

Patrimonio e mercato immobiliare: lo sviluppo delle statistiche OMI

DI MAURIZIO FESTA*, FILIPPO LUCCHESE**, SAVERIO SERAFINI***

1 Le statistiche OMI: il percorso evolutivo

Con la nascita nel 2001 dell'Osservatorio del Mercato Immobiliare (d'ora in avanti OMI) nell'ambito dell'Agenzia del territorio¹ si concretizza una risposta, fino a quel momento scarsa, frammentata e sostanzialmente limitata all'interno del perimetro amministrativo, all'ineluttabile esigenza di costruire un'adeguata informazione economica sul patrimonio immobiliare nazionale e sui flussi di mercato che lo riguardano (compravendite, locazioni, mutui ipotecari).

Prima di allora², l'interesse da parte dell'Amministrazione ad analizzare il mercato immobiliare, originato dalla necessità di stimare il valore dei beni ai fini fiscali, condusse negli anni '90, presso la Direzione Generale Catasto e Servizi Tecnici Erariali, ad istituire l'Osservatorio dei Valori Immobiliari, nell'intento di fornire agli Uffici Tecnici Erariali (quindi alla stessa Amministrazione) dati articolati e il più possibile aggiornati relativamente ai valori di mercato del patrimonio immobiliare. Questi dati erano disponibili quindi esclusivamente per uso interno all'Amministrazione e solo dal 1997 furono resi disponibili all'utenza esterna, facendo emergere nuove esigenze di completezza, attendibilità e affidabilità delle informazioni rese dall'Osservatorio e gettando le basi per il futuro OMI.

Fin dal primo momento, il filo conduttore lungo il quale è maturata l'esperienza dell'OMI in questi venti anni è stato quello di concorrere alla trasparenza del mercato immobiliare³, con l'ambizione di poter divenire un autorevole riferimento istituzionale nel panorama della produzione dell'informazione economica sul mercato immobiliare. In tal senso, l'azione dell'OMI è stata intesa come il contributo volto al superamento di barriere all'informazione economica sul mercato immobiliare attraverso la valorizzazione statistica degli archivi in possesso dell'amministrazione finanziaria, a partire da quelli del catasto, della pubblicità immobiliare e del registro. L'elaborazione dei dati acquisiti in procedimenti amministrativi attraverso strumenti di carattere scientifico-statistico, dunque, è stata interpretata come il valore aggiunto determinante nell'ottica della icasticità dell'informazione.

Il ruolo dell'OMI si è man mano consolidato e parallelamente si è manifestata l'esigenza di ampliare e migliorare le informazioni e i dati forniti, sia all'interno del perimetro amministrativo, sia all'esterno.

Le finalità alla base di questa evoluzione possono essere considerate essenzialmente tre:

* M. Festa, Responsabile Ufficio Statistiche e studi sul mercato immobiliare della Direzione Centrale Servizi stimati e Osservatorio del Mercato Immobiliare (DC SEOMI)

** F. Lucchese, Funzionario Ufficio Statistiche e studi sul mercato immobiliare della DC SEOMI

*** S. Serafini, Funzionario Ufficio Statistiche e studi sul mercato immobiliare della DC SEOMI

¹ Come si è detto l'OMI fu costituito presso l'Agenzia del territorio, incorporata a partire dal 1 dicembre 2012 nell'Agenzia delle entrate. Il decreto istitutivo delle Agenzie fiscali (art.64, comma 3 e 3 bis, del decreto legislativo del 30/07/1999 n. 300 - Riforma dell'organizzazione del Governo - modificato da con Decreto-legge del 02/03/2012 n. 16 e con Decreto-legge del 06/07/2012 n. 95) attribuisce quindi all'Agenzia delle entrate la competenza sulla gestione dell'Osservatorio del Mercato Immobiliare.

² Cfr. G. Guerrieri *Lo sviluppo dell'Osservatorio del Mercato Immobiliare (OMI): un po' di storia*, in questo numero dei Quaderni, §§ 2 e 3.

³ *Ibidem*, § 3.

- costituire adeguato supporto nella prospettiva di una radicale riforma del catasto;
- fornire adeguato supporto alle attività interne di valutazione immobiliare;
- rappresentare un autorevole punto di riferimento quale fonte informativa dei fenomeni socio-economici connessi al patrimonio e al mercato immobiliare.

Se il primo punto non ha ancora oggi trovato attuazione⁴, gli altri due sono diventati il fulcro dell'intera attività dell'Osservatorio e nell'ambito di essi sono stati definiti processi e regole per i servizi erogati e i prodotti concepiti, sono state sviluppate procedure informatiche, progettate nuove attività e tracciate le linee di evoluzione. Il lavoro di costruzione, prima, e di sviluppo e affinamento, poi, dell'OMI è stato dunque incentrato essenzialmente sulle fonti informative (le banche dati) e sui modelli di analisi statistica e di interpretazione dei dati.

Quanto alle fonti informative si tratta, come si è già detto, di banche dati amministrative, dunque strutturate in funzione di obiettivi molto diversi da quelli di carattere statistico. Da un lato, dunque, si è manifestata l'esigenza di migliorare la qualità dei dati di origine, obiettivo non facile e, in ogni caso, perseguibile solo in presenza di una crescente consapevolezza dell'importanza della statistica nel buon funzionamento del settore; dall'altro, si è trattato di elaborare modelli interpretativi in grado di dedurre, dai dati disponibili, il maggior numero possibile di informazioni utili alla comprensione dei fenomeni socio-economici esplorati.

Nei primi vent'anni di questo lungo tragitto ci sono stati alcuni passaggi particolarmente significativi che di seguito, in sintesi cronologica, si ripercorrono.

- **2003**: è stata presentata e pubblicata la prima edizione del *Rapporto Immobiliare* annuale (dal 2009 suddiviso in due Rapporti annuali dedicati al settore residenziale⁵ e agli immobili a destinazione terziaria, commerciale e produttiva⁶), dove per la prima volta il mercato immobiliare italiano veniva analizzato, per dimensione, composizione e dinamica, attraverso i dati estratti dagli archivi amministrativi del Catasto e della Pubblicità immobiliare, riconoscendo a tali dati la validità informativa che tuttora si approfondisce e si amplia.
- **2006**: è stato introdotto⁷ il regime fiscale del cosiddetto "prezzo-valore", cioè la disgiunzione della base imponibile dal prezzo dichiarato di una compravendita, misura che ha trasformato un dato privo di valenza statistica (un prezzo dichiarato di convenienza fiscale, appunto) in un'informazione preziosa (il corrispettivo pagato nella compravendita) per la comprensione delle dinamiche reali di mercato. È seguita un'innovazione altrettanto rilevante, introdotta nella legge di conversione del c.d. decreto Bersani: per tutte le altre fattispecie di cessioni a titolo oneroso, non disciplinate con il prezzo-valore (e cioè, principalmente ma non solo, quelle aventi ad oggetto unità immobiliari non abitative), il legislatore torna al valore venale *tout court*, eliminando di fatto il riferimento, anche ai soli fini del controllo, ai valori fondati sulle rendite catastali⁸.
- **2009**: fu presentato e pubblicato il primo volume di Immobili in Italia (giunto oggi alla sua settima edizione), e con esso si sono gettate le basi per la costituzione di una banca dati integrata contenente informa-

⁴ Cf. G. Guerrieri *Lo sviluppo ...*, op.cit. § 3, in particolare nota 9.

⁵ Dal 2011 il Rapporto è stato pubblicato in collaborazione con l'Associazione Bancaria Italiana (ABI), che cura l'*Housing Affordability Index*.

⁶ Fin dal 2009 il Rapporto è stato realizzato con la collaborazione di Assilea (Associazione nazionale delle società di leasing).

