



UNIONCAMERE

LA METODOLOGIA DEL NUOVO
SISTEMA INFORMATIVO EXCELSIOR

La metodologia del nuovo sistema informativo Excelsior

1	Il punto di partenza: il Registro delle Imprese integrato con gli archivi INPS	4
2	Gli aggiornamenti del campo di osservazione del Sistema informativo Excelsior	5
3	Il trattamento delle serie storiche di fonte INPS.....	6
3.1	L'elaborazione delle serie storiche dei modelli UNIEMENS	7
3.2	Il confronto tra i flussi di entrata originali di fonte INPS e quelli ricostruiti per Excelsior	8
4	Il modello previsionale mensile delle entrate	9
5	La strategia di rilevazione dell'indagine continua	13
6	Struttura del questionario mensile e classificazioni utilizzate	14
6.1	Struttura del questionario mensile.....	14
6.1.1	Sezione 1.....	14
6.1.2	Sezione 2.....	15
6.1.3	Sezione 3.....	15
6.2	Il sistema delle classificazioni utilizzate.....	16
6.2.1	La classificazione dell'attività economica.....	16
6.2.2	La classificazione dei tipi di contratto	17
6.2.3	La classificazione dei livelli di istruzione e dei titoli di studio.....	17
6.2.4	La classificazione delle professioni	18
7	La stima delle caratteristiche dei flussi in ingresso dell'indagine continua	19
7.1	Nuove funzionalità di stima dell'indagine continua	19
7.2	Il trattamento dei questionari di indagine	21
7.2.1	La codifica delle figure professionali	21
7.2.2	La stima delle professioni e delle caratteristiche richieste	21
7.2.3	I controlli sui dati quantitativi e qualitativi del questionario	22
8	Strumenti di divulgazione dell'indagine continua	23
8.1	I volumi statistici mensili	23
8.2	I bollettini mensili	23
8.3	La dashboard mensile.....	24
8.3.1	Sperimentazione di una metodologia di stima della probabilità di assunzione delle imprese.....	25
9	Le modalità di elaborazione su base annuale	26
9.1	Struttura del questionario annuale (Long form)	26
9.1.1	Sezione 4.....	27
9.1.2	Sezione 5.....	27

9.1.3	Sezione 6.....	27
9.2	L'elaborazione dei flussi quantitativi.....	27
9.3	L'elaborazione delle caratteristiche dei flussi quantitativi.....	29
9.4	L'elaborazione delle sezioni specifiche del questionario annuale	30
10	Strumenti di divulgazione dell'indagine annuale	32
10.1	Volumi statistici annuali	32
10.2	Bollettini annuali.....	34
10.3	Il sistema di interrogazione online	34
11	Il modello previsionale a medio termine.....	35
11.1	La formulazione del modello	35
11.2	La stima della "replacement demand"	37
11.3	La stima dei fabbisogni occupazionali	37
11.3.1	Determinazione dei fabbisogni per professione e titolo di studio.....	38
11.3.2	Determinazione dei fabbisogni nella Pubblica Amministrazione.....	38
11.4	la stima dell'offerta di diplomati e laureati	39

A vent'anni dalla sua nascita il Sistema informativo Excelsior si conferma una delle fonti più utilizzate per seguire le dinamiche quali-quantitative della domanda di lavoro. Per sfruttarne al meglio le potenzialità, a partire dal 2017, nel tradizionale impianto del Sistema informativo Excelsior sono state introdotte diverse innovazioni metodologiche utili a cogliere con maggiore precisione la dimensione dei flussi di entrata dei lavoratori nelle imprese, rendendoli ancor più direttamente fruibili al perseguimento degli obiettivi delle politiche attive del lavoro. Alle informazioni su scala annuale, che continuano ad offrire una serie di indicazioni sulle caratteristiche dei flussi, e del modello previsionale di medio periodo, viene affiancata un'indagine mensile che ha lo scopo di fornire le tendenze periodiche a breve del mercato del lavoro alimentando, contemporaneamente, la più ampia indagine annuale. Il motore di questa innovazione risiede nelle potenzialità legate all'integrazione degli archivi amministrativi ed in particolare del Registro delle Imprese delle Camere di Commercio che da qualche anno integra le informazioni occupazionali provenienti da fonte INPS.

1 Il punto di partenza: il Registro delle Imprese integrato con gli archivi INPS

Come anticipato in premessa, una delle più importanti novità del Sistema informativo Excelsior è legata ad un più stretto utilizzo, nel processo elaborativo dei dati, delle fonti amministrative a seguito degli sviluppi informativi del Registro imprese degli ultimi anni che ha integrato i dati occupazionali di fonte INPS.¹

Ciò ha consentito di perseguire i seguenti obiettivi:

- una puntuale² ricostruzione del campo d'osservazione con ridefinizione delle imprese e del relativo stock dei dipendenti;
- una puntuale ricostruzione dei flussi mensili di imprese e dipendenti consolidati nel periodo precedente a quello di elaborazione, potendo inoltre isolare quelli di brevissimo periodo o, per la loro natura amministrativa, non significativi³;
- la possibilità di ricostruire - attraverso opportune procedure statistiche che integrano i risultati dell'indagine con l'analisi dei flussi mensili consolidati - i flussi futuri delle principali forme contrattuali utilizzate dalle imprese ad un livello territoriale molto disaggregato che consente di arrivare sino ai singoli CPI⁴ (Centri per l'impiego).

In questo modo il dato quantitativo espresso dall'indagine non deriva più esclusivamente dal riporto all'universo dei dati di indagine, ma dall'interazione tra il dato amministrativo ed i risultati dell'indagine campionaria presso le imprese. Tale processo consente da un lato di esprimere degli indicatori mensili coerenti con l'entità dei flussi espressi dalle fonti amministrative, anticipando nel contempo delle tendenze

¹ Integrazione che riguarda in particolare il modello mensile UNIEMENS, una denuncia obbligatoria inviata mensilmente all'INPS dai datori di lavoro del settore privato che svolgono le funzioni di sostituti d'imposta e deve essere inviato all'INPS entro l'ultimo giorno del mese successivo a quello di competenza. Il modello consente di avere informazioni puntuali su stock e flussi generati da ogni singola azienda con riferimento ai lavoratori dipendenti e ai collaboratori registrati in "gestione separata".

² Per puntuale si intende l'anagrafica di ogni singola impresa.

³ Sono escluse dalla valutazione dei flussi, in armonia con quanto rilevato dal questionario d'indagine, i contratti inferiori a 20 giorni lavorativi che non consentirebbero in prospettiva alcuna attuazione di politiche attive nel breve periodo. Sono altresì riconosciuti e de-duplicati i contratti ravvicinati riferiti allo stesso lavoratore nei confronti di una medesima impresa frutto di duplicazioni amministrative ed inquadrabili come "false entrate".

⁴ I Centri per l'impiego (CPI) sono strutture pubbliche coordinate dalle Regioni che favoriscono sul territorio l'incontro tra domanda e offerta di lavoro e attuano iniziative e interventi di politiche attive del lavoro.

sulle caratteristiche delle figure professionali richieste, e dall'altro di alimentare le informazioni della più ampia indagine annuale che può contare su un ampio campione d'indagine attraverso il consolidamento delle informazioni sull'entità e delle caratteristiche dei flussi raccolte nei diversi mesi dell'anno; i risultati di quest'ultima alimenteranno, di volta in volta, le proiezioni del modello a medio termine.

2 Gli aggiornamenti del campo di osservazione del Sistema informativo Excelsior

L'universo di riferimento del Sistema informativo Excelsior per l'anno 2017 è costituito dalla totalità delle imprese private dei settori industriali e dei servizi iscritte al Registro delle Imprese delle Camere di Commercio che risultavano attive alla data del 31.12.2016 e che avevano avuto almeno un dipendente medio nel corso del 2016 (fonte INPS).

Le unità di rilevazione e analisi considerate sono l'impresa e l'unità locale provinciale (ULP); quest'ultima è una entità convenzionale definita come l'insieme delle unità locali della stessa impresa operanti in una provincia, i cui relativi addetti corrispondono alla somma degli addetti operanti nella provincia.

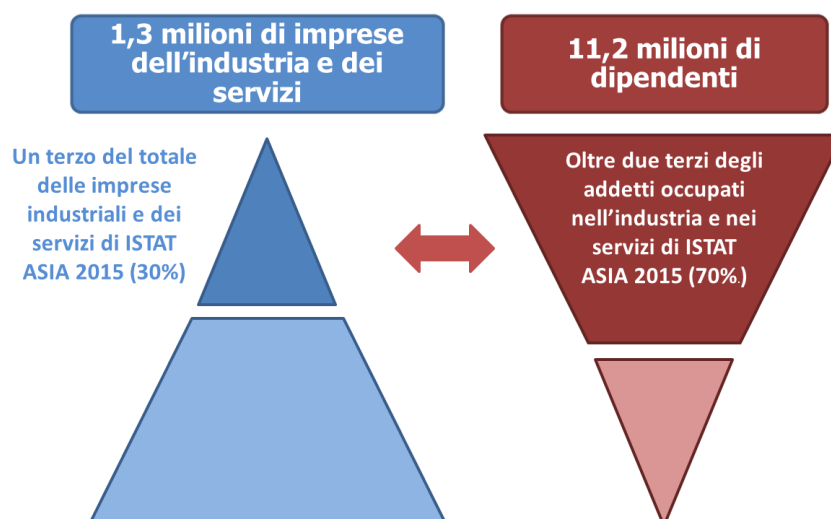
Tenuto conto delle caratteristiche delle imprese registrate nel Registro Imprese, sono esplicitamente escluse:

- le unità operative della pubblica amministrazione;
- le aziende pubbliche del settore sanitario (aziende ospedaliere, ASL, ecc.);
- le unità scolastiche e universitarie pubbliche;
- le organizzazioni associative;
- gli studi professionali non iscritti al Registro imprese⁵.

Inoltre, pur risultando iscritte al Registro Imprese, sono escluse dal campo di osservazione anche le imprese appartenenti al settore agricolo-zootecnico, che nelle passate annualità erano rilevate con una specifica indagine a causa delle particolarità dei fabbisogni professionali richiesti.

Da tale insieme sono state inoltre escluse le imprese cessate nel frattempo e, contemporaneamente, sono state inserite le imprese nate dopo il 31.12.2016, con particolare attenzione a quelle di grande dimensione o potenzialmente tali, suscettibili cioè di esprimere quote rilevanti di nuova occupazione. La composizione dell'universo di riferimento delle imprese con almeno 500 dipendenti viene continuamente aggiornata in relazione ai grandi eventi di trasformazione utilizzando le segnalazioni e le informazioni più recenti ricavate dal Registro delle Imprese utilizzando i bilanci depositati presso le Camere di commercio e le visure camerali⁶.

Questa definizione complessiva del campo d'osservazione ha condotto a considerare, nel 2017, 1.279.600 imprese, con un numero complessivo di dipendenti pari a 11.233.000⁷.



I numeri evidenziati consentono di affermare che Excelsior, pur cogliendo un terzo del complesso delle imprese presenti nell'Archivio statistico delle imprese attive (ASIA), riesce a coprire la maggior parte dello stock occupazione stabile del paese⁸.

3 Il trattamento delle serie storiche di fonte INPS

Gli archivi INPS costituiscono la fonte amministrativa da cui ricavare le informazioni di stock già integrate anagraficamente nel Registro delle Imprese; oltre a questo esse consentono, opportunamente rielaborate, la ricostruzione dei flussi d'ingresso (contratti) delle imprese che vengono attivati nel corso del tempo⁹.

La fonte INPS si articola in realtà in una pluralità di forniture trimestrali derivanti da archivi differenti e collegati tra loro da chiavi di raccordo. Tali chiavi sono per definizione univoche e fisse nel tempo in modo da garantire la ricostruzione in serie storica; ciò vale anche in presenza di informazioni criptate e ridotte in forma anonima per ragioni di privacy.

Le diverse forniture ricevute, opportunamente trattate in sequenza, hanno consentito di creare una serie storica su base mensile dei flussi in ingresso a partire da gennaio 2013. Tale serie, continuamente alimentata a ogni nuova fornitura trimestrale, attualmente arriva a dicembre 2017 e si è resa disponibile nel corso del 2017 sino al giugno dell'anno stesso.

La ricostruzione delle serie storiche INPS, in vista della successiva proiezione per il periodo t+12, ha la finalità di ricostruire il quadro delle entrate secondo modalità coerenti con quanto rilevato dall'indagine Excelsior. Essa viene svolta distintamente per gli addetti dipendenti, i lavoratori in somministrazione ed i collaboratori.

Il principale filone di attività, esemplificativo per il trattamento dei dati INPS, riguarda quello dei dipendenti: una tipologia particolare di dipendenti è rappresentata dai lavoratori in somministrazione che

⁹ Si precisa quindi che il dato rappresentato in Excelsior non rappresenta il numero di teste che vengono avviate al lavoro, ma il numero di occasioni di lavoro (contratti offerti) che sono rese disponibili da parte delle imprese. Questa lettura è quella più funzionale all'attuazione delle politiche attive sul breve periodo, va aggiunto che, nel caso di contratto a tempo indeterminato, le due misurazioni tendono a coincidere ricordando anche che il fenomeno viene ulteriormente ridotto dal numero minimo dei giorni lavorativi presi in considerazione per un avviamento nel campo d'osservazione di Excelsior.

nella fonte INPS risultano quali dipendenti in carico alle agenzie di somministrazione. Grazie a specifici codici di contribuzione si individuano i singoli record che vengono fatti confluire in un sottoinsieme di lavoratori con contratto di somministrazione (interinali) oggetto di una gestione separata e specifica, un secondo filone riguarda invece i collaboratori coordinati e continuativi.

3.1 L'elaborazione delle serie storiche dei modelli UNIEMENS

Una prima parte dell'attività di elaborazione serve a determinare a livello di impresa la serie storica mensile dei dipendenti diretti dell'impresa, individuando e sanando le duplicazioni e i dati mancanti, al fine di contabilizzare lo stock e i flussi di attivazione-cessazione dei rapporti tra lavoratori e impresa secondo criteri¹⁰ il più possibile vicini a quelli utilizzati dall'indagine Excelsior.

In sintesi, questa fase di trattamento può essere riassunta nei seguenti passaggi:

- creazione archivio serie storica dei record relativi a ciascuna -matricola d'impresa;
- gestione sovrapposizioni e deduplicazione dei record relativi alla matricola;
- creazione vettori mensili presenza, giornate ed eventi;
- trattamento incoerenza tra presenze e giornate lavorate;
- individuazione eventi di attivazione e cessazione e gestione incongruenze;
- individuazione e gestione interruzioni parziali di serie di presenze (*missing data*);
- determinazione totale dipendenti per matricola e impresa¹¹;
- analisi serie storica e gestione valori anomali e/o mancanti per impresa.

Particolarmente delicate, nell'ambito delle fasi sopraelencate, risultano le procedure di trattamento UNIEMENS per determinare i flussi mensili di attivazioni per ciascuna impresa, trattando i record relativi al singolo lavoratore in azienda e analizzando gli eventi di attivazione e cessazione al fine di individuare gli eventi reali (quindi al netto, ad esempio, dei passaggi di matricola all'interno della stessa impresa) e non occasionali, cioè superiori alle 20 giornate lavorate.

Questa fase di trattamento si articola nei seguenti passaggi:

- determinazione reali eventi di attivazione al netto di situazioni temporanee;
- calcolo dei flussi attivi mensili per impresa;
- classificazione delle imprese in merito ai comportamenti di assunzione.

Grazie agli elementi di stratificazione caratteristici di ciascuna impresa (classe dimensionale, settore, territorio, forma giuridica) è possibile ricavare, così, la serie storica mensile della struttura di occupazione e dei flussi per livello di inquadramento, classe di età e genere a partire da gennaio 2013 e articolata secondo il massimo dettaglio dei livelli informativi di stratificazione previsti da Excelsior.

