

I valori relativi alla percentuale di coppie di osservazioni concordi/discordi/*tied*⁷⁹ (Tavola V.3) indicano una buona capacità predittiva del modello logistico.

TAVOLA V.3 ASSOCIAZIONE TRA PROBABILITÀ STIMATE E VALORI OSSERVATI

Coppie	Percentuale
Concordi	85,6%
Discordi	14,2%
<i>Tied</i>	0,2%

Numero coppie: 3.565.540

La Tavola V.4 riporta i parametri stimati per le singole variabili.

⁷⁹ Per tutte le possibili coppie di osservazioni con valori differenti della Y (0 o 1, dove 1 è l'evento di interesse) una coppia è classificata come *concorde* se la probabilità stimata dal modello per l'osservazione con $y=0$ è minore di quella stimata per $y=1$; in caso contrario è classificata come *discorde*, mentre nel caso in cui la probabilità sia uguale è classificata come *tied*.

TAVOLA V.4 MODELLO LOGISTICO RELATIVO ALLA PROBABILITÀ DI AVVIO IN RITARDO (Y=1: avvio in ritardo SI)

Variabile	Modalità	β	Livello di significatività	
Intercetta		-4,15	***	
Log costo		0,41	***	
Delibera	No FAS	-1,35	***	
	Delibere precedenti 36/2002 - 17/2003	1,97 0,00	***	
Tipo di intervento	Altro	1,41	*	
	Nuova realizzazione	2,93	***	
	Recupero-restauro	-5,35	***	
	Servizi	0,00		
Area	Centro	1,78	*	
	Nord	4,42	***	
	Sud	0,00		
Durata prevista	Fino ad 2 anni	0,80	***	
	3 anni	0,35	***	
	4 anni	-0,31	+	
	5 anni e più	0,00		
Log costo * Tipo intervento	Altro	-0,18	**	
	Nuova realizz.	-0,17	***	
	Recupero-restauro	0,44	***	
	Servizi	0,00		
Log costo * Area	Centro	-0,10	*	
	Nord	-0,27	***	
	Sud	0,00		
Area * Durata prevista	Centro	Fino a 2 anni	0,18	
	Centro	3 anni	0,01	
	Centro	4 anni	0,29	
	Centro	5 anni e più	0,00	
	Nord	Fino a 2 anni	-1,23	***
	Nord	3 anni	-0,63	***
	Nord	4 anni	-0,10	
	Nord	5 anni e più	0,00	

*** : $p < 0,001$

** : $0,001 \leq p < 0,01$

* : $0,01 \leq p < 0,05$

+ : $p \geq 0,05$

Analizzando in dettaglio le singole variabili:

- **Costo:** la probabilità che l'intervento parta in ritardo, in generale, aumenta con la dimensione economica dell'intervento. L'effetto di questa variabile varia però per area geografica e tipo di realizzazione: l'effetto è infatti più basso al Nord e al Centro rispetto al Sud, come indicano i valori negativi dei parametri associati a questa variabile;
- **Delibera di riferimento:** gli interventi che non prevedono risorse FAS hanno, nel complesso, una minore probabilità di ritardo rispetto agli interventi delle delibere recenti (36/2002 e 17/2003). Al contrario gli interventi relativi alle delibere precedenti hanno probabilità più elevate;
- **Tipo di intervento:** rispetto ai servizi, le opere di recupero e restauro hanno, in generale, una maggiore probabilità di partire nei tempi previsti mentre le nuove realizzazioni e le altre tipologie vengono avviate più tardi del previsto. Tuttavia l'effetto associato all'interazione con l'importo indica che il divario tra le diverse categorie si riduce all'aumentare del costo complessivo: il parametro associato all'interazione tra le varie tipologie di intervento ed il logaritmo dell'importo è infatti di segno opposto a quello associato all'effetto semplice;
- **Area geografica:** la stima relativa alla localizzazione geografica (considerata come effetto semplice) sembrerebbe evidenziare una maggior difficoltà a partire nei tempi previsti al Centro-Nord piuttosto che al Sud. Il valore dei parametri associati all'interazione di questa variabile con il costo complessivo e la durata prevista indica, però, che le differenze si riducono per interventi economicamente più rilevanti o, per il Nord, per interventi con durate previste più brevi (meno di 3 anni);
- **Durata prevista:** rispetto agli interventi con durata superiore a 3 anni, gli interventi più brevi (fino a 2 anni) hanno una maggior probabilità di ritardo. Per le durate intermedie il parametro è negativo ma non significativo.

