

**“ESPERIENZE E PROBLEMI DI VALUTAZIONE DEGLI EFFETTI
DI POLITICHE PUBBLICHE IN ITALIA”**

Workshop con ricercatori impegnati in attività di valutazione

Conference Centre dell'Università di Trento, Sardinia (Tn), 29 febbraio-1 marzo 2008

APPUNTI SULLO WORKSHOP

1^ mezza giornata

Costa:

- definizione di politiche pubbliche: OK. Ma non assoluta di assenza di valutazione su altri ambiti.
- Importante valutazione di (de)regolazione: es. legge sul fumo; distribuzione controllata di eroina;
- Contro fattuale: cruciale in fase 3 di valutazione del farmaco (sperimentazione con randomizzazione) e in fase 4 di applicazione a larga scala. Più lenta, ma visibile, in ambiti epidemiologici o simili.
- Categorizzazione in sanità: sperimentazione con randomizzazione; technology assessment ex-ante per trasferibilità su larga scala; valutazione ex-post (a) di implementazione di requisiti di processo in fase di attuazione su ampia scala e (b) valutazione osservazionale (che mima il paradigma fattuale).
- OK. Ma fuga in avanti con “bilanci sociali”!

- Ritardo italiano rispetto a tradizione liberale (convenienza), scandinava (accountability) e di paesi in via di sviluppo (per aver eccesso a risorse). Esempio: caso Di Bella.
- Un sistema sociale decide su regole del gioco: peso di (i) beneficiabilità, (ii) libertà di scelta, (iii) effetti distributivi. Assenza di queste regole condivisibili e trasparenti rende opaco il sistema. Valutazione di impatto è utile soprattutto per (i).
- Dubbio sul fatto che spesa per trasferimenti sia un disincentivo (“pay for performance” è un caso al margine).
- Manca la segnalazione di “assenza di domanda” da parte di portatori di interessi (es. prevenzione).

- Storia: doing things chepaer, doing better, doing right, doing right things right.
- Povertà di fonti informative per ricerca valutativa (servono dati longitudinali multipurpose)

Samek:

- Esperienza di valutazione di fondi strutturali perché (i) c'è l'obbligo; (ii) ci sono i soldi; (iii) c'è selezione di valutatori indipendenti.
- Dubbio su ambito ristretto di “politiche pubbliche”. Come si fa a valutare macro-politiche in maniera decentemente rigorosa, o comunque dire che non si può valutare limitandosi al monitoraggio.
- Importante l'approccio controfattuale (non sempre è possibile; è costoso).
- OK interazione con analisi di valutazione: casi in cui non si è fatto niente; importante capire perché ciò è accaduto.
- OK diverso da budget, attento AD apprendimento.

- Ritardo. Mancanza di dati longitudinali e di dati integrati amministrativi e di accessibilità.
- Domanda: fattori chiave cultura formalistica + committenza che non di che cosa parla, e quindi fa domande valutative improprie (megalomania cognitiva), in tempi impropri e soldi inadeguati. Es. UE-DG Regio. NB. Anche valutatori esterni dovrebbero concorrere e fare domande. + ambiguità di ruoli di valutazione e di assistenza tecnica (richiesta da AAPP, che pagano), sicché valutazione assume di fatto ruolo di consulenza.
- problema di interazione di discipline dal lato dell'offerta.

Billari:

