto combinati secondo la tecnologia disponibile, ossia

$$Y_t^* = Z_t Y(K_t^*, N_t^*)$$

Abbiamo così ottenuto la funzione di offerta aggregata neo-classica, che dunque indica la massima produzione efficiente che può erogare il sistema economico, detta anche prodotto potenziale.

12.2.2 La famiglia rappresentativa

Consideriamo una famiglia rappresentativa, titolare dei fattori produttivi (lavoro e capitale) dal cui impiego nella produzione ottiene i rispettivi redditi, in base al salario unitario e al tasso d'interesse annuo lordo.

Assumiamo che la dotazione di capitale K_t sia data e costante. Per mantenere fede all'ipotesi di economia stazionaria, ogni anno la famiglia deve rimettere a disposizione dell'impresa lo stesso stock di capitale, cioè l'ammortamento A_t è del 100% ossia $A_t = K_t$. La famiglia dispone di un reddito netto dei fattori (YN) pari a

$$YN_t = Y_t - A_t = N_t W_t + K_t r_t$$

OFFERTA DI FATTORI

Secondo i principi standard di massimizzazione dell'utilità, l'offerta di lavoro di una famiglia nasce dal confronto tra l'utilità del consumo C_t per ottenere il quale occorre lavorare percependo il salario W_t , e l'utilità del tempo libero L_t , per il quale il salario rappresenta il costo-opportunità. Ovviamente c'è un massimo di tempo fisiologico disponibile, ovvero $T \leq 24 \cdot 365$. Dunque, dopo aver scelto le ore di lavoro N_t , il tempo libero sarà dato da $L_t = T - N_t$.

$$\begin{aligned} \max \quad & U_t = U(C_t, L_t) \\ \text{t.c.} \quad & C_t = N_t W_t + K_t r_t \quad L_t + N_t = T \end{aligned}$$

Risolviamo il problema di ottimo ponendo il SMS uguale al prezzo relativo dei beni

$$\begin{aligned} SMS_{L,C} &:= U'(C_t)/U'(L_t) = W_t \\ C_t &= N_t W_t + K_t r_t \end{aligned}$$

La soluzione standard prevede che l'offerta di lavoro sia crescente rispetto al salario. Se la ricchezza della famiglia aumenta, essa domanda sia più beni di consumo sia più tempo libero, ovvero riducono l'offerta di lavoro.

Il **tempo di lavoro ottimale** della nostra famiglia rappresentativa N_t^S , dato il salario W_t , è l'equivalente della forza lavoro (se trattato in senso aggregato).

DOMANDA DI CONSUMO

Siccome l'utilità cresce al crescere del consumo, ogni famiglia sceglie il massimo consumo possibile, cioè tutto il reddito netto. Quindi la **domanda di consumo ottimale** sarà:

$$C_t = N_t^S W_t + K_t r_t$$

Il risultato è coerente con l'assunzione di stato stazionario: non c'è risparmio né accumulazione di capitale. La **domanda aggregata** è dunque:

$$Y_t^d = C_t + A_t$$

12.2.3 L'equilibrio generale macroeconomico

Le relazioni tra le variabili che ci interessano possono essere espresse in forma lineare (trasformandole in logaritmi), e, per distinguerle da quelle originarie, indicheremo le variabili con lettere minuscole.

IL MERCATO DEL LAVORO

Domanda di lavoro: $n_t^d = n_o^d - n_w^d w_t$

Offerta di lavoro: $n_t^s = n_o^s + n_w^s w_t$

Entrambe le funzioni presentano un parametro di elasticità rispetto al saggio di salario reale (denotato dal pedice w) e un parametro indipendente (denotato dal pedice o), il quale misura l'effetto delle variabili esogene che influenzano domanda e offerta diverse dal salario:

- n_o^d comprende lo stock di capitale k_t delle imprese e la PTF z_t .
- n_o^s comprende la ricchezza delle famiglie.

L'equilibrio di mercato richiede l'uguaglianza tra domanda e offerta:

$$n_o^d - n_w^d w_t = n_o^s + n_w^s w_t$$

da cui si ricava il valore del salario (variabile endogena) che realizza l'equilibrio di mercato:

$$w_t^* = \frac{n_o^d - n_o^s}{n_w^d + n_w^s}$$

Siamo dunque in regime di **piena occupazione**, che indichiamo con n_t^* . Attenzione: con piena occupazione si intende che sono occupati tutti coloro che desiderano lavorare in base al salario di mercato vigente w_t^* .