Estendendo i risultati del modello agli interventi non ancora in ritardo e considerando in ritardo gli interventi con probabilità stimata maggiore di 0,5, risulta in ritardo il 77 per cento degli interventi, mentre solo il restante 23 per cento parte nei tempi previsti.

Il modello di durata per la previsione del ritardo nell'avvio

Per poter prevedere la data di inizio dell'esecuzione dei lavori per gli interventi in ritardo non ancora avviati, sono stati utilizzati modelli di durata o *hazard models* (Modello di durata in Figura V.1). Questo tipo di modelli permette di sfruttare sia dati di interventi già avviati, per cui la data di avvio è già nota e dunque il ritardo è effettivo, sia dati di interventi non ancora avviati, per cui il ritardo è un ritardo accumulato e si sa solo che la data effettiva sarà successiva all'ultima data osservata. Nel primo caso si parla di osservazioni *non censurate*, mentre nel secondo di osservazioni *censurate*.

In questo tipo di modelli viene parametrizzato il rischio che un certo evento si verifichi in un certo istante di tempo t , a condizione che non si sia verificato fino all'istante $t-\Delta t$; indirettamente si ottiene una stima della durata del tempo t fino all'accadimento dell'evento.

La funzione di rischio è definita come:

$$h(t) = \lim_{\Delta t \rightarrow 0} \Pr(t \leq T < t + \Delta t \mid T \geq t) \quad (4)$$

dove T è la variabile casuale che misura il tempo di accadimento dell'evento. La funzione $h(t)$ si può interpretare come la probabilità istantanea che si verifichi un certo evento che ancora non si è verificato.

Nei modelli di durata si assume che il logaritmo della funzione di rischio $h(t)$ sia una funzione lineare del tempo t e di una serie di variabili esplicative, $X_1 \dots X_n$, che rappresentano le caratteristiche degli interventi:

$$\log[h(t)] = f(t, x_1, \dots, x_n) = \beta_0 t + \beta_1 x_1 + \dots + \beta_n x_n \quad (5)$$

A particolari ipotesi sulla dipendenza del rischio $h(t)$ dal tempo corrispondono assunzioni sulla distribuzione di probabilità della variabile casuale tempo T ⁸⁰. Si

⁸⁰ Cox (1972) ha proposto un modello 'semiparametrico' alternativo che non presuppone alcuna assunzione per la distribuzione di probabilità della variabile tempo T .

parla in questo caso di modelli parametrici, e le distribuzioni più comuni sono la esponenziale o la Weibull.

Poiché siamo interessati a stimare la durata del ritardo, che nel nostro caso equivale alla variabile casuale T , è possibile considerare l'effetto delle variabili esplicative sulla T , invece che sul rischio. Possiamo, quindi, scrivere:

$$\log T = \alpha_0 + \alpha_1 x_1 + \dots + \alpha_n x_n + \sigma \varepsilon \quad (6)$$

Questa formulazione, che viene utilizzata nel nostro sistema, non permette di stimare direttamente un valore di T , ma i parametri della sua distribuzione, in quanto le osservazioni di T possono essere censurate. Gli α_i sono i parametri incogniti associati alle variabili esplicative X_i^{81} , α_0 è l'intercetta e σ è il parametro associato alla variabile casuale ε , che coglie l'allontanamento del valore di $\log T$ dalla parte lineare del modello.

A partire dai risultati del modello è possibile ottenere una stima della durata del ritardo di avvio utilizzando il valore mediano⁸² della distribuzione di T . La data di avvio prevista è pertanto ottenuta aggiungendo alla data dichiarata alla stipula la durata del ritardo stimata.