⁷ Con l'art. 1, comma 497 della l. 23 dicembre 2005, n. 266 (finanziaria 2006) entra in vigore il cosiddetto regime fiscale "prezzo-valore" che introduce, a partire dal primo gennaio 2006, per le sole cessioni nei confronti di persone fisiche che non agiscano nell'esercizio di attività commerciali, artistiche o professionali, aventi ad oggetto immobili ad uso abitativo e relative pertinenze, la facoltà di determinare l'imponibile ai fini dell'imposta di registro sulla base del valore catastale (valore), indipendentemente dal corrispettivo pattuito e indicato nell'atto (prezzo).

⁸ Infatti l'art. 35, comma 23-ter, del decreto legge 223/2006, convertito con modificazioni nella Legge 248/2006, ha inserito in calce all'art. 52, D.P.R. 131/1986 il comma 5-bis ai sensi del quale «le disposizioni dei commi 4 e 5 [che escludono dall'accertamento il valore o il corrispettivo dichiarato se calcolato sulla base delle rendite catastali] non si applicano relativamente alle cessioni di immobili e relative pertinenze diverse da quelle disciplinate dall'art. 1, comma 497, della legge 23 dicembre 2005 n. 266, e successive modificazioni».

zioni sulle principali caratteristiche degli immobili, sul loro effettivo utilizzo e sulle variabili fiscali e patrimoniali dei proprietari, con il fine di colmare il gap informativo degli archivi disponibili. Il volume, realizzato dal Dipartimento delle Finanze e dall'OMI con il contributo del partner tecnologico SOGEI, è incentrato sull'analisi delle informazioni contenute nella banca dati integrata del patrimonio immobiliare⁹ (gestita dal Dipartimento delle Finanze e progettata assieme all'OMI) e permette, tra l'altro, di comprendere i fenomeni sociali, economici e culturali legati a una delle componenti principali della ricchezza degli italiani, famiglie e imprese, ovvero gli immobili e, in particolare, le abitazioni; obiettivo del volume è fornire a studiosi, addetti ai lavori e *policy-maker* uno strumento indispensabile per comprendere appieno come questo patrimonio viene utilizzato, quale valore esprime, quali differenziali geografici lo caratterizzano. Il volume ha inoltre incluso, nelle diverse edizioni finora pubblicate, approfondimenti tematici, quali quello sulla tassazione immobiliare, con confronti anche di respiro internazionale, sul gettito e sul *tax gap* delle imposte immobiliari.

- **2012:** l'Istat, nell'ambito di un progetto europeo mirato alla produzione di indici dei prezzi degli immobili residenziali, ha avviato la pubblicazione dell'indice dei prezzi delle abitazioni (IPAB) acquistate dalle famiglie sia per fini abitativi sia per investimento; alla sua realizzazione ha concorso l'OMI, collaborando con l'ISTAT per implementare la "*base dati delle compravendite immobiliari*", ottenuta dall'incrocio degli archivi del Catasto, della Pubblicità immobiliare, del Registro e dell'OMI, contenente le informazioni necessarie sia alla costruzione del primo indice dei prezzi delle abitazioni in Italia sia allo sviluppo e al miglioramento qualitativo delle statistiche fornite dall'OMI.
- **2015:** l'OMI ha portato a termine un nuovo processo di produzione dei dati statistici sul mercato immobiliare basata sulla disponibilità di un'evoluta piattaforma di *predictive analytics* e *data science* che opera direttamente sugli archivi dei microdati, consentendo l'esplorazione sempre più di dettaglio dei fenomeni connessi al patrimonio e al mercato immobiliare; questo importante sviluppo, da un lato, ha permesso di superare, riducendo la filiera di lavorazione, i limiti di macchinosità nella definizione e nello sviluppo di elementi innovativi e le inevitabili difficoltà nella individuazione delle potenzialità dei contenuti informativi degli archivi; dall'altro, ha consentito di indagare a fondo le informazioni originarie disponibili, di rilevarne criticità e potenzialità inesprese, di verificare possibili nuovi incroci e relazioni e di definire algoritmi di selezione e riorganizzazione delle informazioni più appropriati ed efficaci.
- **2016:**
 - o sono state diffuse, nell'ambito del Rapporto immobiliare – settore residenziale, le statistiche afferenti le locazioni di abitazioni a seguito dell'implementazione della *base dati delle locazioni immobiliari*, contenente le informazioni sugli immobili oggetto di nuova locazione; tale base dati è il risultato dell'integrazione tra l'archivio dei modelli presentati per la registrazione dei contratti di locazione, provenienti dall'archivio del registro, nei quali viene specificato l'oggetto della locazione e sono contenute le informazioni che riguardano tipologia, durata e valore economico del negozio, gli archivi catastali e la banca dati OMI;
 - o è stato approvato il D.lgs. 72/2016 che introduce nel Testo Unico Bancario (D.lgs. 1° settembre 1993, n. 385), il Capo I bis e, in questo, l'articolo 120-sexiesdecies che assegna all'Osservatorio del Mercato Immobiliare istituito presso l'Agenzia delle entrate il ruolo del *controllo statistico sul mercato immobiliare residenziale* e il compito di *effettuare le opportune comunicazioni ai fini dei controlli di vigilanza ma-*

⁹ Tale banca dati ha natura statistica e non amministrativa in quanto è formata da una estrazione, con correzione e completamento dei dati in base ad una serie piuttosto ampia di ipotesi, delle informazioni tratte dalla banca dati del Catasto edilizio urbano, dalla banca dati delle quotazioni OMI, dalle dichiarazioni dei redditi (UNICO e 730) e dai CUD, dalle banche dati dei versamenti delle imposte sulla proprietà degli immobili (ICI, IMU e TASI), dalla banca dati del registro.

cro-prudenziale. Da qui ha avuto inizio una costante collaborazione con il Servizio stabilità finanziaria del Dipartimento economia e statistica della Banca d'Italia per la definizione di specifici indicatori finalizzati al monitoraggio del mercato immobiliare residenziale, seguendo i dettati della raccomandazione¹⁰ dell'*European Systemic Risk Board* (ESRB), utili al controllo dei rischi connessi alle dinamiche del mercato stesso.

- **2017:** è stato operato un mutamento sostanziale nelle procedure di estrazione, selezione ed elaborazione dei dati relativamente al numero e alla superficie di unità immobiliari compravendute.
- **2018:**
 - o è stata attivata, con provvedimento del Direttore dell'Agenzia, la *diffusione gratuita sui servizi telematici* Fisconline ed Entratel dei dati delle compravendite immobiliari nel dettaglio comunale, con l'obiettivo di migliorare la fruibilità dei dati a beneficio sia della ricerca scientifica sia della trasparenza e concorrenza del mercato¹¹.
 - o si è aggiunto alla produzione statistica dell'OMI il *Rapporto Mutui Ipotecari*, ulteriore tassello nel panorama dell'informazione economica diffusa dall'Osservatorio, che conclude un lavoro pluriennale di progettazione e analisi sugli archivi delle iscrizioni ipotecarie, gestite nell'ambito della pubblicità immobiliare, con l'obiettivo di analizzare, traducendo a fini statistici le informazioni disponibili, il complesso dei finanziamenti concessi a fronte di garanzia ipotecaria su immobili e di fornire informazioni sulla destinazione di utilizzo di tali finanziamenti.
- **2020-2021:** dal 7 dicembre 2020 è attivo il nuovo servizio "*Consultazione dei valori dichiarati nelle compravendite immobiliari*", che rappresenta un ultimo contributo, in ordine di tempo, al quadro informativo sugli aspetti dei valori economici del mercato, rendendo disponibile la consultazione di dati puntuali sui prezzi di compravendita dei beni immobili nelle singole transazioni.

Le diverse tappe dello sviluppo sopra sinteticamente rappresentate hanno comportato un progressivo ampliamento e perfezionamento delle basi dati, che si è tradotto in una ricca offerta editoriale concentrata sui principali temi del patrimonio e del mercato immobiliare (*stock*, compravendite, locazioni, mutui ipotecari) e nella produzione di un'ampia informazione statistico-economica, indipendente e liberamente accessibile, al servizio del Paese.