Il processo così descritto prevede una frequenza di aggiornamento trimestrale, legato alla disponibilità delle fonti INPS.

L'elaborazione delle serie storiche dei modelli Emens riferiti ai "lavoratori in somministrazione" e ai "collaboratori" ricalcano a grandi linee i processi già descritti a proposito dei dipendenti.

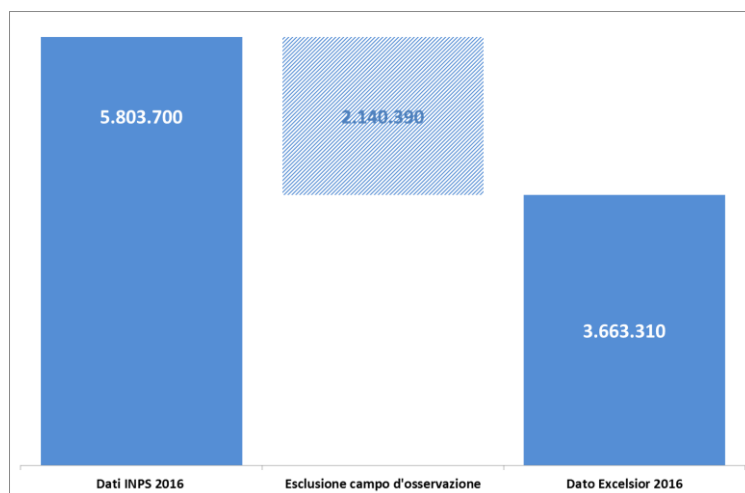
¹⁰ Si escludono ad esempio i flussi relativi a trasformazioni di contratto che riguardano uno stesso lavoratore nell'ambito della medesima impresa - o false riprese del rapporto di lavoro, determinate da comunicazioni mancanti o incomplete rese dall'impresa o ancora i passaggi di matricola all'interno della stessa attività.

¹¹ Una stessa impresa può avere più matricole INPS assegnate.

3.2 Il confronto tra i flussi di entrata originali di fonte INPS e quelli ricostruiti per Excelsior

Il complesso delle innovazioni introdotte nel Sistema Informativo Excelsior con la finalità di renderlo sempre più uno strumento informativo a supporto delle politiche attive del lavoro e dell'orientamento professionale e formativo, ha avuto un importante impatto per il dimensionamento dei flussi di entrata¹² rilevati, ora coerenti con quanto registrato dalla fonte amministrativa INPS, considerata al netto dei fenomeni non osservati per definizione dal Sistema informativo Excelsior.

Confronto dei flussi – Anno 2016



Fonte: Elaborazioni su dati INPS e Unioncamere – ANPAL, Sistema Informativo Excelsior

I passaggi che conducono dal dato iniziale di 5,8 milioni di attivazioni ai 3,6 milioni finali sono riassumibili in una serie di operazioni ascrivibili alla riduzione del campo d'osservazione:

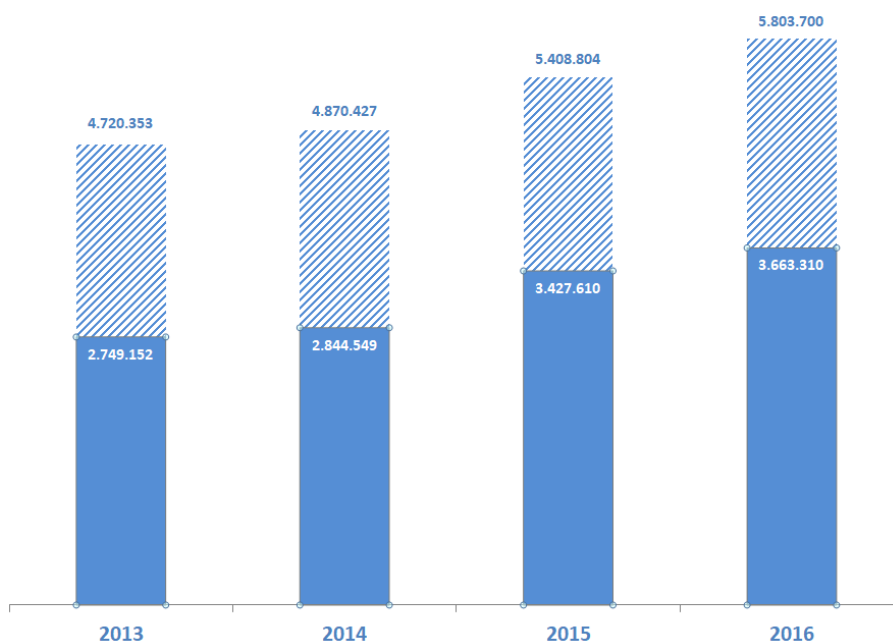
- esclusione del settore agricolo, gli studi professionali ed i soggetti, anche non profit, che non risultano iscritti nei registri delle Camere di Commercio;
- depurazione della quota di entrate espressa dalle imprese senza dipendenti, ovvero quelle con meno di 0,5 dipendenti in media¹³;
- esclusione dei flussi relativi a trasformazioni di contratto che riguardano uno stesso lavoratore nell'ambito della medesima impresa - o false riprese del rapporto di lavoro, determinate da comunicazioni mancanti o incomplete rese dall'impresa;
- riconoscimento e depurazione dei contratti di brevissimo periodo, ovvero quelli di durata inferiore ai 20 giorni lavorativi.

L'applicazione di questa armonizzazione con i dati INPS ricondotti al campo d'osservazione Excelsior porta il volume degli ingressi rilevati dal Sistema informativo a livelli ben più contenuti rispetto a quelli comunicati ufficialmente dall'INPS, cogliendone comunque la parte privata più stabile e strutturata che risulta ricostruibile anche per i periodi antecedenti il 2016, come riportato nel diagramma seguente:

¹² Si precisa che nel grafico, per favorire il confronto con il dato INPS (Osservatorio sul precariato) ci si riferisce alle attivazioni di contratti di lavoro dipendente, inclusi quelli in somministrazione.

¹³ Media ricalcolata trimestralmente sugli ultimi 12 mesi.

Confronto dei flussi – Serie storica 2013-2016



Fonte: Elaborazioni su dati INPS e Unioncamere – ANPAL, Sistema Informativo Excelsior, 2017

4 Il modello previsionale mensile delle entrate

Disponendo di una serie storica dei flussi su base mensile aggiornata con cadenza trimestrale, è stato realizzato un modello previsionale per consentire una proiezione di breve periodo delle stime delle attivazioni di contratti da parte delle imprese, in coerenza con il sottoinsieme che l'indagine Excelsior intende rilevare.

Alla luce della ampia disaggregazione dei domini su cui è necessario stratificare i risultati di Excelsior, l'impianto sviluppato non presuppone l'elaborazione di un unico modello, ma di una serie di modelli applicati a differenti domini di stratificazione dell'universo di riferimento. Infatti le tecniche di costruzione delle serie storiche illustrate nel precedente paragrafo e, in particolare, la possibilità di relazionare i dati di flusso a livello di singola impresa con gli elementi di stratificazione desunti dal Registro Imprese, consentono di ricostruire delle segmentazioni dettagliate definendo ciascuno dei domini con una propria serie storica per la quale costruire ed applicare un modello¹⁴.

Il continuo accantonamento di una serie storica di indagini mensili e la progressiva sovrapposizione delle stesse con dati di riscontro desumibili da fonti amministrative ha guidato la scelta di sviluppare un modello di tipo autoregressivo che valuti il contributo delle differenti indagini per la determinazione delle stime di un dato complessivo coerente con le grandezze realmente osservate.

La ricerca della migliore soluzione per la creazione di un modello previsionale delle entrate è tuttora in corso ed in continua evoluzione; la strategia attualmente applicata è quella di creare ed alimentare un modello VectorAutoregressive (VAR), potendo attraverso questa modellistica:

- tenere conto della serie storica della banca dati dei flussi amministrativi;

¹⁴ Fanno eccezione e quindi non rientrano in domini aggregati le circa 1.650 imprese con almeno 500 dipendenti, per le quali si prevede una stima ed un monitoraggio puntuale.

- tenere conto di opportune variabili esogene che risultino sufficientemente tempestive nel cogliere i momenti di svolta dovuti a un cambiamento congiunturale.

L'interazione tra le due dimensioni del problema consente di unire alla capacità di analisi di trend e stagionalità basata sulla profondità della serie storica dei flussi di origine amministrativa l'immediatezza e la tempestività delle indagini mensili o di altri indicatori esterni capaci di cogliere i momenti di svolta del mercato. In particolare i dati utilizzati nella modellistica per la stima delle entrate sono stati:

- la serie storica delle entrate INPS (microdati), disponibile con frequenza mensile dal 2013;
- la serie storica dei microdati delle previsioni delle entrate (e indicatori standardizzati da queste derivabili) di Excelsior, disponibile su base mensile¹⁵ dal 2013 fino a giugno 2018;
- clima di fiducia delle imprese (ISTAT) disponibile su base mensile a partire dal 2013 fino a marzo 2018.

Mentre le serie INPS e Excelsior sono disponibili rispettivamente a livello di impresa e microdato, e dunque sono riaggregabili ex-post a livello settoriale, i dati grezzi relativi al clima di fiducia delle imprese sono disponibili per 32 domini definiti dall'incrocio di:

- 8 settori (manfatturiero, costruzioni, commercio, turismo, trasporti, informatica, servizi alle imprese e servizi),
- 4 ripartizioni geografiche (Nord-Ovest, Nord-Est, Centro, Sud).

Ciò ha condizionato la stima del modello stabilendo un vincolo sul livello di riagggregazione dei dati di base per la costruzione della matrice di input del modello VAR che è comunque stato stimato separatamente per dipendenti (al netto dei lavoratori in somministrazione), lavoratori in somministrazione¹⁶ e collaboratori.

I dati di indagine Excelsior e quelli amministrativi INPS sono stati quindi riaggregati in modo coerente con i domini ricavabili dagli indicatori del clima di fiducia delle imprese creando - per ogni variabile di riferimento - una serie storica mensile settoriale e per ripartizione territoriale, in modo funzionale alla costruzione di un modello VectorAutoregressive (VAR). Tale modello descrive l'evoluzione dinamica di un insieme di variabili endogene sulla base della dinamica comune passata delle serie stesse. Questo tipo di approccio è molto diffuso nell'ambito del *forecasting* macroeconomico in quanto i modelli VAR si sono dimostrati molto accurati e al tempo stesso parsimoniosi, consentendo di tener conto sia della dinamica temporale delle singole variabili che delle interazioni tra le stesse.

Di seguito si richiamano brevemente gli aspetti essenziali della metodologia VAR. Consideriamo due variabili (x e y). Un VAR di primo ordine (ovvero un VAR che contiene un solo ritardo delle variabili) può essere scritto come

$$y_t = \delta_1 + \varphi_{11}y_{t-1} + \varphi_{12}x_{t-1} + \varepsilon_{1t}$$

$$x_t = \delta_2 + \varphi_{21}y_{t-1} + \varphi_{22}x_{t-1} + \varepsilon_{2t}$$

¹⁵ Per la serie storica Excelsior dei microdati delle previsioni delle entrate stimate solo su base campionaria è stata creata una serie storica ibrida ottenuta utilizzando le indagini trimestrali condotte dal primo trimestre 2013 al primo trimestre 2017 (compresi) e le indagini mensili svolte da aprile a dicembre 2017 con previsioni da aprile 2017 fino a giugno 2018. Per rendere temporalmente omogenea (serie mensile), la serie i dati delle indagini trimestrali è stata effettuata la scomposizione delle entrate trimestrali per mese utilizzando la percentuale di composizione nota dagli archivi INPS. Il modello VAR richiede inoltre una elaborazione dei dati di input per ciascun mese di previsione; tale elaborazione necessita dell'elaborazione preventiva dell'indagine mensile secondo la tecnica dei riporti all'universo di tipo tradizionale (ovvero effettuati solo su base campionaria), oltre all'aggiornamento della serie del clima di fiducia.

¹⁶ La serie dei lavoratori in somministrazione non è scomponibile a livello settoriale e dunque è stata stimata solo a livello aggregato.

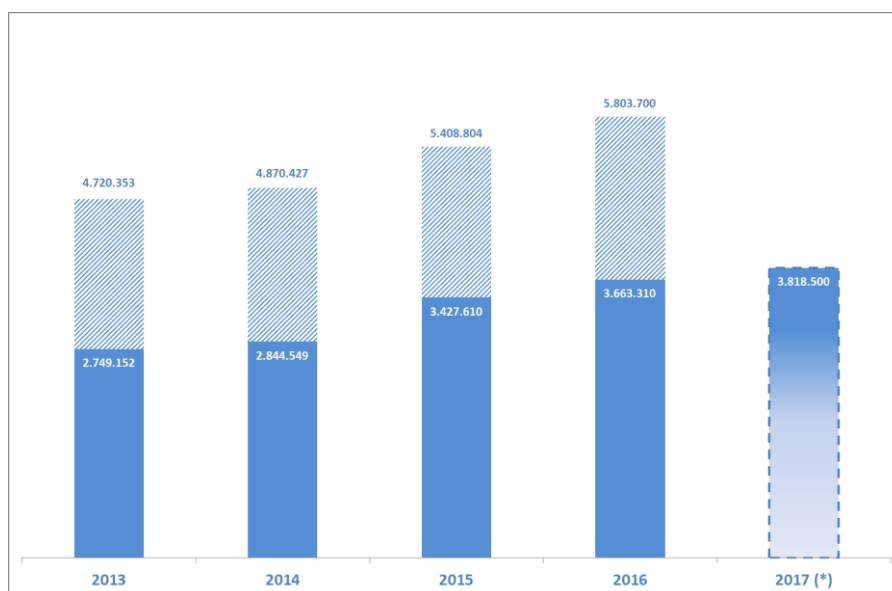
In forma vettoriale, e più in generale, il modello può essere rappresentato come segue:

$$Y = \delta + \Phi_1 Y_{t-1} + \varepsilon$$

dove Y è un vettore $K \times 1$ di variabili, δ , è un vettore $K \times 1$ di parametri (costanti), Φ è una matrice $K \times K$ di parametri, ε il vettore degli errori.

