



upB

ufficio
parlamentare
di bilancio

NOTA DI LAVORO

Quando il *gap* si fa incerto: stime alternative del potenziale e dell' *output gap* nell'economia italiana

di Cecilia Frale e Sergio De Nardis

La collana Note di lavoro comprende approfondimenti tecnici e istituzionali elaborati dall'UPB, eventualmente in collaborazione con esperti esterni. Le Note di lavoro intendono approfondire specifici profili rilevanti per i contributi istituzionali dell'UPB.

I contenuti delle Note riflettono le opinioni degli autori e non necessariamente quelle dell'UPB o degli Istituti di appartenenza degli esperti esterni.

La versione elettronica può essere scaricata da: www.upbilancio.it

Ufficio parlamentare di bilancio
Via del Seminario, 76
00186 Roma
segreteria@upbilancio.it

Quando il *gap* si fa incerto: stime alternative del potenziale e dell'*output gap* nell'economia italiana

di Cecilia Frale e Sergio De Nardis¹

Abstract

In questo lavoro si presenta una stima dell'*output gap* basata su di un filtro multivariato che estrae il segnale da indicatori di *output* (PIL, tasso di disoccupazione e capacità produttiva) e da una curva di Phillips. In particolare, si propone una serie di modelli a componenti inosservate che differiscono per il *set* informativo e la specificazione della componente di *trend*, che presenta un diverso grado di accostamento ai dati osservati. I risultati empirici mostrano che, al modificarsi della specificazione del modello, la stima dell'*output gap* e del potenziale variano in un intervallo molto ampio. I vari modelli, valutati secondo una pluralità di misure (*fit* statistico, grado di revisione delle stime e capacità previsiva dell'inflazione), risultano ottimali solo in relazione ad una specifica misura o un intervallo temporale e nessuno di essi è il migliore in termini assoluti. Inoltre, modelli che prima della recente crisi economica sembravano ottimali, appaiono aver perso capacità esplicativa negli ultimi anni. Queste evidenze concorrono a sottolineare, come implicazione di *policy*, la necessità di basare le scelte di politica economica su una pluralità di modelli o comunque di definire un modello sufficientemente flessibile da incorporare diverse ipotesi economiche (come cicli anomali o isteresi) e di affiancare le stime ottenute con opportune misure di errore.

Classificazione JEL: C32, C53, E32.

Keywords: Modelli a componenti inosservate; *output gap*; curva di Phillips; previsioni dell'inflazione.

¹ Cecilia Frale, UPB, cecilia.frale@upbilancio.it; Sergio De Nardis, UPB, sergio.denardis@upbilancio.it. Gli autori ringraziano Davide Delle Monache e Roberto Golinelli per i suggerimenti che hanno contribuito a migliorare la qualità di questo lavoro. Si ringraziano inoltre Michele Lenza, Tommaso Proietti, Carlo Favero, Marco Cacciotti e Federico Giammusso per le interessanti discussioni sul tema della stima dell'*output gap*.

Introduzione

I concetti di prodotto potenziale e *output gap* (OG, vale a dire la deviazione percentuale del prodotto effettivo rispetto al potenziale) hanno acquisito negli ultimi anni forte rilevanza nel contesto delle regole UE. Dalla loro definizione dipendono, infatti, la valutazione della posizione ciclica dell'economia e le implicazioni che ne derivano per il processo di sorveglianza fiscale nell'ambito del Patto di stabilità e crescita. Tuttavia, tali variabili non sono osservabili e devono essere oggetto di stime caratterizzate da un elevato livello di incertezza, differenziandosi in modo significativo a seconda del modello utilizzato. Ad esempio, la metodologia adottata dalla Commissione europea, concordata con i paesi membri, fornisce misure del potenziale e dell'OG che differiscono, soprattutto negli ultimi anni, in modo sensibile rispetto a quelle proposte dalle altre organizzazioni internazionali (OCSE e FMI).

La metodologia adottata dalla Commissione europea si basa sul metodo della funzione di produzione, per cui il PIL potenziale deriva dalla combinazione dei fattori della produzione (lavoro e capitale) e della *Total factor productivity*, espressi al loro livello di *trend*. Le componenti di *trend* dei fattori della produzione sono ottenute tramite l'applicazione di filtri statistici (univariati del tipo Hodrick e Prescott o bivariati del tipo Kalman). In particolare, per la stima del NAWRU (Non-Accelerating Wage Rate of Unemployment) che entra nell'equazione del lavoro potenziale, si applica un filtro di Kalman su un modello bivariato che include la serie del tasso di disoccupazione e la curva di Phillips. Il *trend* della TFP viene invece estratto con un filtro di Kalman bayesiano su un modello bivariato che include il cosiddetto residuo di Solow e un indicatore di capacità produttiva utilizzata.²

Anche l'OCSE segue l'approccio della funzione di produzione, ma con alcune differenze rispetto al metodo utilizzato dalla Commissione europea. Ad esempio i due metodi si discostano in modo considerevole per il concetto di NAIRU che per l'OCSE corrisponde a quel tasso di disoccupazione di equilibrio coerente con un'inflazione stabile e uguale all'obiettivo ufficiale delle autorità monetarie ed include pertanto le aspettative di inflazione³.

Infine, il FMI utilizza metodi diversi per la stima dell'*output gap* in base al paese considerato, facendo riferimento sia a filtri multivariati (considerando ad esempio PIL, tasso di disoccupazione e prezzi al consumo) che alla funzione di produzione⁴.

A conferma dell'incertezza che circonda queste valutazioni, si riscontra un'elevata variabilità non solo tra i diversi approcci, ma anche nell'ambito di una stessa

² Per maggiori dettagli si veda: Karel Havik et al. (2014), "The Production Function Methodology for Calculating Potential Growth Rates & Output Gaps", European Economy, Economic Papers n. 535.

³ Per maggiori dettagli si veda: Ollivaud P., Turner D., (2014) "The Effect of the Global Financial Crisis on OECD Potential Output" OECD Economics Department Working Papers No. 1166.

⁴ Per maggiori dettagli si veda: Alichì A., Bizimana O., Domit S., Fernandez Corugedo E., Laxton D., Tanyeri K., Wang H., e Fan Zhang F. "Multivariate Filter Estimation of Potential Output for the Euro Area and the United States", IMF working paper, WP/15/253.

metodologia tra i vari *vintage* di stime, essendo quest'ultime riviste nel corso del tempo talvolta in modo sostanziale (è, in particolare, il caso delle stime OCSE e CE, (tab.1)).

La recente crisi economica sembra, avere reso ancora più complessa la determinazione di tali variabili. Il prolungato deterioramento ciclico può, infatti, avere avuto effetti di isteresi, danneggiando le capacità produttive potenziali delle economie e accrescendo, di conseguenza, la difficoltà di distinguere tra componente temporanea (ciclo) e strutturale (*trend*) della crescita del PIL⁵. D'altro canto, si deve anche considerare la possibilità che i metodi di estrazione dell'*output gap* elaborati sulla base degli andamenti economici del passato, caratterizzati da frequenze cicliche *standard*, potrebbero risultare non adatti nelle condizioni congiunturali anomale (indotte da *shock* ciclici ripetuti o comunque più persistenti dello *standard* storico) che hanno caratterizzato gli ultimi anni. Se si vuole, quindi, tenere conto della maggiore ampiezza del ciclo intorno al *trend* che avrebbe contrassegnato l'ultima crisi ci si dovrebbe basare su una stima del prodotto potenziale la cui dinamica è più stabile nel tempo e meno prociclica⁶. In effetti, ricercatori della BCE hanno recentemente mostrato, con riferimento all'area dell'euro, che utilizzando forme funzionali per il *trend* diverse da quelle comunemente adottate in passato, compatibili con una componente ciclica più ampia rispetto alle precedenti esperienze, si ottengono stime del potenziale dell'area euro che si discostano in modo considerevole dai valori proposti dalle istituzioni internazionali e che hanno migliori capacità predittive dell'inflazione negli anni della crisi⁷.