2 I recenti sviluppi delle analisi statistiche e l'OMI

Le tendenze in atto negli ultimi anni in ambito statistico seguono essenzialmente tre direttrici.

La prima è lo sfruttamento sempre più diffuso degli archivi amministrativi, cioè dei dati raccolti per finalità non statistiche, i quali stanno progressivamente affiancando alcune indagini più tradizionali, penalizzate da una serie di fattori: i costi molto elevati, sia per gli istituti di statistica che per la collettività, la complessità organizzativa, i tempi lunghi di elaborazione nonché i limiti rispetto all'analisi longitudinale.

La seconda è l'irruzione dei cosiddetti *Big Data*, cioè di quei dati caratterizzati da un'incomprimibile combinazione di volume, velocità e varietà, alla cui base vi è la digitalizzazione ogni giorno più pervasiva, sia degli oggetti sia delle relazioni, e da cui scaturiscono nuove potenziali fonti informative, non sempre agevoli da gestire e da valorizzare.

¹⁰ La raccomandazione dell'*European Systemic Risk Board* (ESRB) del 31.10.2016, aggiornata il 21.03.2019, definisce una serie di indicatori (sezione 1 – raccomandazione A e B) che si riferiscono in particolare agli acquisti di immobili residenziali da parte di famiglie "consumatrici" effettuati con l'ausilio di mutui ipotecari.

¹¹ Sugli stessi canali telematici Entratel e Fisconline a partire dal 1° semestre del 2016, le quotazioni di mercato dell'OMI e le zonizzazioni comunali sono fruibili gratuitamente in formato elaborabile (file .csv e .kml) all'utenza, operatori del settore e semplici cittadini, per l'intero territorio Nazionale, per una specifica Area territoriale, per una Regione, una Provincia o un Comune.

La terza, che attiene alla fase di diffusione, ma di riflesso anche alla produzione, grazie alle possibilità di integrazione che offre, è la liberalizzazione dei dati nota sotto il nome di *Open Data*, con cui si intende la disponibilità da parte dei titolari a condividere i dati in loro possesso in forma libera, gratuita e, soprattutto, riutilizzabile. Sullo sfondo di questo scenario si pone la questione della riservatezza dei dati personali, cui la società e il legislatore prestano un'attenzione crescente e le cui esigenze di tutela rischiano di entrare in aperto conflitto rispetto ad alcune possibili evoluzioni di questi fenomeni.

La prima di queste tendenze è senza dubbio quella più risalente nel tempo. L'uso degli archivi amministrativi a scopi statistici è una pratica di lunga tradizione in molte parti del mondo, soprattutto in Nord Europa, ma anche in Canada e negli Stati Uniti. La novità recente sta nel fatto che in questi ultimi anni il fenomeno ha conosciuto una forte accelerazione. Tante sono le ragioni che hanno innescato la transizione verso i dati di fonte amministrativa, prima tra tutti l'innovazione tecnologica, che, sia pure con ritardo, ha investito anche la sfera delle attività pubbliche, comportando un sostanziale miglioramento nella qualità e nella accessibilità degli archivi.

In Europa una spinta importante in questa direzione è venuta dal regolamento statistico (Reg CE n.223/2009), che dedica l'art.24 alle esigenze di accesso ai dati amministrativi ai fini "dello sviluppo, della produzione e della diffusione di statistiche europee". Come si è già accennato, alla base di questa scelta strategica c'è un favorevole rapporto costi-benefici, che in realtà può variare significativamente a seconda del tipo e del campo di indagine. In termini molto generali si può dire che gli effetti positivi nell'utilizzo statistico dei dati amministrativi sono i seguenti: semplificazione e riduzione degli adempimenti di cittadini e imprese; contenimento dei costi per la funzione statistica; aumento dell'accuratezza delle stime e incremento della capacità informativa dell'*output*. Riguardo agli ultimi due punti, tanti sono i fattori che concorrono. Attraverso gli archivi amministrativi possono infatti rendersi accessibili nuove variabili, spesso relative a fenomeni non ancora esplorati, e solitamente si riesce ad avere una maggiore copertura della popolazione d'interesse, nonché un più elevato dettaglio territoriale e settoriale. La fonte amministrativa permette inoltre, almeno a volte, di incrementare la frequenza e la profondità temporale dei dati e, soprattutto, dà, in molti casi, la possibilità ai ricercatori di seguire le stesse unità nel tempo, il che offre occasioni di studio molto preziose. Infine, non bisogna trascurare che il dato amministrativo ha una serie di impieghi derivati molto importanti in statistica come, ad esempio, nei procedimenti di stima indiretta, nella costituzione di liste per la conduzione di indagini statistiche o nella validazione dei dati raccolti attraverso queste ultime¹².

D'altra parte, si deve porre particolare attenzione nell'utilizzare i dati amministrativi, in quanto vi è una serie di limitazioni di cui occorre essere consapevoli e che include: l'assenza o la scarsità di controllo sulla qualità dei dati, la possibilità di avere variabili o record mancanti, le possibili differenze nella definizione dei concetti che potrebbero portare a problemi di distorsione, nonché di copertura, la tempestività dei dati (essendo la raccolta dei dati fuori dal controllo dell'istituto di statistica, è possibile che, a causa di eventi esterni, parte o tutti i dati non vengano ricevuti in tempo) e infine alcuni costi associati con i dati amministrativi, ad esempio quelli relativi ai sistemi informatici eventualmente necessari per pulire e integrare i dati stessi.

Per rendere più tangibile la difficoltà nell'utilizzo di queste informazioni si possono fare alcuni esempi concreti relativi all'archivio dei contratti di locazione impiegato dall'OMI per la produzione di alcune statistiche e in cui confluiscono i dati di varie fonti amministrative: il registro (presso il quale i contribuenti devono assolvere gli adempimenti fiscali legati alla stipula di un nuovo contratto e che restituisce dati sulle parti e sull'oggetto della transazione), il catasto (che fornisce informazioni sulla natura, sulla rendita, sulla consistenza del bene locato), la banca dati OMI (che è un archivio derivato con le quotazioni e i canoni relativi alla zona omogenea di appartenenza dell'immobile) e, infine, le dichiarazioni (limitatamente al dato relativo all'utilizzo del bene come abita-

¹² R. Monducci, *Dati amministrativi per le analisi socio-economiche e la valutazione delle politiche pubbliche*, Convegno Inps-Siep, 2018.

zione principale). Il primo problema riguarda la copertura della popolazione di interesse (non essendovi obbligo di registrazione per contratti di durata inferiore ai 30 giorni ed essendo rilevante il fenomeno dell'evasione totale¹³). Il secondo, comune alle compravendite, è che l'unità di rilevazione non è l'unità statistica oggetto di analisi: le registrazioni e le trascrizioni che alimentano gli archivi si riferiscono a contratti che possono contenere una pluralità di immobili, con la conseguenza che il trattamento di alcune variabili non differenziate all'origine (in particolare il prezzo o il canone) costituisce una seria criticità. Vi è poi la questione della nomenclatura, nel senso che le definizioni utilizzate in ambito amministrativo possono non coincidere con quelle che servono alla statistica: nello specifico esiste uno iato tra la tipologia fiscale dei contratti così come classificata nei modelli di registrazione e quella civilistica prevista dalle norme di diritto privato e su cui si vorrebbero riportare le variabili rilevate. Non è solo la complessità della legislazione civilista e fiscale, ma anche la sua instabilità a complicare il lavoro dello statistico. Quest'ultima caratteristica, infatti, solleva un ulteriore problema, che è quello della discontinuità delle serie temporali, peraltro legata a modifiche non solo normative ma anche puramente amministrative (in questo senso le locazioni costituiscono un caso di scuola, stante la molteplicità di modelli dichiarativi susseguirsi nel tempo). Un altro tema concerne la tempestività, cioè il rischio della obsolescenza dei dati, i quali accumulano inevitabilmente una serie di ritardi prima di poter essere elaborati e diffusi: quello tra la stipula e la registrazione, quello dei tempi tecnici legati al controllo sulla liquidazione da parte degli uffici e di fornitura e alle verifiche di qualità e alla gestione dei flussi da parte dei sistemi informatici, senza dimenticare che per alcune variabili i dati si formano molto prima (nel corso della trattativa tra le parti) rispetto a quando appaiono poi nel contratto. L'aspetto più problematico riguarda, infine, la qualità di alcuni dati che sono di grande interesse per lo statistico, ma che non essendo, al contrario, rilevanti nel procedimento amministrativo non ricevono la stessa attenzione rispetto ad altri (in generale tutti i dati che non hanno incidenza sul regime fiscale). Ciò rende necessaria un'attività più o meno intensa di *data cleaning*¹⁴ e persiste il rischio di compromettere la validità di alcuni risultati.