L'applicazione della modellistica ai dati integrati del Registro delle imprese e dell'INPS ricondotti al campo d'osservazione Excelsior consente di confrontare il volume degli ingressi rilevati dal Sistema informativo Excelsior con la ricostruzione di quelli comunicati ufficialmente dall'INPS ricostruibile per i periodi antecedenti il 2016 e quindi stimata per il 2017¹⁷:

Confronto dei flussi – Serie storica 2013-2017



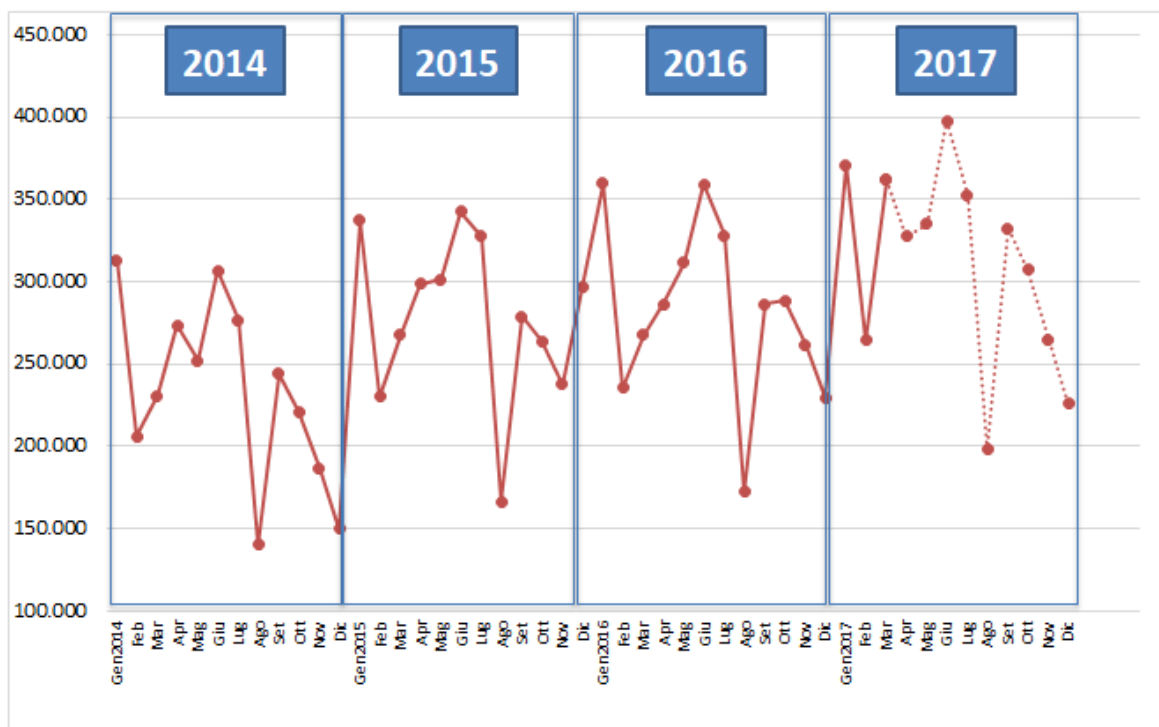
(*) dato previsionale

Fonte: Elaborazioni su dati INPS e Unioncamere – ANPAL, Sistema Informativo Excelsior, 2017

Ovviamente, oltre che per l'anno nel suo complesso, la stima avviene per le singole mensilità, in relazione alle esigenze di raccordo con la cadenza dell'indagine. Si deve inoltre evidenziare che se i parametri del modello vengono ristimati ogni 3 mesi, in concomitanza con la periodicità dell'integrazione dei dati INPS, lo stesso viene applicato ai dati mensilmente per tener conto del periodico aggiornamento delle variabili esogene:

Confronto dei flussi – Serie storica mensile 2014-2017

¹⁷ Il corrispondente valore per l'anno 2017, ottenuto considerando l'elaborazione dei flussi ricavati dall'INPS sino al mese di marzo e la loro integrazione con quanto desunto dalle indagini Excelsior a cadenza mensile.



Per la natura e articolazione delle componenti utilizzate, il modello VAR non prevede la possibilità di effettuare una proiezione di previsioni maggiormente dettagliate rispetto ai 32 domini base usati per determinare i marginali: pertanto per tutte le componenti la disaggregazione massima è ottenuta utilizzando degli algoritmi che sfruttano la distribuzione deterministica nota dell'ultima fornitura INPS disponibile.

In base alle esigenze del Sistema informativo Excelsior i dati vengono elaborati suddividendo i risultati ottenuti in modo da garantire degli output articolati per :

- dettaglio territoriale corrispondente ai Centri Provinciali per l'Impiego e relativi livelli superiori (province, regioni, grandi ripartizioni);
- dettaglio settoriale, corrispondente a 98 aggregazioni di divisioni/gruppi ATECO e relativi livelli superiori;
- dettaglio dimensionale di impresa calcolate sui dipendenti relativo a 7 aggregazioni di classi dimensionali (1-9 dipendenti, 10-49 dipendenti, 50-99 dipendenti, 100-249 dipendenti, 250-499 dipendenti, 500-999 dipendenti, 1000 dipendenti e oltre);
- dettaglio di livello inquadramento¹⁸, 3 aggregazioni (dirigenti, impiegati e operai);
- dettaglio relativo alle componenti demografiche dei flussi quali genere e classi di età del personale in entrata (attualmente in fase di sperimentazione).

Al termine del processo di modellizzazione e ripartizione si ottengono mensilmente le grandezze di riferimento delle entrate previste distinte secondo la tipologia di contratto (dipendenti – distinti a loro volta in dirigenti, impiegati e operai - lavoratori in somministrazione e collaboratori) e l'articolazione richiesta. Questi risultati costituiscono la base per le successive stime e caratterizzazioni dei flussi mensili effettuate tramite l'indagine campionaria, la cui metodologia viene richiamata successivamente.

¹⁸ Solo per la componente di flusso dei dipendenti.

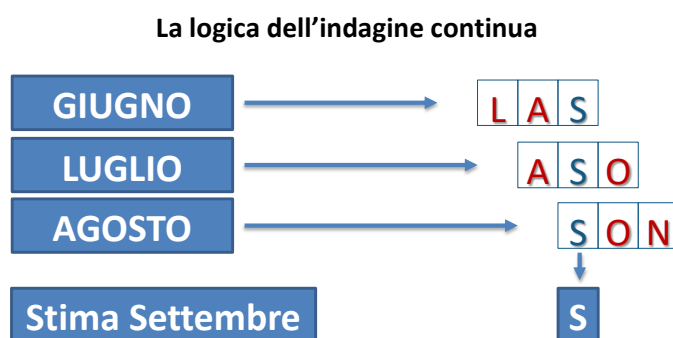
5 La strategia di rilevazione dell'indagine continua

L'evoluzione della metodologia di rilevazione ha riguardato sia la cadenza, sia le modalità di somministrazione, passando dalla precedente e più tradizionale metodologia CATI (*Computer Aided Telephone Interview*) a cadenza trimestrale/annuale ad una più flessibile rilevazione CAWI (*Computer Aided Web Interview*) a periodicità mensile.

Rispetto alle indagini svolte in passato, che prevedevano un'indagine annuale e delle rilevazioni trimestrali, la strategia utilizzata nel nuovo Excelsior è quella di realizzare una rilevazione mensile, nella quale si formulano stime per un trimestre previsionale mobile, cioè in ogni indagine l'orizzonte temporale si estende ai tre mesi successivi al mese in cui si effettua la rilevazione (e non più all'intero anno solare in corso, come avveniva in passato).

Quindi, se l'indagine viene svolta nel corso del mese di giugno, essa si riferisce alle previsioni occupazionali relative al trimestre luglio-settembre, con la specificazione del dettaglio per ciascuno dei tre mesi; quella svolta in luglio avrà come periodo di riferimento il trimestre agosto-ottobre, e così via.

In tal modo i dati relativi a ciascun mese indagato ottengono il contributo informativo da 3 rilevazioni:



Tale scelta presenta due vantaggi immediati rispetto al passato:

- la rilevazione è sempre integralmente previsionale¹⁹;
- il periodo di previsione è limitato a tre mesi, orizzonte temporale per il quale le imprese riescono a formulare valutazioni affidabili riguardo ai nuovi contratti attivati.

Grazie alla nuova impostazione seguita, nel periodo tra aprile e dicembre 2017 sono state raccolte circa 340mila interviste raccolto - tra le 40 e le 50mila al mese - con un elevato ritorno campionario (32%)²⁰ rispetto al numero di complessiva dei contatti, risultato pari a 850mila imprese.

¹⁹ Nell'indagine annuale una parte delle assunzioni era avvenuta in un periodo antecedente a quello di rilevazione, ma non sempre questo costituiva un vantaggio: se già avvenute le assunzioni erano certe, tuttavia se lontane nel tempo potevano non essere tenute in considerazione al momento dell'intervista.

²⁰ Poco più di 300mila sono, invece, i questionari utilizzati nell'indagine annuale 2017, un esito ben superiore ai tradizionali 100mila delle passate edizioni. Il numero delle interviste utilizzate per elaborare i risultati dell'indagine annuale considera, oltre a quelle acquisite nel periodo aprile-agosto 2017, anche quelle acquisite nell'ultima parte del 2016 per le previsioni di gennaio-marzo 2017. Tutto ciò considerato, quindi, per le previsioni annuali vengono utilizzate le stime relative agli ultimi 12 mesi, comprendendo così l'intero arco temporale dell'anno.

Interviste raccolte nel corso del 2017

Settore	Totale
Industrie alimentari, delle bevande e del tabacco	10.276
Industrie tessili, dell'abbigliamento, del cuoio e delle calzature	12.371
Industrie del legno e del mobile	10.058
Industrie della carta, cartotecnica e della stampa	5.047
Industrie dei gioielli e delle pietre preziose	1.679
Industrie estrattive	1.014
Industrie della lavorazione dei materiali da costruzione (vetro, ceramica, pietra)	4.435
Industrie elettriche ed elettroniche	8.146
Industrie della lavorazione dei metalli e dei prodotti in metallo	20.642
Industrie della fabbricazione e manutenzione di macchinari e apparecchiature	18.284
Industrie chimico-farmaceutiche	3.150
Industrie della gomma e delle materie plastiche	3.488
Altre industrie manifatturiere	1.476
Public utilities	7.750
Costruzioni	37.940
Commercio e riparazione di autoveicoli e motocicli	8.784
Commercio all'ingrosso	17.233
Commercio al dettaglio	22.180
Servizi turistici, di alloggio e ristorazione	30.193
Servizi informatici e delle telecomunicazioni	13.964
Servizi avanzati di supporto alle imprese	21.769
Servizi operativi di supporto alle imprese e alle persone	21.788
Servizi di trasporto, logistica e magazzinaggio	19.898
Servizi finanziari e assicurativi	10.342
Altri servizi alle imprese	3.215
Istruzione e servizi formativi privati	4.341
Sanità, assistenza sociale e servizi sanitari privati	9.624
Servizi ricreativi, culturali e altri servizi alle persone	13.916
TOTALE Nazionale	343.003

6 Struttura del questionario mensile e classificazioni utilizzate

6.1 Struttura del questionario mensile

Il questionario proposto mensilmente alle imprese attraverso indagine CAWI si articola in 4 sezioni:

- sezione 1: situazione dell'occupazione (dipendente e non) alla data dell'ultimo aggiornamento delle fonti amministrative (al momento 30 settembre 2017)
- sezione 2A: attivazioni di contratto per figura professionale e/o eventuali cessazioni previste nel trimestre successivo al mese di rilevazione
- sezione 2B: caratteristiche delle figure professionali previste in entrata nel corso del trimestre successivo al mese di rilevazione
- sezione 3: alternanza scuola lavoro

Di seguito viene specificato il contenuto di ogni sezione.

6.1.1 Sezione 1

Nella sezione 1 del questionario viene innanzitutto proposto lo stock di occupati alle dipendenze²¹, distinguendoli per livello di inquadramento comprensivi degli stagionali, secondo quanto risultante dall'ultimo aggiornamento degli archivi anagrafici integrati Registro Imprese-INPS effettuato sulla base della

²¹ Il dato presentato può essere eventualmente modificato ed una specifica sezione, attivata solo in presenza di società cooperative, è stata prevista per la presenza di soci lavoratori. Qui vengono rilevati sia i soci lavoratori che i dipendenti non soci secondo la loro funzione all'interno dell'impresa (direttiva, impiegatizia e tecnica, operaia).

metodologia del sistema informativo Excelsior descritta in precedenza. La presentazione in anteprima di questi dati contestualizza l'impresa sulla propria situazione interna favorendo la compilazione delle successive sezioni.

Viene poi richiesto all'impresa di indicare il numero complessivo di occupati, specificando il numero, oltre che degli stessi dipendenti noti, anche dei lavoratori somministrati, dei collaboratori e degli altri lavoratori non alle dipendenze.

6.1.2 Sezione 2

Nella sezione 2A viene richiesto all'impresa di indicare se sono previste cessazioni di rapporto di lavoro nel trimestre di previsione, i motivi di cessazione e l'eventuale presenza di tipologie contrattuali in uscita non presenti negli stock indicati in sezione 1.

Viene poi richiesto all'impresa di fornire il numero di attivazioni di contratto previste per il trimestre successivo, con la richiesta di specificare la denominazione attribuita dall'impresa alle diverse figure professionali in entrata e il relativo numero di unità, il mese in cui avverranno tali ingressi e l'area aziendale nella quale se ne prevede l'inserimento.

La sezione 2B riguarda le caratteristiche delle figure professionali segnalate in entrata da parte dell'impresa ed è quindi proposta solo alle imprese che prevedono attivazioni di contratto nel trimestre successivo; la rilevazione delle caratteristiche professionali viene effettuata distintamente per ogni figura professionale richiesta.

Per ciascuna figura della quale si programma l'entrata viene chiesto di indicare:

- una descrizione dei compiti della figura prevista in entrata, degli strumenti che utilizzerà, delle conoscenze tecniche che dovrà possedere, dei soggetti con cui dovrà relazionarsi, del tipo di soluzioni che dovrà applicare (standard, creative o innovative) e se la figura stessa dovrà coordinare altre persone o essere coordinata;
- il tipo di competenze (trasversali) richieste per quel profilo professionale e l'importanza di ciascuna;
- il livello di istruzione e il titolo di studio specifico richiesto;
- altri caratteri della figura da assumere quali, ad esempio, l'età, l'esperienza specifica precedente, la difficoltà di reperimento della figura, i motivi della difficoltà e le azioni previste per trovare la figura in questione, l'assunzione della figura per sostituire o meno una analoga figura presumibilmente in uscita dall'impresa, la presenza o meno nell'impresa di altre figure professionali simili a quelle in assunzione, la decisione o meno di assumere personale proveniente da paesi esteri;
- l'associazione tra figura professionale e genere (figure maschili o femminili più adatte per l'impresa);
- le forme contrattuali previste e i motivi dell'eventuale utilizzo di contratti a tempo determinato (es. attività stagionali);
- la necessità di ulteriore formazione (mediante corsi specifici interni o esterni all'impresa o con affiancamento a personale interno all'impresa o con altre modalità);
- il livello di inquadramento (dirigente, impiegato, operaio), nel caso si tratti di personale alle dipendenze.

6.1.3 Sezione 3

La sezione 3 contiene una serie di quesiti sul tema dell'alternanza scuola lavoro: numero di studenti ospitati nel corso dell'anno precedente e previsioni per l'anno corrente. Viene inoltre chiesto all'impresa se conosce il Registro nazionale per l'alternanza scuola lavoro e viene verificato l'eventuale interesse a sviluppare accordi con le scuole tramite il supporto della Camera di commercio del territorio.

6.2 Il sistema delle classificazioni utilizzate

Il sistema informativo Excelsior adotta due principali tipologie di classificazioni:

- classificazioni relative alle caratteristiche dell'impresa, tra cui assume rilievo quella riferita all'attività economica
- classificazioni relative alle caratteristiche della figura professionale.

Per quest'ultima sono rilevanti in particolare le classificazioni relative a:

- tipo di contratto
- livello di istruzione e titoli di studio
- professioni e figure elementari.

6.2.1 La classificazione dell'attività economica

Per i propri fini di rilevazione, elaborazione ed esposizione, l'indagine Excelsior da sempre si è dotata di una propria classificazione dei settori economici referenziata alla classificazione ATECO dell'ISTAT che espone in modo snello e articolato le attività fino a un livello pari a un incrocio tra le divisioni e i gruppi ATECO.

Tale classificazione è organizzata su 6 livelli gerarchici, dal più aggregato distinto in 5 settori (Industria, Costruzioni, Commercio, Turismo e servizi) fino a un massimo di 98 sub-settori.

Di seguito si elencano i 28 settori sulla base dei quali vengono presentati i principali risultati dell'indagine:

- Industrie alimentari, delle bevande e del tabacco
- Industrie tessili, dell'abbigliamento, del cuoio e delle calzature
- Industrie del legno e del mobile
- Industrie della carta, cartotecnica e della stampa
- Industrie dei gioielli e delle pietre preziose
- Industrie estrattive
- Industrie della lavorazione dei materiali da costruzione (vetro, ceramica, pietra)
- Industrie elettriche ed elettroniche
- Industrie della lavorazione dei metalli e dei prodotti in metallo
- Industrie della fabbricazione e manutenzione di macchinari e apparecchiature
- Industrie chimico-farmaceutiche
- Industrie della gomma e delle materie plastiche
- Altre industrie manifatturiere
- Public utilities
- Costruzioni
- Commercio e riparazione di autoveicoli e motocicli
- Commercio all'ingrosso
- Commercio al dettaglio
- Servizi turistici, di alloggio e ristorazione
- Servizi informatici e delle telecomunicazioni
- Servizi avanzati di supporto alle imprese
- Servizi operativi di supporto alle imprese e alle persone
- Servizi di trasporto, logistica e magazzinaggio
- Servizi finanziari e assicurativi
- Altri servizi alle imprese
- Istruzione e servizi formativi privati
- Sanità, assistenza sociale e servizi sanitari privati
- Servizi ricreativi, culturali e altri servizi alle persone

6.2.2 La classificazione dei tipi di contratto

La tipologia di contratto è rilevata per le figure professionali che le imprese prevedono di inserire e presenta tre livelli diversamente articolati.