Sempre in tema di accresciute difficoltà di identificazione delle componenti di ciclo/*trend*, occorre anche considerare che le variabili delle inchieste congiunturali basate sulle valutazioni degli operatori, impiegate ad esempio nel metodo della Commissione per depurare dal ciclo la dinamica della produttività totale dei fattori, possono essere diventate, negli ultimi anni, meno informative, avendo la recente crisi modificato giudizi e aspettative degli agenti economici circa i livelli ritenuti normali dell'attività produttiva⁸.

⁵ Sugli effetti di isteresi (ciclo negativo che riduce il livello del PIL potenziale) e super-isteresi (ciclo negativo che riduce la crescita del PIL potenziale) si vedano Blanchard O., Cerutti E. e Summers L. (2015), "Inflation and Activity: Two Explanations and their Monetary Policy Implications", MF Working Paper, n. 15, novembre, e Fatàs A. e Summers L. (2015), "The Permanent Effect of Fiscal Consolidations", Cepr Discussion Paper Series n. 10902, ottobre. Per l'Italia, fenomeni di isteresi nella serie del PIL vengono rilevati da Proietti, T. (2002), "Some Reflections on Trend-Cycle Decompositions with Correlated Components", EUI working paper ECO n. 2002/23. Stime successive sull'Italia, che adottano la stessa metodologia del lavoro precedentemente citato, hanno messo in evidenza l'intensificazione del fenomeno nel corso dell'ultima crisi, si veda in proposito, Ufficio parlamentare di bilancio (2016) "Rapporto sulla programmazione di bilancio 2016", aprile, pp. 20-31.

⁶ Si veda su questo Ledvai J., Salto M. e Thum-Thysen A. (2015) "Structural unemployment vs. NAWRU: Implications for the assessment of the cyclical position and the fiscal stance", European Commission, Economic Papers 552/June. Forti dubbi sull'attendibilità dell'evoluzione pro-ciclica delle stime del potenziale, elaborate da diverse istituzioni per i paesi avanzati, sono avanzati da Coibion O., Gorodnichenko Y. e Ulate M. (2017), "The cyclical sensitivity in estimates of potential output", Nber Working Paper Series, n. 23.580, July.

⁷ Jarocinski M. e Lenza M. (2016), "An inflation-predicting measure of the output gap in the euro area", *ECB working paper series* n. 1966, settembre.

⁸ Si vedano in proposito: European Commission (2016) "New normal? The impact of the financial crisis on business and consumer survey data" in European Business Cycle Indicators-3rd quarter 2016, Technical

Tab. 1 – Stime dell'*output gap* delle varie istituzioni internazionali in *vintages* successivi

	OCSE			FMI			CE		
	Giu 17	Nov 16	Giu 16	Apr 17	Ott 16	Apr 16	Mag 17	Nov 16	Mag 16
2002	1,2	1,1	1,0	-0,4	-0,4	-0,4	1,1	1,1	0,5
2003	0,3	0,2	0,1	-1,1	-1,1	-1,1	0,2	0,1	-0,4
2004	0,5	0,5	0,4	-0,3	-0,3	-0,3	0,5	0,5	0,4
2005	0,8	0,7	0,7	0,0	0,0	0,0	0,6	0,6	0,7
2006	2,0	2,0	2,0	1,5	1,5	1,5	1,8	1,8	2,4
2007	2,7	2,7	2,8	2,7	2,7	2,7	2,4	2,4	3,9
2008	1,2	1,2	1,3	1,8	1,8	1,8	1,2	1,1	3,0
2009	-4,4	-4,5	-4,2	-3,2	-3,2	-3,2	-4,0	-4,0	-2,3
2010	-2,7	-2,9	-2,4	-1,3	-1,3	-1,3	-2,0	-2,0	-0,2
2011	-1,9	-2,3	-1,5	-0,5	-0,5	-0,5	-1,3	-1,4	0,9
2012	-4,4	-4,9	-3,9	-2,8	-2,8	-2,8	-2,9	-2,9	-1,4
2013	-5,8	-6,3	-5,0	-4,1	-4,1	-4,1	-4,2	-4,1	-2,8
2014	-5,3	-5,9	-4,8	-4,1	-4,1	-4,1	-3,8	-3,7	-2,8
2015	-4,5	-5,1	-3,7	-3,3	-3,3	-3,3	-2,8	-2,6	-1,9
2016	-3,4	-4,2	-2,4	-2,4	-2,5	-2,5	-1,7	-1,6	-0,8

Per cercare di tenere conto in un unico *framework* di questi elementi di incertezza nel caso italiano, nel presente lavoro si procede a stimare un modello a componenti inosservate della tipologia proposta dagli economisti della BCE⁹, in cui l'*output gap* è rappresentato da una componente ciclica che è comune a un *set* di variabili di attività economica (PIL, tasso di disoccupazione e capacità produttiva). Specifici *trend* identificano, invece, la componente strutturale di ognuna di tali variabili. Il modello viene completato da una curva di Phillips in base alla quale l'inflazione evolve in funzione dell'*output gap*, di una componente di *trend* della dinamica dei prezzi al consumo e del prezzo del petrolio.

Esiste un'ampia letteratura che stima l'*output gap* tramite modelli a componenti inosservate, che inizia con modelli bivariati di PIL e disoccupazione o PIL e inflazione, cui si aggiungono valutazioni degli effetti di politica monetaria includendo il tasso di interesse reale nell'equazione dell'*output gap*. Ulteriori sviluppi hanno riguardato la costruzione di modelli trivariati (PIL disoccupazione e inflazione o, alternativamente, utilizzo della capacità produttiva). Più recentemente, è stata proposta una versione

paper 011, October; Bruno G., Crosilla L., e Margani P. "Inspecting the relationship between business confidence and industrial production: evidence based on Italian survey data", paper presented at the 2016 CIRET conference, https://www.ciret.org/media/ciret_papers/copen25; Malgarini M., (2011) "Industrial production and Confidence after the crisis: what's going on?", CIRET/KOF/HSE Workshop on National Business Cycles in the Global World, Moscow, September 16-17; Fantacone S., Garalova-Stieg P. e Malgarini M., "Is Business confidence still a good indicator for industrial production? Evidence from the EC survey", paper presented at the Confindustria CSC seminar on 12 May, available at: http://www.confindustria.it/wps/wcm/connect/www.confindustria.it/5266/d6d2c5f8-dff3-48d6-9d0a-7af4326088d4/Fantacone-Garalova-Malgarini.pdf?MOD=AJPERES&CONVERT_TO=url&CACHEID=d6d2c5f8-dff3-48d6-9d0a-7af4326088d4.

