

Questo testo si occupa principalmente di tre questioni: come si spiega la crescita economica? Quali sono le cause dei cicli economici? A cosa è dovuta l'inflazione?

La crescita economica è la tendenza del PIL pro capite ad aumentare costantemente, in media, in tutti i paesi del mondo. La teoria della crescita cerca di spiegare perché ciò accade e perché le varie nazioni hanno tassi di crescita diversi. Piccole differenze tra i tassi di crescita possono determinare notevoli differenze di tenore di vita tra paesi diversi e periodi diversi.

Numerose variabili economiche tendono a fluttuare lentamente nel corso del tempo; la tendenza di una variabile economica all'inerzia prende il nome di *persistenza*. Diverse serie temporali tendono a variare contemporaneamente: questo movimento congiunto prende il nome di *coerenza*. La tendenza di diverse serie temporali al co-movimento e alla persistenza sta alla base dei *cicli economici*; essi sono importanti perché, nonostante il tenore di vita aumenti mediamente nel corso del tempo, il PIL pro capite può rimanere per anni al di sotto del livello di trend per effetto di una recessione.

La teoria dell'inflazione si propone di spiegare, per esempio, perché negli Stati Uniti i prezzi sono aumentati a un tasso medio annuo del 4,5% a partire dalla fine della seconda guerra mondiale. I periodi d'inflazione molto elevata, cioè di *iperinflazione*, sconvolgono la vita economica; ma anche un'inflazione di entità più modesta ha effetti negativi sull'economia, soprattutto se non era stata prevista da coloro che percepiscono un reddito fisso. Una volta che l'inflazione ha preso piede in un sistema economico è difficile da debellare, perché le politiche che riducono l'inflazione provocano anche un temporaneo aumento della disoccupazione.

Sebbene la crescita, i cicli economici e l'inflazione siano argomenti distinti, i fattori che stanno alla base dei tre fenomeni sono collegati tra loro. Per studiare questi argomenti gli economisti costruiscono modelli economici, che sono sistemi di equazioni.

L'economia mondiale è un insieme di economie nazionali, ciascuna delle quali può essere studiata come un sistema chiuso o aperto. L'economia nazionale può essere suddivisa in due settori: quello pubblico e quello privato, e nell'ambito di quest'ultimo si può distinguere il settore delle famiglie da quello delle imprese.

Il principale indicatore della capacità produttiva di un sistema economico è il PIL, il cui ammontare può essere calcolato in tre modi: con il metodo del reddito, con il metodo della spesa e con il metodo del prodotto (o del valore aggiunto). In un'economia chiusa il PIL è pari alla somma del consumo e dell'investimento; in un'economia aperta esso è pari alla somma del consumo, dell'investimento e delle esportazioni nette. Il PIL, il consumo, l'investimento e le esportazioni nette sono esempi di flussi, cioè di variabili misurate per un dato periodo di tempo. Nel 1993 il PIL pro capite degli Stati Uniti era circa pari a 25 000 dollari; l'80% di esso veniva destinato al consumo e il restante 20% all'investimento.

La ricchezza di un sistema economico è uno stock, cioè una variabile misurata in un determinato momento nel tempo. La ricchezza è costituita da attività reali e attività finanziarie: un'attività reale è un bene tangibile, come per esempio un terreno o una macchina; un'attività finanziaria è la promessa da parte di un altro soggetto economico di cedervi determinati beni in futuro. Nel 1993 la ricchezza nazionale degli Stati Uniti ammontava a circa 99 000 dollari per ogni abitante del paese.

Esistono due tipi di investimento: quello lordo e quello netto. L'investimento lordo è pari all'investimento netto (la spesa destinata ad accrescere lo stock di capitale) più gli ammortamenti (la spesa destinata a sostituire il capitale usurato). Esiste un legame tra stock di ricchezza e investimento netto, in quanto l'investimento netto va ad accrescere la ricchezza.

I dati economici vengono raccolti sotto forma di serie storiche. Per analizzare le serie storiche, si separa la loro *componente a bassa frequenza* (il trend) dalla *componente ad alta frequenza* (o componente *ciclica*). I tre principali metodi per rimuovere il trend da una serie temporale sono: l'interpolazione lineare, la detrendizzazione flessibile e la differenziazione delle serie.