- Attenzione a politiche di famiglie con figli e povertà. Record di bassa natalità, lunga permanenza di figli in famiglia, deboli politiche di contrasto alla povertà.
- C'è ritardo anche in parte esplorativa/di monitoraggio. Meno esigenti su definizione di "politiche pubbliche"
- In ambito familiare/demografico, cruciale il fatto dei tempi lunghi (variabili che sono connesse dinamicamente).
- Ritardo sì, ma – per politiche familiari – comune a Germania e Francia (e a UE). Diverso il contesto anglosassone e quello scandinavo.
- Decisiva valutazione d'impatto, sennò prevale ideologia (su famiglie in particolare).
- Colpevole è cultura anti-empirica, e in particolare facoltà Giurisprudenza.
- Assenza di cultura di committenti/valutatori che sappiano valutare i valutatori.
- Altro fattore di domanda: ciclo politico è troppo breve.
- L'assenza di dati, e la loro scarsa accessibilità, vero killer della valutazione. Monopolio Istat in produzione e, in fatto, utilizzo dei dati.
- Interdisciplinarietà contrastata da settori scientifico-disciplinari universitari (e conseguente rigidità in programmi di insegnamento universitario).
- Chi valuta i valutatori? Mancano protocolli e standard. C'è quindi offerta di qualità eterogenea. *Replication studies?*
- Mancano infrastrutture di ricerca (es. dati longitudinali per coorte). Serve competizione in mercato della ricerca. *UK Data Archive?*

Stame:

- Paradigma contro fattuale. D'accordo, ma limitazione a certe politiche pubbliche, che sono però rare e poco rilevanti. Ma ci sono politiche più complesse, con tanti altri soggetti, per le quali si possono avere risultati cognitivi più modesti. *Definizione limitata di interventi valutabili con approccio controfattuale?*
- Importanza di analisi di implementazione.
- Serve ragionamento su approcci più vasti: valutazione basata su teoria (teorie del cambiamento); approccio multi-metodo.
- Ritardo è complessivo rispetto alla cultura della valutazione e di *policy analysis*.
- Causa: cultura di pubblica amministrazione (norma/procedura e non risultato, e quindi non apprendimento).

Cappi:

- Poleis: su lavoro, fp, sviluppo locale; assistenza tecnica, monitoraggio, valutazione di effetti.
- Definizione di politica valutabile è importante per **programmare/disegnare** la politica.
- Enfasi di AAPP è su allocazione/gestione di risorse. Politiche disegnate in modo che siano poco valutabili. Serve agire sul **ciclo di valutazione**: (i) dimensione di problema su cui intervenire; (ii) simulazione ex-ante; (iii) disegno adeguato della politica; (iv) infrastrutture informative; (v) valutazione d'impatto.
- Superare ritardo è processo lungo e impegnativo.
- Cause:
 - cultura antiempirica e formalistica di PA,
 - domanda valutativa espressa in maniera inadeguata, generica.
- Cause: scarsa completezza/qualità e poca accessibilità dei dati. Sistema informativo cruciale per finalità cognitive. *Es.: collocamento mirato per disabili; mancanza di informazioni su disabilità.*
- FSE autorità superiore; provincia valutato che chiede valutazione a consulente esterno, con possibile rischi salva deontologia di consulenza. *NB. Depersonalizzare: si valuta la politica.*

Battistin:

- Concetto di evento controfattuale non è un metodo, ma una definizione (logica/filosofica) che permette di esplorare scelte di comportamento. Ricostruzione di evento controfattuale è possibile a livello di gruppo.
- Maggiore conoscenza del contesto istituzione, migliore interazione con attori politico-amministrativi, migliore la possibilità di svolgere la stima dell'effetto.
- Chiarire meglio ambito di applicabilità di approccio controfattuale: anche interventi strutturali sono valutabili in tal modo, con variante ex-ante (in particolare non unicità: serve per altri interventi analoghi; marginale reversibilità Approccio
- Esperienza inglese: in 10 anni è cresciuta.
- Migliorare soprattutto lato di domanda (sensibilizzare decision makers)
- Prepararsi a tsunami/domanda:
 - formazione universitaria;
 - infrastruttura di dati (Es. 1: inglese: Training Partnership Act con sperimentazione; Es. 2: Messico e politica del latte importato da USA, inefficace).