⁹ Si tratta del lavoro citato di Jarocinski M. e Lenza M. (2016).

bayesiana del modello bivariato di PIL e disoccupazione¹⁰. L'approccio delle componenti inosservate ha alcuni vantaggi rispetto al più complesso metodo di stima basato sulla funzione di produzione utilizzato dalla Commissione europea e dall'OCSE¹¹. Innanzitutto la stima dell'*output gap* che ne deriva è compatibile con il concetto di potenziale di Okun¹², ovvero quel livello massimo di *output* che una economia può sostenere senza generare spinte inflazionistiche. Inoltre, è relativamente semplice da interpretare poiché basato su un numero limitato di variabili. Infine, dato che la stima non può discostarsi troppo dai dati osservati può essere particolarmente adatto a catturare situazioni di *shock* che producono effetti permanenti sull'economia con una riduzione del potenziale, come in coincidenza di severe cadute dell'*output*. La principale limitazione è quella di essere un filtro statistico e come tale di essere, da un lato, affetto dal problema dell'*end-point-bias*¹³ e, dall'altro, di non fornire una immediata interpretazione economica.

In questo lavoro per tenere conto dei fattori di incertezza ricordati in precedenza (scomposizione ciclo/*trend* e capacità informativa delle variabili cicliche delle inchieste) si considerano vari modelli di stima dell'*output gap*. Questi si differenziano, in primo luogo, per diverse tipologie del *trend* il cui andamento riflette il grado di influenza esercitato dal PIL effettivo. In particolare, si adottano due specificazioni in cui la dinamica del *trend* si abbassa negli ultimi anni col peggiorare del PIL effettivo, implicando effetti di isteresi relativamente maggiori, e una specificazione in cui l'andamento del *trend* risulta, invece, più stabile nel tempo, risentendo in minor misura di tali effetti indotti dal ciclo. In secondo luogo, per ciascuna modalità di specificazione del *trend*, si stimano diverse versioni del modello a seconda che includano o meno la variabile, desunta dalle inchieste congiunturali, del grado di utilizzo della capacità produttiva.

¹⁰ Per i modelli bivariati, si veda Kuttner K.N. (1994) "Estimating potential output as latent variable", *Journal of Business and Economic Statistics* 12: 361-368. Per gli effetti della politica monetaria in tali modelli si veda Gerlach S., e Smets F. (1999), "Output gaps and monetary policy in the EMU area", *European Economic review* 43: 801-812. Per lo sviluppo dei modelli trivariati si veda Apel M., e Jansson P. (1999). "System estimates of potential output and the NAWRU", *Empirical Economics* 24: 373-388. La stima bayesiana è sviluppata da Planas, C., Rossi, A., e Fiorentini, G. (2008), "Bayesian analysis of the output gap", *Journal of Business and Economic Statistics*, 26(1): 18-32. A quest'ultimo approccio si ispirano nel loro lavoro Jarocinski M. e Lenza M. (2016). Per l'uso della variabile capacità produttiva nella stima dell'*output gap* si veda Graff M., e Sturm J-E, (2010) "The Information Content of Capacity Utilisation Rates for Output Gap Estimates", CESIFO WORKING PAPER NO. 3276, dicembre.

¹¹ Per una discussione dei metodi e delle problematiche di stima di OCSE e Commissione europea si veda Fioramanti M., Padrini F. e Pollastri C. "La stima del PIL potenziale e dell'*output gap*: analisi di alcune criticità", Nota di lavoro UPB, 1/2015. Per una rassegna delle diverse metodologie, con applicazioni al caso dell'Italia, si veda Bassanetti A., Caivano M. e Locarno A., (2010) "Modelling Italian potential output and the output gap", Banca d'Italia Working Paper, N. 771.

¹² Okun (1962), "Potential GNP: Its Measurement and Significance," in *Proceedings of the Business and Economic Statistics Section*, pp. 98-104 (Washington: American Statistical Association).

¹³ È la situazione per cui i filtri statistici sono particolarmente influenzati dalle osservazioni alla fine del campione e pertanto soggetti a revisione quando nuove osservazioni risultano disponibili. Tuttavia, anche i metodi basati sulla funzione della produzione, finiscono spesso per impiegare filtri statistici nel trattamento delle variabili, e pertanto non sono esenti da questo problema.

In estrema sintesi i risultati dell'analisi svolta sono riassumibili nei seguenti punti.

- Forme alternative del modello di stima, nella specificazione del *trend* e nel *set* informativo, producono un ampio intervallo di valori per l'*output gap*, confermando il largo margine di incertezza che caratterizza questa variabile.
- L'incertezza non viene sostanzialmente superata attraverso l'adozione di un criterio di valutazione della bontà dei vari modelli. Criteri di natura meramente statistica (massima verosimiglianza o stabilità dei parametri) conducono, infatti, a scelte di modelli diverse rispetto a un criterio più di carattere economico basato sulla bontà della capacità previsiva dell'inflazione.
- Anche la scelta basata su un particolare criterio può, però, non essere stabile nel tempo. Modelli che prima della recente crisi economica sembravano ottimali, appaiono avere perso capacità esplicativa negli ultimi anni. Nel periodo 2012-16 il modello che prevale per bontà previsiva è quello caratterizzato da un *trend* meno pro-ciclico e, quindi, da un *output gap* più ampio. Tale superiorità non è, però, verificata se si considera la capacità previsiva su un più ampio periodo di tempo. Ciò sembra costituire un segnale delle differenti caratteristiche della fase ciclica degli anni recenti rispetto all'evoluzione *standard* del passato.
- L'evidenza circa la capacità predittiva dei vari modelli riguardo all'inflazione pone in luce anche un deterioramento, nel periodo 2012-16, del segnale proveniente dal grado di utilizzo della capacità produttiva. Anche questa sembra un'indicazione del comportamento anomalo dell'ultimo ciclo economico.

Dato l'elevato grado di incertezza, l'implicazione di *policy* di queste evidenze è di basare le scelte di politica economica non su un'unica stima di *trend* e *output gap*, ma su una pluralità di modelli o comunque di definire un modello sufficientemente flessibile da incorporare diverse ipotesi economiche (come cicli anomali o isteresi) e di affiancare le stime ottenute con opportune misure di errore.

Il lavoro è organizzato come segue. Dopo la descrizione del modello, riportata nel paragrafo 1, si presentano i risultati della stima per l'Italia. Il terzo paragrafo è dedicato alla valutazione della bontà dei modelli alternativi e alcune brevi riflessioni concludono il lavoro.

1. Il modello econometrico utilizzato

Il modello utilizzato per la stima del prodotto potenziale e dell'*output gap* è nella classe dei cosiddetti modelli a componenti inosservate e prevede la scomposizione di alcune variabili di *output* (PIL, tasso di disoccupazione (UR) e capacità produttiva (CP)) nella componente ciclica e in quella strutturale. Il modello è poi completato da una equazione che lega la dinamica dei prezzi all'*output gap* e a fattori esogeni di spinta dal lato dei costi (curva di Phillips).