Grande attenzione viene posta, in ogni caso, a questa fase: a livello europeo si è sviluppato un *framework* di riferimento per la valutazione della qualità dei dati che riguarda tredici dimensioni riferibili a tre grandi ambiti (la fonte, i metadati e le caratteristiche intrinseche dei dati).

Tornando alle considerazioni generali, è evidente che l'introduzione di queste fonti debba essere graduale e non implichi la definitiva scomparsa delle indagini campionarie, anche perché alcune dimensioni, soprattutto nell'ambito della statistica sociale, possono essere indagate solo con il contatto diretto presso le stesse unità di analisi (famiglie o imprese). La direzione nella quale sono orientati gli istituti di statistica, almeno a livello di sistema, non è infatti la sostituzione tra le fonti ma la loro integrazione, nel tentativo di trovare le soluzioni più efficienti che consentano di cumulare i vantaggi di ciascuna risorsa superandone i limiti. I censimenti sono il banco di prova di questa trasformazione, che si è in parte già realizzata. Dal 2018 i censimenti sono infatti diventati permanenti e non coinvolgono più la totalità delle famiglie bensì solo un campione selezionato di esse, e grazie all'integrazione dei dati raccolti con quelli provenienti dalle fonti amministrative consente di restituire a cadenza annuale informazioni rappresentative dell'intera popolazione. L'ulteriore estensione di questo approccio richiede una revisione profonda nei processi di produzione delle varie statistiche in termini di piena integrazione logico-fisica dei microdati di varia provenienza, con cambiamenti conseguenti nei sistemi informativi e organizzativi che necessitano a loro volta di orizzonti temporali ampi e di una serie di pre-condizioni tra cui gli investimenti in tecnologia e in formazione. Un ruolo sempre più importante rivestono, inoltre, gli uffici statistici

¹³ Sul tema: ISTAT – *L'economia non osservata nei conti nazionali - Anni 2015-2018*, pubblicazione ISTAT 2020

¹⁴ *Data cleaning* è quel processo alla base di qualsiasi operazione di *data mining* (estrazione di informazioni dai dati) volto a garantire, con una certa soglia di affidabilità, la correttezza di una base dati.

interni alle amministrazioni grazie alla conoscenza specialistica dei metadati e del contesto di riferimento, presupposti essenziali per la trasformazione dei dati amministrativi in informazione statistica.

Su questo fronte, e in particolare su un'integrazione sempre più stretta dei vari archivi interni ma anche esterni, l'OMI è impegnato da sempre, avendo fin dalle origini come missione istituzionale la valorizzazione statistica degli archivi amministrativi detenuti dall'Agenzia delle entrate.

La seconda grande tendenza in ambito statistico è forse anche quella più dibattuta e riguarda la digitalizzazione sempre più penetrante delle società moderne, vale a dire la produzione e accumulazione a ritmi impetuosi di dati che, seguendo la classificazione proposta dall'UNECE¹⁵, hanno essenzialmente tre origini: quelli generati direttamente dagli individui e dalle loro relazioni (*human-generated data* quali, ad esempio, i dati da *social media*, *blog*, *sms*, *e-mail*, ecc.); quelli che scaturiscono attraverso attività, per lo più economiche, di vario genere (*process-mediated data* quali sistemi transazionali, commerciali e bancari tradizionali, *e-commerce*, carte di credito) e infine quelli in cui non vi è intervento dell'uomo, cioè i *machine-generated data*, tipicamente ciò che ora va sotto il nome di *Internet of things*, come sensori fissi (*home-automation*, sensori ambientali/meteorologici, sistemi per il controllo del traffico, ecc.) e mobili (dispositivi mobili, sensori su automezzi, immagini satellitari). La complessità, le dimensioni e la frequenza di questi dati ha portato a coniare il termine, non sempre utilizzato a proposito, di *Big Data*, che andrebbe riferito soltanto a quei dati che soddisfano almeno uno di tre specifici criteri, ovvero volume elevato (più di 50 *Terabyte* o crescita annua maggiore del 50%), velocità (raccolta e analisi dei dati *real-time*) e varietà (fonti eterogenei e dati non solo strutturati), e che, in quanto tali, necessitano, per essere raccolti, verificati, elaborati e diffusi, di soluzioni informatiche non convenzionali, diverse cioè dai tradizionali *database* di tipo relazionale. Molte sono le sfide che questi dati pongono a chi si occupa di statistica, in quanto possono sommare uno o più dei seguenti fattori critici che nelle tradizionali fonti statistiche o amministrative risultano maggiormente sotto controllo: l'alto grado di eterogeneità, la non indipendenza o identica distribuzione, la mancanza di una struttura predefinita, la misurazione attraverso dispositivi non uniformi, l'interferenza dei "rumori", la provenienza non da esperimenti progettati ma da una miscela di molte diverse popolazioni, la presenza non sempre evidente di osservazioni mancanti e distorte. Da qui la necessità di nuovi approcci ai concetti teorici e di sviluppi metodologici, oltre al riconoscimento della essenziale attività di pre-elaborazione dei dati¹⁶.

Finora nel trattamento dei *Big Data*, peraltro spesso detenuti da soggetti privati, hanno prevalso gli aspetti pratici e commerciali rispetto a quelli scientifici e sono stati privilegiati i profili informatici e computazionali rispetto a quelli più prettamente statistici. Non esiste pertanto ancora una cornice teorica e metodologica ben definita. I paradigmi inferenziali sono in evoluzione verso modelli statistici non parametrici e non lineari (come la *spline regression*), verso approcci di stima locale (come i *classification trees*) e si è rilanciata una serie di tecniche pensate per ridurre la complessità delle informazioni come il *clustering* o la *PCA* (*principal component analysis*). Inoltre, ancor più che nell'uso degli archivi amministrativi, occorre verificare le varie dimensioni qualitative dei dati, in particolare dal punto di vista della rappresentatività, fondamentale per poter produrre stime affidabili, e della accuratezza, intesa come qualità intrinseca dei valori acquisiti. Rispetto alle potenzialità che molti riconoscono ai *Big Data*, i benefici effettivi apportati sono ad oggi piuttosto contenuti, almeno in ambito di statistiche ufficiali. Quella attuale è infatti una fase di apprendimento, in cui si cerca di "familiarizzare" con i dati e di collaudare tecniche adeguate, soprattutto sul fronte, ancora embrionale, dei modelli basati su algoritmi (*machine learning*¹⁷).

¹⁵ UNECE, *Conference of European statisticians: What does "Big Data" mean for official statistics?*, 2013

¹⁶ Società italiana di Statistica, *Statement della SIS su Statistica, Scienza dei dati e Big data*, 2018

¹⁷ P. Secchi, *On the role of statistics in the era of big data: a call for a debate*. *Statistics & Probability Letters*, volume 136, 2018; p. 10–14.