Mura:

- Esperienza di osservatorio Sardegna: poche valutazioni, e solo regionali.
- Importante l'aspetto di analisi di implementazione accanto alla valutazione. Committente vuole capire effetto quantificato e perché: tra i perché diversità di implementazione.
- Tra cause del ritardo, processo politico con cambio di committente (chi riceve i risultati dello studio non ha condiviso la definizione degli obiettivi delle da valutare).
- Peso di formazione giuridica di AAPP: nel formulare domanda; nell'accettare/interpretare un risultato probabilistico.
- Dal lato dell'offerta, pochi gli esperti. E anche qui un problema culturale: altre offerte di valutazione di un modello strutturale con variabili di controllo (banalmente, di regressione).

- Problema dei dati: innescando una buona domanda, si stimola raccolta di dati insieme con la realizzazione della politica.

Ercoli:

- Commissione bilancio Senato e esperienza in CIL e per CSS.
- Rassegna di incentivi alle imprese: 488 su incentivi all'occupazione ha effetti positivi; le altre politiche senza effetti; forse serve ampliare l'ottica valutando interazioni fra diverse politiche.
- In Italia non c'è un modello né di beneficiabilità né di accountability. Quadro confuso.
- In ruolo di processo politico, distinguere fra governo centrale e governi locali. Questi ultimi hanno attenzione a disegno della politica, molto meno il governo centrale. Cause? (i) Vincoli di risorse. (ii) Opposizione a sperimentazione. (iii) Tempi politici troppo stretti. (iv) Difficoltà d implementazione (che portano a politiche rozze, senza selettività).

Anastasia:

- Non c'è rischio di domande esigenti da parte di committenti.
- Funzione di Veneto Lavoro:
 - monitoraggio;
 - messa a disposizione di microdati per valutazione.
- Problemi di PA:
 - eccesso di risorse e esigenza di impiegare una somma, deleteri;
 - una qualche idea della valutazione e di cosa funzioni o no, i funzionari se la fanno. Ma non la esplicitano per il condizionamento politico. Manca atteggiamento di "laicità", di non personalizzazione e di apprendimento.
 - esercizi di valutazione utili, e verranno dall'esterno. Utili per mettere a fuoco: (i) processo di implementazione (aspetti organizzativi); (ii) retroazione sugli obiettivi (si ripensa la politica e si precisa l'obiettivo).

Casovola:

- Essere più espliciti nel dire che ci sono certe politiche che sono valutabili in maniera rigorosa. Ma è importante valutare anche altre politiche, es. di quelle di sviluppo locali: con che strumenti?
- Indietro nel formulare le *policies*, innanzitutto (più che nel valutare).
- Funzionari pubblici non hanno tare genetiche, sono mediamente come tutta la popolazione. Contano gli incentivi per i funzionari pubblici: sono valutati/premiati su base di coerenza con procedure.
- C'è anche un problema di potere.
- Problema di Fondi Strutturali riguarda politiche complesse, difficilmente valutabili. Eccezione è FSE, che riguarda politiche teoricamente ben valutabili ma poco trasparenti, condizionate da domande di Commissione Europea e con un mercato dei valutatori poco competitivo (concorrono i peggiori).

Pellegrini:

- Politiche da valutare: esigenza di valutare tutte le politiche: quelle multi-obiettivo e multi-azione (es. politiche regionali da valutare a livello europeo): quelle infrastrutturali (anche perché ripetibili; o perché servono per vagliare plausibilità di valutazione ex-ante).
- Siamo in ritardo? In politiche di incentivo alle imprese, letteratura internazionale è scarsa; ritardo in Italia non è particolare (anche perché in Usa non ci sono politiche di questo tipo).
- Problema è non interazione fra studi di valutazione e decisione di politici (es. 488: cambiata, poi chiusa per [pochi] casi giudiziari eclatanti). (i) C'è processo di disegno e gestione dell'intervento che non include, e si chiude, con valutazione. (ii) Quindi mancano i dati, che vengano cercati faticosamente ex-post. (iii) Sensibilità di dirigenti di AAPP è cresciuta, ma non chiedere e giudicare la valutazione [e con pochi incentivi per questi]. (iv) Quindi, cattivi valutatori sono indotti a barare.