Rappresentando le due rette su un piano cartesiano $n \times w$, si può notare che:

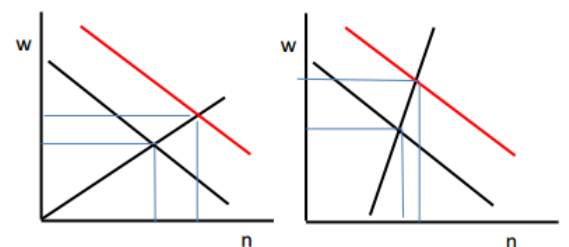
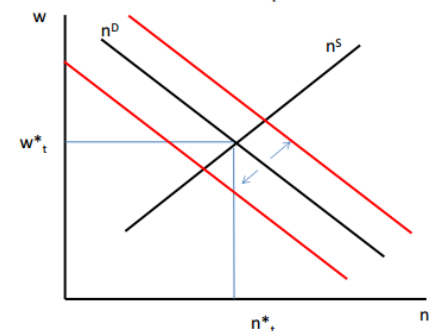
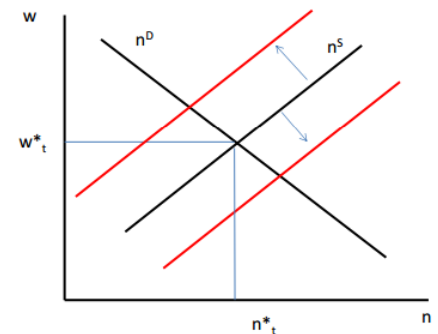
- All'aumentare di n_o^d , ossia all'aumentare di stock di capitale k_t e della PTF z_t , le imprese sono disposte ad aumentare il salario w_t^* e/o l'occupazione n_t^* . La composizione tra effetto salario ed effetto occupazione dipende ovviamente dall'elasticità della curva di offerta: se l'elasticità è alta (retta molto inclinata), prevale l'effetto salario, dunque le imprese sono disposte ad aumentare l'occupazione ma le famiglie richiedono un sostanzioso aumento di salario per lavorare di più; viceversa se l'elasticità è bassa.
- All'aumentare di n_o^s , ossia all'aumentare del reddito capitale delle famiglie, queste sono disposte a lavorare di meno e per un salario maggiore, dunque diminuisce l'occupazione n_t^* ma aumenta il salario w_t^* .

IL MERCATO DEL CAPITALE

Ricordiamo che in questo mercato l'offerta coincide con la dotazione di capitale che le famiglie concedono alle imprese e che, per supposizione della NMC, essa è fissa a k_t^* .

Domanda di capitale: $k_t^d = k_o^d - k_r^d r_t$

Dotazione/offerta di capitale: k_t^*



Qui, al solito, k_o^d è la variabile esogena, mentre k_r^d è l'elasticità rispetto al tasso di interesse reale r_t . Quindi l'uguaglianza tra domanda e offerta risulta:

$$k_o^d - k_r^d r_t = k_t^*$$

da cui si ricava che il massimo/ottimo tasso d'interesse che le imprese possono pagare ai possessori del capitale è

$$r_t^* = \frac{k_o^d}{k_r^d} - \frac{k_t^*}{k_r^d}$$

Rappresentando le due rette su un piano cartesiano $k \times r$, si può notare che:

- All'aumentare di k_o^d , ossia all'aumentare della PTF z_t , aumenta la domanda di capitale e, data la disponibilità fissa, ciò si traduce interamente in un aumento del tasso d'interesse.
- All'aumentare del capitale k_t^* , diminuisce il tasso d'interesse di equilibrio r_t^* .

IL MERCATO DEL PRODOTTO

Ora dobbiamo determinare il livello della produzione decisa dalle imprese, ovvero la loro **offerta aggregata**.

Funzione di produzione: $y_t(n_t, k_t) = z_t + a k_t + b n_t$

Se la produzione gode di rendimenti di scala costanti, come ipotizzato all'inizio, allora vale che $a + b = 1$. Ora inseriamo nella funzione di produzione la quantità ottima di fattore lavoro n_t^* e di capitale k_t^* (che è l'unica possibile, essendo fissa), ottenendo che:

$$y_t^* = z_t + a k_t^* + b n_t^*$$

A fronte di questa offerta di prodotto abbiamo la sua domanda. Ora in teoria dovremmo uguagliare, al solito, domanda e offerta di prodotto al fine di ottenere la situazione di equilibrio. Tuttavia in questo caso non serve in quanto possiamo utilizzare il seguente risultato.

Legge di Walras: In un sistema economico generale, se tutti i mercati meno uno sono in equilibrio, e i bilanci di tutti gli agenti sono in pareggio, allora anche il restante mercato deve essere in equilibrio.