Sarebbe a rigore improprio parlare di *Big Data* con riferimento all'attività dell'Osservatorio, anche se con il paesaggio ai microdati già descritto nel primo paragrafo si è sperimentato un momento di forte discontinuità, che ha posto l'OMI a contatto diretto con problemi di volume, velocità e varietà, sollecitando l'adozione di nuove tecniche e strumenti con cui esplorare e analizzare i dati. È indubbio, inoltre, che almeno nel panorama nazionale, non vi siano basi di dati sul mercato e sul patrimonio immobiliare paragonabili, per ricchezza e complessità, a quelle gestite dall'Osservatorio.

Infine, l'altro grande fenomeno che, sul fronte della divulgazione del dato statistico, sta caratterizzando l'attuale epoca riguarda gli *Open Data*, che potremmo definire come quei dati che i titolari, spontaneamente o in adempimento a obblighi normativi, mettono a disposizione della comunità e che possono essere utilizzati da chiunque senza licenza, a condizione di citarne la fonte e permetterne il riutilizzo¹⁸. Esistono diversi gradi di apertura dei dati, che vengono misurati attraverso una scala proposta nel 2009 da Tim Berners-Lee, creatore del *world wide web*. Il punteggio massimo è di cinque stelle e viene assegnato quando i dati (1) sono disponibili su web con una licenza aperta, (2) in formato strutturato, (3) attraverso un tipo di file non-proprietario, (4) vengono identificati con gli standard del W3C (RDF and SPARQL) per identificare i dati (5) e sono collegati con altri dati forniti da differenti soggetti per collocarli in un contesto più ampio. L'accresciuta sensibilità, sia in ambito pubblico che privato, per la trasparenza ha notevolmente incrementato la "liberalizzazione" dei dati nel corso degli ultimi decenni. Questa tendenza costituisce un tassello essenziale di una doppia prospettiva auspicata da molti: l'*open knowledge*, cioè la conoscenza condivisa, con le ricadute che ciò comporta sull'innovazione e sullo sviluppo economico, e l'*open government*, o amministrazione aperta, cioè una modalità di esercizio del potere tesa a favorire la partecipazione dal basso e il rafforzamento dei meccanismi democratici. Anche in questo caso il processo è tutt'altro che maturo e si può ben dire che quella in corso è ancora una fase di sperimentazione¹⁹. L'interoperabilità tra i dati è mediamente bassa e le iniziative di pubblicazione procedono in ordine sparso, per cui, nonostante lo sforzo dei portali deputati alla raccolta e alla catalogazione, i dati risultano allo stato attuale scarsamente fruibili. Esiste anche una certa resistenza culturale, non sempre i vertici organizzativi mostrano sensibilità su questi temi, così come la conoscenza da parte degli utenti non è ancora molto diffusa.

L'Osservatorio da anni mette a disposizione, in formato elaborabile, alcuni dei dati che sono alla base dei report trimestrali e dei rapporti annuali. Un ulteriore significativo passo in questa direzione è stato compiuto nelle due ultime edizioni del volume *Immobili in Italia*, in cui la versione cartacea è stata affiancata da una pagina web, curata dal Dipartimento delle finanze, con dati navigabili e scaricabili.

Nell'ambito degli *Open Data* un contributo specifico può giungere dal già citato servizio per la "Consultazione dei valori dichiarati nelle compravendite immobiliari", attivo dal 7 dicembre 2020.

3 *Open Data* in campo immobiliare: le esperienze di Gran Bretagna e Francia

Come descritto più in dettaglio in un precedente articolo²⁰ al quale si rimanda, Inghilterra e Galles sono stati tra i primi Paesi a intraprendere la strada degli *Open Data*. Nel 2010 venne svolta una pubblica consultazione per capire quali dati suscitassero maggior interesse e come fosse possibile valorizzarli al fine di massimizzare il benessere socio-economico nel lungo termine. Tra i vari *dataset* da "aprire" al pubblico fu individuato quello relativo ai prezzi di compravendita registrati presso il *Her Majesty Land Registry* (HMLR), un organismo pub-

¹⁸ Sul tema si possono citare alcuni interessanti studi commissionati dall'Unione Europea come ad esempio "Creating value through Open Data" oltre il primo monitoraggio svolto in Italia sul riutilizzo dei dati di fonte pubblica già disponibili, "Open data 200: Italy"

¹⁹ S. De Francisci *Open Data nella statistica ufficiale: ruolo, opportunità e implicazioni* in Convegno 3a Giornata della trasparenza in Istat, 2019

²⁰ M. Festa e S. Serafini, *La trasparenza del mercato immobiliare nell'era dei big data*, Quaderni dell'Osservatorio, 2017.

blico istituito nel 1862 che ha la responsabilità delle funzioni catastali e di pubblicità immobiliare. Nel marzo 2012 è cominciata la pubblicazione con cadenza mensile dei prezzi relativi alle compravendite oggetto di nuove registrazioni, mentre in due fasi successive è stato reso disponibile l'archivio storico prima fino al 2009 poi fino al 1995. Nel 2013 i dati hanno assunto la forma di *Open Data* a cinque stelle, mentre nel 2014, con l'aiuto di società esterne, è stato approntato uno strumento per generare report su misura selezionando tra vari possibili parametri. Infine nel 2015 è stato aggiunto l'archivio *additional price paid data*, che si riferisce ad alcune tipologie particolari di compravendite inizialmente escluse dalla pubblicazione. I dati forniti per ogni transazione includono gli identificativi catastali, l'indirizzo, il codice postale, il prezzo effettivo, la data di compravendita nonché alcune informazioni di base sulle caratteristiche dell'immobile, ad esempio se è di nuova costruzione, la tipologia cui appartiene (monofamiliare, bifamiliare, appartamento, altro), il diritto che è stato acquisito, se piena proprietà (*freehold*) o diritto di superficie (*leasehold*), quest'ultimo molto diffuso in Inghilterra.

Per ottenere ulteriori informazioni, occorre versare le tariffe previste per la consultazione delle note di trascrizione e dei relativi titoli. Sono escluse le compravendite parziali (cioè di parti o quote di proprietà), le compravendite di diritti diversi dal diritto di proprietà o dal diritto di superficie ultrasettennale, le compravendite in esecuzione di sentenze, le successioni, le donazioni e gli altri trasferimenti a titolo gratuito, le compravendite di più immobili intesi come parti di un portafoglio e infine i cosiddetti *right to buy*, cioè lo schema giuridico con cui gli inquilini di immobili pubblici possono diventare proprietari a condizioni agevolate. In un archivio a parte, l'*additional price paid dataset*, sono state successivamente inserite alcune tipologie di transazioni in un primo tempo escluse quali le compravendite a favore di persone non fisiche, quelle a scopo di affitto indeterminate tramite ipoteca (*buy to let mortgage*), nonché le compravendite assistite da mutuo e con potere di vendita da parte del creditore ipotecario (*power of sale*).

Va precisato che il dato fin dal 1990 era consultabile, sia pure per singola transazione, pagando le tariffe previste per l'accesso ai pubblici registri dell'HMLR. Inoltre a partire dal 2005 gli stessi dati venivano forniti anche in blocco, sempre dietro corrispettivo, per lo più ad utenti professionali ed erano pertanto nel tempo confluiti su portali privati. In occasione dell'attivazione di questo servizio a pagamento era stata consultata l'*Authority* competente, la quale aveva ritenuto che il dato in questione non fosse, almeno direttamente, un dato personale o biografico in quanto riferito in primo luogo all'oggetto immobile. La valutazione interna di impatto effettuata alla vigilia della pubblicazione confermò tale posizione, paventando, tuttavia, alcuni rischi in termini di possibile incremento del *direct marketing*, formulando una serie di raccomandazioni in gran parte accolte nella versione finale del progetto, tra le quali:

- rendere ben esplicite le fattispecie di compravendite non inserite tra i dati da pubblicare,
- mettere in evidenza il periodo di riferimento dei dati,
- illustrare chiaramente la natura non personale dei dati, prevedere una procedura specifica per presentare reclami e per segnalare errori,
- ripetere la valutazione d'impatto a distanza di mesi.