Ichino:

- Esperienza E-R: 1.000 € a famiglia con reddito ISE sotto 10.333 se hanno un figlio a scuola secondaria superiore; 300 € a chi sopra la soglia ha media del 7 (obiettivo: ridurre abbandono scolastico). Costo: 20 milioni di €.
- Contatto con funzionario di E-R: intelligente, sensibile a valutazione. Province mettono insieme i dati a livello regionale. Ma non c'è voto rilevato sotto la soglia.
- Problema non è di domanda e offerta. Ma acquisizione di dati amministrativi non è pensata in funzione di processi di apprendimento: non rilevazione di voto sotto la soglia; non abbinamento possibile con dati ISE sopra la soglia ottenibili da Ministero di Finanze.
- Problema è spersonalizzare: scrivere che E-R ha fatto buona valutazione ed ha imparato (ma ha sbagliato politica!), con rischio di mettere però in difficoltà esperienza di E-R.

Cipollone:

- Che cosa si fa della valutazione? Ciò induce reazioni comportamentali (discorsivi ex ante, e quindi falsifica risultati. (i) Es. Invalsi-Elias: scuole nel Sud lasciano studenti copiare (o copiare la soluzione dettata dall'insegnante).
- Raccolta di dati non solo contestuale, ma anche preventivamente definita in relazione a caratteristiche oggettive di politica. (ii) Es. bonus-bebé e possibili effetti su salute dei bambini nel 1^o anno di vita in famiglie povere.
- Cultura formalistica in PA? Cultura è che scrivere una norma assume che essa produca effetti, non ci si domanda che reazioni comportamentali induce.

Rettore:

- Controfattuale è una logica, non un metodo.
- Cruciali non sono i settori scientifico-disciplinari, ma modalità di reclutamento (e incentivi a queste) dei ricercatori universitari.
- Politica sociale non è un trattamento (Stame): trattamento non è solo fattore causale; per identificare l'effetto devo misurare l'effetto di tutte le altre variabili rilevanti. Domanda di tutte le variabili causali è plausibile, ma non è quella che a noi interessa. Per rispondere alla domanda di interesse non serve misurarsi con la complessità.

NB: Risposta di Stame. Contro-risposta di UT: - eterogeneità ⇒ stratificazione

- effetto: - NO, tutto si chiude

- SI, si approfondisce causal chain.

Sestito:

- A carenza di domanda di committenza pubblica, c'è da parte di consulenti/valutatori capacità di indurre miglioramento nella formulazione di domande?
Postilla per Pellegrini (su 488): non è che il valutatore si è posto quesito su risultati scentrato?
- Posizione di AI: domanda c'è, il *trigger* sono i dati.
Posizione di BA: domanda non c'è; e importante produrre e distribuire dati.
Ma perché nel primo caso non si producono i dati, e nel secondo caso perché si producono?
- Posizione di Casavola: problema è disegno di *policies*. Ma non può essere *trigger* da parte di valutatori?

Barca:

- 488 positiva, ma perché è stata chiusa. Perché? Domanda a Pellegrini:
 - effetti positivi su occupazione, non su cambiamenti nel mercato di credito (vero obiettivo!);
 - perché studi non hanno trovato eco su stampa (sensibile a scandali, pur rari: 3-4%).

Pellegrini:

- cambiamento di priorità politiche (fuoco su innovazione);
- processo di selezione di imprese, al quale si preferisco meccanismo automatico di credito d'imposta (con 50% di fraudolenti tra ispezionati).

Mancini:

- Diversità di approcci a tema dei dati: (i) disponibilità; (ii) accessibilità.
- AI: uso integrato di dati amministrativi.
- FB: creazione di infrastrutture informative.
- Istat ha produzione soprattutto con *surveys*, con scopi di comparazione europea.