In termini formali il modello è il seguente:

$$y_t^n = \mu_t^n + b^n(L)\varphi_t + \varepsilon_t^n \quad (\text{output}) \quad n=1,..N \quad \varepsilon_t^n \sim NID(0, \sigma_{\varepsilon^n}^2) \quad (1)$$

$$\pi_t = \mu_t^\pi + a(L)\varphi_t + \gamma(L)X_t + \varepsilon_t^\pi \quad (\text{curva di Phillips}) \quad \varepsilon_t^\pi \sim NID(0, \sigma_{\varepsilon^\pi}^2)$$

in cui

y_t^n sono le N variabili di *output* (PIL, tasso di disoccupazione, capacità produttiva);

$b^n(L)$ e $a(L)$ sono dei polinomi nell'indicatore Lag; ε_t^n e ε_t^π sono gli errori supposti Gaussiani e incorrelati.

π_t è il tasso di inflazione *core* (al netto di prodotti energetici ed alimentari freschi), le cui oscillazioni rispetto al trend (μ_t^π) dipendono dalla dinamica dell'*output gap* e dal prezzo del petrolio (X_t), considerata come variabile *cost-push*, con effetti contemporanei e ritardati espressi tramite i coefficienti $\gamma(L)$ nell'operatore *lag*.

Le N variabili di *output* (y_t^n) derivano quindi dalla somma di un ciclo comune (φ_t) definito da un processo autoregressivo del secondo ordine:

$$\varphi_t = \theta_1\varphi_{t-1} + \theta_2\varphi_{t-2} + k_t \quad k_t \sim NID(0, \sigma_k^2)$$

e di specifici *trend* (μ_t^n) tali che:

$$\mu_t^n = \mu_{t-1}^n + \beta_t^n + \eta_t^n \quad \omega_t \sim NID(0, \sigma_{\xi^n}^2)$$

i cui errori k_t e η_t^n sono supposti Gaussiani e incorrelati.

In particolare per il PIL si introducono le seguenti restrizioni: $b^1(L) = 1$ e $\sigma_{\varepsilon^1}^2 = 0$ e l'equazione (1) diventa:

$$y_t^1 = \mu_t^1 + \varphi_t \quad (1a)$$

cosicché φ_t corrisponde allo scostamento dell'*output* dal suo *trend*, venendo così a rappresentare l'*output gap*.

Per tenere conto della possibilità di *trend* più o meno flessibili, si considerano le seguenti specificazioni alternative di μ_t che permettono un accostamento più o meno elevato ai dati osservati:

1. *Local level model* (LLT): con livello $\mu_t^1 = \mu_{t-1}^1 + \beta_t^1 + \eta_t$ e *drift* (che fornisce la pendenza del *trend*) $\beta_t^1 = \beta_{t-1}^1 + \xi_t$ variabili nel tempo ed errori incorrelati $\eta_t \sim NID(0, \sigma_\eta^2)$ e $\xi_t \sim NID(0, \sigma_\xi^2)$.

2. *Random walk plus drift* (RW): ottenuto imponendo, nella equazione del modello LLT, $\sigma_{\xi}^2 = 0$; ovvero con *drift* β costante nel tempo.
3. *Integrated random walk* (IRW): ottenuto imponendo, nella equazione del modello LLT $\sigma_{\eta}^2 = 0$; ovvero con *drift* β_t variabile nel tempo.

Come detto, le diverse specificazioni del *trend* rappresentano situazioni di maggiore/minore accostamento ai dati. In particolare, il modello IRW, produce un *trend* smussato e liscio influenzato principalmente dai punti di massimo e minimo della serie, mentre il *trend* RW si accosta di più ai dati osservati nei vari sotto-periodi.

I *trend* adottati per le altre variabili di *output* e dell'inflazione seguono invece un processo RW.

Infine, alla determinazione dell'inflazione concorre, come variabile *cost-push*, il prezzo del petrolio (X_t) con effetti contemporanei e ritardati espressi tramite i coefficienti $\gamma(L)$ nell'operatore *lag*.

2. I risultati empirici

Il modello (1) è stato applicato ai dati italiani del periodo 1985Q1 -2016Q4¹⁴. La serie del PIL è quella della contabilità trimestrale Istat destagionalizzata a prezzi costanti. L'inflazione stimata è quella cosiddetta *core*, ovvero dei prezzi al consumo senza alimentari freschi ed energetici, di fonte Istat¹⁵. La serie della capacità produttiva è il grado di utilizzo degli impianti come misurato dalle inchieste Istat. Il prezzo del petrolio è misurato come quotazione *crude spot* di fonte World Bank, espressa in euro.

Il modello, scritto nella forma Spazio-Stato, è stato stimato col metodo della Massima Verosimiglianza tramite il filtro di Kalman e l'algoritmo di *smoothing* ad esso associato¹⁶. I risultati, riportati nella tabella 2, mostrano che i parametri dell'*output gap*, θ_1 e θ_2 , risultano significativi in tutte le specificazioni e relativamente stabili, eccezion fatta per IRW che differisce maggiormente dagli altri modelli. Maggiore variabilità si ha invece per i parametri associati all'*output gap* nelle altre due equazioni. L'equazione dell'inflazione

¹⁴ I modelli che includono la serie della capacità produttiva utilizzata sono stimati a partire dal 1986Q1 essendo tale variabile disponibile da questa data.

¹⁵ La serie del PIL precedente al 1996 è costruita utilizzando la contabilità nazionale trimestrale NACE Rev. 1 ribassata sui prezzi 2010 e collegata alla contabilità nazionale Nace rev. 2. L'indice dei prezzi al consumo al netto degli energetici prima del 1996 è stato calcolato utilizzando i dati OECD Main economic indicator. Le serie entrano nel modello tramite delle trasformazioni: la serie del PIL entra nel modello in logaritmi, quella del CPI in differenze logaritmiche, il tasso di disoccupazione secondo una trasformazione logit e la capacità produttiva in livelli. Per quest'ultima variabile test di radici unitarie di tipo Dickey-Fuller sono rifiutati al 95 per cento di confidenza.

¹⁶ Per una completa trattazione della metodologia di stima si veda Harvey 1989, "Forecasting structural time series and the Kalman filter", Cambridge University press, Cambridge. La stima effettuata utilizza la libreria SsfPack 2.2 fornita da Koopman SJ, Shepard N, e Doornik JA (1999), "Statistical algorithms for models in state space using SsfPack22", Econometric Journal 2 113-166 del software Ox Metrics 6.0.

include anche il prezzo del petrolio contemporaneo e ritardato, i cui coefficienti appaiono significativi nei modelli che non includono la capacità produttiva.

I risultati in termini di componenti stimate, riportati nella tabella 3 e nei grafici seguenti, mostrano che, al modificarsi della specificazione del modello, la stima dell'OG e del potenziale variano in un intervallo molto ampio¹⁷. Le maggiori differenze si hanno nelle fasi di espansione e, soprattutto, nel periodo della recente crisi alla fine del campione. In particolare i modelli più flessibili, IRW e LLT, producono nell'ultimo periodo una stima del prodotto potenziale più bassa, e quindi un OG che si chiude più rapidamente con la ripresa dell'*output*, rispetto al modello RW che ipotizza una crescita potenziale più stabile nel periodo della crisi (fig 1). Tutti i modelli identificano la doppia caduta dell'*output gap* della recente crisi economica, che appare più consistente nel 2012 che nel 2008, ma in un *range* stimato molto ampio che ad esempio nel 2016 va da -0,5 (IRW) a -5,3 per cento (RW-CP). Anche per gli anni di previsione¹⁸ i diversi modelli ipotizzano un restringimento dell'OG negativo più o meno rapido; nel modello IRW il PIL torna sostanzialmente al suo livello potenziale già nel 2017¹⁹. Tali risultati riflettono fondamentalmente le diverse ipotesi circa la pendenza del *trend* nei vari processi di *random walk* adottati. Nel modello RW, l'assunzione di costanza del *drift* rende la dinamica di questo processo (ovvero la pendenza del *trend*) relativamente stabile nel tempo (gli *shock* in tale modello sono solo sul livello del *trend*) e, quindi, meno prociclica ed esposta ai fenomeni di isteresi (o più precisamente di super-isteresi, intesa come abbassamento persistente, in fasi di recessione, del tasso di crescita dell'*output* potenziale). L'opposto si verifica nei modelli LLT e IRW nei quali la variabilità temporale del *drift* tende a indurre modifiche nella pendenza del *trend* (cioè del tasso di crescita del potenziale) che possono risultare anche persistenti in fasi recessive prolungate (ciò si verifica in particolar modo nel caso del modello IRW, nel quale gli *shock* interessano esclusivamente la pendenza del *trend*).