Oltre a favorire quegli utenti che prima pagavano un prezzo di accesso per singola consultazione, la pubblicazione in formato *open* ha stimolato nuove iniziative imprenditoriali²¹. Sono nate diverse *start-up* nel mondo immobiliare. Un esempio, tra i tanti, è *GetAgent*, sito che valuta l'efficienza degli intermediari utilizzando i prezzi di offerta, i prezzi effettivi di compravendita e i tempi impiegati per la conclusione dell'affare. Lo stesso HMLR tiene eventi con la comunità *Open Data* per monitorare queste iniziative e incoraggiarne di ulteriori. Il *feedback* ricevuto dagli imprenditori è molto positivo, anche se alcuni auspicano un'ulteriore apertura, con la messa a disposizione di maggiori caratteristiche sugli immobili. L'opinione condivisa è che questi dati acquisti-

²¹ <http://odimpact.org/case-united-kingdoms-hm-land-registry-price-paid-data.html>

no reale valore di *business* solo se combinati con altri dati pubblici e privati. Per quanto attiene agli effetti della maggiore trasparenza, nel contesto britannico ci si attende che la divulgazione dei prezzi effettivi possa aiutare a “lubrificare” il mercato; la previsione degli stessi intermediari è che i venditori siano indotti a fissare prezzi più realistici, mettendo al contempo gli acquirenti più a loro agio, per così dire, consentendo di arrivare a vendite più rapide a prezzi più equi. Dal punto di vista dell’HMRL la qualità dei dati, che è inevitabilmente legata alla qualità degli *input* forniti dalle parti o dai *solicitor*, dovrebbe avere un miglioramento grazie alle segnalazioni degli utenti e all’attenzione generale di cui ora sono oggetto.

Un’esperienza analoga, ma più recente, è in corso in Francia dove, a partire da aprile 2019, è attivo un servizio di consultazione dei prezzi dichiarati ai fini fiscali nei trasferimenti a titolo oneroso, sia in forma massiva che attraverso una mappa navigabile. La diffusione sotto forma di *Open Data* viene effettuata in attuazione di un nuovo articolo del *Livre des procédures fiscales* (LPF), introdotto nel 2018 dalla Legge organica di riforma amministrativa, finalizzata a creare uno “Stato al servizio di una società improntata alla fiducia” (ESSOC). Tale articolo prevede che “*le autorità fiscali devono mettere gratuitamente a disposizione del pubblico, in formato elettronico, le informazioni in loro possesso sui valori degli immobili dichiarati in occasione dei trasferimenti avvenuti negli ultimi cinque anni*”, al fine di “*contribuire alla trasparenza dei mercati fondiari e immobiliari*”. I dati vengono aggiornati su base semestrale e riguardano vendite, aste, espropriazioni e permutate di beni immobili provenienti dagli atti notarili e del catasto. Sono esclusi i dati relativi ai dipartimenti di Basso Reno, Alto Reno, Mosella e Mayotte, che hanno regole particolari di pubblicità immobiliare. Per ogni trasferimento, le informazioni rese disponibili riguardano data e natura giuridica del trasferimento, prezzo (che non include le spese di agenzia immobiliare, le spese notarili o l’eventuale valore dei beni mobili), indirizzo e una serie di dati catastali descrittivi dell’immobile. Le condizioni generali relative all’uso delle informazioni sono stabilite all’art 112 A-3 del LPF e prevedono che il trattamento non possa avere come oggetto o effetto di consentire la reidentificazione delle persone coinvolte e che questi dati non possano essere sottoposti a indicizzazione sui motori di ricerca online. Trattandosi di un obbligo legale, le persone interessate non possono opporsi alla diffusione dei dati che li riguardano, né è possibile per loro eliminarli, ma conservano il diritto di accesso e di rettifica. Esiste un’applicazione georeferenziata creata da Etalab, dipartimento del DINUM (*Direction interministérielle du numérique*, che coordina la politica di apertura e condivisione dei dati pubblici) e che rappresenta su mappa i dati opportunamente rielaborati, mentre i dati grezzi, anch’essi disponibili su web, provengono direttamente dalla Direzione generale delle finanze pubbliche.

I dati sui valori dichiarati erano precedentemente disponibili su *impots.gouv.fr*, ma richiedevano l’accesso attraverso un codice fiscale. Il governo francese ha inteso andare oltre, rendendo pubblico questo database. Gli auspici alla base di questa scelta sono elencati di seguito.

- Offrire una maggiore trasparenza sul funzionamento del mercato immobiliare e ridurre le asimmetrie informative tra le parti. Queste ultime, infatti, non sono più obbligate a rivolgersi a terze parti (agenzia immobiliare, notaio, ecc.) per ottenere queste informazioni.
- Limitare i trasferimenti di dati tra amministrazioni e comunità. Dato che le informazioni sono ora pubbliche, visibili e consultabili online da tutti, la direzione generale delle finanze pubbliche non deve più dedicare tempo all’invio del database DVF ad altre amministrazioni o comunità.
- Favorire l’innovazione. La pubblicazione di questi dati dovrebbe consentire alle società di riutilizzare i dati per creare nuovi prodotti e servizi e rendere così più rapide e fluide le transazioni.

4 La nuova frontiera dei dati elementari

Il tema della fornitura diretta di dati non elaborati è andato man mano imponendosi, come già accennato, anche nel contesto italiano; si tratta, evidentemente, di un “fronte” del tutto nuovo rispetto all’intero sviluppo che

l'Osservatorio ha conosciuto nel suo percorso evolutivo e che è stato sintetizzato nel primo paragrafo di quest'articolo.

Come si è visto, infatti, l'attività ormai pluridecennale dell'Osservatorio si è incentrata da sempre sull'analisi e sull'interpretazione statistica e scientifica dei dati, nell'intento di offrire un'informazione strutturata e in grado di aiutare la comprensione dei contesti e delle dinamiche reali.

Negli ultimi anni, tuttavia, si è percepita questa nuova esigenza, maturata essenzialmente in ambito giuridico e da lì rapidamente affermatasi, come sempre più spesso accade, quale ineludibile principio di carattere generale, e che si usa concettualizzare attraverso la locuzione "diritto alla trasparenza".

Attenzione, non si tratta di per sé di un concetto nuovo, perché la trasparenza è da sempre stata non soltanto un'esigenza avvertita dall'opinione pubblica nei confronti dell'amministrazione, ma soprattutto uno dei parametri di riferimento di una corretta azione amministrativa; lo stesso Osservatorio, come sottolineato ampiamente in precedenza, si è da sempre posto, tra gli obiettivi di fondo, quello di fornire un contributo di trasparenza rispetto al mercato immobiliare.

L'elemento di novità, in questo caso, va individuato, piuttosto, nella nuova accezione che il concetto via via è andato assumendo.

In precedenza, infatti, la trasparenza si caratterizzava concretamente per una connotazione almeno duplice:

- da un lato, il diritto di accesso a dati e documenti di interesse per il cittadino;
- dall'altro, l'esplicitazione dei processi posti in atto dall'amministrazione.

Più che un obiettivo in sé, dunque, essa assumeva il carattere di uno strumento, finalizzato a garantire l'efficacia e il buon funzionamento della pubblica amministrazione nell'interesse dei cittadini amministrati.

L'Osservatorio, ad esempio, ha da sempre aderito in pieno a questo principio, esplicitando fonti dei dati e modelli di elaborazione all'interno delle cosiddette "Note Metodologiche"²², documenti tecnici di supporto alle pubblicazioni statistiche, utili per una loro corretta comprensione ed interpretazione.