La rimozione del trend ci permette di individuare relazioni tra le serie storiche che esistono a una certa frequenza, ma non a un'altra. L'intensità di una relazione statistica tra due serie storiche si misura mediante il coefficiente di coerenza. Una variabile che ha un alto coefficiente di coerenza col suo andamento passato è persistente; due variabili tra le quali esiste un alto coefficiente di coerenza sono coerenti. I co-movimenti possono essere positivi (quando due serie si muovono insieme nella stessa direzione) oppure negativi (quando due serie si muovono in direzioni opposte). Se una variabile è correlata positivamente con il PIL è detta prociclica, se invece è correlata negativamente con il PIL è detta anticiclica. All'origine dei cicli economici vi è la tendenza di molte serie storiche economiche a essere persistenti e correlate con il PIL.

L'occupazione e la disoccupazione sono due diverse misure del livello di attività economica. Il tasso di occupazione indica la percentuale della popolazione che svolge un'attività lavorativa remunerata, mentre il tasso di disoccupazione indica la percentuale della forza lavoro che sta cercando un impiego. Poiché negli Stati Uniti la forza lavoro è cresciuta dalla fine della seconda guerra mondiale a oggi, sono aumentati contemporaneamente sia il tasso di occupazione sia il tasso di disoccupazione. Il tasso di disoccupazione è anticiclico e fortemente correlato con il PIL.

L'inflazione è il tasso di variazione di un indice dei prezzi. I principali indici dei prezzi sono tre; noi utilizzeremo uno di questi, il deflatore del PIL, perché è il più completo ed è disponibile per una lunga serie di anni. Il deflatore del PIL è dato dal rapporto tra l'ammontare del PIL calcolato ai prezzi correnti e l'ammontare del PIL calcolato con i prezzi in vigore nell'anno base. Negli Stati Uniti l'inflazione è stata prociclica prima della seconda guerra mondiale, ma dalla fine della guerra ha avuto un andamento anticiclico.

Il concetto di equilibrio permette di spiegare come si determinano il livello di occupazione, la produzione e il salario reale. Per salario reale si intende il salario misurato in unità di beni prodotti. Le imprese massimizzano il loro profitto e richiedono più lavoro quando il salario reale diminuisce. La relazione tra salario reale e quantità di lavoro domandata prende il nome di curva di domanda di lavoro. Le famiglie massimizzano la loro utilità e offrono più lavoro quando il salario reale aumenta. La relazione tra salario reale e quantità di lavoro offerta prende il nome di curva di offerta di lavoro. Secondo gli economisti classici il livello di occupazione è determinato dal punto d'intersezione tra le curve di domanda e di offerta di lavoro.

Recentemente gli economisti della scuola del ciclo economico reale, hanno affermato che la teoria classica non dovrebbe essere accantonata troppo frettolosamente; a loro parere la Grande Depressione fu un evento anomalo mentre, in generale, il modello classico spiega in modo soddisfacente il funzionamento dell'economia.

Teoria quantitativa della moneta, teoria classica del livello dei prezzi e teoria classica della domanda aggregata sono denominazioni diverse di una stessa teoria. Tale teoria si basa sul presupposto che l'uso della moneta negli scambi comporta un costo e un beneficio. Il costo è rappresentato dal fatto che le risorse che si tengono sotto forma di denaro liquido potrebbero essere utilizzate per acquistare altri beni. Il beneficio è l'utilità che deriva dalla possibilità di effettuare vendite e acquisti in momenti diversi, grazie all'impiego della moneta. Secondo la teoria quantitativa della moneta, la domanda di moneta è proporzionale al PIL e la costante di proporzionalità k prende il nome di propensione a detenere moneta.

Per passare da una teoria della domanda e dell'offerta di moneta a una teoria della domanda aggregata, si parte dal presupposto che per l'intera comunità la quantità di moneta domandata deve sempre coincidere con la quantità offerta. Se il livello dei prezzi scende, il valore reale dello stock di moneta in circolazione aumenta e le singole famiglie si trovano ad avere un eccesso di denaro contante. Poiché le famiglie cercheranno di spendere l'eccesso di contanti, la domanda aggregata di beni crescerà.