Il primo livello è così articolato:

- contratto alle dipendenze
- contratto in somministrazione (ex-interinale)
- contratto di collaborazione coordinata e continuativa
- altri contratti di lavoratori non alle dipendenze (es. partita IVA)

Per i primi tre elementi si ha una esatta corrispondenza con i dati di fonte amministrativa (INPS) dai quali si ricavano le stime delle entrate.

Il secondo livello si applica solo per i contratti alle dipendenze per i quali è possibile un'analisi maggiormente dettagliata. Tra questi si distingue ulteriormente:

- contratto a tempo indeterminato (a tutele crescenti)
- contratto a tempo determinato
- apprendistato
- contratto a chiamata.

Le suddette tipologie non sono direttamente riconducibili a dati di fonte amministrativa. Infatti i dati INPS rilevano tali caratteristiche in campi differenti (con incroci e sovrapposizioni) e non sono esaustive di tutte le casistiche. Pertanto la ripartizione delle entrate per questo livello viene interamente effettuata sui dati di indagine che, presentando solo tali modalità, le rendono esaustive e alternative senza problemi di sovrapposizione e/o lacune.

Infine un terzo livello interviene nel caso di utilizzo di contratti a tempo determinato. Qui, in sede di indagine si richiede all'impresa di specificare la motivazione circa l'utilizzo di tale tipologia contrattuale, indicando una tra le seguenti:

- prova di nuovo personale
- sostituzione temporanea di personale (per maternità, aspettativa, ferie, malattia)
- copertura di un picco di attività
- carattere stagionali

A maggior ragione in questo caso tali tipologie non sono direttamente riconducibili a dati di fonte amministrativa e la ripartizione delle entrate per questo livello viene interamente effettuata sui dati di indagine che, presentando solo tali modalità, le rendono esaustive e alternative.

6.2.3 La classificazione dei livelli di istruzione e dei titoli di studio

La classificazione utilizzata, con talune aggregazioni di titoli relativi a corsi di studio simili, fa riferimento con alcuni adattamenti a quella utilizzata dall'ISTAT e organizzata su tre livelli gerarchici: livello di istruzione, indirizzo di studio e titolo di studio specifico.

Il primo livello gerarchico prevede quattro livelli di istruzione:

- nessuna qualifica (*alias* scuola dell'obbligo, considerata come livello minimo di istruzione);
- corsi regionali di formazione professionale e istruzione professionale di Stato (fino a 4 anni);
- diploma superiore, che comprende i diplomi rilasciati dagli istituti tecnici e le maturità della scuola media superiore. Tra questi, viene evidenziata la quota in possesso di una specializzazione post-diploma, ossia quella conseguita attraverso i corsi di Istruzione e Formazione Tecnica Superiore (IFTTS) e i corsi di Istruzione Tecnica Superiore (ITS);

- titolo universitario, che comprende i titoli di studio rilasciati dall'università dopo corsi di durata anche inferiore ai cinque anni (come era il caso dei diplomi di laurea ante-riforma o delle nuove classi di laurea di primo livello).

All'interno di ogni livello di istruzione vengono poi associati specifici indirizzi di studio e titoli di studio.

I titoli di studio riferiti al livello di istruzione universitario e ai diplomi vengono considerati così come classificati all'interno di specifici indirizzi formativi dal Ministero dell'Istruzione, dell'Università e della Ricerca.

Per i livelli formativi corrispondenti all'istruzione e alla formazione professionale i dati vengono esposti secondo un insieme di indirizzi riconducibili agli indirizzi della scuola media superiore.

6.2.4 La classificazione delle professioni

Il Sistema Informativo Excelsior si è dovuto confrontare, sin dal suo esordio nel 1997, con il tema della classificazione delle figure professionali operanti nelle aziende.

Funzionalmente agli scopi dell'indagine, si è provveduto alla elaborazione di una classificazione delle figure professionali che fosse al contempo:

- utilizzabile e comprensibile dagli imprenditori e dalle aziende intervistate, in quanto basata sul linguaggio e sulle terminologie da questi stessi utilizzati;
- aggiornabile, in modo da poter recepire costantemente le naturali evoluzioni del mondo del lavoro;
- confrontabile comunque con le altre fonti /classificazioni ufficiali;
- corretta, perché incentrata sull'osservazione di almeno 3 delle principali caratteristiche/parametri che concorrono alla sua definizione.

L'indagine Excelsior si è inizialmente basata su un insieme di professioni condiviso: il dizionario delle professioni della classificazione ISTAT.

Lo svolgimento delle successive indagini ha poi condotto, da una parte, a semplificare e, dall'altra, a integrare tale insieme:

- molte figure sono state eliminate dall'insieme adottato da Excelsior perché obsolete, mai richieste dalle imprese o difficilmente distinguibili da altre figure esistenti;
- alcune figure sono state introdotte ex-novo perché emergenti, richieste dalle imprese o segnalate da altre fonti specialistiche.

Attualmente la nomenclatura delle figure elementari Excelsior conta oltre 4.100 voci elementari, delle quali mediamente poco più di 800 vengono effettivamente citate almeno una volta nel corso di una *wave* mensile. E' altresì disponibile una tavola di raccordo tra le figure contenute nel "dizionario" Excelsior e la classificazione ISTAT CP 2001.

A fini espositivi, infatti, le professioni elementari sono state classificate secondo la classificazione delle professioni ISTAT 2011, che consente sia la coerenza con una classificazione di livello europeo, dato il raccordo esistente tra la CP 2011 e la classificazione ISCO 2008, sia la possibilità di associare un significato univoco alle descrizioni delle figure attraverso l'introduzione di una definizione delle classi di appartenenza delle figure stesse e, conseguentemente, favorisce un miglior controllo della corrispondenza tra descrizione della figura da parte dell'impresa e descrizione codificata.

7 La stima delle caratteristiche dei flussi in ingresso dell'indagine continua

7.1 Nuove funzionalità di stima dell'indagine continua

Come anticipato precedentemente l'indagine non è più concentrata in un periodo dell'anno e limitata a un campione predefinito, ma è sempre attiva lungo tutto l'anno e sottoposta a un *panel* mensile di imprese: tale panel è sub-stratificato per garantire la distribuzione delle interviste a livello di territorio provinciale, settore di attività e classe dimensionale e ruota rinnovandosi di mese in mese con l'obiettivo di contattare almeno una volta nel corso dell'anno²² tutte le imprese del campo di osservazione di indagine cercando di minimizzare, nel contempo, il fastidio statistico dei soggetti da intervistare.

A quanto detto si deve inoltre aggiungere un altro elemento fondamentale, ovvero che la stima del modello dei flussi precedentemente esposta, beneficia della serie storica mensile dei micro-dati delle previsioni campionarie delle entrate, nonché di indicatori standardizzati da queste derivabili. Di conseguenza è bene chiarire che i flussi rilevati dall'indagine mensile, riferiti ad un determinato mese, vengono utilizzati per due finalità, di seguito descritte.

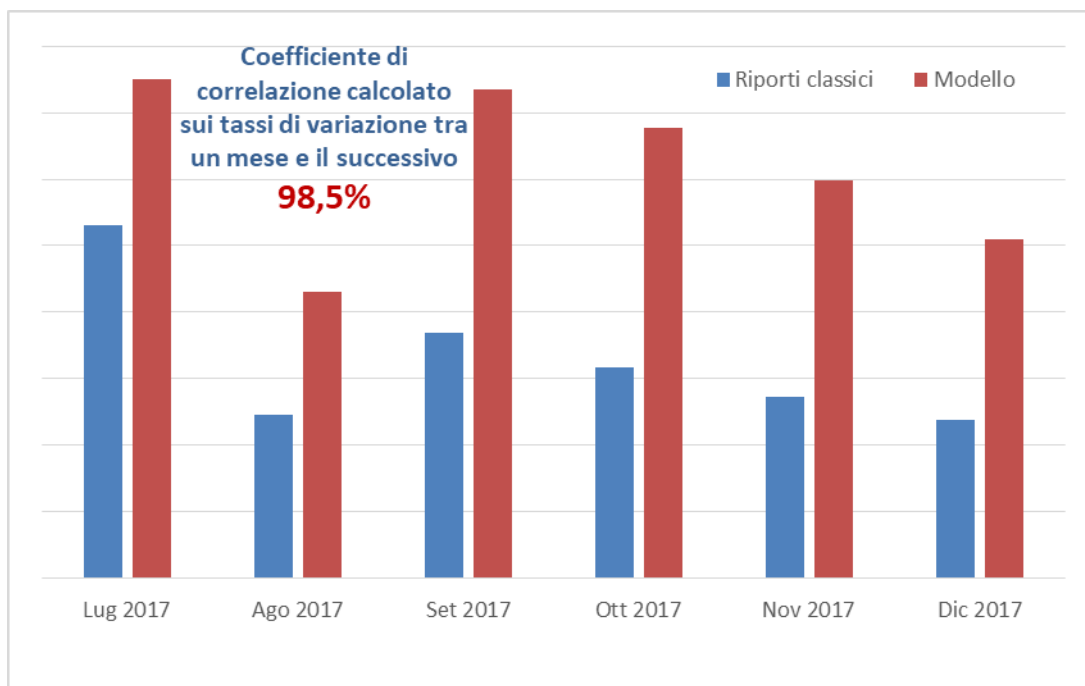
- Una prima finalità riguarda le procedure di riporto classico²³ assimilabili ad un campionamento stratificato senza ripetizione il cui stimatore corretto del totale \hat{Y} dei flussi è

$$\hat{Y} = \sum_{i=1}^n \frac{y_i}{\pi_i}$$

Tale stima, essendo riferita al periodo previsionale dei flussi del modello e poiché disponibile il mese precedente tale stima, può essere utilizzata come variabile anticipatoria che, come espressione ravvicinata delle intenzioni degli imprenditori, condensa nell'indicatore "*black box*" tutta una serie di contingenze che sarebbe assai complesso esprimere esplicitamente dal punto di vista settoriale e territoriale cogliendo eventuali "*turning point*" non intercettabili da una modellistica esclusivamente autoregressiva. Il grado di correlazione di questa variabile con il dato deterministico amministrativo dell'INPS è ben evidente in questo grafico esemplificativo:

²² All'obiettivo di contattare tutte le imprese si affianca quelle di intervistarle non più di una volta ogni 6 mesi; tale orientamento è il compromesso derivante dalla limitazione del disturbo statistico e dalla necessità di copertura degli strati corrispondenti alle imprese di medio/grande dimensione.

²³ Si intendono procedure che sfruttano esclusivamente i dati provenienti dalle indagini per le stime dei flussi di assunzione.



Ovviamente la relazione anticipatoria dei flussi viene modellizzata all'interno di ciascun sub-modello.

- Una seconda finalità riguarda l'utilizzo dei risultati dell'indagine per la caratterizzazione²⁴ dei flussi che, avendo a disposizione i risultati di una modellizzazione delle entrate che include, oltre che le caratteristiche economiche delle imprese richiedenti, anche diverse caratteristiche²⁵ dei profili richiesti rilevati in sede d'indagine, si presta per effettuare delle stime di tipo vincolato che hanno come riferimento le grandezze ottenute con il modello.

L'obiettivo è quello di ottenere uno stimatore che sia più efficiente di quello di Horvitz-Thompson utilizzando uno stimatore indiretto che garantisca un netto miglioramento dell'accuratezza delle stime. Tale stimatore a ponderazione vincolata (o calibrato) risulta indicato allo scopo anche grazie alla sua duttilità di impiego determinando i pesi di riporto all'universo in modo che le stime delle medie (totali) delle variabili ausiliarie ottenute con questi pesi sia uguali alle corrispondenti medie (totali) di popolazione.

$$\hat{Y} = \sum_{k=1}^n w_k y_k$$

$$\sum_{k \in S} w_k x_k = X$$

rispettando un sistema di vincoli sintetizzato dalla relazione

In tal modo si è in grado di sfruttare l'informazione sui valori totali delle variabili ausiliarie riferita all'universo oggetto d'indagine per produrre stime coerenti²⁶ (vincolate).

²⁴ Ovvero la scomposizione dei flussi di entrata nelle figure professionali e relative caratteristiche.

²⁵ Si ricorda che sia il dettaglio di livello inquadramento (dirigenti, impiegati e operai) che quello relativo alle componenti demografiche dei flussi quali sesso ed classi di età del personale in entrata sono elementi deterministici dei flussi delle forniture INPS integrati con il Registro delle imprese nonché elementi raccolti in fase d'indagine.

²⁶ L'impiego dello stimatore vincolato a variabili ausiliarie note da una fonte amministrativa risulta, inoltre, particolarmente utile per correggere l'impatto delle mancate risposte.

7.2 Il trattamento dei questionari di indagine

7.2.1 La codifica delle figure professionali

Data la centralità dell'informazione raccolta, e in funzione degli obiettivi dell'indagine, una specifica attenzione viene dedicata alla codifica delle figure professionali in ingresso dichiarate dalle imprese, operazione che risulta preliminare a qualsiasi attività di trattamento ed elaborazione dell'indagine.

La cadenza mensile dell'indagine necessita di adottare tempistiche rapide in tutte le operazioni di trattamento dei dati, senza però rinunciare all'affidabilità dei risultati; a tale proposito, per agevolare le attività di codifica delle professioni è stato sviluppato un sistema semi-automatico basato su algoritmi probabilistici di analisi del testo che si fondano di informazioni raccolte in forma libera nei seguenti campi del questionario:

- figura richiesta;
- compiti svolti;
- strumenti utilizzati;
- conoscenze tecniche richieste;
- figure con cui si relazionerà all'interno e all'esterno dell'azienda.

L'applicazione delle procedure di analisi del testo basate su vocabolari²⁷ di parole e regole chiave consente di analizzare le relazioni tra le descrizioni dei campi sopra citati e le figure elementari²⁸.

L'esito di questo processo automatico consiste nella proposta di una serie di codifiche di figure ordinate per il proprio *scoring* di associazione rispetto ai testi raccolti in forma libera; se tale *scoring* non supera un valore-soglia prestabilito, e comunque valutando il criterio di affidabilità specifico dell'operazione automatica, esperti codificatori confermano o correggono la codifica proposta, analizzando e valorizzando l'insieme di informazioni rilevate all'interno del questionario.

7.2.2 La stima delle professioni e delle caratteristiche richieste

- Gli elementi di qualificazione delle professioni e delle loro caratteristiche raccolti nelle diverse rilevazioni mensili costituiscono la principale fonte da cui Excelsior attinge questo tipo di informazioni²⁹; tale processo non è, tuttavia, privo di memoria: infatti, la reiterazione dell'indagine negli ultimi anni ha consentito di stratificare una serie di informazioni su professioni e caratteristiche creando un dizionario "cumulativo" delle figure rilevate che viene alimentato e arricchito sistematicamente da ogni nuova indagine. In tal modo si ha a disposizione di un ampio "bagaglio" informativo - continuamente aggiornato - a cui attingere in occasione delle diverse elaborazioni che consente di dettagliare per ogni dominio d'indagine le figure rilevate in tutte le indagini che insistono sul medesimo periodo di rilevazione³⁰/proiezione e di qualificarne le relative caratteristiche.

²⁷ Si tratta di dizionari di relazioni che non comprendono solo dei lemmi ma anche una serie di parametri di associazione tra le figure proposte e le principali caratteristiche rilevate per il profilo richiesto: attività economica dell'impresa, livello di inquadramento, livello di istruzione e titolo di studio, conoscenza informatica e delle lingue, area funzionale di inserimento.