L'influenza giuridica del modello anglosassone, recepito dalla disciplina europea, ha tuttavia progressivamente trasformato questa valenza strumentale del concetto di trasparenza in una valenza assoluta, non solo per l'estensione dei soggetti passivi interessati (dalle amministrazioni agli organi costituzionali, dalle istituzioni politiche alle società partecipate, ecc.), ma anche per la trasversalità dei campi di applicazione: *"...dal diritto di accesso ai documenti si è passati ... a una accessibilità totale, per promuovere la partecipazione e per favorire il controllo diffuso del potere pubblico"*, senza neanche il *"bisogno di indicare un interesse giuridicamente rilevante"*²³.

Come si vede, dunque, questo nuovo indirizzo tende a prevalere su ogni possibile contestualizzazione, come ad esempio valutazioni sull'impatto atteso da determinate scelte in uno specifico ambito o settore; l'unico limite posto dalla giurisprudenza è rappresentato, di fatto, dalla tutela della riservatezza dei dati personali.

In considerazione di ciò, l'Osservatorio ha ritenuto di dover sviluppare, come detto in precedenza²⁴, questo nuovo filone inerente la consultazione di alcuni dati originari non filtrati da elaborazioni statistiche, relativamente ai prezzi di compravendita degli immobili; si tratta di un'operazione abbastanza complessa, che ha richiesto una serie di approfondimenti e verifiche, soprattutto per preservare il citato, necessario equilibrio tra il diritto alla trasparenza e il diritto alla riservatezza.

²² Le Note metodologiche illustrano le procedure di estrazione, selezione ed elaborazione dei dati alla base delle statistiche sulle compravendite e sulle locazioni immobiliari e sui mutui ipotecari. Sono pubblicate alla pagina web:

<https://www.agenziaentrate.gov.it/portale/web/guest/schede/fabbricatiterreni/omi/pubblicazioni/note-metodologiche-2018>

²³ S. Cassese, *Evoluzione della normativa sulla trasparenza*, Sinapsi, 8, n. 1, pp. 5-7, 2018.

²⁴ Si ripete che il servizio di Consultazione dei valori immobiliari dichiarati è attivo dal 7 dicembre 2020.

Il servizio, fornito con accesso autenticato all'interno di un'apposita sezione del sito internet dell'Agenzia delle entrate, consente di accedere ad una mappa con grafi stradali in cui sono georiferite (posizionando un "segnalatore" denominato POI – *point of interest* - nel punto esatto del grafo stradale di una mappa) le singole transazioni relative ad un determinato periodo e, solo in presenza di un minimo di cinque transazioni, di consultare delle schede contenenti le informazioni sui prezzi dichiarati per le compravendite immobiliari - non georiferite puntualmente, ritenendo così di riuscire a garantire l'anonimizzazione dei soggetti coinvolti nel rispetto delle norme sulla *privacy*.

Tali informazioni si andranno ad affiancare così alle quotazioni, che rappresentano intervalli medi di valori per zona comunale.

La costruzione di un sistema di informazione caratterizzato dalla diffusione periodica dei micro dati afferenti ai prezzi di compravendita dei beni immobili, inciderebbe, con riferimento ai criteri di matrice anglosassone in uso in ambito internazionale, proprio su una delle componenti sulla base delle quali viene valutata la trasparenza del mercato immobiliare nazionale, costituita proprio dalla disponibilità dei *market fundamentals data*. Va citato poi il possibile impatto che si determinerà sugli strumenti delle valutazioni immobiliari, a partire dal mercato dei mutui ipotecari (mediante la valutazione dei tecnici alle banche sul valore del collaterale offerto in garanzia), fino alle consulenze degli esperti indipendenti dei fondi immobiliari eventualmente interessati ad effettuare investimenti.

La riorganizzazione delle banche dati e i nuovi strumenti di accesso e gestione hanno rappresentato l'indispensabile supporto per questo progetto; sarà ora di grande interesse valutare e gestire, necessariamente *ex post*, l'impatto che questa innovazione produrrà sul mercato immobiliare italiano, a partire dagli esempi appena citati.

5 Le prossime sfide per le statistiche OMI

Il lungo percorso seguito dall'OMI in questi primi vent'anni di attività è stato innegabilmente caratterizzato dalla costruzione di un sistema informativo strutturato al servizio della conoscenza del patrimonio e del mercato immobiliare. È stato un percorso complesso, quasi mai svolto con le risorse ottimali, frequentemente minato da criticità e ostacoli, che hanno spesso rallentato, ma non impedito, il raggiungimento di importanti traguardi. Il cammino ovviamente deve proseguire e possono delinearsi alcuni nuovi obiettivi che comporteranno un intenso programma di lavoro incentrato su evoluzioni e sviluppi orientati al miglioramento concreto e al consolidamento del contributo alla conoscenza e alla trasparenza di un mercato relevantissimo sul piano economico e sociale.

In concreto, si possono sintetizzare i nuovi obiettivi dell'OMI nei seguenti punti:

- ampliare il campo dell'informazione e abbreviare i tempi di diffusione
- migliorare la qualità delle basi dati
- valutare e gestire gli effetti relativi alle nuove modalità di fruizione delle informazioni (*Open Data*).

La prima è sicuramente la sfida più importante e riguarda i contenuti. Il campo di interesse deve essere ampliato per rispondere a nuove esigenze conoscitive. Per analizzare e comprendere tali esigenze si è rivelato fondamentale anche il dialogo costante intrapreso con l'utenza, come l'interazione assidua con la società scientifica e gli istituti di ricerca, pubblici e privati²⁵.

Diverse sono le necessità emerse e, prima tra tutte, quella di ampliare la produzione delle attuali statistiche al patrimonio dei terreni, agricoli ed edificabili, in termini sia di *stock*, sia di flusso delle compravendite. A tal fine

²⁵ In qualche misura i Forum, di cui si dà conto in questo numero dei Quaderni, hanno offerto un significativo contributo.

sta per essere concluso l'ampliamento della base dati delle compravendite immobiliari anche ai terreni, integrandolo con tutte le informazioni disponibili dagli archivi censuari e cartografici del catasto, che consentirà successivamente la pubblicazione di un primo *report* sul mercato dei terreni.

Ancora allo studio è invece la costituzione di una nuova base dati dello *stock* immobiliare che possa consentire analisi dei fenomeni legati alle attività edilizie (nuove costruzioni, ristrutturazioni, manutenzioni ordinarie e straordinarie) che intervengono sulla trasformazione del patrimonio immobiliare e sull'uso del suolo.

Inoltre l'OMI è già impegnato in diversi progetti (con ISTAT, Banca d'Italia e altri istituti di ricerca) volti ad arricchire l'informazione economica con nuovi aspetti indagabili nelle basi dati, quali ad esempio gli acquisti della prima casa, l'analisi della domanda e dell'offerta delle abitazioni, gli acquisti delle abitazioni per investimento (*buy to let*), l'analisi del *loan to value* nei mutui ipotecari per l'acquisto di abitazioni, la definizione di un indice dei prezzi degli immobili non residenziali, l'analisi della composizione ed evoluzione dei patrimoni delle famiglie, la ricostruzione dello *stock* delle abitazioni locate a partire dai flussi dei nuovi contratti e delle comunicazioni di proroga, risoluzione e cessazione.

Sul campo della riduzione dei tempi di diffusione delle statistiche, si sta procedendo, grazie alle nuove basi dati implementate, alla previsione di trimestralizzare alcune informazioni fino ad oggi disponibili solo su base annua, a partire dalle locazioni delle abitazioni e dai mutui ipotecari che assistono le compravendite residenziali. Non è inoltre escluso che alcune statistiche relative al mercato immobiliare possano avere, nel prossimo futuro, una diffusione con periodicità mensile, con riferimento quantomeno ai principali parametri descrittivi (NTN e STN).