¹⁷ Per confronto si è stimato anche un modello LLT senza il prezzo del petrolio che tuttavia non presenta differenze di rilievo rispetto al modello base e pertanto è stato ommesso per brevità.

¹⁸ I valori previsti delle variabili sono ottenuti sulla base dei parametri stimati applicando il filtro di Kalman per gli anni futuri.

¹⁹ Per avere dei termini di confronto di questi risultati con le valutazioni disponibili sull'OG italiano (riportate nella tabella 1) si osserva che le stime ufficiali della Commissione europea si avvicinano maggiormente al risultato del modello IRW prevedendo una chiusura dell'OG già nel 2018. La serie stimata dall'OCSE presenta un *output gap* che si chiude più lentamente e si avvicina maggiormente al risultato del modello RW-CP. Infine, la stima del FMI si avvicina dal 2012 in poi al risultato del modello LLT-CP.

Tab. 2 – Parametri stimati e loro significatività nelle diverse specificazioni del modello (1)

		LLT	IRW	RW	LLT-CP	IRW-CP	RW-CP
OG	θ_1	1,77 ***	1,44 ***	1,81 ***	1,74 ***	1,81 ***	1,79 ***
	θ_2	-0,79 ***	-0,52 ***	-0,82 ***	-0,76 ***	-0,82 ***	-0,81 ***
	σ_{θ}	2,75 ***		3,31 ***	2,36		2,58
trend PIL	σ_{θ}	0,06 **	0,03 ***		0,02 ***	5,31	
trend UR	b_{11}	-3,60 ***	-0,93 ***	-2,98 ***	-1,93 **	-6,91 ***	-1,60 ***
	b_{21}	-2,17 *	-1,46 ***	-2,17 ***	-2,64 ***	-1,18	-2,35 ***
	σ_{θ}	57,84 ***	91,25 ***	63,72 ***	69,66 ***	58,52 ***	73,64 ***
trend infl	a_1	0,75 **	0,11 *	0,66	0,38 **	0,91 ***	0,35
	a_2	-0,68 **	-0,02	-0,61	-0,31 *	-0,86 ***	-0,29
	σ_{θ}	1,33	1,68 ***	1,42	1,46 *	1,40 ***	1,47
	γ_0	0,005 ***	0,006 ***	0,005 ***	0,001 *	0,001	0,001
	γ_1	0,003 *	0,003 *	0,003 *	0,001	0,001	0,001
trend CP	b_{13}				0,26 ***	0,45 **	0,25 ***
	b_{23}				-0,15 **	-0,41 **	-0,16 **
	σ_{θ}				0,11	0,13 *	0,11 ***
	logLik	1252	1239	1242	1865	1839	1859
	BIC	-2446	-2424	-2431	-3657	-3610	-3650
	Akaike	-2480	-2455	-2463	-3700	-3650	-3689

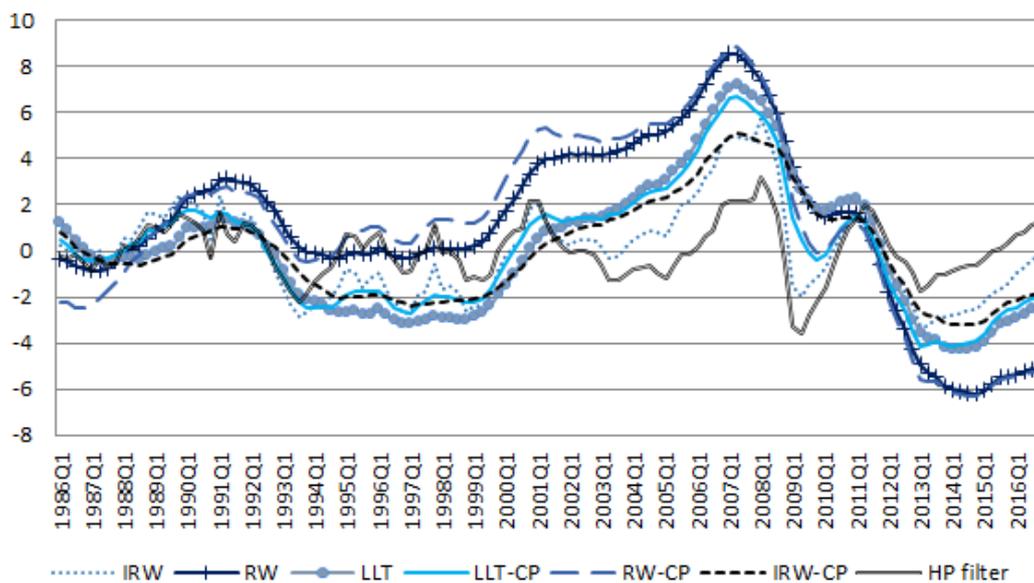
LogLik è la verosimiglianza (in logaritmi). Il confronto è possibile solo tra modelli annidati e quindi nei due gruppi con o senza capacità produttiva. BIC (Schwarz) e AIC (Akaike) sono i criteri per la selezione di un modello tra una classe di modelli con un diverso numero di parametri. (***) indica significatività del parametro al 99 per cento, (**) al 95 per cento e (*) al 90 per cento. Le varianze sono moltiplicate per 100.000