²⁸ Figure professionali ISTAT 2011

²⁹ Si precisa che in via sperimentale sono già stati acquisiti dati provenienti dagli annunci dei principali portali di ricerca del personale finalizzati alla costruzione di meta-questionari univocamente riconducibili a domini o imprese del campo d'osservazione, nonché l'acquisizione e la valutazione di elementi desumibili da fonti amministrative quali le COB (Comunicazioni obbligatorie).

³⁰ Si ricordi che la strategia di rilevazione dell'indagine continua fa sì che su ciascun mese insistono le previsioni provenienti da tre rilevazioni.

La ratio di questa impostazione sta nella considerazione che, se la composizione delle attivazioni per figura può risentire di una certa variabilità nell'anno³¹, le caratteristiche ad essa associate e *desiderate* dall'impresa (es: titolo di studio, esperienza richiesta, ...) sono invece più strutturali e che pertanto l'utilizzo di un serbatoio storicizzato più ampio può consentire un maggiore supporto alle stime delle caratteristiche delle assunzioni e, di conseguenza, una maggiore stabilità temporale nei risultati da un mese all'altro.

Tuttavia, per cogliere eventuali tendenze nella variazione dei desiderata il dizionario garantisce un ricambio prefissato delle indagini che lo alimentano, aggiornandosi con quelle più recenti a scapito di quelle più datate.

7.2.3 I controlli sui dati quantitativi e qualitativi del questionario

Le risposte delle imprese, ottenute tramite il questionario acquisito attraverso una piattaforma web dedicata (con tecnica C.A.W.I.), sono soggette a controlli di coerenza con riferimento alle diverse informazioni raccolte sia sui dati quantitativi (numero di addetti, numero di entrate ecc.) sia su quelli qualitativi (professioni segnalate, titoli di studio ecc.). Questi controlli avvengono sia durante l'acquisizione via web dell'intervista che nella fase di trattamento preliminare dei questionari, non appena conclusa la fase di importazione dei dati.

La fase di controlli che avviene *online*, in presenza del soggetto compilatore, prevede una serie di controlli formali preimpostati nelle diverse sezioni della piattaforma C.A.W.I. quali esemplificativamente la quadratura e la compatibilità³² dei dati quantitativi: per i primi vengono imposti dei vincoli per i secondi degli *alert* di compilazione. Occorre peraltro trovare il giusto equilibrio su questi aspetti, che richiedono all'impresa intervistata di correggere le informazioni fornite, per evitare che l'eccesso di vincoli e controlli possa scoraggiare la compilazione del questionario rischiando che esso non venga completato.

Un'ulteriore fase di controlli, che avvenendo *off-line* può essere ancor più puntuale, prevede di incrociare altri elementi, anche contenuti in più sezioni del questionario, al fine di verificarne la coerenza³³.

³¹ Pertanto è opportuno utilizzare quelle effettivamente richieste in un dato periodo.

³² Alcune funzionalità della piattaforma, quelle sui controlli di totali e subtotali di grandezze deterministicamente legate tra loro, hanno compiti di ridondanza bloccante e non consentono il proseguimento dell'intervista sinché non si rientra in coerenza. Per evitare errori di compilazione anche sulle variabili non deterministicamente ricavabili quali per esempio il numero complessivo delle attivazioni di contratto, una funzione lineare di compatibilità delle entrate rispetto lo stock dei dipendenti, avverte con un alert i casi di possibile errata compilazione. In quest'ultimo caso però non vi è ridondanza bloccante e si consente la prosecuzione dell'intervista in caso di conferma del dato, in questo caso si rimanda ad una successiva fase desk l'analisi dell'informazione acquisita online.

³³ In caso di impossibilità a determinare l'informazione corretta, perché assente o anomala, il dato viene stimato, in caso di outlier su valori quantitativi ciò avviene all'interno di cluster dinamici definiti rispetto a territorio, settore attività economica e classe dimensionale sulla base di scostamenti rispetto al valore medio μ e in funzione dello scarto quadratico medio σ , per cui risultano outlier i valori y_i che non rispettano la condizione:

$$\bar{y} - 2\sigma < y_i < \bar{y} + 2\sigma$$

Questa, come altre casistiche, possono generare dei missing data:

- per eliminazione esito delle procedure di trattamento degli outliers.
- per mancata acquisizione del dato in sede di intervista;
- per eliminazione come esito delle procedure di controllo di coerenza post acquisizione di intervista;

A secondo della tipologia e quindi della provenienza del caso, le tecniche di missing data estimation operano secondo due possibili modalità:

- stima della risposta con il valore medio del cluster, per i dati quantitativi;
- assegnazione casuale di una risposta ponderata sulla distribuzione del cluster, per i dati qualitativi.

8 Strumenti di divulgazione dell'indagine continua

La diffusione dei dati mensili avviene attraverso la produzione di tre tipologie di strumenti:

- i volumi (o rapporti) statistici di indagine con cadenza mensile, a livello nazionale, regionale e provinciale;
- i bollettini a livello nazionale, regionale e provinciale, intesi quali report sintetici a carattere divulgativo contenenti le principali informazioni ottenute dall'indagine;
- l'interfaccia web per l'interrogazione dei dati mensili.

I primi due prodotti sono di tipo statico ed hanno come obiettivo di offrire una vista standardizzata sul dato mensile periodicamente prodotto. I rapporti statistici offrono un dettaglio che ha come target un'utenza più scientifica, mentre i bollettini cercano di mettere in evidenza le tematiche più direttamente collegate alle politiche attive del lavoro, avendo come target gli stakeholder del mercato del lavoro.

L'interfaccia d'interrogazione dei dati è invece lo strumento più innovativo introdotto in Excelsior e consente, oltre alla navigazione del database delle previsioni mensili attraverso particolari viste infografiche, di svolgere il compito di individuazione della potenziale domanda di lavoro, avendo attenzione specifica ad un possibile utilizzo da parte dei Centri per l'impiego.

8.1 I volumi statistici mensili

I volumi statistici consentono un monitoraggio frequente dei fenomeni oggetto d'indagine. Tali output presentano una serie di tavole con i dati più significativi tra quelli prodotti alla conclusione di ciascuna indagine mensile. I volumi sono strutturati su diverse tematiche che permettono di analizzare le entrate previste nel periodo in esame per gruppo professionale, per livello di istruzione e per settore di attività. Tale analisi viene condotta attraverso la lettura di alcune caratteristiche riguardanti i fenomeni di rilievo presenti nel questionario di indagine, quali, a livello indicativo: le aree aziendali di inserimento, i profili professionali con più entrate, le attività lavorative in cui è richiesta esperienza e quelle per le quali le imprese dichiarano difficoltà nel trovare personale, la classe di età e il genere richiesto per le figure in entrata e le relative competenze, la ripartizione dei contratti previsti in entrata.

Inoltre, sia il volume nazionale che i volumi regionali presentano al loro interno tavole con dati a livello provinciale. Sono infine previste alcune tavole che mettono a confronto le previsioni delle entrate a livello mensile e a livello trimestrale.

8.2 I bollettini mensili

Questi output hanno un taglio maggiormente divulgativo e sono finalizzati all'evidenziazione delle tematiche più vicine alle politiche attive del lavoro; ogni mese vengono prodotti un bollettino a carattere nazionale, 20 bollettini regionali e 104 bollettini provinciali³⁴.

Attraverso tavole, grafici e commenti i bollettini trattano in modo sintetico e con un efficace impatto visivo le numerose tematiche riguardanti le previsioni occupazionali del mese/trimestre in esame: opportunità di lavoro previste, gruppi professionali e livelli di istruzione delle figure in entrata, principali professioni richieste e relativa difficoltà di reperimento, aree aziendali di inserimento, tipi di contratto previsto, settore di attività economica, tipo di esperienza, assunzioni di giovani e di donne, caratteristiche e competenze delle figure in entrata, classe dimensionale delle imprese assumentesi.

Nel bollettino nazionale le tematiche precedentemente indicate vengono affrontate con ancora maggior dettaglio analizzando più approfonditamente le professioni e le difficoltà di reperimento, sia in relazione al totale delle entrate previste, che per il segmento giovanile. Vengono inoltre predisposti specifici focus su:

³⁴ E' esclusa la provincia di Aosta per la quale vi è coincidenza con il relativo bollettino regionale.

aree aziendali di inserimento, caratteristiche dei lavoratori previsti in entrata, competenze richieste, settori di attività di inserimento, titoli di studio richiesti e tipo di contratto proposto alle figure in entrata.

8.3 La dashboard mensile

Come già accennato precedentemente, la cadenza mensile dell'indagine e le informazioni da questa ricavabili, oltre ad essere funzionali alla realizzazione dell'elaborazione e dell'osservatorio statistico su base annuale, svolgono un ruolo più congiunturale di supporto alle politiche attive del lavoro per l'orientamento professionale e formativo che ha come *target* i Centri per l'Impiego.

A questo scopo è stata creata un'applicazione che si compone di diversi cruscotti, organizzati per tematiche, che consentono all'utente di navigare i dati raccolti secondo diversi livelli di disaggregazione geografica, settoriale e dimensionale con una esposizione anche in forma grafica. Attraverso una serie di schermate correlate l'utilizzatore può accedere a differenti modalità di esplorazione dei dati di previsione mensile o trimestrale, visualizzando di volta in volta:

- le imprese che prevedono assunzioni distinte per settore di attività, classe dimensionale;
- le imprese, di cui al punto precedente, declinate in un maggior dettaglio per settore economico e per ranking territoriale;
- le assunzioni previste distinte nelle principali professioni richieste;
- le assunzioni, di cui al punto precedente, declinate in maggior dettaglio secondo diverse caratteristiche: distribuzione per settore di attività economica, per titolo di studio, per area aziendale, dimensione di impresa, tipo di contratto e così via.

Per tutte le modalità di esplorazione sono possibili operazioni di selezione territoriale permettendo all'utente di filtrare i dati a livello provinciale spingendosi fino al dettaglio del singolo Centro per l'Impiego.

Sfruttando la possibilità di personalizzare i cruscotti sulla base delle varie esigenze si ha la possibilità di effettuare operazioni di drill-down filtrando l'intera dashboard sulla base della singola professione, del settore economico o del titolo di studio semplicemente cliccando sulla relativa *label* dei diversi grafici.

Tra le peculiarità dello strumento scelto sono anche la possibilità di esporre informazioni frutto di join secondo chiavi comuni di basi dati differenti e non omogenee; ciò consente, ad esempio, di integrare i dati con informazioni sugli stock occupazionali e di impresa ottenuti per mera totalizzazione dei dati esito delle fasi di trattamento statistico delle fonti amministrative descritto nei paragrafi precedenti. Infine è costantemente garantita la possibilità di visualizzare dati secondo granularità differenti attraverso gerarchie predefinite o scelte dall'utente.

La domanda che si ritiene possa essere assolta dallo strumento della *dashboard*³⁵ mensile, oltre che l'esplorazione dei dinamica risultati proposti in forma statica nei bollettini, è quella di fornire un elenco di probabili datori di lavoro interessati ad un determinato profilo professionale. La risposta a questa domanda è modellizzabile tramite tecniche statistiche che consentono di generalizzare i risultati di Excelsior per stimare la probabilità di assumere di ogni impresa.

Tale applicazione viene sfruttata, in particolare, anche per esporre visivamente i risultati del modello di *scoring*, estraendo liste delle imprese con associate le maggiori probabilità di assumere in generale o una specifica professione per il territorio o il settore di interesse.

³⁵ Lo strumento è stato costruito con l'ausilio del software Tableau, uno dei più diffusi sul mercato dei BI and Analytics platforms, caratterizzato da un'elevata facilità nell'esplorazione e visualizzazione dei dati attraverso cruscotti interattivi di facile condivisione tra vari utenti.

8.3.1 Sperimentazione di una metodologia di stima della probabilità di assunzione delle imprese

Come anticipato, è stato realizzato in via sperimentale, per ora a esclusivi fini interni, un modello di *scoring* con l'obiettivo di generalizzare i risultati di Excelsior per stimare la probabilità di assumere di ogni impresa. La stima avviene sulla base delle caratteristiche delle imprese che dichiarano di assumere³⁶ una determinata figura professionale, considerando la probabilità di assunzione da parte di altre imprese che condividano le stesse caratteristiche. Queste informazioni possono essere molto utili in prospettiva per i centri per l'impiego, che potrebbero migliorare l'efficienza del proprio servizio.

Lo strumento utilizzato per stimare questa probabilità è un modello di regressione non lineari probit multivariato specificatamente disegnato per variabili dipendenti binarie:

$$\Pr(y = 1|x_1, \dots, x_n) = \Phi(\beta_0 + \beta_1x_1 + \dots + \beta_px_p)$$

I coefficienti *probit* vengono stimati con il metodo della massima verosimiglianza che dà luogo a stime consistenti e si distribuisce normalmente in grandi campioni. Le informazioni che alimentano il modello derivano dall'incrocio della base informativa Excelsior con quella dell'archivio INPS, per ogni impresa rispondente vengono registrate sia le assunzioni che la variabile dicotomica assume/non assume connesse, in caso di assunzione, con il profilo della figura. In aggiunta vengono calcolate una serie di variabili quali:

- Settore di appartenenza
- Classe dimensionale
- Età dell'impresa
- Ragione sociale dell'impresa
- Età media dei dipendenti
- Quota dipendenti over 60
- N. entrate complessive registrate nell'anno precedente (variabile categorica 5 classi)
- Dummies temporali che identificano il mese dell'assunzione
- Area geografica (provincia).

Sulla base di queste variabili viene applicato il modello *probit* che stima la probabilità di assumere una determinata figura professionale per le imprese appartenenti all'universo di osservazione di Excelsior. Il modello viene stimato separatamente per figura professionale, ciò significa che se una impresa assume due figure professionali diverse avrà due diverse probabilità di assunzione una per ciascun gruppo professionale³⁷.

I coefficienti stimati dal modello *probit* sono successivamente utilizzati per prevedere la probabilità di assunzione di ogni impresa non appartenente al campione Excelsior sulla base delle variabili elencate precedentemente che caratterizzano ogni impresa. Il modello consente dunque di associare ad ogni impresa la probabilità di assumere una determinata figura professionale dando luogo ad uno *scoring* su cui effettuare il ranking delle imprese in relazione alle probabilità di assunzione. Le imprese sono successivamente ordinate su scala territoriale e ordinate in termini di *scoring*, alla fine di questo processo per ogni unità territoriale considerata (ad esempio provincia) il modello consente di individuare:

- Quali imprese hanno la maggiore probabilità di assunzione
- Quali imprese hanno la maggiore probabilità di assunzione di una determinata figura professionale.³⁸

³⁶ Si tratta delle rispondenti all'indagine nelle varie wave mensili.

³⁷ 2 digit della classificazione CPI ISTAT.

³⁸ Il passaggio dai 2-digit alle 4-digit CPI ISTAT è stato ottenuto applicando allo *scoring* assegnato a ciascuna impresa un fattore di ponderazione dato dalla distribuzione congiunta per gruppo di attività ATECO /figure professionali a 4-digit CPI risultati dalle stime di Excelsior.

9 Le modalità di elaborazione su base annuale

Le elaborazioni su base annuale relative a dipendenti, lavoratori in somministrazione e collaboratori, prendono forma e consistenza da quanto derivante dalla rilevazione continua che ha come suo primo scopo quello di fornire le tendenze periodiche sul mercato del lavoro e come secondo scopo, ma non secondario, di alimentare la base di conoscenza della più ampia e significativa elaborazione annuale.

I flussi di informazioni che l'indagine continua mette a disposizione per l'elaborazione su base annuale sono fondamentalmente due:

- le stime mensili derivanti dall'applicazione della modellistica VAR;
- le interviste raccolte mensilmente dall'indagine continua.