Mura:

- Dati non creati se non rispondono a esigenze amministrative.

NB. Due problemi distinti: - *accessibilità,*
- *miglioramento di dati amministrativi e loro potenziamento.*

2^ mezza giornata

Cipollone:

- Non basta clausola valutativa. Chiaro disegno di valutazione deve essere ex-ante deve essere fatto insieme a disegno di politica (coinvolgendo anche in questa fase gli esperti, in un ragionamento ex-ante).
- Contenuti di ricerca valutativa già in formazione (anche Sc. Pol. e Giurisprudenza; in Scuola superiore PA) e in reclutamento funzionari.

- Fare proposta per rendere liberale normativa su privacy. Proposta in Libro bianco? Sì.
- No a Istituto pubblico ad hoc. Serve consapevolezza diffusa.

Costa:

- Su aspetti valoriali e di interesse, vedi ieri su conoscenza del contesto sociale.
- Clausole valutative sì, ma che non sia un gran “girare di carte” (rischio forte: esempio VIA).
- Ruolo professionale? Parole d’ordine tipo “sicurezza” importanti se entrano in cultura del management; non basta creare nuova burocrazia o nuovo servizio.
- Formazione: aiutare mondo di ricerca a consapevolezza su obiettivi valutativi, più che puntare formazione tecnica specifica (domanda di questa viene di conseguenza). Importante finanziamento di ricerca applicata anche su fronte valutativo.
- Cruciale problema di dati: (i) completezza e (ii) accessibilità (NB. Es. poco uso di Indagini Multiscopo; serve piano specifico per valorizzare queste indagini.)
- No Istituto pubblico di valutazione. Meglio network di centri specialistici (che insieme costituiscono l’“Istituto”).
- Network scientifico-professionale nazionale di valutatori importante per dare indipendenza a valutatori, messa in gioco da diversificazione “federalistica” o da gruppi di interesse. Non c’è “peer review”.

Domanda di PS e risposta: (i) network ha favorito benchmarking al rialzo e reputazione; (ii) specializzazione non ha favorito concorrenza.

Domanda di PI e risposta: (i) effetto primo coinvolgimento in disegno questionario e in disegno longitudinale (obiettivi e fondi di Piano sanitario nazionale; protocollo di ricerca fra Regione Piemonte, capofila, e Istat, e distribuzione di dati a tutte le Regioni; (ii) elaborazione di base di dati longitudinali distribuiti come anonimi.

*NB. Mancini: - organizzare la domanda per interferire su disegno di indagini !!
- impegno di Istat a favorire accessibilità*

Samek:

- Clausole valutative: sì, ma ritualistiche in esperienza di Fondi Strutturali!!
- Dubbio su figure professionali specifiche.
- Fondi strutturali: spostare interesse? Servono tutte e due dimensioni. E formalmente anche “grottesca” valutazione d’impatto.
- Cruciale crescita di consapevolezza culturale, con dibattito pubblico e con formazione “ad hoc” in università.

Es. 1: “Trento valutazione”, anche per confronto interno!

Es. 2: network con ambito internazionale.

- No Istituto nazionale, monopolistico.
- Dati:
 - cruciale accessibilità;
 - parità anche per privati.

*Domanda di PS: - commesse specifiche vs. reclutamento qualità di valutatori che mirano a pubblicazioni di qualità. MS: difficoltà non a reclutare, ma a tenere ricercatori.
- vincoli di committente a rendere disponibili i dati per “replication studies”.*

Zuliani:

- Fino al 1995 Istat aveva difficoltà ad avere accesso ai dati di Ministero delle Finanze.

- Contano anche di stakeholders. - in Compstat due rappresentanti del MUR;
 - presenza in circoli di qualità.
- Uso di (i) dati standard Istat è poco usato, (ii) uso di protocolli è bassissimo. Per valutare, importano soprattutto dati amministrativi, e poi dati ad hoc, e ultimi dati Istat.
- Per favorire valutazione: - accentuare sperimentali di politica;
 - prevedere quota di budget vincolata a valutazione.
- Su Istituto pubblico: previsione che Istat possa fare scuola di formazione con SSPA e Università. Puntare su questa per fare innanzitutto scuola di valutazione e massa critica?