Tab. 3 – Output gap in base a diverse specificazioni del modello di stima

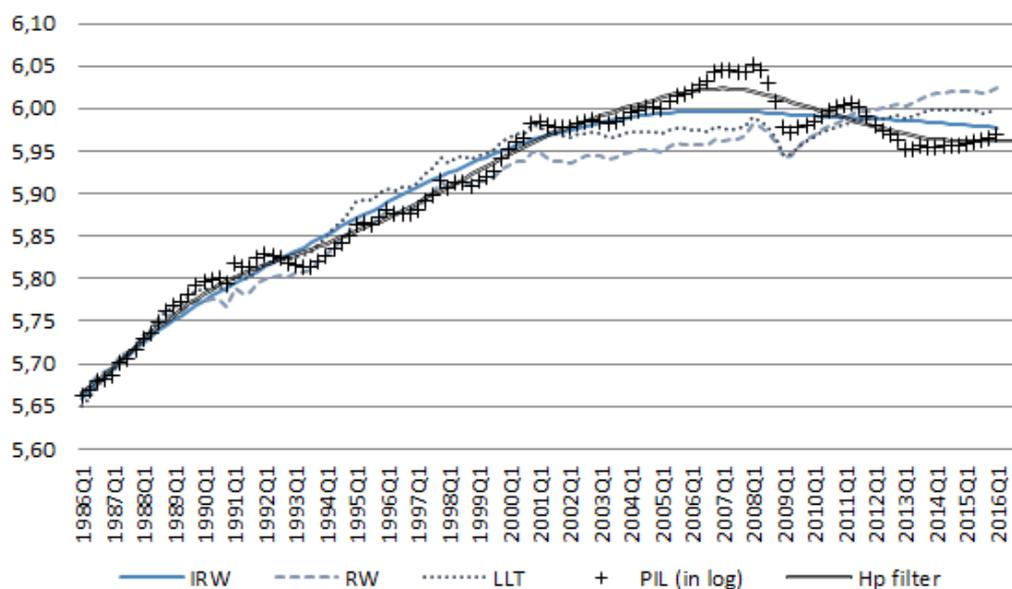
	LLT	RW	IRW	LLT-CP	RW-CP	IRW-CP
2010	2,0	1,6	0,3	0,6	0,6	1,4
2011	1,5	0,7	1,2	0,8	0,3	0,9
2012	-1,9	-3,0	-1,6	-2,4	-3,4	-1,3
2013	-3,9	-5,4	-3,1	-4,0	-5,7	-2,9
2014	-4,2	-6,1	-2,7	-4,0	-6,2	-3,2
2015	-3,4	-5,7	-1,8	-3,0	-5,7	-2,6
2016	-2,7	-5,2	-0,5	-2,2	-5,3	-2,0
2017	-2,4	-5,0	0,1	-1,7	-5,0	-1,7
2018	-1,9	-4,3	0,1	-1,1	-4,1	-1,3

Fig. 1 – Output gap e potenziale in base a diverse specificazioni del modello di stima (1)

Output gap



Prodotto potenziale



(1) Nei grafici viene riportata come *benchmark* la scomposizione tra ciclo e *trend* come risulta dall'applicazione del filtro di Hodrick e Prescott sulla serie del PIL.

3. Valutazione dei modelli alternativi

Data la notevole varietà dei risultati, si può cercare di ricorrere a un criterio di valutazione per la scelta del modello da preferire²⁰. Si considerano di seguito tre differenti criteri: 1) il grado di accostamento ai dati; 2) la stabilità dei parametri stimati, 3) l'adeguatezza nell'identificare la fase ciclica sulla base della capacità previsiva dell'inflazione. I differenti criteri, tuttavia, non forniscono indicazioni univoche, conducendo a privilegiare forme funzionali diverse.

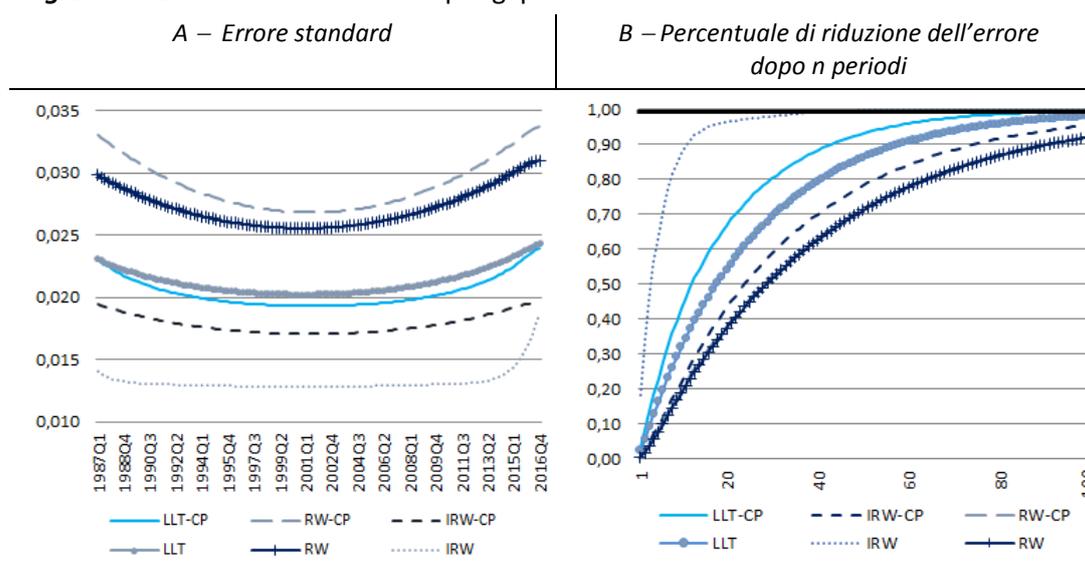
In termini di verosimiglianza (primo criterio) il modello che è maggiormente confermato dai dati è LLT, con o senza capacità produttiva (tab. 1), anche considerando i due criteri BIC e AIC²¹. Se, invece, per tale criterio ci si basa sullo *standard error* dell'OG, ottenuto con l'algoritmo di *smoothing*, il modello migliore risulta quello con *trend* integrato (IRW, fig.2a) anche se in questo caso l'errore intorno alla stima puntuale si amplia alla fine del campione. La possibilità di costruire delle bande di confidenza intorno al valore puntuale dell'*output gap* risulta in questo ambito particolarmente utile considerata la difficoltà di misurazione del fenomeno. La stima di un intervallo piuttosto che di un singolo valore potrebbe essere maggiormente funzionale per le decisioni di politica economica.

Un ulteriore elemento di valutazione dei modelli è l'entità delle revisioni della stima (secondo criterio), misurata in termini di numero di osservazioni necessarie per avere stabilità nei parametri. In questi termini il modello migliore tra quelli analizzati risulta IRW (fig.2b). In questo caso infatti già aggiungendo le prime osservazioni si hanno revisioni forti, ma che conducono velocemente alla stima finale e già dopo una decina di trimestri si raggiunge una riduzione dell'errore del 90 per cento (solo, comunque, nel caso in cui non si include la variabile capacità produttiva). Tra i restanti modelli LLT-CP ha una velocità di convergenza migliore, infatti si ha una riduzione dell'errore di stima del 90 per cento dopo circa 40 osservazioni (10 anni) mentre ne servono quasi il doppio per RW e RW-CP. In termini strettamente statistici quindi il modello definito LLT-CP sembra in complesso il migliore (sia per LogLik che in termini di variabilità delle stime).

²⁰ Con riferimento alla questione dei criteri di selezione si può osservare che il metodo della Commissione europea, pur non prevedendo una misura di valutazione del risultato del modello, privilegia, nella scelta discrezionale dei parametri di inizializzazione (del NAWRU e della TFP), l'obiettivo di minimizzare le revisioni delle stime. Recentemente è stato introdotto un sistema di controllo circa la plausibilità dei risultati (il cosiddetto *Plausibility tool*) che produce una banda di accettazione per identificare situazioni anomale nella stima dell'*output gap*. Per maggiori dettagli si veda: Hristov A., Raciborski R. e Vandermeulen V., (2017) "Assessment of the Plausibility of the Output Gap Estimates", European Economy Economic Briefs, Aprile. Nel presente lavoro si segue una procedura di selezione simile a quella seguita da Jarocinski e Lenza (2016). Questi autori selezionano tra i più modelli caratterizzati da diverse specificazioni del trend, effettuando confronti sulla base della verosimiglianza e della capacità predittiva dell'inflazione.