Dunque poiché i mercati dei due fattori (lavoro e capitale) sono in equilibrio, lo è anche quello del prodotto. Proviamo comunque a verificarlo con le nozioni di contabilità nazionale: poiché i mercati dei fattori sono in equilibrio, sappiamo che il valore del PIL risulta pari al RNL:

$$Y_t^* = N_t^* W_t^* + K_t^* R_t^*$$

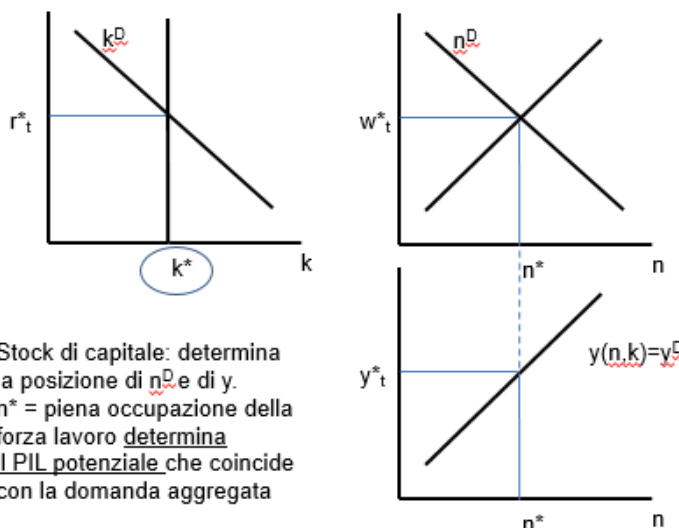
Sappiamo che le famiglie utilizzano parte del RNL per il consumo $C_t = N_t^* W_t^* + K_t^* r_t^*$ e la parte rimanente per il rimpiazzo del capitale A_t . Di conseguenza la domanda aggregata risulta:

$$Y_t^d = C_t + A_t$$

e, siccome $A_t = K_t$, ricordando che $R_t = 1 + r_t$, abbiamo che

$$Y_t^d = C_t + K_t^* = N_t^* W_t^* + K_t^* r_t^* + K_t^* = Y_t^*$$

Abbiamo dunque provato il seguente risultato.



Legge di Say: In un sistema in equilibrio generale, quale che sia il livello ottimale di produzione deciso dalle imprese, è sempre assorbito dalla domanda aggregata: **l'offerta crea la domanda**.

Si noti che essendo $A_t = K_t$, si ha che A_t rientra comunque nella domanda di prodotto, ovverosia non ci sono perdite. Si ha dunque l'uguaglianza risparmio-investimento.

Morale: Il sistema neo-classico di base, partendo dall'analisi del processo di scelta ottimale dei consumatori e delle imprese, è in grado di determinare contemporaneamente i livelli di equilibrio del PIL, della sua componente di consumo, del livello di occupazione, dei redditi dei fattori e di mettere in luce le relazioni che intercorrono tra i vari mercati.

12.3 La distribuzione del reddito

Secondo la NMC le quote distributive del reddito dei due fattori (lavoro e capitale) sono determinate in relazione alla funzione di produzione, ovverosia dal relativo contributo alla produzione totale misurato dal prodotto marginale di ciascun fattore. Infatti, dato che

$$RNL_t = N_t W_t + K_t R_t \quad W_t = PML \quad R_t = PMK$$

la distribuzione funzionale risulta:

$$\text{Lavoro:} \quad \frac{N_t}{RNL_t} PML$$

$$\text{Capitale:} \quad \frac{K_t}{RNL_t} PMK$$

Dunque alla base della teoria della distribuzione del reddito neo-classica sta l'idea che ciascun fattore riceve una quota del prodotto sociale commisurata al contributo che esso ha dato per crearlo.

Ciò rispetto i principi di efficienza e di equità sociale. Al contrario di come invece sostenevano teorie, come quelle marxiste, per le quali la distribuzione dei redditi era il risultato di uno sfruttamento sociale da parte di una classe e a danno di altre.

Per la NMC, dallo studio del problema distributivo occorre eliminare qualsiasi connotazione socio-politica. Le quote di reddito dipendono solo dalla tecnologia, che determina in ultima analisi le produttività marginali. Poiché la tecnologia è un fenomeno strutturale, che tende a variare molto lentamente nel tempo, le quote destinate a capitale e lavoro devono necessariamente rispecchiare tale tendenza.

Per la NMC, tuttavia, le variazioni della PTF non generano cambiamenti rilevanti sulle quote distributive quanto possono invece fare le seguenti situazioni:

- Situazione di monopolio: la quota di reddito capitale aumenta a scapito di quella del lavoro.
- Presenza di organizzazioni sindacali: esercitano il loro potere di contrattazione per spostare parte del reddito a favore del lavoro.
- Intervento dello Stato: crea una distribuzione del reddito dopo le imposte diversa da quella prima delle imposte, e generalmente più favorevole alle classi meno abbienti.

12.4 Fluttuazioni macroeconomiche

Come anticipato all'inizio, cominciamo ad introdurre ulteriori elementi al modello macroeconomico al fine di poter spiegare i seguenti due fatti:

1. Le continue oscillazioni dell'attività economica aggregata.
2. La presenza di un certo numero di lavoratori che risulta non occupato.