La teoria classica della domanda aggregata si può unire alla teoria classica dell'offerta aggregata per spiegare il modo in cui si determinano il prodotto interno, il livello di occupazione, il salario reale, il salario nominale e il livello dei prezzi. Secondo la teoria classica, se la quantità di moneta in termini nominali raddoppia, tutte le variabili nominali raddoppieranno, mentre le variabili reali resteranno immutate; questo è il motivo per cui si parla di neutralità della moneta.

La teoria classica del livello dei prezzi risulta appropriata per i paesi in cui l'inflazione è molto elevata, come per esempio l'Argentina, il Brasile e Israele, mentre non va altrettanto bene per i paesi con un tasso d'inflazione ridotto, come gli Stati Uniti, il Giappone e il Regno Unito. Ciò è dovuto al fatto che in realtà la propensione a detenere moneta non è una costante. Nei paesi ad alta inflazione le grosse variazioni dell'offerta di moneta fanno passare inosservate le variazioni di k , ma nei paesi a bassa inflazione ciò non avviene. La lacuna principale della teoria classica è rappresentata dalla incapacità a spiegare perché durante la Grande Depressione i prezzi abbiano avuto un andamento prociclico, cioè perché siano scesi contemporaneamente con la caduta del PIL.

Possiamo utilizzare gli strumenti della teoria della domanda e dell'offerta per studiare la distribuzione delle risorse nel tempo. L'investimento è un modo per trasferire beni dal presente al futuro e le imprese decidono quanto investire con l'obiettivo di ottenere il massimo profitto possibile. Il risparmio consiste nel rinviare il consumo e le famiglie decidono quanto risparmiare con l'obiettivo di ottenere la massima utilità possibile. Il tasso d'interesse è il prezzo relativo di beni presenti e beni futuri; esso viene fissato nel mercato dei capitali in corrispondenza del punto in cui la quantità di investimenti domandata e la quantità di risparmio offerta coincidono.

Nel corso del ciclo economico l'investimento è soggetto a notevoli variazioni, mentre il consumo è relativamente stabile. I nuovi economisti classici della scuola del ciclo economico reale ritengono che le fluttuazioni dell'investimento siano dovute a variazioni della produttività conseguenti alle invenzioni di nuove tecnologie. Secondo gli economisti keynesiani molte variazioni dell'investimento sono dovute a fattori emotivi e mutamenti di opinione degli investitori, definiti da Keynes *animal spirits*. I rappresentanti di entrambe le scuole di pensiero ritengono che si possa utilizzare il modello della domanda e dell'offerta di capitali per spiegare il modo in cui si determina il tasso d'interesse.

Il modello della domanda e dell'offerta di capitali ci aiuta a comprendere gli effetti del boom demografico del dopoguerra sul tasso d'interesse e sull'investimento. Si può prevedere che il progressivo invecchiamento della generazione del *baby boom* determinerà un aumento del tasso d'interesse e un calo dell'investimento. Lo stesso modello può essere utilizzato per studiare gli effetti dei debiti contratti dagli Stati Uniti nel mercato mondiale dei capitali e l'interazione fra deficit pubblico e disavanzo commerciale.

L'approccio moderno alla disoccupazione individua un tasso naturale di disoccupazione risultante dai flussi in entrata e in uscita dalla forza lavoro. Gli economisti di scuola keynesiana sono convinti che il tasso effettivo di disoccupazione sia generalmente diverso dal tasso naturale perché, a seguito di uno shock delle variabili nominali, per esempio una riduzione dell'offerta di moneta, il salario nominale non è in grado di riportare l'economia al livello di equilibrio con sufficiente rapidità. La rigidità del salario nominale, in presenza di una caduta del livello dei prezzi, produce un rialzo temporaneo del salario reale e, quindi, una riduzione dell'occupazione.

Le opinioni degli economisti circa l'opportunità di un intervento statale finalizzato alla stabilizzazione del tasso di disoccupazione nel ciclo economico sono divergenti. Appare comunque plausibile che le disparità del livello occupazionale registrate nei vari paesi siano imputabili all'adozione di politiche economiche diverse. Il dibattito sulla natura e l'efficacia di queste politiche è ancora particolarmente vivace.

Le attività finanziarie presentano diversi gradi di liquidità. Le attività finanziarie più liquide sono quelle da cui si ricava maggiore utilità nelle transazioni. Affinché gli individui siano portati a detenere anche le attività meno liquide, il rendimento in termini di interesse sale al diminuire del grado di liquidità. Nel modello da noi costruito facciamo l'ipotesi che esistano solo due tipi di attività: i titoli, ossia le attività fruttifere, e la moneta, ossia l'insieme delle attività liquide infruttifere.