²¹ I criteri BIC – Bayesian information criterion o Criterio di Schwarz e AIC – criterio di Akaike valutano la verosimiglianza di un modello introducendo una penalizzazione in base al numero di parametri stimati così da evitare problemi di *overfitting* del modello. Essi risultano come segue: $BIC = -2\ln(L) + k \ln(n)$ e $AIC = 2k - 2\ln(L)$ in cui $\ln(L)$ è la log-verosimiglianza, k è il numero di parametri ed n il numero di osservazioni.

Fig.2 – Errore standard dell'output gap e revisione delle stime ²²



Risultati sostanzialmente diversi si ottengono invece quando si considera il terzo criterio, ossia la capacità dei vari modelli di prevedere dell'inflazione (tab. 4). A tal fine si è condotto un esercizio di validazione basato sulla stima ricorsiva dei modelli sul campione 2002Q2-2016Q4. In particolare, ogni modello è stato stimato su un intervallo temporale mobile che ad ogni passo, partendo dal 2002Q2, si incrementa di una osservazione. A ogni passo si prevede l'inflazione annualizzata fino a 4 trimestri in avanti e si confrontano i valori previsti con il dato osservato. Come misura sintetica dell'errore complessivo si considera il Mean Square Error (MSE) ²³.

Inoltre, i diversi modelli vengono messi a confronto con un semplice *benchmark*, rappresentato da un modello *random walk* sui livelli della serie dei prezzi al consumo, comunemente utilizzato come modello di riferimento *naive* per prevedere l'inflazione ²⁴.

²² La revisione delle stime è calcolata tramite un *fixed-point-smoother* rapportando la varianza *real time* con quella di *steady state* che si ottiene considerando un campione infinito. Si veda Anderson, B. D.O., and Moore J. B. (1979), "Optimal Filtering" (Prentice-Hall: Englewood Cliffs, NJ) e de Jong, P. (1989), "Smoothing and Interpolation with the State Space Model", *Journal of the American Statistical Association*, 84, 1085-1088.

²³ Risultati analoghi si ottengono considerando come misura di errore il Mean Absolute Error. Queste evidenze vengono omesse per brevità.

²⁴ Sulle problematiche connesse alla previsione dell'inflazione si veda Stock and Watson (2007), "Why has the U.S. inflation become harder to forecast?", *Journal of Money, Credit and Banking*, vol 39, febbraio.

Tab. 4 – Errore di previsione (MSE) dell’inflazione annualizzata e test Diebold – Mariano h trimestri in avanti su diversi campioni (1)

		h=1	h=2	h=3	h=4	Media
LLT	2002-2016	0,09	0,24	0,44	0,64	0,36
	<i>DM test</i>	<i>0,08</i>	<i>0,09</i>	<i>0,04</i>	<i>0,03</i>	
	2002-2007	0,05	0,18	0,37	0,62	0,31
	2008-2011	0,14	0,35	0,56	0,61	0,42
	2012-2016	0,09	0,22	0,43	0,71	0,36
RW	2002-2016	0,13	0,36	0,62	0,86	0,49
	<i>DM test</i>	<i>0,35</i>	<i>0,19</i>	<i>0,01</i>	<i>0,01</i>	
	2002-2007	0,05	0,18	0,37	0,61	0,30
	2008-2011	0,29	0,81	1,35	1,68	1,03
	2012-2016	0,08	0,17	0,28	0,44	0,24
IRW	2002-2016	0,05	0,14	0,30	0,49	0,25
	<i>DM test</i>	<i>0,07</i>	<i>0,07</i>	<i>0,05</i>	<i>0,04</i>	
	2002-2007	0,05	0,10	0,21	0,35	0,18
	2008-2011	0,05	0,16	0,31	0,40	0,23
	2012-2016	0,06	0,17	0,41	0,77	0,35
LLT-CP	2002-2016	0,12	0,23	0,37	0,52	0,31
	<i>DM test</i>	<i>0,01</i>	<i>0,03</i>	<i>0,03</i>	<i>0,04</i>	
	2002-2007	0,05	0,09	0,16	0,27	0,14
	2008-2011	0,07	0,16	0,24	0,27	0,18
	2012-2016	0,27	0,48	0,76	1,08	0,65
RW-CP	2002-2016	0,15	0,28	0,45	0,65	0,38
	<i>DM test</i>	<i>0,00</i>	<i>0,01</i>	<i>0,01</i>	<i>0,01</i>	
	2002-2007	0,05	0,10	0,20	0,33	0,17
	2008-2011	0,11	0,20	0,31	0,36	0,24
	2012-2016	0,33	0,58	0,93	1,35	0,80
IRW-CP	2002-2016	0,23	0,39	0,58	0,76	0,49
	<i>DM test</i>	<i>0,00</i>	<i>0,00</i>	<i>0,00</i>	<i>0,00</i>	
	2002-2007	0,11	0,17	0,28	0,43	0,25
	2008-2011	0,28	0,47	0,65	0,74	0,54
	2012-2016	0,34	0,59	0,91	1,24	0,77
<i>Benchmark</i>	2002-2016	0,04	0,10	0,15	0,21	0,13
	2002-2007	0,03	0,07	0,13	0,24	0,12
	2008-2011	0,03	0,08	0,15	0,21	0,12
	2012-2016	0,08	0,17	0,18	0,18	0,15

(1) In tabella viene riportato in grassetto il minimo errore tra modelli, escluso il *benchmark*. Il valore appare sottolineato se l'errore risulta inferiore anche al *benchmark*. Il test Diebold-Mariano è un test di significativa differenza tra le previsioni ed è condotto confrontando ogni modello con il *benchmark*. Quindi un valore pari o inferiore a 0,05 indica che gli errori di previsione sono significativamente diversi per un livello di confidenza pari o superiore al 95 per cento. Confronti basati sul MSE vanno considerati con cautela quando il numero di parametri stimati è molto diverso tra i modelli considerati.

La bontà dei modelli in termini di capacità previsiva dell'inflazione non è stabile nel tempo, ma dipende dal periodo considerato. In generale, gli errori di previsione dell'inflazione risultano più ampi nell'ultimo periodo (2012-16) contrassegnato dalla caduta dell'inflazione a livelli storicamente molto bassi. Sull'intero campione 2002-2016 il modello IRW risulta il migliore, in media e per tutti gli orizzonti e l'errore è statisticamente inferiore a quello degli altri modelli secondo il test di Diebold e Mariano²⁵ (con un errore massimo del 7 per cento). Se, però, si spezza il campione e, in particolare, si isola il periodo della recente crisi economica il risultato cambia in modo sostanziale. Fino al 2011 il modello LLT-CP è il migliore. Dal 2012 in poi, ovvero nel periodo dell'ultima recessione e successiva debole ripresa, il modello che commette meno errore (in media e per ciascun passo di previsione, eccetto il primo) è quello RW, ovvero il modello caratterizzato da una dinamica del *trend* più stabile e meno pro-ciclica. Quest'ultima evidenza potrebbe costituire una conferma dell'anomalia che ha contrassegnato la recente fase economica (ciclo notevolmente più ampio dello *standard*) e delle conseguenti implicazioni per la stima del potenziale (dinamica del *trend* meno pro-ciclica).