La spiegazione base delle fluttuazioni macroeconomiche fornita dalla NMC deriva dal fatto che un sistema economico perfettamente efficiente produce ogni anno il suo PIL ottimale (o potenziale)

$$y_t^* = z_t + ak_t^* + bn_t^*$$

L'andamento del PIL potenziale non è direttamente osservabile, in quanto esso è il risultato di stime statistiche. Esso, per lo stesso motivo, non coincide esattamente con quello del PIL effettivo, tuttavia le differenze sono ritenute errori statistici relativamente piccoli e non sistematici. Dunque il PIL potenziale è considerato sia una valida spiegazione teorica, sia una buona approssimazione empirica, del fenomeno della crescita fluttuante.

Le variazioni in termini percentuali in un periodo t rispetto al precedente si possono calcolare agevolmente, per approssimazione, come somma delle variazioni delle componenti:

$$\hat{y}_t^* = \hat{z}_t + a\hat{k}_t^* + b\hat{n}_t^*$$

Poiché in questa visione la causa ultima del ciclo economico è legata ai fattori produttivi di natura reale, mentre si tende a minimizzare l'importanza dei fenomeni di natura monetaria o finanziaria, la teoria in questione prende il nome di **teoria del ciclo economico reale**. Per questa teoria le fluttuazioni avvengono in un contesto di mercati costantemente in equilibrio. Inoltre, le fluttuazioni non sono deviazioni dal *trend*, in quanto, per questa teoria che rifiuta la concezione di *trend lineare*, l'economia è sempre sul *trend* della propria capacità produttiva potenziale, solo che esso stesso varia in maniera irregolare a causa degli shock ai fattori produttivi.

I cicli economici sono stati in larga parte il risultato di cambiamenti inaspettati causati dal tasso di crescita della produttività dei fattori z ; gli altri fattori variano molto meno. Effettivamente, si può notare empiricamente che la PTF è la componente di variazione del PIL più rilevante anche in Italia.

Le ragioni più importanti alla base di variazioni così repentine della PTF sono attribuite a:

- Cambiamenti della tecnologia con la quale vengono combinati i fattori della produzione e le risorse nell'attività di produzione dei beni e servizi finali (es.: tecnologia che riduce gli scarti di lavorazione).
- Miglioramenti nelle capacità (*skills*) dei lavoratori.
- Introduzione di nuovi prodotti o processi (es.: introduzione del computer).
- Mutamenti dei prezzi internazionali delle materie prime (es.: prezzo del petrolio).
- Interventi regolativi della pubblica amministrazione (es.: legislazione ambientale più restrittiva).

All'aumentare della PTF aumentano la domanda di lavoro, la domanda di capitale e la produzione. Complessivamente (con offerta di lavoro elastica), le variazioni della PTF sono in grado di spiegare le variazioni del PIL e le variabili che risultano procicliche: occupazione, redditi dei fattori e consumo.

12.5 Il tasso di disoccupazione

Dunque per uno shock negativo della PTF diminuiscono sia l'occupazione, sia il PIL. Ma attenzione: nel modello analizzato finora, il mercato del lavoro è sempre in equilibrio. Ciò significa, tra le altre cose, che chi desidera lavorare al salario di mercato riesce sempre a trovare un impiego, mentre chi non lavora sceglie di non farlo perché considera quel salario troppo basso.

In altre parole, il modello neo-classico di base non ammette la possibilità di **disoccupazione involontaria**. Ciò sembra tuttavia scontrarsi con una chiara evidenza quotidiana.

12.5.1 La misurazione dei flussi nel mercato del lavoro

Lavoratori: Empiricamente si nota che il tasso di partecipazione alla forza lavoro (forza lavoro / popolazione attiva) è molto volatile. Ciò lascia intravedere una notevole vivacità in entrata e in uscita dal gruppo delle persone che periodicamente decidono se partecipare o meno al mercato del lavoro.

Imprese: L'ordine di grandezza del cosiddetto *gross job turnover* (GJT), vale a dire il numero complessivo di posti di lavoro creati e distrutti dalle imprese sul totale dei posti di lavoro presenti in un certo arco temporale, è attorno al 20% all'anno.

12.5.2 Una disoccupazione compatibile con l'equilibrio

La presenza di disoccupazione involontaria in un mercato in equilibrio generale competitivo è spiegata dalla presenza di "frizioni" durante l'incontro tra domanda e offerta di lavoro: **disoccupazione frizionale**.

Queste frizioni possono essere data della variazione costante di domanda e offerta, dai costi di acquisizione delle informazioni sulla presenza di posti di lavoro disponibili, dai costi di mobilità, ecc.