Billari:

- Problema di stabilità di Miniseri di riferimento per famiglie, bambini, immigrati, ecc..
- Attenzione a dire che tutto si risolve con dati amministrativi. Non vale per famiglie!
- Formazione:
 - a livello generale, cultura empirico-sperimentale (anche formazione di giornalisti, ecc.);
 - all'università, colpa specifica degli statistici (corsi di statistica chiusi; ruolo negativo di settori scientifico-disciplinari).
- Su dati: OK con Ugo; ma processi devono essere trasparenti. Commissione dovrebbe dire priorità (es. meglio Multiscopo cross-section grande ogni 5 anni o campione più piccolo+panel?). In prospettiva, i dati non sono "dati", ma ridisegnati.
- Ruolo di stampa: opera di sensibilizzazione da fare (es. Stella).

Stame:

- Accountabilty vs. apprendimento: OK. Ma quanti sono gli attori di apprendimento: valutatori, *policy makers*, amministratori (dello specifico contesto). In questo importante è analisi di implementazione, per politiche che cambia nel tempo.
- Questione del ruolo professionale: sarebbe importante avere reclutamenti di persone specificamente competenti di valutazione.
- In formazione; (i) esigenza di rafforzare dimensione valutativa in tutte discipline; (ii) esigenza di avere persone specificamente formate a competenza professionale valutativa.
- Istituto massa critica: paura di carrozzone; paura di monopolio. Abbiamo bisogno di pluralismo. Strumenti possono esser diversi (interessante proposta di Scuola di Zuliani).
- Idem per risultati di clausola valutativa: es. RMI.

Cappi:

- Prioritario ripensare il disegno della politica (che oggi porta al ribasso), che prevede strutturalmente esigenza di valutative (e anche di basi di dati). Solo così clausola valutativa non viene ridotta a funzione ritualistica.
- Indicazioni suggerite in C1 vanno lette come concorrenti, integrate (non specifiche).
- Rafforzamento di formazione: specificamente su analisi contro fattuale (Scuola di Martini);
- Trattare nel libro bianco il problema della *privacy*.

NB. Norma specifica nazionale derogatoria di "privacy" e di imposizione di "replication", in leggi con spesa pubblica.

NB bis. PS: - accessibilità ma non solo; ampliare/organizzare dati.

- IFAU: non monopolista; organizzatore dati; assicura dati replication studies.

NB ter. AI: box in libro bianco su buone pratiche e vincoli da dati.

Battistin:

- Quadro ideale: - fare sempre *demonstrations*,
- con randomizzazione (per aree in UK).
- Cultura di valutazione da far crescere: - non basta (e può essere pericolosa) clausola valutativa;
- formazione importante (di funzionari pubblici; [es. inglese: *randomization by area* e *propensity score matching*]);
- no singolo Istituto pubblico, meglio pluralità di istituti che fanno network (tipo Germania).
- Partire da programmi importanti, valutarli al meglio (possibilmente con esito +), per indurre un processo di *learning by doing*.
- Dati: usare al meglio i dati amministrativi; ma non bastano. Serve infrastruttura di base + *surveys*.