Oltre a ciò, un ulteriore elemento da porre in evidenza riguarda il ruolo della capacità produttiva utilizzata nel contribuire a prevedere l'inflazione. Il modello LLT-CP, che utilizza la serie dell'indagine Istat sul grado di utilizzo degli impianti, commette un errore medio di previsione inferiore agli altri modelli fino al 2011, ma diventa meno preciso negli anni più recenti. Ciò potrebbe essere legato alla circostanza che il segnale derivante dalle inchieste Istat sul grado di utilizzo della capacità produttiva è divenuto meno correlato con l'effettivo andamento del ciclo. Tale fenomeno potrebbe derivare da una serie di condizioni realizzatesi a seguito della recente crisi economica. Ad esempio la severità della caduta dell'*output* potrebbe avere modificato il livello ritenuto "normale" per gli imprenditori che oggi hanno aspettative più contenute sulla produzione rispetto al passato. Accanto a ciò, occorre tenere conto che nel corso della crisi si sono avuti intensi processi selettivi e le imprese sopravvissute sono quelle più efficienti che probabilmente riescono a sfruttare al meglio gli impianti in dotazione²⁶. Tale mutamento di composizione, renderebbe la valutazione del grado di utilizzo non del tutto confrontabile con il passato. Infine, si deve considerare che durante la crisi molte imprese hanno rimandato a tempi migliori il rinnovamento degli impianti, talché la relativamente rapida crescita del grado di utilizzo osservata nella fase di ripresa rifletterebbe, più che un fenomeno ciclico, le conseguenze di tale rinvio.

²⁵ Diebold FX., e Mariano RS. (1995), "Comparing predictive accuracy", *Journal of Business and Economic Statistics*, 13: 253-263.

²⁶ Sul mutamento di composizione della popolazione delle imprese dopo la crisi, si veda Linarello A., Petrella A. "Productivity and reallocation: evidence from the universe of Italian Firms". Banca d'Italia (2016), "Questioni di Economia e Finanza" n. 353, settembre.

Conclusioni

In questo lavoro è stato applicato un modello a componenti inosservate per stimare il potenziale e l'*output gap* dell'Italia nel campione 1985-2016, considerando diverse specificazioni del *set* informativo e della forma funzionale per il *trend*. In sintesi, i risultati confermano l'elevato livello di incertezza che caratterizza la stima di tali variabili. In particolare:

- forme alternative del modello di stima, nella specificazione del *trend* e nel *set* informativo, producono un intervallo molto ampio di valori per l'*output gap*;
- la valutazione della bontà dei vari modelli è soggetta alla scelta del criterio adottato che non conduce, però, all'individuazione univoca di un modello da preferire agli altri. Il modello LLT sembra in complesso il migliore in termini statistici. Considerando, invece, la capacità previsiva dell'inflazione nell'intero periodo 2002-2016, la specificazione IRW risulta preferibile;
- tuttavia, questa relativa superiorità nel predire l'inflazione muta nel tempo. Considerando il periodo più recente di duplice recessione, caduta del tasso di inflazione su livelli molto bassi e lenta ripresa dell'economia (2012-16), il modello che prevale per bontà previsiva è quello RW, caratterizzato da un *trend* meno pro-ciclico;
- inoltre, l'evidenza circa la capacità predittiva dei vari modelli riguardo all'inflazione pone in luce anche un deterioramento del segnale ciclico nel periodo 2008-2011 per cui l'errore di previsione è in tutti i modelli, eccetto il *benchmark*, superiore agli altri intervalli temporali. Successivamente al 2011 anche il *benchmark* commette un errore notevolmente superiore alla media storica;
- l'anomalia della recente fase economica potrebbe anche essere posta in relazione con l'indebolimento della capacità esplicativa della variabile grado di utilizzo degli impianti.

Tali risultati confermano che la stima del prodotto potenziale e dell'*output gap*, già normalmente circondata da incertezza, è divenuta ancora più complessa dopo la recente crisi economica. Il principale punto interrogativo è, evidentemente, costituito dal fatto di non conoscere se, e in caso affermativo quanto, le ampie cadute cicliche dell'*output*, succedutesi dal 2008, abbiano inciso negativamente anche sul *trend* dell'economia (sul suo tasso di crescita oltre che sul livello). Il *framework* utilizzato in questa Nota di lavoro, nel mettere in luce le conseguenze derivanti da una simile incertezza, fornisce indirettamente indicazioni su come si può cercare di affrontare questo problema. Un organismo di *policy* che fa dipendere le prescrizioni di politica economica da variabili non osservabili e peraltro affette da elevata incertezza non può basare le proprie scelte su una singola stima. Occorre che disponga di una varietà di stime, corredate di bande di confidenza, ottenute con approcci metodologici differenti oppure, alternativamente, se si vuole definire un unico modello, questo dovrà essere sufficientemente flessibile da poter incorporare nella propria struttura ipotesi economiche diverse circa la natura del

ciclo e del *trend*²⁷. La scelta di una delle stime tra le diverse ottenute, che come mostra questa Nota possono essere anche molto distanti, può basarsi, ad esempio, sul valore centrale dell'intervallo²⁸, oppure può essere desunta (come nel presente lavoro) combinando criteri statistici con criteri legati alla specificità della fase ciclica, particolarmente nei casi in cui sono molto incerte lunghezza e profondità della fase negativa del ciclo. In quest'ottica, lo strumento del *plausibility tool*²⁹ adottato dalla Commissione europea per verificare l'attendibilità delle proprie stime di OG va nella direzione di valutare l'incertezza associata alle stime, ma in modo da considerarsi insufficiente data la portata del problema. Con l'adozione di un simile strumento, infatti, si verifica unicamente se la stima dell'OG rientra in un intervallo di accettazione che non è una vera e propria banda di confidenza essendo basato su un modello di regressione ausiliario che prende comunque per data l'ipotesi di *trend* (oltre che di OG) del modello della Commissione. Invece, è proprio su questa variabile che, come mostrano anche i risultati di questa Nota, si è concentrata negli ultimi anni la maggiore incertezza.

Con riferimento agli sviluppi futuri del *framework* presentato in questo lavoro, una strada che potrebbe essere percorsa riguarda la scelta di forme funzionali flessibili, oltre che per il PIL, anche per le altre variabili di *output* e per l'inflazione. In particolare la verifica dell'ipotesi di anomalia della recente fase di crisi economica potrebbe essere effettuata stimando un modello con parametri variabili nel tempo. Inoltre, per verificare in modo rigoroso l'ipotesi di isteresi si potrebbe considerare un modello con ciclo e trend correlati come proposto da Proietti e coautori³⁰. Infine, la specificazione della curva di Phillips potrebbe beneficiare di una parametrizzazione più flessibile che consideri ad esempio una varianza di tipo stocastico come recentemente proposto da Stock e Watson³¹.

²⁷ Si veda, ad esempio, Proietti T., Musso A. and Westermann T. (2007), "Estimating Potential Output and the Output Gap for the Euro Area: a Model-Based Production Function Approach, *Empirical Economics*", 33, 85-113.

²⁸ Si veda, ad esempio, Banca d'Italia, "Relazione annuale sul 2016", pp. 57-58.

²⁹ Si veda nota 20.

³⁰ Si veda il lavoro citato nella nota 27.

³¹ Stock H.J, Watson M. (2010), "Modeling Inflation After the Crisis", NBER Working Paper No. 16488, October.