La disoccupazione così generata va a formare il cosiddetto **tasso naturale di disoccupazione**.

Questo genere di disoccupazione è sempre presente, tuttavia, per quanto banale, i disoccupati non sono sempre gli stessi. Basti immaginare la disoccupazione naturale come una cosa più o meno lunga per uno spettacolo, la quale si è formata per un problema di coordinamento tra flussi in entrata e flussi in uscita.

Il concetto di disoccupazione frizionale è dunque distinto da quello di un vero e proprio squilibrio del mercato del lavoro. Perciò per risolvere questo problema non bisogna intervenire sulle determinanti fondamentali del mercato (salario, domanda di lavoro, tecnologia,...) bensì sui tempi di attesa, agendo ad esempio sui sistemi di collocamento.

12.6 Moneta e prezzi

Tutti i fenomeni visti finora non richiedevano l'utilizzo della moneta, ora è giunto il momento di introdurla. Tre funzioni

1. **Mezzo di scambio:** lo scambio tra venditore e compratore è reso possibile dal trasferimento di denaro.
2. **Unità di conto:** permette di misurare e confrontare valori economici (un certo bene costa tot volte rispetto un altro bene, o il prezzo di quel bene è cambiato durante il tempo).
3. **Riserva di valore:** permette di trasferire potere di acquisto dal presente al futuro, dunque permette di risparmiare denaro, posticipare il consumo ad un periodo futuro.
(Centrale per Keynes, ma non per i neoclassici)

La moneta non è l'unica riserva di valore: azioni, obbligazioni, immobili, altre attività reali.

Ricchezza = complesso delle riserve di valore.

Liquidità: facilità di conversione di un valore patrimoniale in mezzo di scambio: moneta è la riserva di valore più liquida e non è una buona riserva di valore, il suo valore è influenzato in maniera significativa dal

livello dei prezzi (valore nominale costante, ma valore reale cambia nel tempo). L'azione e gli immobili invece devi trovare qualcuno che te li acquisti.

Lo stock di moneta:

- **Circolante:** denaro contante.
- **Depositi a vista:** saldi dei conti correnti bancari (assegni).
- **Altri strumenti:** libretti di risparmio, e altre attività che possono essere trasformate facilmente in contanti o depositi.

Maggiore è l'economia sommersa, maggiore è la quantità di denaro circolante. La moneta viene utilizzata come riserva di valore sia perché è liquida, sia per cose illegali.

Ci sono due interrogativi su cui abbiamo sorvolato fino ad ora:

1. Come avvengono gli scambi nell'economia che abbiamo studiato?
2. L'uguaglianza tra domanda e offerta di prodotto non ci dice nulla su quale sia il prezzo di equilibrio.

Ovviamente la risposta a questi interrogativi risiede nel concetto di **moneta**. Fino ad ora abbiamo trattato solo variabili reali, ovverosia rapportate al PIL (ad esempio dicendo che il salario è di 16, intendevamo che era 16 "chili" di PIL). Introducendo la moneta come mezzo di pagamento nasce il concetto di **prezzo monetario**, ossia il costo in termini di **unità di conto** (la moneta data).

Ricordando le formule del valore di equilibrio dei redditi dei fattori e indicando con P_t il valore di un "chilo" di PIL in euro, stiamo dicendo che:

$$P_t W_t = P'_t Y(N_t)$$

$$P_t R_t = P'_t (K_t)$$

L'irrelevanza dell'unità di conto, e del suo livello, è un postulato fondamentale del calcolo economico razionale. Il modello neo-classico è costruito sull'ipotesi di **assenza di illusione monetaria**, ossia che tutti gli agenti siano unicamente interessati alle grandezze reali, cioè in ultima istanza a quanto prodotto si può produrre e consumare piuttosto che a quanta moneta vale.

Però non c'è dubbio che, per quanto "finta", la moneta sembra avere effetti ben reali e concreti (es.: inflazione). Gli economisti classici stabilirono l'idea della "moneta velo" secondo cui da un lato si riconosce che essa agevoli in modo essenziale il funzionamento di un'economia di mercato, e dall'altro si vuole dimostrare che esso potrebbe farne a meno.

Noi ci limiteremo a vedere come la moneta e i prezzi monetari sono integrabili nel sistema economico della NMC mediante la **teoria quantitativa della moneta**.

Da cosa dipende la quantità di moneta all'interno di un sistema economico? Dall'incontro tra la domanda e l'offerta di moneta.