La combinazione di queste due componenti, opportunamente rielaborate in funzione della stabilizzazione del montante dei flussi nel momento in cui avviene l'elaborazione del dato annuale, consente di ottenere le stime dei flussi cumulati per l'intero anno.

A queste due componenti, che sono fondamentali per la stima e la caratterizzazione dei flussi, va aggiunta una terza componente:

- la parte di rilevazione dedicata alle domande del questionario annuale che consentono la segmentazione delle imprese e che sono legate esclusivamente all'indagine "long form".

9.1 Struttura del questionario annuale (Long form)

La struttura del questionario annuale è del tutto simile a quella del questionario mensile per quanto riguarda la parte relativa alla rilevazione dei flussi e delle loro caratteristiche comprendendo:

- sezione 1: situazione dell'occupazione (dipendente e non) alla data dell'ultimo aggiornamento delle fonti amministrative (al momento 30 settembre 2017)
- sezione 2A: attivazioni di contratto per figura professionale e/o eventuali cessazioni previste nel trimestre successivo al mese di rilevazione
- sezione 2B: caratteristiche delle figure professionali previste in entrata nel corso del trimestre successivo al mese di rilevazione
- sezione 3: alternanza scuola lavoro.

A queste si affiancano, con una periodicità standard di 3 wave³⁹, equivalenti a circa 135-150mila questionari, alcune sezioni che sono più funzionali alla creazione di prodotti informativi, anche editoriali, che corrispondono alla diffusione dei dati annuali:

- sezione 4: investimenti in formazione di personale interno relativi all'anno precedente
- sezione 5: tirocini e stage attivati relativi all'anno precedente
- sezione 6: informazioni relative all'azienda utili alla loro segmentazione tematica⁴⁰.

³⁹ Nella sperimentazione del 2017 ci si è limitati a 2 wave, quella di luglio e quella di agosto, per un totale di circa 90mila interviste acquisite, ma per il triennio 2018-2020 sono già state previste 3 wave.

⁴⁰ Ovvero che consente di segmentare le imprese rispetto i temi di approfondimento annuale direttamente legati alle pubblicazioni tematiche di Excelsior quali, solo a titolo di esempio, "Imprese Culturali e Creative" o "Imprese Green"

Di fatto la struttura del questionario è funzionale alla normale prosecuzione dell'indagine mensile svolta nelle tre wave richiamate, riservando ad un secondo sottoinsieme di informazioni la rilevazione di fenomeni che non devono essere investigati in maniera continua perché rilevati a consuntivo o perché elementi di caratterizzazione sufficientemente stabile nel corso di un anno.

9.1.1 Sezione 4

La sezione 4 rileva l'attività formativa promossa dall'impresa nel corso dell'anno precedente a quello di indagine a favore dei propri dipendenti, la finalità principale del processo formativo e le sue modalità di svolgimento.

9.1.2 Sezione 5

Nella sezione 5 si pone l'attenzione sull'attività svolta dall'impresa in termini di tirocini o stage ospitati, richiedendo il numero di quelli attivati nell'anno precedente a quello di indagine, la quota che di questi ultimi che viene poi trasformata in assunzioni, quanti tra i tirocinanti sono laureati/laureandi, nonché il di cui per i quali è stato previsto un compenso e la durata dei tirocini stessi.

9.1.3 Sezione 6

Nell'ultima sezione del questionario, a conclusione del set di domande, vengono rilevate una serie di informazioni relative alle caratteristiche dell'impresa quali: la vendita di prodotti/servizi all'estero, i processi di innovazione di prodotto/servizio avviati nel corso dell'anno appena trascorso, le scelte aziendali in termini di investimenti in prodotti / tecnologie a maggior risparmio energetico, nonché sull'utilizzo di internet per migliorare il business dell'impresa.

9.2 L'elaborazione dei flussi quantitativi

Come richiamato in precedenza l'indagine annuale viene alimentata dalle rilevazioni mensili che includono al loro interno, in termini procedurali, l'applicazione di un modello VAR ai dati integrati del Registro delle imprese e dell'INPS, già esposta nel capitolo 4.

Tale procedura consente di stimare cumulativamente il volume degli ingressi rilevati dal Sistema informativo con la periodicità con la quale il modello VAR viene alimentato dai dati INPS⁴¹:

$$Y = \delta + \Phi_1 Y_{t-1} + \varepsilon$$

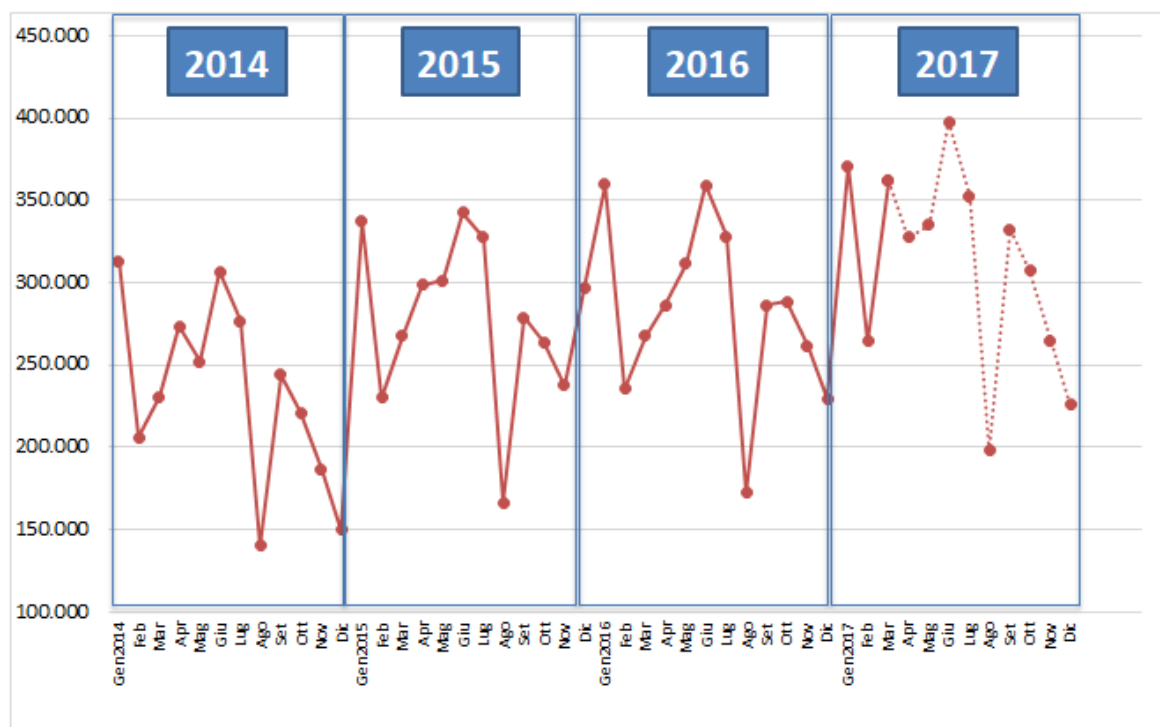
dove Y è un vettore $K \times 1$ di variabili, δ , è un vettore $K \times 1$ di parametri (costanti), Φ è una matrice $K \times K$ di parametri, ε il vettore degli errori.

Ciò consente una ricostruzione deterministica dei dati sui flussi che va dai 3 a 6 mesi antecedenti il periodo previsionale esteso sino a fine anno. Ciò è particolarmente rilevante perché, dato il vincolo di esposizione dei risultati annuali, questi vengono elaborati a partire da settembre⁴² dell'anno di stima.

⁴¹ I dati relativi ai flussi INPS sono disponibili con un lag temporale da 6 a 3 mesi dal momento della loro effettiva realizzazione.

Conseguentemente a ciò una parte delle previsioni dei flussi annuali, indicativamente comprese tra 3 e 6 mesi, sono di fatto rilevate a consuntivo e non sono oggetto di stima. In particolare per il 2017⁴³:

Confronto dei flussi – Serie storica mensile 2014-2017



Come si può dedurre, nel mese di settembre i mesi di gennaio, febbraio e marzo erano già disponibili a consuntivo nel database di input del modello VAR e conseguentemente la stima cumulativa dei flussi per i risultati annuali di Excelsior 2017 è avvenuta solo per i restanti 9 mesi.

Come più volte richiamato, i flussi relativi ai lavoratori dipendenti risultano la parte più corposa dell'analisi del sistema informativo Excelsior ma, fermo restando il campo di osservazione, vengono prese in considerazione ulteriori tipologie contrattuali:

- i rapporti di lavoro con contratto di collaborazione coordinata e continuativa, che vengono elaborati a partire dagli archivi INPS dei contribuenti della "gestione separata" con un processo di armonizzazione analogo a quello descritto per i dipendenti⁴⁴ e che corrispondono ad un flusso pari a 93mila unità;
- le attivazioni con contratti di lavoro autonomo previste dalle imprese con dipendenti, su base annua pari a circa 180mila unità⁴⁵.

Si arriva così ad un valore complessivo pari ad oltre 4 milioni di ingressi con una quota di imprese assuntori intorno al 60%⁴⁶:

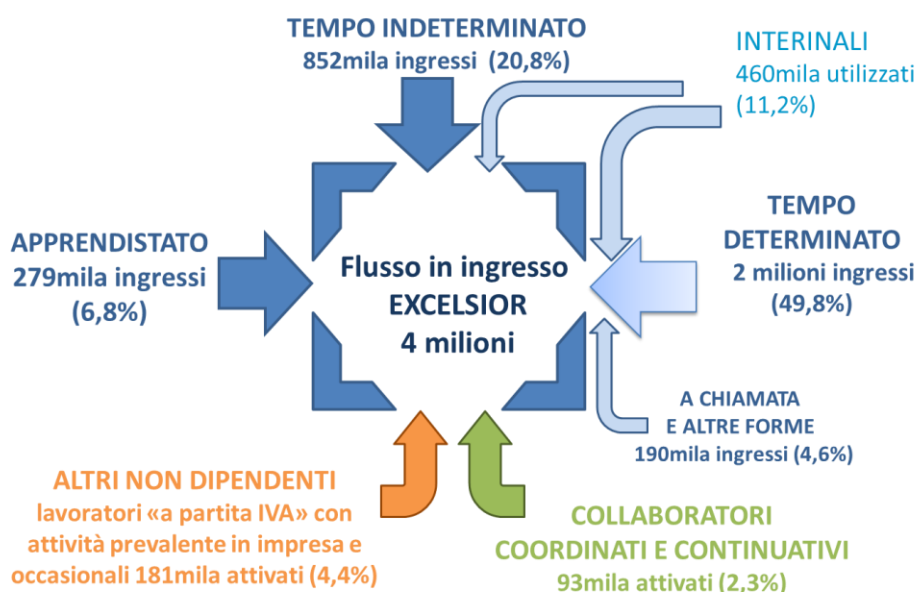
⁴² Settembre è anche il primo mese utile dopo la serie di tre wave "annuali", ovvero contenenti elementi di rilevazione ulteriori rispetto alla rilevazione dei flussi e più funzionali alla segmentazione delle imprese utile alla creazione di prodotti informativi, anche editoriali, che corrispondono alla diffusione dei dati annuali.

⁴³ Il corrispondente valore per l'anno 2017, ottenuto considerando l'elaborazione dei flussi ricavati dall'INPS sino al mese di marzo e la loro integrazione con quanto desunto dalle indagini Excelsior a cadenza mensile.

⁴⁴ Rispetto ai contribuenti registrati con questa forma contrattuale dall'INPS, si escludono alcune tipologie quali gli amministratori di società, che non costituiscono oggetto di interesse per la rilevazione Excelsior.

⁴⁵ Contratti professionali a partita iva o di natura occasionale.

Articolazione dei flussi in ingresso nel 2017



Fonte: Unioncamere – ANPAL, Sistema Informativo Excelsior, 2017

9.3 L'elaborazione delle caratteristiche dei flussi quantitativi

La seconda componente dalla quale le stime annuali dei flussi attingono informazioni è, come già anticipato, l'insieme delle interviste raccolte mensilmente dall'indagine continua che, stante la cadenza dell'elaborazione dell'indagine annuale, vengono raccolte sino ad agosto/settembre di ciascun anno:

Settore	Totale
Industrie alimentari, delle bevande e del tabacco	8.955
Industrie tessili, dell'abbigliamento, del cuoio e delle calzature	12.270
Industrie del legno e del mobile	9.512
Industrie della carta, cartotecnica e della stampa	4.683
Industrie dei gioielli e delle pietre preziose	1.537
Industrie estrattive	1.015
Industrie della lavorazione dei materiali da costruzione (vetro, ceramica, pietra)	4.218
Industrie elettriche ed elettroniche	8.367
Industrie della lavorazione dei metalli e dei prodotti in metallo	19.240
Industrie della fabbricazione e manutenzione di macchinari e apparecchiature	16.963
Industrie chimico-farmaceutiche	3.302
Industrie della gomma e delle materie plastiche	3.586
Altre industrie manifatturiere	1.422
Public utilities	7.817
Costruzioni	31.511
Commercio e riparazione di autoveicoli e motocicli	7.782
Commercio all'ingrosso	15.378
Commercio al dettaglio	18.115
Servizi turistici, di alloggio e ristorazione	22.837
Servizi informatici e delle telecomunicazioni	12.374
Servizi avanzati di supporto alle imprese	19.322
Servizi operativi di supporto alle imprese e alle persone	18.652
Servizi di trasporto, logistica e magazzinaggio	18.698
Servizi finanziari e assicurativi	9.391
Altri servizi alle imprese	2.971
Istruzione e servizi formativi privati	4.020
Sanità, assistenza sociale e servizi sanitari privati	9.529
Servizi ricreativi, culturali e altri servizi alle persone	9.898
TOTALE Nazionale	303.365

⁴⁶ Nel 2016 i flussi annuali erano valutati in 960mila attivazioni di contratti con una quota di imprese assuntivi pari al 19,5%. Tali numeri sono in parte giustificabili dal metodo di rilevazione non continuo, dalla lunghezza dell'orizzonte previsionale e dalla probabile de-duplicazione nel computo delle attivazioni dei contratti a termine.

Nel 2017 sono stati poco più di 300mila i questionari utilizzati nell'indagine annuale, un esito ben superiore alle tradizionali 100mila interviste delle passate edizioni annuali di Excelsior. Il numero delle acquisizioni utilizzate per il calcolo dei risultati dell'indagine annuale considera, oltre a quelle acquisite nel periodo aprile-agosto 2017, anche quelle acquisite nell'ultima parte del 2016 per le previsioni di gennaio-marzo 2017⁴⁷.

In un anno standard, che raccoglie interviste mensili a partire da gennaio e non da aprile come nel 2017, si stima che il numero delle interviste raccolte, funzionali alle stime annuali, possa superare le 400mila unità.

L'utilizzo dei risultati dell'indagine per la caratterizzazione dei flussi, ovvero la scomposizione dei flussi di entrata nelle figure professionali e relative caratteristiche, può beneficiare per la sua stima dei risultati della modellizzazione dei flussi.

Infatti la stima dei flussi quantitativi di cui al paragrafo 8.2 è resa disponibile, oltre che secondo le caratteristiche economiche delle imprese richiedenti, anche per diverse caratteristiche dei profili richiesti rilevati in sede d'indagine quali il livello di inquadramento (dirigenti, impiegati e operai) e le componenti demografiche di genere ed età del personale.

La correlazione diretta tra alcune caratteristiche delle figure professionali delle stime campionarie ed i risultati della modellizzazione dei flussi suggeriscono di utilizzare una particolare classe di stimatori a ponderazione vincolata (o calibrato) risulta indicato allo scopo anche grazie alla sua duttilità di impiego determinando i pesi di riporto all'universo in modo che le stime delle medie (totali) delle variabili ausiliarie ottenute con questi pesi sia uguali alle corrispondenti medie (totali) di popolazione.

$$\hat{Y} = \sum_{k=1}^n w_k y_k \qquad \sum_{k \in S} w_k x_k = X$$

rispettando un sistema di vincoli sintetizzato dalla relazione

In questo modo è possibile sfruttare l'informazione riferita ad alcune variabili ausiliarie riferite all'universo oggetto d'indagine per produrre stime coerenti (vincolate) ai marginali prefissati da modellistica; a ciò si deve aggiungere che tale procedura è anche particolarmente utile per la correzione dell'impatto delle mancate risposte, particolarmente indicato in un'indagine di tipo CAWI.