Dal punto di vista dello statistico c'è poi un'ulteriore sfida, ad oggi, però, solo poco più che un auspicio: riguarda la qualità dei dati e si riferisce alla prima raccolta dei dati stessi. Si tratterebbe, in sintesi, di coniugare le esigenze dell'operatività amministrativa con quelle dell'informazione statistica mediante la partecipazione alla definizione e alla revisione dei modelli di acquisizione dei dati amministrativi (catasto, pubblicità immobiliare, registro), affinché si tenga conto in modo strutturale e fin dal principio della rilevanza statistica delle informazioni rilevate. Le analisi condotte sugli archivi in questi anni, infatti, hanno fatto emergere diverse carenze informative, che potrebbero essere superate con semplici accorgimenti tecnici nei modelli adottati nei procedimenti amministrativi (come, ad esempio, l'adozione del controllo di esistenza negli archivi catastali dell'unità immobiliare oggetto del procedimento o la segnalazione di valori anomali in fase di accettazione).

La qualità, infatti, non è solo una questione di errori presenti nelle basi dati, da intercettare a monte per essere ridimensionati, ma si tratta anche di disporre di modelli strutturati di acquisizione dei dati che consentano una migliore lettura dell'informazione acquisita (ad esempio, le caratteristiche degli immobili presenti nei modelli catastali, seppur acquisite puntualmente, non costituiscono basi dati strutturate per essere integrate, se non con complesse e macchinose elaborazioni, con altri archivi). Inoltre, in presenza di testi semi-strutturati come gli atti pubblici notarili, può rivelarsi uno strumento molto utile il *text mining*, cioè un processo di derivazione dal testo di informazioni di alta qualità le cui tecniche sono sempre più affinate. Un esempio in questo senso sono le informazioni relative al classamento energetico degli immobili, oppure quelle volte a individuare transazioni che non avvengono tra soggetti terzi (transazioni tra familiari), dove il corrispettivo potrebbe deviare dall'ordinario prezzo di mercato.

Inoltre, l'implementazione di nuovi campi nei modelli di acquisizione dei dati e/o una miglior definizione di quelli già esistenti arricchirebbero il contenuto informativo, con immediato vantaggio per le analisi statistiche su quelle basi dati (un esempio su tutti può essere la specificazione della quota relativa all'immobile oggetto di locazione parziale nei modelli RLI per la registrazione dei contratti).

A queste sfide che attendono l'Osservatorio nei prossimi anni se ne aggiunge una terza, che riguarda la gestione del progetto per la divulgazione dei prezzi dichiarati nelle compravendite immobiliari. Si può prefigurare,

infatti, il rischio che la disponibilità di dati reali georeferenziati forniti da una fonte autorevole come l'OMI possa finire per sostituire, di fatto, le quotazioni immobiliari quale "bussola" di riferimento per gli operatori; si ridurrebbe così, da un lato, la percezione del rapporto tra valore patrimoniale del bene e prezzo, utile nell'ottica della razionalità delle scelte, e si indurrebbe, dall'altro, la tentazione di attribuire valenza generale a informazioni certamente attrattive per via della loro intrinseca concretezza, ma, proprio per questo, espressione di condizioni specifiche e peculiari non facilmente interpretabili e generalizzabili in assenza di opportune ipotesi e valutazioni di tipo scientifico-statistico²⁶.

L'esperienza, poi, anche in riferimento a quanto accaduto in passato in ambiti del tutto diversi, insegna come importare modelli esogeni imponga sempre approfondite valutazioni, per evitare che delle opportunità si trasformino in criticità.

L'analisi del contesto internazionale conferma in pieno la delicatezza e la centralità di questo tema; basti pensare al fatto che proprio i criteri anglosassoni di valutazione della trasparenza del mercato immobiliare, come ad esempio il noto indice GRETI²⁷, pongano da sempre in posizione di eccellenza Paesi, come gli Stati Uniti, che si sono rivelati poi tra i più fragili e in cui gli operatori (cittadini, intermediari, finanziatori), pur a fronte di una notevole disponibilità di informazioni "originarie" e di dettaglio, si sono rivelati straordinariamente esposti all'irrazionalità delle dinamiche speculative (la crisi deflagrante del 2007 negli Stati Uniti ne è la più palese evidenza).

Anche in considerazione di ciò è in atto, presso l'Osservatorio, un'approfondita riflessione proprio sulle prospettive tecniche ed operative delle quotazioni immobiliari²⁸, strumento strategico e fondamentale; è chiara infatti la consapevolezza della centralità di questo tema non tanto sul ruolo dell'Osservatorio stesso, quanto soprattutto su un'efficace ed equilibrata gestione delle dinamiche di un mercato estremamente complesso e strategico per gli equilibri economici e finanziari del Paese.

Come si è avuto modo di leggere in queste pagine, dunque, una lunga strada è stata percorsa dall'OMI in questi venti anni e il tracciato già delineato per il futuro appare altrettanto lungo e complesso, con la consapevolezza che bisognerà farsi trovare sempre pronti a cambiare passo quando nuove o mutate condizioni lo richiederanno.

Bibliografia e sitografia

Alleva G., *The future of Statistics: challenges for understanding new phenomena* in *Keynote speech*, ITACOSM 2019, 5-7 giugno, Firenze.

Cassese S., *Evoluzione della normativa sulla trasparenza*, Sinapsi, 8, n. 1, 2018.

De Francisci S., *Open Data nella statistica ufficiale: ruolo, opportunità e implicazioni* in *Convegno 3a Giornata della trasparenza* in Istat, 2019

Festa M., Serafini S., *La trasparenza del mercato immobiliare nell'era dei big data*, I Quaderni dell'Osservatorio, 2017.

Ghirardo E., Papa F., *Quotazioni OMI: pregi e limiti*, I Quaderni dell'Osservatorio, 2020.

Guerrieri G., *Lo sviluppo dell'Osservatorio del Mercato Immobiliare (OMI): un po' di storia*, I Quaderni dell'Osservatorio, 2020.

ISTAT, *L'economia non osservata nei conti nazionali - Anni 2015-2018*, 2020.

²⁶ A proposito della complessa relazione tra il concetto di trasparenza ed un mercato particolare come quello immobiliare si veda l'approfondita disamina sviluppata nell'articolo di M. Festa e S. Serafini *La trasparenza...*, op. cit.

²⁷ Cfr. JLL *Global Real Estate Transparency Index*, vari anni.

²⁸ Cfr. E. Ghirardo, F. Papa. *Quotazioni OMI: pregi e limiti*, in questo numero dei Quaderni, §§ 4 e 5.

JLL *Global Real Estate Transparency Index*, vari anni, 1998-2020.

Monducci R., *Dati amministrativi per le analisi socio-economiche e la valutazione delle politiche pubbliche*, Convegno Inps-Siep, 2018

Secchi P., *On the role of statistics in the era of big data: a call for a debate*. *Statistics & Probability Letters*, volume 136, 2018

Società italiana di Statistica, *Statement della SIS su Statistica, Scienza dei dati e Big data*, 2018

UNECE, *Conference of European statisticians: What does "big data" mean for official statistics?*, 2013

http://dati.istat.it/Index.aspx?DataSetCode=DCSP_IPAB

<https://www.agenziaentrate.gov.it/portale/web/guest/schede/fabbricatiterreni/omi/pubblicazioni>

<https://www1.finanze.gov.it/finanze3/immobili/#/>

<https://www.gov.uk/government/statistical-data-sets/price-paid-data-downloads>

<http://odimpact.org/case-united-kingdoms-hm-land-registry-price-paid-data.html>

<https://app.dvf.etalab.gouv.fr/>

<https://www.data.gouv.fr/fr/datasets/5c4ae55a634f4117716d5656/>