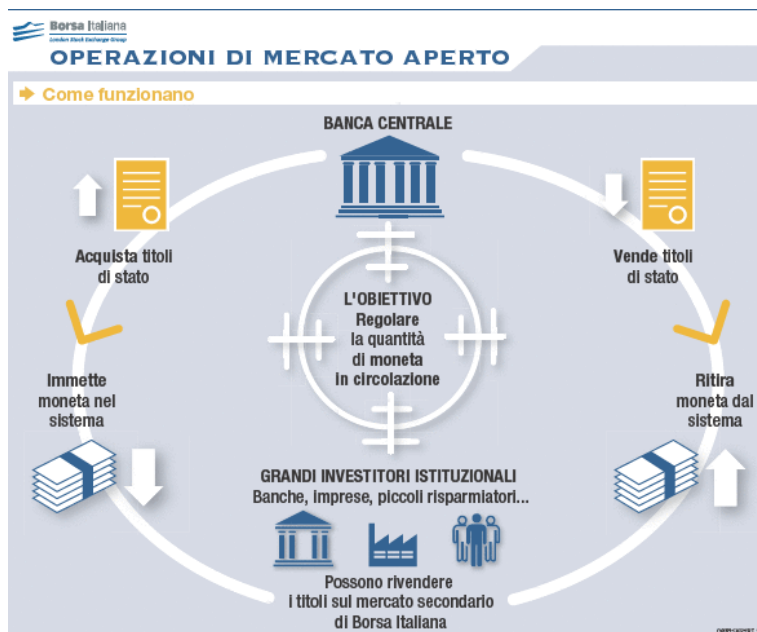
12.6.0 Il ruolo delle banche centrali

Le banche centrali supervisionano il sistema bancario, regolamentano la quantità di moneta (offerta di moneta) e possono agire come prestatori di ultima istanza in situazioni particolari, ossia salvano altre banche, ma ciò non ci interessa.

Offerta di moneta e operazioni di mercato aperto.

Aumento dell'offerta: acquisto di titoli di stato sul mercato obbligazionario (*quantitative easing*).

Riduzione dell'offerta: vendita di titoli di stato.



Banche e offerta di moneta

I depositi a vista sono gestiti da banche e sono parte dello stock di moneta. Quindi le banche possono influenzare l'offerta di moneta.

Il caso più semplice: il **sistema bancario a riserva totale** (ossia ammette solo depositi e non fa prestiti).

Circolante = 100 euro. Viene fondata la Banca Uno che accoglie solo depositi (non fa prestiti).

Depositi raccolti ma non impiegati = riserve.

Attivo e passivo di Banca Uno

Attivo		Passivo	
Riserve	100 euro	Depositi	100 euro

Banca Uno decide di usare parte dei depositi per fare prestiti, mantenendo una riserva minima: sistema bancario a riserva frazionaria.

Attivo		Passivo	
Riserve	10 euro	Depositi	100 euro
Impieghi	90 euro		

Tasso di riserva: quota del totale trattenuto come riserva.

La banca centrale determina la riserva obbligatoria: quota minima.

Nel regime precedente l'offerta di moneta era sempre di 100 euro (in depositi).

Con il nuovo regime l'offerta di moneta aumenta: i depositanti hanno 100 euro in depositi, ma chi si è indebitato ha 90 euro in contanti.

Offerta di moneta = 190 euro.

Se colloca a riserva solo una frazione della raccolta la banca crea moneta.

Nota: l'economia non è più ricca, ma solo più liquida: si creano attività ma anche debito.

I debitori di Banca Uno usano i 90 euro per acquistare beni da altre persone e queste depositano il circolante presso Banca Due.

Se il tasso di riserva è lo stesso, attivo e passivo di Banca Due sono:

Attivo		Passivo	
Riserve	9 euro	Depositi	90 euro
Impieghi	81 euro		

I debitori di Banca Due possono a loro volta usare gli 81 euro per acquistare beni da altre persone e queste possono depositare il circolante in Banca Tre ecc.

Deposito originario = 100 euro

Prestiti Banca Uno = 90 euro (0.9×100)

Prestiti di Banca Due = 81 euro (0.9×90)

Prestiti di Banca Tre = 72,90 euro (0.9×81)

...

Offerta totale di moneta = 1000 euro

Moltiplicatore monetario = quantità di moneta generata da un sistema a partire da ogni euro depositato. Nel nostro caso 10.

Se R è il tasso di riserva.

Moltiplicatore monetario = $1/R$

Nell'esempio $R=0.10$, moltiplicatore = 10

Dunque al crescere di R diminuisce la quota di depositi che la banca impiega e quindi l'ammontare di moneta creata dal sistema bancario.

IL CONTROLLO MONETARIO

La Banca centrale deve tenere conto dell'influenza delle attività bancarie sull'offerta di moneta.