9.4 L'elaborazione delle sezioni specifiche del questionario annuale

Una terza componente deve essere ricompresa nelle stime dell'indagine annuale, quella relativa all'elaborazione delle domande del questionario "long form" annuale che consentono la segmentazione tematica delle imprese:

- sezione 4: investimenti in formazione di personale interno relativi all'anno precedente
- sezione 5: tirocini e stage attivati relativi all'anno precedente
- sezione 6: informazioni relative a possibili elementi di stratificazione delle imprese ti tipo tematico.

⁴⁷ Il 2017 è il primo anno sperimentale e di conseguenza, per coprire il periodo tra gennaio e marzo, si sono dovute necessariamente utilizzare le proiezioni dell'indagine trimestrale classica di Excelsior; negli anni successivi durante l'intero anno verrà utilizzata la tecnica dei trimestri mobili e si potrà disporre dei dati mensili per tutto l'anno.

L'elaborazione di queste sezioni consente la segmentazione delle imprese rispetto le tematiche di approfondimento direttamente collegate alle pubblicazioni tematiche di Excelsior annuale quali:

- La formazione continua, i tirocini formativi e l'alternanza scuola lavoro nelle imprese italiane;
- I fabbisogni occupazionali delle imprese esportatrici;
- I fabbisogni occupazionali delle imprese innovatrici;
- I fabbisogni di competenze digitali delle imprese;
- La domanda di "green jobs" nelle imprese italiane;
- Le imprese culturali.

In questo caso, non essendo in presenza di dati amministrativi o di una modellistica che consenta di vincolare i risultati a valori noti o pre-stimati, non è possibile utilizzare, come per i flussi, uno stimatore vincolato, e ci si rivolge quindi alle tecniche più classiche tipo Horvitz-Thompson.

In termini generali la procedura di inferenza scelta, sulla base della strategia campionaria, è quella di uno stimatore stratificato il cui stimatore corretto del totale \hat{Y} per il campionamento senza ripetizione è:

$$\hat{Y} = \sum_{i=1}^n \frac{y_i}{\pi_i}$$

dove y_i sono le osservazioni campionarie e π_i le probabilità di inclusione nel campione delle unità della popolazione a cui si riferiscono le osservazioni.

Poiché nel campionamento stratificato la probabilità dell'unità i -esima di essere inclusa nel campione dello strato h è

$$\pi_{hi} = \frac{n_h}{N_h}$$

pari cioè al tasso di sondaggio nello strato, allora si ricava immediatamente che lo stimatore corretto del totale sarà:

$$\hat{Y}_{ST} = \sum N_h \bar{y}_h$$

Come si vede, lo stimatore del campionamento stratificato è uguale alla somma di tanti stimatori per espansione quanti sono gli strati; le operazioni di inferenza sul singolo strato vengono condotte applicando il rapporto campionario rilevato alla numerosità delle unità contenute nello strato stesso secondo l'universo di riferimento.

Il campione raccolto nell'edizione del 2017, utile ai fini dell'elaborazione delle sezioni "annuali", è stato pari a circa 90mila interviste:

Settore	Totale
Industrie alimentari, delle bevande e del tabacco	2.932
Industrie tessili, dell'abbigliamento, del cuoio e delle calzature	3.261
Industrie del legno e del mobile	2.525
Industrie della carta, cartotecnica e della stampa	1.187
Industrie dei gioielli e delle pietre preziose	344
Industrie estrattive	199
Industrie della lavorazione dei materiali da costruzione (vetro, ceramica, pietra)	1.206
Industrie elettriche ed elettroniche	2.274
Industrie della lavorazione dei metalli e dei prodotti in metallo	5.377
Industrie della fabbricazione e manutenzione di macchinari e apparecchiature	4.916
Industrie chimico-farmaceutiche	641
Industrie della gomma e delle materie plastiche	959
Altre industrie manifatturiere	286
Public utilities	999
Costruzioni	10.541
Commercio e riparazione di autoveicoli e motocicli	1.995
Commercio all'ingrosso	4.383
Commercio al dettaglio	5.690
Servizi turistici, di alloggio e ristorazione	8.057
Servizi informatici e delle telecomunicazioni	4.056
Servizi avanzati di supporto alle imprese	6.304
Servizi operativi di supporto alle imprese e alle persone	5.516
Servizi di trasporto, logistica e magazzinaggio	5.502
Servizi finanziari e assicurativi	2.396
Altri servizi alle imprese	834
Istruzione e servizi formativi privati	841
Sanità, assistenza sociale e servizi sanitari privati	2.214
Servizi ricreativi, culturali e altri servizi alle persone	3.925
TOTALE Nazionale	89.360

corrispondente a quanto raccolto in sole 2 wave tra luglio ed agosto 2017; si tratta di una numerosità di interviste simile a quanto raccolto in passato per l'indagine annuale standard.

Per il triennio 2018-2020 si prevede di dedicare 3 wave alle sezioni aggiuntive "annuali" per un totale di 120-135mila acquisizioni in ciascuno dei 3 anni.

10 Strumenti di divulgazione dell'indagine annuale

10.1 Volumi statistici annuali

I volumi statistici annuali forniscono informazioni utili per orientare al meglio le politiche attive del lavoro e per agevolare l'incontro tra domanda e offerta di lavoro in Italia: i prodotti realizzati presentano una collezione delle tavole più significative tra quelle prodotte nella fase di spoglio delle domande presenti nel questionario di indagine (nella versione annuale). I volumi sono organizzati in sezioni/tematiche che permettono di analizzare, per settore di attività, professione e per livello di istruzione, più in dettaglio rispetto ai volumi mensili, gli orientamenti delle imprese, i fabbisogni occupazionali e le figure professionali richieste. Vengono anche analizzati in modo approfondito gli indirizzi di studio e le competenze richieste. Infine vengono approfonditi i temi della formazione del personale e dei percorsi di "Alternanza scuola lavoro", che stanno assumendo sempre più importanza nel dialogo tra mondo della formazione e mondo delle imprese.

Le sezioni di cui si compongono il volume nazionale e i volumi regionali e provinciali sono le seguenti:

SEZ. 1: QUALI SONO LE PROFESSIONI RICERCATE DALLE IMPRESE?

SEZ. 2: QUALI SONO LE COMPETENZE RICERCATE DALLE IMPRESE?

SEZ. 3: QUALI SONO GLI INDIRIZZI DI STUDIO RICHIESTI DALLE IMPRESE

SEZ. 4: ALTRE CARATTERISTICHE DELLE ENTRATE PREVISTE DALLE IMPRESE

SEZ. 5: PREVISIONI E ORIENTAMENTI DELLE IMPRESE IN ITALIA

SEZ. 6: ATTIVITA' FORMATIVE DELLE IMPRESE

Le aggregazioni settoriali utilizzate nei volumi statistici sono esaustive dei 28 settori evidenziati nel paragrafo 5.2.1 per il volume nazionale e regionale, ma a livello provinciale potrebbero non rappresentare adeguatamente la specificità economica del territorio: per ovviare a questa criticità vengono predisposte delle tavole dinamiche settoriali specifiche⁴⁸ per ogni singola provincia, enfatizzando le relative potenzialità economiche.

In aggiunta al volume nazionale, corredato con un'ampia sezione testuale di commento, e dei volumi territoriali, vengono prodotti una serie di volumi "tematici" che hanno come focus argomenti e temi di particolare interesse, che per il 2017 hanno riguardato:

- Le previsioni occupazionali di personale immigrato;
- La formazione continua, i tirocini formativi e l'alternanza scuola lavoro nelle imprese italiane;
- I fabbisogni occupazionali delle imprese esportatrici;
- I fabbisogni occupazionali delle imprese innovatrici;
- Le caratteristiche e i fabbisogni occupazionali delle nuove imprese, con particolare riferimento alle imprese giovanili;
- I fabbisogni di competenze digitali delle imprese;
- Laureati e lavoro;
- Diplomatici e lavoro;
- Il lavoro dopo gli studi: la domanda e l'offerta di laureati e diplomati;
- La domanda di formazione tecnica superiore delle imprese italiane;
- La domanda di green jobs nelle imprese italiane;
- Le imprese culturali.

⁴⁸ I settori di specializzazione provinciale sono basati su un algoritmo basato sul quoziente di specializzazione applicato agli occupati. Tale indice si ricava dal rapporto tra il quoziente del numero di dipendenti occupati in una determinata attività i nella provincia j e il totale dei dipendenti dell'area e il quoziente tra numero dei dipendenti nazionale nell'attività i ed il totale dei dipendenti in Italia. In formule:

$$Q = \frac{O_{ij} / O_j}{O_i / O_{..}}$$

dove

O_{ij} = ammontare degli occupati nel settore i -esimo nella provincia j -esima

O_j = totale degli occupati nella provincia j -esima

O_i = totale degli occupati in Italia nel settore i -esimo

$O_{..}$ = totale degli occupati a livello nazionale.

In una provincia il cui quoziente di specializzazione risulta maggiore di 1 è mediamente più specializzata in un determinato settore e tale specializzazione cresce al crescere dell'indice. Fissata una soglia convenzionale in relazione al quoziente di specializzazione e una soglia minima di numerosità dei soggetti appartenenti al dominio, è possibile definire gli elementi della matrice settori di attività per ciascuna delle provincie italiane.

10.2 Bollettini annuali

Analogamente a quanto viene prodotto mensilmente, è prevista la realizzazione di bollettini il cui orizzonte temporale si riferisce all'intero anno. Tenendo conto di una maggiore consistenza e significatività dei dati, questi rapporti permettono di trattare la maggior parte delle tematiche presenti nell'indagine.

L'articolazione ricalca quella dei prodotti mensili: un bollettino nazionale, 20 bollettini a livello regionale e 104 bollettini provinciali.

Il bollettino nazionale fornisce chiavi di lettura dell'evoluzione della realtà italiana (lo scenario economico-sociale che sta attraversando il paese e le trasformazioni in atto) cercando di cogliere i bisogni emergenti delle imprese e le professioni del futuro.

In aggiunta alle variabili oggetto di analisi nei rapporti mensili, i bollettini annuali territoriali trattano in modo più approfondito temi come le opportunità di lavoro per i giovani under 30, per la componente femminile, le professioni maggiormente richieste sul territorio, le competenze richieste per le figure assunte e la formazione del personale già presente in azienda.

10.3 Il sistema di interrogazione online

Le modalità di diffusione dei risultati del sistema informativo Excelsior comprendono anche il rilascio di banche dati interrogabili liberamente dagli utenti. Vengono messe a disposizione on-line (all'indirizzo <http://excelsior.unioncamere.net>, nella sezione Banca Dati, Excelsior On Line) tre categorie di archivi interrogabili in maniera dinamica:

- "settori": comprendente le previsioni di assunzione di dipendenti e di entrate tout-court per settore economico e caratteristiche associate;
- "professioni": comprendente le previsioni di assunzione di dipendenti e di entrate tout-court per professione e caratteristiche associate (le professioni sono espresse secondo la classificazione ISTAT, fino alla 4a cifra);
- "titoli di studio": comprendente le previsioni di assunzione di dipendenti e di entrate tout-court per indirizzo di studio e caratteristiche associate.

Le variabili di stratificazione presenti negli archivi sono le seguenti:

- relativamente all'impresa: settore di attività economica prevalente, ambito territoriale di appartenenza, classe dimensionale (in termini di dipendenti) e attività di carattere artigianale o non artigianale;
- relativamente alle figure professionali richieste dalle imprese: professione, titolo di studio, livello di inquadramento (solo per le assunzioni di dipendenti), classe di età, esperienza richiesta, tipologia della formazione richiesta, difficoltà di reperimento, genere, area funzionale e competenze trasversali.

A seconda del focalizzarsi dell'attenzione sul settore di attività economica, sulla professione o sul titolo di studio, le altre variabili di stratificazione sono indagate a livelli di dettaglio più o meno spinti.

Il sistema predisposto per la consultazione degli archivi appena descritti consente la più ampia possibilità e facilità di interrogazione. Nella sua impostazione si sono infatti curate funzioni di interfaccia "amichevole" con l'utente e la possibilità di risposta flessibile a qualsivoglia richiesta, superando la rigidità di interrogazioni limitate a tavole predefinite e non modificabili. Il sistema utilizzato supporta altresì l'esportazione di tabelle in formato di foglio elettronico e HTML.

11 Il modello previsionale a medio termine

A completamento dei processi previsionali del Sistema Informativo Excelsior, che vede l'elaborazione mensile come funzionale alle indicazioni di breve periodo nonché alla composizione dell'indagine statistica annuale, viene innestata un'attività di stima a medio termine che sfrutta le serie storiche delle indagini annuali Excelsior, relative alle caratteristiche della domanda di lavoro dipendente privato nonché le fonti disponibili per la domanda nel comparto pubblico⁴⁹, per la previsione dei fabbisogni occupazionali aprioristicamente definiti come somma algebrica di due componenti distinte:

- Expansion/reduction demand, intesa come la variazione degli stock di occupati totali prevista per i diversi settori economici;
- Replacement demand, come componente occupazionale derivante dalla necessità di sostituire le persone in uscita dal mercato del lavoro. In questa analisi sono state considerate le uscite per pensionamento e per mortalità, mentre non è stato per il momento considerato il possibile effetto delle sostituzioni richieste per spostamenti intersettoriali o interaziendali degli attuali occupati.

I dati utilizzati nell'analisi previsionale derivano integralmente da fonti Istat. In particolare sono stati utilizzati i dati dei Conti Economici Nazionali (Aprile 2017) per occupati, retribuzioni e output. Tutte le variabili sono espresse in termini reali, ove necessario è stato applicato il deflatore dell'output derivato dalle tavole relative alla produzione. Le serie storiche relative alle esportazioni provengono dalla base dati Coeweb; i dati sono in questo caso aggiornati a Luglio 2017.

Le serie storiche sono state ricostruite riaggregando i dati settoriali Istat (classificati secondo Ateco 2007) nei macrosettori Excelsior. Le previsioni sono state elaborate sia sugli occupati dipendenti che sugli occupati totali, ottenendo gli indipendenti come differenza tra le due serie.

11.1 La formulazione del modello

Le stime e le previsioni sono state effettuate utilizzando un Vector Autoregressive model (VAR) che descrive l'evoluzione dinamica di un certo numero di variabili a partire dalla dinamica evolutiva di ciascuna di esse e dalla dinamica comparata di tutte le variabili. Questo approccio è molto diffuso nell'ambito del forecasting macroeconomico in quanto i modelli VAR si sono dimostrati molto accurati e al tempo stesso poco dispersivi in quanto consentono di tener conto sia della dinamica temporale delle singole variabili che delle interazioni tra le stesse. Possiamo brevemente riassumere la metodologia VAR come segue: consideriamo due variabili y_1 e y_2 , un VAR di primo ordine (ovvero un VAR che contiene un solo ritardo delle variabili) può essere scritto come:

$$\begin{aligned}y_t &= \delta_1 + \varphi_{11}y_{t-1} + \varphi_{12}x_{t-1} + \varepsilon_{1t} \\x_t &= \delta_2 + \varphi_{21}y_{t-1} + \varphi_{22}x_{t-1} + \varepsilon_{2t}\end{aligned}$$

⁴⁹ Le statistiche di riferimento per l'analisi iniziale dei dati sono quelle relative ai posti di lavoro nella Pubblica Amministrazione elaborate dalla Ragioneria Generale dello Stato che, annualmente, rilevano la struttura del personale in servizio ed i movimenti relativi alle entrate e alle cessazioni del rapporto di lavoro (aggiornate fino al 2015). Per ulteriori approfondimenti sono state utilizzate – laddove disponibili – altre fonti statistiche, di natura più settoriale, per esempio, quelle del MIUR sul personale insegnante e non insegnante o quelle del personale di alcuni comparti nelle regioni a statuto speciale. Sulla base delle informazioni statistiche sopra citate, analizzate in serie storica, è stata definita una procedura che, a partire dai dati tendenziali rettificati con specifiche norme e disposizioni di legge, permetta di formulare previsioni sulla possibile evoluzione del quadro occupazionale ed in particolare sui flussi relativi alle entrate e alle uscite di personale relative ai tre macro-comparti del pubblico impiego, ovvero la Pubblica Amministrazione in senso stretto (Ateco 84), l'istruzione (Ateco 85) e la sanità (Ateco 86).