NB. GC: - *esempio di condizioni sperimentali per valutazione del farmaco;*
- *su caso di*

NB. FB: - *su dati: best practices UK Data Archive, Svezia, Danimarca;*
- *poi tappe italiane (Compstat, revisione codice privacy Istat);*
- *poi migliorare i dati amministrativi/da surveys.*

Mura:

- Clausola valutativa sì, ma formulata da chi è sufficientemente formato. Poi (i) amministrazione deve avere all'interno risorse per fare bandi e selezionare; (ii) affidamento all'esterno; (iii) uso dei risultati per decisioni.
 - Dati buoni: cruciali. E cruciale "replicabilità" (anche per depersonalizzare valutazione).
- NB. AM: *mette la sordina a clausole valutative (propriamente di attuazione di una politica, di monitoraggio).*

Ercoli:

- Non c'è ritardo su ambiti prossimi: - valutazione finanziaria (di costi) di una legge,
- di deve collocare in un sistema di *checks and balances*;
- non basta un piccolo ufficio, serve massa critica.
- Secondo aspetto: assetto istituzionale di AIR c'è, forse usabile.
- Obbligo di valutazione per leggi importanti (o per ambiti e misure).

NB. PS: *ipotesi di sessioni valutative per ambiti e misure circoscritte, in Commissioni. Cruciale approccio bipartisan.*

NB. AS: *Government Accounting Office, con dentro "Programme evaluation division". Ben diverso da Corte dei Conti.*

NB. FB: *domanda di parlamentari? Interessante apprezzare esistenza e caratteristiche di questa domanda. Risposta: logica Tremonti di fondi "ad hoc" distribuiti selettivamente vs. logica Visco di sgravi automatici.*

Casavola:

- Rafforzare obblighi più standardizzati di “reporting”. (Oggi per quasi nessuna politica si è in grado di produrre serie storica di erogazioni di spesa.) Poi da migliorare progressivamente.
- Le competenze: siccome valutazione ha senso se poi è utilizzata, competenze “interne” sono indispensabili. ⇒ Implicazioni su modalità di reclutamento (difficile per resistente di PA, difficile per ostilità di consulenti).
- Dati:
 - disponibilità di informazione strutturata su dati amministrativi disponibili;
 - formazione che serve per valutazione: (i) decisione ex ante anticipate per *policies* da valutare (con funzione di apprendimento); (ii) per valutazioni decise ex-post, dati amministrativi da migliorare in completezza, qualità e accessibilità in termini selettivi.
- Parlamento luogo prezioso per avanzamenti. Luogo ricco e potente.

Anastasia:

- Di valutazione c'è bisogno estremo. E di fatto tutto è “sperimentale” per vincoli di bilancio (es. apprendistato ipoteticamente con formazione ipoteticamente obbligatoria).
- Dati amministrativi: ci sono più dati che valutatori. E prospettiva è rosea. Oggi, 1 marzo, c'è dichiarazione di assunzione obbligatoria *on-line*. E dati miglioreranno. Cruciale accessibilità dei dati.
- Dati amministrativi sono una “brutta bestia”. Tra dati amministrativi e loro uso c'è attività, cruciale, di preparazione per fini conoscitivi. ⇒ Messa in qualità generale e *linkage* (forse troppo costosa). Serve una “intelligenza” ex-ante ed ex-post.
- Offerta e/o domanda. Devono crescere tutte e due. Crescita anche di offerta retroagisce su domanda. Occorre creare ridondanza di valutazioni perché processo politico ne sia influenzato. Ruolo di università per alimentare dibattito, in quanto svincolata da committenza.

Pellegrini:

- Clausola valutativa non basta. Serve processo complessivo di disegno della politica fino alla valutazione, con obbligo di tenerne conto (*sunset clause*).
- Funzionari alfabetizzati su valutazione? Non basta. Servono nuclei di valutazione con alta competenza in PA centrale (tenuto conto anche spostamento di competenze a Regioni).
- Sui dati: sfruttare stakeholders in processo di produzione dei dati da parte dell'Istat.

Ichino:

- Risposta più “bottom up”: dare buoni esempi di valutazione e pubblicizzarli molti; partire da cose concrete.
- C'è domanda latente, non necessariamente buona. Da orientare con esempi concreti (es. 1 Cipollone in Invalsi; es. 2 Ichino E-R).
- Impresione che Governo e Parlamento sono molto indietro: modelli simulazione statistici, senza reazioni comportamentali. (Es.