Può tenere sotto controllo l'offerta di moneta con diversi strumenti:

1. **Operazioni di mercato aperto:** acquisto o vendita di titoli di stato.
Se crea moneta con l'acquisto, aumenteranno sia il circolante che i depositi e quindi l'offerta di moneta da parte delle banche. Se vende accade il contrario.
2. **Riserva obbligatoria:** può intervenire sul tasso di riserva obbligatorio riducendo o aumentando il moltiplicatore monetario.
3. **Tasso di sconto:** tasso di interesse applicato ai prestiti concessi alle banche. Le banche ricorrono a prestiti quando le riserve accantonate sono inferiori a quelle obbligatorie (ondate di prelievi, troppi impieghi). Quindi, quando la BC concede un prestito aumentano le riserve. Se il tasso di sconto aumenta si riducono le riserve e quindi la possibilità di creare moneta.
Se il tasso di sconto si riduce, le banche ridurranno le riserve (aumentando gli impieghi) e aumenterà anche la domanda di prestiti.
I prestiti possono essere concessi anche per aiutare le banche in difficoltà.

La BC non ha il controllo diretto dell'offerta di moneta.

Le famiglie decidono autonomamente di depositare il circolante e ciò crea le condizioni per l'aumento o la riduzione della quantità di moneta.

Le banche decidono autonomamente l'entità delle riserve (per la parte discrezionale) e degli impieghi.

I comportamenti di famiglie e banche possono però essere monitorati per mettere a punto strategie di intervento.

12.6.1 Domanda e offerta di moneta

Se introduciamo la moneta come unità di conto/mezzo di pagamento cosa succede?

Per prima cosa, nel sistema economico dovrà essere disponibile un certo ammontare di banconote, che chiamiamo **offerta di moneta** M_t^s . In secondo luogo, dovrà esserci un'istituzione pubblica che produce le banconote e ne garantisce l'autenticità, ossia la **banca centrale**. In terzo luogo occorre stabilire quale deve essere il valore totale della moneta nel sistema economico. Qui la teoria prescrive di guardare al fabbisogno di moneta che il pubblico richiede per effettuare i pagamenti connessi ai propri piani economici (consumo per le famiglie, acquisto dei fattori per le imprese): questo fabbisogno è la **domanda di moneta** M_t^d .

Indichiamo con P_t il prezzo in euro di un "chilo" di PIL nel corso dell'anno t ; si tratta dell'indice generale dei prezzi (IGP) o deflatore del PIL.

Partiamo dal caso più semplice in cui ogni compravendita è effettuata in contanti. Allora ogni singolo operatore necessita di una quantità di moneta esattamente uguale al proprio piano economico. Vale a dire:

- Famiglie: piano di consumo + ammortamenti $M_t^{dFAM} = P_t C_t + P_t A_t$
- Imprese: piano di produzione + ammortamenti $M_t^{dIMP} = P_t W_t N_t + P_t R_t K_t$

L'ipotesi di assenza d'illusione monetaria impone che i piani monetari siano identici a quelli reali, quindi sia le famiglie che le imprese soddisfano i propri vincoli di bilancio anche in valore monetario:

$$P_t C_t + P_t A_t = P_t Y_t = P_t W_t N_t + P_t R_t K_t$$

da cui si deduce che vale anche

$$M_t^{dFAM} = P_t Y_t = M_t^{dIMP}$$

vale a dire che in un'economia in cui la moneta è solo un mezzo di pagamento e tutte le transazioni sono in contanti, la domanda di moneta è uguale al valore monetario del PIL.

Questo risultato può essere modificato per tener conto di un aspetto importante dei sistemi monetari moderni: la possibilità di effettuare compravendite senza utilizzare contanti (es.: lettere di credito commerciali). Dunque, poiché il sistema monetario può far sì che solo una parte delle compravendite debba essere effettuata in contanti, possiamo introdurre questa caratteristica nella teoria quantitativa come **velocità di circolazione della moneta**, vale a dire la "frequenza" dell'impiego della moneta nelle compravendite misurata come:

$$V_t = \frac{P_t Y_t}{M_t}$$

Nel primo caso che abbiamo esaminato (tutte le compravendite sono in contanti) si ha che $V_t = 1$. Se per esempio si avesse che solo una compravendita su due avvenisse in contanti, si avrebbe che $V_t = 2$. La velocità di circolazione è vista come un fenomeno legato all'organizzazione del sistema economico, dunque non soggetto a mutamenti di breve periodo.

Dal lato della domanda, avremo una quantità di moneta in proporzione al PIL, corretta per la velocità di circolazione

$$M_t^d = \frac{P_t Y_t}{V_t}$$

L'equilibrio monetario richiede che l'offerta di moneta sia uguale alla domanda. Ponendo $M_t^d = M_t^s$, si deriva la cosiddetta **equazione quantitativa della moneta** (o **equazione degli scambi monetari**):

$$M_t^s V_t = P_t Y_t$$

12.6.2 Offerta di moneta, domanda aggregata e indice generale dei prezzi

Tuttavia la teoria quantitativa ripercorre il percorso in senso opposto, ossia: essendo l'offerta di moneta una variabile indipendente, controllata direttamente da un'istituzione autonoma ed esterna al mercato (esogena), da essa partiamo per spiegare la domanda di moneta per arrivare infine al valore monetario del PIL. Il problema che dunque ci poniamo ora è: in che misura l'offerta di moneta rimane in linea con la domanda, dati un certo IGP, il PIL reale e la velocità di circolazione della moneta?