Per spiegare come mai le famiglie decidono di detenere, oltre ai titoli, anche una certa quantità di moneta, assumiamo che il valore reale della moneta fornisca utilità. Detenere moneta ha un costo legato alla rinuncia del rendimento ottenuto dai titoli. Questo rendimento potrebbe essere utilizzato per acquistare una maggiore quantità di beni e servizi. Le famiglie, tuttavia, fanno un raffronto tra i costi in termini di rinuncia al rendimento e l'utilità marginale derivante dal possesso di moneta in termini reali. La funzione di domanda di moneta che ne risulta generalizza la teoria quantitativa della moneta, mettendo in relazione la velocità di circolazione della moneta (ossia l'inverso della propensione a detenere moneta) e il tasso di interesse. L'evidenza empirica è a sostegno della teoria dell'utilità della moneta e smentisce la teoria quantitativa che considerava costante la velocità di circolazione. In base ai dati raccolti negli Stati Uniti nell'ultimo secolo, infatti, la velocità di circolazione della moneta mostra variazioni parallele a quelle subite dal tasso di interesse.

Affinché il mercato monetario sia in equilibrio, domanda e offerta di moneta devono essere uguali. L'innalzamento del reddito provoca un aumento della domanda di moneta. Le famiglie detengono infatti più liquidità per far fronte a un maggior numero di transazioni. L'aumento del tasso di interesse fa sì che la quantità di moneta domandata diminuisca, in quanto detenere moneta genera costi maggiori in termini di rinuncia di reddito derivante dal rendimento dei titoli. La rappresentazione grafica del reddito in funzione del tasso di interesse assume la forma di una curva inclinata positivamente. La curva, chiamata LM, descrive le combinazioni di reddito e tasso di interesse in corrispondenza delle quali il mercato monetario è in equilibrio.

All'origine del sistema bancario moderno vi è l'esigenza dei mercanti medievali di depositare il loro patrimonio in monete d'oro o d'argento presso gli orefici. Poiché non tutti i mercanti chiedevano di prelevare tali depositi nello stesso momento, gli orefici di quel tempo presero a concedere prestiti a interesse, creando così depositi che iniziarono a circolare sotto forma di note o atti di deposito. Il sistema bancario odierno rispecchia nei meccanismi fondamentali quello medievale con la differenza che divise metalliche e biglietti cartacei hanno sostituito l'oro come mezzo ultimo di pagamento.

La banca centrale controlla la quantità di moneta fissando il tasso di sconto e il coefficiente di riserva obbligatoria per le banche commerciali. Il principale strumento a sua disposizione è tuttavia costituito dalle operazioni di mercato aperto, ossia l'acquisto e la vendita di titoli del debito pubblico. Le operazioni di mercato aperto determinano la variazione delle passività in essere della banca centrale. Le passività dell'autorità monetaria sono chiamate base monetaria e sono detenute in parte dalle banche, sotto forma di depositi, e in parte dal pubblico, sotto forma di banconote in circolazione. L'offerta di moneta è composta dal circolante in mano al pubblico e dai depositi di famiglie e imprese presso le banche commerciali. La banca centrale non può intervenire direttamente sull'offerta di moneta ma ha la possibilità di controllare la base monetaria. Poiché l'offerta di moneta è un multiplo della base monetaria, la banca centrale è dunque in grado di controllare per via indiretta l'intera quantità di M1 presente nel sistema economico.

Il tasso d'interesse reale è pari al tasso d'interesse nominale meno il tasso d'inflazione atteso. Gli investimenti dipendono dal tasso d'interesse reale, mentre il risparmio è sensibile sia al tasso d'interesse reale sia al reddito. Il tasso d'interesse reale per cui il risparmio eguaglia gli investimenti varia a seconda del livello del reddito e la funzione delle combinazioni di reddito e tassi d'interesse in corrispondenza delle quali risparmio e investimento si eguagliano è chiamata curva IS. La curva IS ha pendenza negativa perché quando il tasso d'interesse reale si riduce le imprese sono più disponibili ad aumentare gli investimenti. Per riportare in equilibrio il mercato dei capitali occorre allora incrementare il livello del reddito. Le variabili in grado di spostare la curva IS sono l'inflazione attesa, la spesa pubblica, la spesa per trasferimenti e l'imposizione fiscale. Un aumento di una delle prime tre variabili o una riduzione della tassazione spostano la curva IS verso destra.