In forma vettoriale la formulazione diviene:

$$Y = \delta + \Phi_1 Y_{t-1} + \varepsilon$$

dove Y è un vettore di variabili, δ è un vettore di parametri, Φ è una matrice di parametri, ε il vettore degli errori. Il numero di ritardi è stato selezionato in base a criteri di efficienza (Akaike e Schwarz): questi criteri identificano il numero ottimale di ritardi cercando di trovare un compromesso tra la complessità del modello (più ritardi implicano un modello più complesso) e la bontà di adattamento dello stesso (in genere più ritardi spiegano meglio i dati).

Coerentemente con la letteratura è stata stimata una forma ridotta di una equazione di domanda di lavoro in cui l'occupazione è stata espressa come funzione dell'output e delle retribuzioni. A queste variabili sono state aggiunte le esportazioni, in forma settoriale per i settori industriali ed in forma aggregata per i servizi, nonché il tasso di crescita dell'economia aggregata.

I settori della sanità e dell'istruzione sono stati stimati, nella loro evoluzione, per l'insieme della componente pubblica e privata. L'andamento degli occupati della componente pubblica di questi settori – così come i “servizi generali della pubblica amministrazione” – è stato poi ristimato con apposite modalità.

Al fine di massimizzare la trasparenza ed interpretabilità è stato considerato come variabile esogena solo il tasso di crescita dell'economia aggregata e su di esso sono state effettuati gli scenari. In particolare, gli scenari sono stati formulati sulla base delle previsioni effettuate a novembre 2017 dall'Istat fino al 2019. Per gli anni 2020-2022 sono state utilizzate le previsioni formulate dal Fondo Monetario Internazionale nel suo World Economic Outlook Aprile 2017. In dettaglio, si prevede per l'economia Italiana un aumento dell'output dell'1,5%, per il 2017 e per gli anni 2018, 2019, 2020, 2021 e 2022 una crescita rispettivamente dell'1,4%, 1,1%, 1%, 0,94% e 0,85%.

Nella formulazione dello scenario positivo è stato considerato il contributo dell'OCSE che nella Economic Survey of Italy dell'Aprile 2016 ha stimato il contributo che l'approvazione di tutte le riforme progettate dal governo (la riforma della pubblica amministrazione, il completamento di quella del mercato del lavoro, la riforma istituzionale ecc.) porterebbe alla crescita economica. Sommarariamente il successo nell'implementazione delle riforme si potrebbe tradurre in una crescita aggiuntiva del PIL dello 0.7% l'anno. La tabella di seguito riassume le assunzioni relative all'andamento del tasso di crescita del PIL aggregato (valori espressi in percentuale).

	2017	2018	2019	2020	2021	2022
Benchmark	1,5	1,4	1,1	1,0	0,94	0,85
Positivo	2,2	2,1	1,8	1,7	1,64	1,55

A livello settoriale è stato considerato come esogeno il valore delle esportazioni. Nella formulazione degli scenari sono state utilizzate le previsioni dell'analisi dei settori industriali effettuato da Prometeia-Banca Intesa (Maggio 2017).

A partire dalle ipotesi di base descritte, sono state quindi determinate le variazioni degli stock di occupati totali, dei dipendenti e degli indipendenti previste per i diversi settori economici (tale variazione corrisponde alla expansion/reduction demand), sia per lo scenario benchmark, sia per lo scenario positivo.

Al termine della fase appena descritta si è pervenuti a dei fabbisogni occupazionali per settore (per ciascuno dei due scenari): si è reso, quindi, necessario disaggregare questi stessi fabbisogni per ripartizione geografica.

Per ogni settore e anno si è utilizzata la distribuzione Excelsior per territorio degli stock di dipendenti e quella di Forze Lavoro (FL) per gli indipendenti, calcolando rispettivamente la media mobile di 7 variazioni degli stock Excelsior (considerando il periodo 2009-2016) e di 6 variazioni degli stock FL (periodo 2010-2016), per poi applicare i tassi ottenuti ai consueti universi di riferimento di Contabilità Nazionale.

11.2 La stima della “replacement demand”

Una volta definite le variazioni degli stock di occupati, ci si è posto l’obiettivo di stimare la componente occupazionale in uscita dal mercato del lavoro che dà luogo alla cosiddetta replacement demand: a questo fine sono state considerate le componenti prevedibili in relazione alle uscite per pensioni e per mortalità.

La base di partenza per la stima dei flussi in uscita per pensionamento sono i dati sulle pensioni liquidate nel 2016 per i lavoratori dipendenti del settore privato e per gli indipendenti (autonomi e collaboratori), per classe di età e genere (di fonte INPS),

Applicando i dati così ottenuti agli stock di dipendenti e indipendenti per età e genere, ricostruiti utilizzando i dati sulla struttura per età degli occupati di Contabilità Nazionale per l’anno 2015, si sono stimati i tassi di pensionamento per età e per genere del 2016.

I tassi così ottenuti sono stati utilizzati per ricalcolare le uscite per pensionamento di quell’anno, moltiplicandoli per i rispettivi stock di riferimento al 31 dicembre 2015.

I tassi di pensionamento calcolati per il 2016 sono poi stati utilizzati come base di partenza per la stima dei tassi degli anni successivi - dal 2017 al 2022 – sempre distinti per età, genere e posizione professionale, applicando i principi stabiliti dalla normativa attualmente vigente, considerando quindi anche il recente adeguamento dei requisiti per l’accesso alla pensione in seguito all’innalzamento della speranza di vita certificato dall’Istat e gli effetti delle norme contenute nelle Leggi di Bilancio 2016 e 2017⁵⁰.

In particolare, per quanto riguarda le stime relative al periodo 2018-2022 si è tenuto conto dei trascinalamenti (in termini di maggiori flussi in uscita) delle novità che erano state introdotte in tema di pensioni nella legge di Bilancio 232/2016 e delle ulteriori modifiche previste dalla legge di Bilancio 205/2017. Le informazioni sui maggiori flussi pensionistici in uscita sono state tratte dai documenti e dalle relazioni illustrative e tecniche che accompagnano le succitate leggi.

Inoltre, per quanto riguarda il 2017, le stime dei tassi per età e genere sono state migliorate sulla scorta dei dati INPS⁵¹ sulle pensioni liquidate nel 2016 e nel 2017 distintamente per lavoratori dipendenti e autonomi, che hanno fornito una indicazione della variazione percentuale di tali flussi di pensioni.

Per quanto riguarda la determinazione delle uscite per mortalità, sono stati utilizzati i tassi di mortalità di fonte ISTAT per sesso e per età relativi all’anno 2016.

11.3 La stima dei fabbisogni occupazionali

A partire dalle due componenti del fabbisogno (expansion e replacement demand) per anno, settore e per ripartizione geografica, sono stati definiti i fabbisogni occupazionali in complesso.

⁵⁰ Legge 11 dicembre 2016, n. 232, Bilancio di previsione dello Stato per l’anno finanziario 2017 e bilancio pluriennale per il triennio 2017-2019; Legge 27 dicembre 2017, n. 205, Bilancio di previsione dello Stato per l’anno finanziario 2018 e bilancio pluriennale per il triennio 2018-2020.

⁵¹ INPS, Monitoraggio dei flussi di pensionamento - Pensioni liquidate alla data del 2 gennaio 2018 con decorrenza entro dicembre 2017 (pubblicato il 18 gennaio 2018).

L'obiettivo successivo è consistito nel ripartire i fabbisogni ottenuti per professione e indirizzo di studio. Per la stima di queste variabili sono state pienamente valorizzate le serie storiche dei flussi occupazionali derivanti dalle indagini Excelsior, rese tra loro previamente compatibili.

Tali basi dati non possono tuttavia fornire un contributo laddove il fabbisogno occupazionale assume un valore negativo. Infatti, la variabile fabbisogno per sua natura può assumere valori sia positivi che negativi. I fabbisogni negativi si possono ottenere quando la "reduction demand" è superiore in valore assoluto alla "replacement demand". In questo caso per determinare il fabbisogno a livello professionale e per titolo di studio sono stati utilizzati i dati in serie storica delle Forze di Lavoro Istat (FL).

11.3.1 Determinazione dei fabbisogni per professione e titolo di studio

La metodologia si differenzia per addetti dipendenti e indipendenti: per i dipendenti, a seconda che il fabbisogno occupazionale assuma un valore positivo o negativo, come già accennato, sono state effettuate proiezioni sulla base delle tendenze delle serie storiche Excelsior o di Forze di Lavoro Istat (FL). Per gli indipendenti è stata utilizzata invece unicamente la base dati Forze di Lavoro Istat (FL).

Per quanto riguarda i fabbisogni positivi dei dipendenti privati sono stati considerati i dati delle ultime 8 indagini annuali Excelsior, mentre per i restanti fabbisogni (fabbisogno negativo dei dipendenti privati e fabbisogno positivo e negativo degli indipendenti) sono stati considerati i dati medi annui delle ultime 7 annualità della Rilevazione FL.

Tale procedimento è stato eseguito in modo indipendente per professioni e titoli di studio, pertanto non è possibile incrociare tra loro queste variabili.

11.3.2 Determinazione dei fabbisogni nella Pubblica Amministrazione

Partendo dagli ultimi dati disponibili (RGS 2015) e tenendo conto delle Leggi e delle norme riguardanti il personale nella Pubblica Amministrazione, attraverso una specifica procedura sono state ricavate in una prima fase le stime relative agli stock e ai flussi per gli anni 2016 e 2017. Esse sono state determinate considerando le serie storiche degli stock e dei flussi degli ultimi anni, per ciascun comparto, a partire dal 2010 e fino al 2015.

In una seconda fase di lavoro sono stati stimati i flussi per il quinquennio 2018-2022, facendo riferimento, oltre che alla qualifica del personale entrato/uscito, anche alle informazioni riguardanti le modalità di assunzione e i motivi di cessazione del rapporto di lavoro

I flussi di uscita del periodo 2018-2022 sono stati determinati tenendo anche conto della composizione per età e per genere dei dipendenti pubblici nei tre macro-comparti di attività (integrando i dati di RGS con l'indagine sulle Forze di Lavoro, Istat). Ciò ha permesso di stimare sia i flussi pensionistici, a partire dalle normative di legge in vigore, ed analogamente le uscite a seguito di cause naturali (mortalità) e quelle relative ad "altre cause" (dimissioni, ecc.), così da ottenere dati ancor più in dettaglio nel modello previsionale riguardo le uscite complessive per ogni anno, distintamente per ogni comparto.

Il flusso delle entrate a partire dall'anno 2018 e fino al 2022 è stato stimato tenendo conto delle normative e dei parametri che attualmente regolano le entrate stesse in relazione alle uscite dell'anno precedente.

Occorre precisare che per la Pubblica Amministrazione la metodologia per la determinazione dei "fabbisogni" avviene con metodologie peculiari rispetto al settore privato. Nel pubblico impiego infatti le uscite complessive della P.A. (ovvero pensionamento per età o per anzianità, decesso, dimissioni o altre cause) rappresentano tutto il potenziale "replacement" sostituibile. In base alla loro consistenza, suddivisa per singolo comparto, e applicando le disposizioni relative alle assunzioni (per esempio: 8 entrate ogni 10

uscite avvenute nell'anno precedente), sono state poi calcolate le entrate. È quindi ragionevole ritenere che le "entrate" nella Pubblica Amministrazione siano coincidenti con i "fabbisogni" di personale. La determinazione annuale dello stock di personale, e dei relativi saldi, è pertanto il risultato della differenza tra entrate e uscite.

Tali entrate/fabbisogni sono poi stati disaggregati per gruppi professionali, a partire dai dati elementari rilevati dall'Istat con le indagini sulle Forze di Lavoro (FL), relative agli ultimi anni disponibili. Sulla base delle rilevazioni Istat - FL e dei relativi dati "elementari" è stato possibile quantificare alcune caratteristiche occupazionali del personale impiegato nella Pubblica Amministrazione, con riferimento sia alla consistenza in complesso del personale, che al segmento di personale occupato nella P.A. da 3 anni o meno (supponendo che tale dato possa essere una proxy della composizione delle entrate più recenti).

Una volta definita la composizione delle assunzioni, sono state stimate quelle per gruppi professionali sulla base dei dati tendenziali e successivamente confrontate ed allineate con le assunzioni relative alle professioni dettagliatamente rilevate all'interno dei dati RGS. Un procedimento simile si è utilizzato anche per la disaggregazione dei fabbisogni ottenuti per livello di istruzione e titolo di studio.

11.4 la stima dell'offerta di diplomati e laureati

Il modello previsionale di offerta è costituito da una serie di moduli sequenziali tra loro per lo sviluppo dei quali vengono utilizzate diverse fonti statistiche, prevalentemente MIUR e ISTAT.

Il modello di stima dei diplomati si fonda sulla proiezione degli iscritti al primo anno sulla base dell'evoluzione prevista per la classe di età dei quattordicenni (di fonte Istat) ai quali si applicano tassi di iscrizione e tassi di conseguimento del diploma calcolati in base alle tendenze storiche delle serie disponibili. Per i percorsi post-diploma vengono invece utilizzate, da un lato, le statistiche degli immatricolati secondo il diploma posseduto, dall'altro parametri tratti dalle indagini ISTAT sui percorsi di studio e lavoro post diploma.

Il modello di previsione dei laureati è stato sviluppato separatamente per tre tipologie di corso: corsi triennali (sia del vecchio che del nuovo ordinamento), corsi lunghi (corsi di 4-5 anni del vecchio ordinamento e corsi a ciclo unico del nuovo) e corsi specialistici (biennali, del nuovo ordinamento). Ciascuna tipologia, a sua volta, viene sviluppata per singolo gruppo di corsi.

Per la previsione dei laureati in uscita dall'università, i dati disponibili sono organizzati su una matrice iscritti al 1° anno/laureati, all'interno della quale si possono evidenziare gli iscritti al 1° anno di alcune annualità complete, per le quali sono disponibili i laureati nei 10 anni successivi all'inizio degli studi. Per le annualità disponibili in modo completo è quindi possibile calcolare un tasso di conseguimento della laurea nell'arco di 10 anni dall'immatricolazione, a sua volta ripartito in tassi specifici a 2, 3, 4 e fino a 10 anni. I laureati degli anni non ancora disponibili saranno quindi la somma degli immatricolati ai quali viene via via applicato il tasso di conseguimento della laurea a "n" anni dall'immatricolazione.

Il rapporto tra i laureati triennali totali e gli iscritti al 1° anno dei corsi specialistici consente inoltre di calcolare il tasso di prosecuzione degli studi dal 1° al 2° ciclo. Questo tasso è ormai stabilizzato da anni e viene utilizzato per stimare coloro che iniziano i corsi specialistici (per gli anni non disponibili); questi a loro volta, vengono detratti dai laureati triennali (lordi) di ciascun anno, determinando quindi i laureati "netti" potenzialmente destinati all'ingresso sul mercato del lavoro. Per la stima dei neo-laureati in ingresso sul mercato del lavoro il modello utilizza i risultati della rilevazione AlmaLaurea, che pur non avendo una copertura totale dell'universo delle università italiane, presenta un dettaglio delle informazioni particolarmente interessante ed è aggiornato annualmente per i diversi tipi e gruppi di corsi.