La tesi centrale di illusione monetaria è che nessuno desidera detenere moneta "inutile" e tenderà di spenderla in beni reali. Viceversa, una insufficienza di moneta rispetto al proprio piano di spesa impone di ridurlo. Di conseguenza, aumenti o riduzioni dell'offerta di moneta, si traducono in aumenti o riduzioni della domanda aggregata.

La **funzione di domanda aggregata** che lega la domanda di prodotto al valore reale della moneta è:

$$Y_t^d = \frac{M_t^s V_t}{P_t}$$

Ora abbiamo una funzione di domanda aggregata autonoma per via della presenza della moneta come mezzo necessario per gli acquisti e per via del fatto che la quantità di moneta è controllata esogenamente dalla banca centrale.

La domanda aggregata è decrescente rispetto all'IGP: un IGP più alto riduce il potere d'acquisto.

Siccome l'offerta aggregata è prefissata e non influenzata da variabili monetarie, questo schema culmina nella determinazione dell'**IGP di equilibrio**, ossia quel valore dell'IGP tale che la domanda aggregata, data la moneta disponibile, è uguale all'offerta. Si deduce anche una teoria della **fluttuazioni dell'IGP** (e quindi anche dell'inflazione): le variazioni della domanda aggregata provocate dall'offerta di moneta non hanno effetti reali, ma solo sul livello dell'IGP.

In caso di un aumento dell'offerta di moneta, si ha un tentativo di disfarsi della moneta in eccesso comprando di più. Ma ciò fa solo aumentare l'IGP: cioè si ha **neutralità dell'offerta di moneta**.

12.6.3 Shock reali, shock monetari e inflazione

La teoria quantitativa della moneta offre anche un'estensione dello studio della fluttuazioni macroeconomiche. La spiegazione neo-classica del PIL potenziale individua le cause delle **fluttuazioni reali** (shock reali o dal lato dell'offerta) che abbiamo visto prima. La teoria della domanda aggregata basata sull'offerta di moneta individua anche le **fluttuazioni monetarie** (shock monetari o sul lato della domanda). Strettamente connessa è l'inflazione/deflazione.

Partiamo dall'offerta di moneta. Essa, a causa dell'incompetenza della banca centrale, può trovarsi al di sopra o al di sotto della domanda, con effetti diretti sull'IGP:

- Eccessi di offerta di moneta provocano eccessi di domanda di prodotto e inflazione (aumento dell'IGP).
- Insufficienze di offerta di moneta provocano insufficiente di domanda di prodotto e deflazione (diminuzione dell'IGP).

Considerando i tassi di variazione (differenza variabili con periodo precedente), per approssimazione dall'equazione degli scambi monetari si ottiene che:

$$\hat{m}_t + \hat{v}_t = \hat{p}_t + \hat{y}_t^*$$

Da cui si ricava direttamente il tasso di variazione dell'IGP

$$\hat{p}_t = \hat{m}_t + \hat{v}_t - \hat{y}_t^*$$

Dunque il tasso di variazione dell'IGP è dato dalla variazione dell'offerta di moneta rispetto alle variazioni della velocità di circolazione e del PIL. In particolare si ha inflazione ($\hat{p}_t > 0$) se l'offerta di moneta aumenta eccessivamente rispetto alle altre due determinanti dell'equilibrio monetario.

Va anche sottolineato il ruolo degli shock reali di offerta, che hanno un impatto negativo sull'inflazione.

Un aumento del PIL, quindi, ha anche il beneficio di ridurre l'inflazione. Dunque questa teoria predice che l'inflazione è anticiclica.

Morale: Sul lungo periodo, la crescita della moneta genera inflazione, ma è neutrale rispetto al PIL (reale).

Tuttavia nel breve periodo esistono situazioni controverse, in cui le variazioni di moneta non sembrano avere alcuna relazione con l'inflazione e per contro sembrano influenzare l'andamento del PIL (e del tasso d'interesse reale).

La teoria quantitativa della moneta può essere così sintetizzata:

1. La velocità di circolazione della moneta è relativamente stabile,
2. quindi, quando la BC cambia M provoca una variazione proporzionale del valore nominale del PIL ($P \cdot Y$)
3. La quantità di moneta non influenza Y
4. Variazioni di M influenzano solo P
5. Quindi aumenti rapidi di M inducono un tasso di inflazione elevato.