L'equilibrio IS-LM corrisponde al punto in cui il livello del reddito e il tasso d'interesse portano il mercato dei capitali in equilibrio e la quantità domandata di moneta eguaglia la quantità offerta. Poiché la posizione della curva LM dipende dal livello dei prezzi, vi sarà un diverso equilibrio IS-LM per ogni valore di P . Quando il livello dei prezzi si riduce la curva LM si sposta verso destra e il reddito di equilibrio risulta più elevato. La funzione definita dalle combinazioni di livelli dei prezzi e redditi di equilibrio costituisce la curva keynesiana di domanda aggregata. Le variabili che determinano la posizione di questa curva sono la spesa pubblica, la spesa per trasferimenti, l'imposizione fiscale, l'inflazione attesa e la quantità nominale di moneta.

Secondo la teoria keynesiana, le fluttuazioni del reddito e del livello dei prezzi sono imputabili alle variazioni dell'offerta e della domanda a livello aggregato. Se i cicli economici sono provocati dai movimenti della domanda aggregata, allora l'andamento del livello dei prezzi sarà prociclico, come in effetti è accaduto nel periodo tra le due guerre. Quando invece le fluttuazioni cicliche sono causate da shock dell'offerta, il livello dei prezzi mostrerà un andamento anticiclico, come si è potuto osservare nel secondo dopoguerra. Resta acceso tra gli economisti il dibattito riguardante l'opportunità dell'intervento statale volto a stabilizzare le fluttuazioni dell'attività economica.

Il tasso di cambio si definisce come numero di unità di moneta estera che possono essere acquistate con un'unità di moneta nazionale. Da questa definizione segue quella dei termini apprezzamento, deprezzamento, rivalutazione e svalutazione. Quando il tasso di cambio cresce si ha un deprezzamento della valuta estera; quando il tasso di cambio subisce un calo si ha un apprezzamento della valuta estera. In regime di cambi fissi, sulla base della definizione di tasso di cambio, un aumento e una riduzione del tasso di cambio vengono definiti rispettivamente svalutazione e rivalutazione.

Abbiamo visto anche il significato dei termini tasso di cambio reale e parità dei poteri d'acquisto assoluta e relativa. La parità assoluta dei poteri d'acquisto implica che il tasso di cambio reale sia identico per tutti i paesi. Si parla invece di parità relativa dei poteri d'acquisto se si ritiene che le variazioni del tasso di cambio reale avvengano in linea con quelle dei tassi di cambio reali degli altri paesi. Le verifiche empiriche non confermano tuttavia nessuna delle due ipotesi, dal momento che paesi diversi producono panieri di beni differenti, e che negli ultimi venticinque anni si sono registrate vistose variazioni relative dei prezzi. Abbiamo discusso, inoltre, la condizione di parità scoperta dei tassi d'interesse, secondo la quale il tasso d'interesse, una volta aggiustato in base alle variazioni attese del tasso di cambio, è identico per tutti i paesi.

I due regimi di cambio, fisso e fluttuante, influenzano l'economia mondiale in modi molto diversi. In regime di cambi fissi la principale modifica da apportare al modello IS-LM riguarda il tasso d'interesse che non è più controllato dalle banche centrali in quanto queste ultime sono impegnate a sostenere il tasso di cambio. Se il tasso di cambio fosse fissato a un livello inadeguato, si potrebbe ingenerare, in un sistema di cambi fissi, un conflitto tra le politiche monetarie nazionali e i target dei tassi di cambio. È questo il conflitto che nel 1973 ha portato al crollo del regime di cambi fissi introdotto con gli accordi di Bretton Woods del 1948. In regime di tassi flessibili ciascun paese può decidere indipendentemente la propria politica monetaria, in tal modo i tassi di interesse possono muoversi in maniera indipendente.

Da questo capitolo, nel quale abbiamo introdotto e discusso alcuni tra i principali concetti dell'economia internazionale, possiamo trarre un'importante lezione: per tutti i paesi tranne uno la politica macroeconomica nazionale è molto differente a seconda che il sistema monetario internazionale adotti un regime di cambi fissi o di cambi flessibili. Il principale vantaggio del regime di cambi fissi consiste nella diminuzione dei rischi sugli investimenti esteri e, di conseguenza, nella promozione degli scambi internazionali. Lo svantaggio principale è essenzialmente quello di impedire ai vari paesi di perseguire una propria politica monetaria e fiscale. Fino a quando l'ordine internazionale si reggerà sull'esistenza di Stati sovrani, riuscirà difficile coniugare queste due esigenze. In assenza di una forma di governo sovranazionale è dunque altamente probabile che il sistema monetario rimanga fondato sulla coesistenza di diverse valute.

Le equazioni alle differenze vengono utilizzate per descrivere l'evoluzione dinamica (nel tempo) di una variabile di stato. La soluzione di un'equazione alle differenze è data dalla serie di numeri che descrivono la variabile di stato nella successione di periodi temporali. L'equilibrio di stato stazionario è una soluzione speciale tra le soluzioni di un'equazione alle differenze e corrisponde a un valore della variabile di stato che rimane costante nel tempo. Gli stati stazionari possono essere stabili o instabili. In caso di stato stazionario stabile, la variabile di stato converge verso l'equilibrio a partire da un qualsiasi valore iniziale. In caso di stato stazionario instabile, la variabile di stato tende invece ad allontanarsi da esso.

La relazione tra il debito pubblico e il disavanzo di bilancio pubblico è espressa da un'equazione alle differenze la cui variabile di stato corrisponde al rapporto tra debito e PIL. Il comportamento dell'equazione dipende dal rapporto tra tasso d'interesse e tasso di crescita. Se il tasso d'interesse supera il tasso di crescita, lo stato stazionario è instabile. Viceversa, se il tasso d'interesse è inferiore al tasso di crescita, lo stato stazionario è stabile.

Fino al 1979 il rapporto debito/PIL degli Stati Uniti era descritto da una equazione alle differenze con equilibrio stabile. Dal 1979 lo stato stazionario è diventato instabile. Prima del 1979 i disavanzi di bilancio degli Stati Uniti erano di entità ridotta e non hanno mai provocato aumenti del debito poiché l'economia cresceva a un tasso superiore al tasso d'interesse. A partire dal 1979 il tasso d'interesse ha superato il tasso di crescita e le autorità di politica economica si sono viste costrette a ridurre il deficit per riportare il bilancio in pareggio. Lo stesso è accaduto in molti altri paesi, tanto da suggerire che il fenomeno possa avere una spiegazione comune. È probabile che tale spiegazione sia legata al rialzo del tasso d'interesse provocato negli USA dall'aumento del deficit. Il rialzo si sarebbe poi trasmesso al resto del mondo attraverso il mercato internazionale dei capitali.

Le teorie neoclassiche della produzione e della distribuzione possono essere utilizzate per misurare le fonti della crescita del PIL pro capite. Il processo di crescita non può essere spiegato solo in termini di accumulazione di capitale e di lavoro. Tra i fattori della crescita assume infatti un ruolo determinante l'efficienza del lavoro. L'accumulazione di capitale e lavoro rimane tuttavia un fattore non secondario, dato che si potrebbe accelerare il processo di crescita con nuovi investimenti in beni capitali. Per verificare se quest'ipotesi possa essere fondata, abbiamo elaborato un modello che include tutte le componenti della crescita e spiega in che modo esse sono in relazione tra loro.

Il modello neoclassico di crescita si basa su tre fatti stilizzati. Il primo è che negli Stati Uniti il PIL pro capite ha registrato nell'ultimo secolo una crescita media dell'1,64%. Il secondo riguarda invece la quota del PIL destinata alla remunerazione del lavoro, che negli ultimi cento anni è rimasta pressoché costante. Infine, anche la quota del PIL destinata al consumo è rimasta costante. Il secondo e il terzo di questi fatti costituiscono i presupposti del modello. La crescita del PIL è invece spiegata dalla logica del modello.

Per individuare le determinanti della crescita economica abbiamo fatto ricorso a un'equazione alle differenze la cui variabile di stato esprime il rapporto tra capitale e lavoro, quest'ultimo misurato in unità di efficienza. La variabile converge verso uno stato stazionario. Il modello prevede un aumento del reddito pro capite solo a condizione che si ammetta un aumento esogeno della produttività. In caso contrario, la crescita si arresta. Da un raffronto tra le due ipotesi emerge, dunque, che la determinante della crescita del PIL pro capite è costituita, in ultima analisi, dagli aumenti esogeni dell'efficienza del lavoro.

La crescita della produttività assume un ruolo centrale in quanto, secondo l'ipotesi di rendimenti di scala costanti, gli aumenti proporzionali del prodotto fanno seguito ad aumenti proporzionali dei fattori capitale e lavoro. Un sistema economico in cui il capitale venga accumulato mantenendo una quantità fissa di input lavoro mostrerà rendimenti marginali decrescenti del capitale; di conseguenza l'output riporterà incrementi proporzionalmente inferiori a quelli del capitale. Il rapporto tra investimento e prodotto rimane in tal caso costante, ma gli incrementi del prodotto da un periodo all'altro sono via via sempre minori. La crescita si arresta, infine, quando lo stock di capitale giunge al valore di stato stazionario. Il modello neoclassico aggira l'ostacolo costituito dalla quantità fissa di ore di lavoro pro capite mediante l'ipotesi che, se misurato in unità di efficienza, il lavoro aumenta in ragione degli incrementi esogeni della produttività.

Una maniera per estendere il modello neoclassico della crescita con gli scambi internazionali di capitali è quella di formulare l'ipotesi che i mercati siano completamente aperti (perfetta mobilità dei capitali) e che dunque gli operatori di un paese possano chiedere e concedere prestiti agli operatori di qualsiasi altro paese. Se questa fosse un'adeguata descrizione della realtà, il PIL pro capite dovrebbe tendere a essere lo stesso per tutti i paesi in conseguenza dei movimenti di capitali effettuati dagli investitori che vanno alla ricerca delle migliori opportunità di rendimento. Tuttavia l'evidenza empirica mostra che gli standard di vita non tendono a convergere.

Un secondo metodo per estendere il modello di crescita neoclassico consiste nell'ipotizzare che i mercati dei capitali siano, al contrario, completamente chiusi (mobilità zero dei capitali). Quest'ipotesi è coerente con il fatto osservato in tutti i paesi che il risparmio tende a eguagliare l'investimento. Tuttavia, in base a questa stessa ipotesi, il modello neoclassico di crescita prevede che i paesi con un elevato tasso di risparmio tendano a sperimentare livelli di PIL pro capite più alti, fenomeno che non è assolutamente confermato dall'evidenza empirica. Inoltre, mentre il modello neoclassico prevede un tasso di crescita del PIL pro capite uguale per tutti i paesi, nella realtà riscontriamo per lunghi periodi di tempo divergenze nei tassi di crescita pro capite che arrivano fino al 7%. Accertata l'incapacità del modello neoclassico di spiegare questi fatti stilizzati, abbiamo rivolto la nostra attenzione a un diverso modello di crescita, il modello *learning by doing*.

Nel modello *learning by doing*, che è stato comunque elaborato a partire dal modello neoclassico, gli aumenti di capitale comportano aumenti proporzionali del PIL. Per spiegare come mai il coefficiente del capitale nella funzione di produzione sia diverso dalla quota di reddito destinata al capitale stesso, il modello considera due diverse funzioni di produzione: una relativa all'intero sistema economico, l'altra riferita alla singola impresa. Nella funzione di produzione individuale il coefficiente del capitale è pari a $1/3$, mentre in relazione all'intera economia (cioè nella funzione di produzione aggregata) il coefficiente è uguale a 1. La diversità scaturisce dal fatto che le singole imprese non riescono a usufruire dei benefici dovuti agli effetti che la propria attività produce sull'acquisizione delle conoscenze a livello dell'intero sistema economico.

La teoria del *learning by doing* offre una spiegazione di carattere endogeno per la crescita, abbandonando dunque l'ipotesi di un aumento esogeno del progresso tecnico. La teoria è in grado di spiegare perché il PIL pro capite non è uguale per tutti i paesi e perché i paesi con livelli di risparmio più alti crescono a un tasso maggiore degli altri. In una versione meno estrema, infine, il modello *learning by doing* appare molto simile al modello di crescita esogena, ma prevede che la convergenza dei tassi di crescita avvenga molto più lentamente.

La teoria dinamica della domanda e dell'offerta aggregata si compone di tre elementi: la teoria dinamica della domanda aggregata, la teoria dinamica dell'offerta aggregata e la teoria dell'aggiustamento dei salari.

Il bisogno di una teoria dinamica dell'aggiustamento dei salari trae origine dal fatto che la teoria statica era in grado di spiegare le deviazioni della disoccupazione dal tasso naturale solo con un salario nominale fisso. Un'ipotesi che si rivela inadatta a spiegare la crescita dei salari nell'ultimo secolo a un tasso annuo del 2-3%. La nostra teoria dell'aggiustamento dei salari è corroborata dall'evidenza empirica relativa agli anni '50 e '60, un periodo nel quale l'inflazione attesa è rimasta pressoché costante. La teoria potrebbe anche essere applicata a dati più recenti, ma in tal caso occorre specificare con più attenzione le determinanti delle aspettative.

La teoria statica della domanda e dell'offerta aggregata risulta incompleta in quanto vi si ipotizza che tanto il salario nominale quanto l'inflazione attesa rimangano fissi. Nella teoria dinamica l'ipotesi che il salario nominale rimanga fisso viene abbandonata. Le modifiche da apportare alla teoria così da rendere endogene anche le aspettative verranno studiate nel capitolo 16.

Per spiegare in che modo il PIL e i prezzi vengono determinati congiuntamente, nel breve e nel lungo periodo, si può utilizzare un modello dinamico completo della domanda e dell'offerta aggregata. Secondo la teoria oggi più accreditata, nel breve termine le aspettative di inflazione sono fisse, mentre nel lungo termine esse vengono riviste. Nel breve periodo la politica monetaria si rivela efficace producendo effetti sull'economia reale, in quanto le aspettative d'inflazione sono già state inserite nei contratti salariali. Nel lungo periodo, invece, i contratti salariali vengono modificati in modo da riflettere i cambiamenti delle aspettative e gli aumenti dell'offerta di moneta finiscono con l'essere interamente assorbiti dall'incremento dei prezzi.

La curva di Phillips descrive la relazione stabile tra inflazione salariale e disoccupazione quando le aspettative sono fisse, ma si rivela inaffidabile quando le aspettative di inflazione vengono determinate endogenamente. La moderna teoria della domanda e dell'offerta aggregata contribuisce a spiegare perché la curva di Phillips apparisse stabile prima del 1970 e perché la relazione che essa descriveva è venuta meno nei due decenni successivi. Secondo questa teoria, la curva di Phillips si è spostata nel momento in cui le autorità di politica economica hanno tentato di trarre vantaggio dal legame tra inflazione e disoccupazione portando la crescita del PIL oltre il suo tasso naturale. Dal momento che l'aumento dei prezzi relativo a questo comportamento veniva anticipato dagli agenti economici, ben presto si è riflesso nelle aspettative spostando verso l'alto la curva di Phillips. Le autorità di politica economica sono oggi pienamente consapevoli di quanto sia importante impedire che si generino attese di inflazione. Tale consapevolezza suggerisce loro di essere assai prudenti nell'impiegare la politica monetaria per controbilanciare gli effetti che gli shock dell'offerta producono sulla disoccupazione e sulla crescita.

L'approccio moderno alla disoccupazione individua un tasso naturale di disoccupazione risultante dai flussi in entrata e in uscita dalla forza lavoro. Gli economisti di scuola keynesiana sono convinti che il tasso effettivo di disoccupazione sia generalmente diverso dal tasso naturale perché, a seguito di uno shock delle variabili nominali, per esempio una riduzione dell'offerta di moneta, il salario nominale non è in grado di riportare l'economia al livello di equilibrio con sufficiente rapidità. La rigidità del salario nominale, in presenza di una caduta del livello dei prezzi, produce un rialzo temporaneo del salario reale e, quindi, una riduzione dell'occupazione.

Le opinioni degli economisti circa l'opportunità di un intervento statale finalizzato alla stabilizzazione del tasso di disoccupazione nel ciclo economico sono divergenti. Appare comunque plausibile che le disparità del livello occupazionale registrate nei vari paesi siano imputabili all'adozione di politiche economiche diverse. Il dibattito sulla natura e l'efficacia di queste politiche è ancora particolarmente vivace.