Stima del ritardo per gli APQ

Per la stima e la previsione della durata del ritardo all'avvio degli interventi è stato utilizzato un modello di durata di tipo parametrico.

Nel caso degli APQ definiamo come durata del ritardo il tempo T che intercorre fra la data di avvio prevista alla stipula e quella effettiva. Per gli interventi avviati si hanno, quindi, durate non censurate, mentre per quelli non avviati le durate sono censurate e vengono calcolate considerando il tempo trascorso fino all'ultima data di monitoraggio.

⁸¹ Esiste una relazione diretta tra i parametri β_i della (5) e gli α_i della (6), che dipende dalla distribuzione assunta per la T . In generale, se β_i misura l'effetto della variabile X_i sulla funzione di rischio e α_i l'effetto della stessa variabile sul tempo trascorso fino al verificarsi dell'evento, a valori elevati del primo, corrispondenti ad una elevata probabilità che l'evento accada, corrisponderanno valori contenuti del secondo, associati a tempi rapidi di accadimento. Cfr. Collett (1994).

⁸² Nel caso di modelli di durata, in cui si considerano anche osservazioni censurate oltre a durate effettive non censurate, è più appropriato utilizzare la durata mediana invece della durata media. Infatti, la durata media di sopravvivenza stimata non sempre è calcolabile nei modelli non parametrici, nel caso in cui sia particolarmente elevata la frequenza di durate censurate.

Per la T si è assunta una distribuzione Weibull, che corrisponde ad una distribuzione del rischio $h(t)$ monotona.

I dati utilizzati si riferiscono ai 3.425 interventi, avviati e non, in ritardo alla data dell'ultima versione di monitoraggio considerata⁸³.

Tra gli interventi con durate non censurate, cioè con ritardi definitivi, il ritardo osservato ha una distribuzione piuttosto larga con media di 226 giorni e deviazione standard di 180 giorni. Considerando anche le durate censurate, cioè i ritardi non ancora definitivi, si ottengono valori confrontabili.

Poiché il modello di durata utilizzato appartiene alla classe di modelli 'a rischi proporzionali', sono state scartate le variabili per cui non è verificata la validità dell'ipotesi di proporzionalità del rischio osservato rispetto alle diverse modalità considerate⁸⁴. Pertanto le variabili utilizzate sono:

- Area geografica di localizzazione dell'intervento (Nord, Centro, Sud);
- Tipo di intervento (Nuova realizzazione, Recupero - restauro, Altro tipo di realizzazione, Servizi);
- Logaritmo del costo dell'intervento;
- Risorse FAS (No, Si);

dove le modalità di riferimento delle variabili categoriche sono indicate in corsivo.

Sono stati considerati sia effetti semplici che interazioni di primo livello.

Risultati del modello

Ad esclusione della variabile relativa alle risorse FAS, tutte le variabili incluse nel modello risultano fortemente significative (Tavola V.5).

⁸³ Dalla stima sono stati esclusi i casi anomali, corrispondenti agli interventi con durate del ritardo superiore al 90° percentile. L'esclusione di queste osservazioni ha determinato un migliore adattamento della distribuzione della T alla Weibull.

⁸⁴ La variabile 'Asse' inizialmente considerata, per cui questa ipotesi non è risultata accettabile, è stata successivamente esclusa dal modello.

TAVOLA V.5 VARIABILI ESPLICATIVE INCLUSE NEL MODELLO DI DURATA

Variabile	Livello di significatività
Area	***
Risorse FAS	+
Log costo	***
Tipo di intervento	***
Log costo * Tipo di intervento	***
*** :	$p < 0,001$
** :	$0,001 \leq p < 0,01$
* :	$0,01 \leq p < 0,05$
+ :	$p \geq 0,05$

In sintesi i parametri del modello (Tavola V.6) mostrano che la durata del ritardo è più lunga al Centro ed al Sud piuttosto che al Nord. Le opere di recupero e restauro hanno ritardi più brevi rispetto ai servizi; il ritardo tende, tuttavia, ad aumentare con il costo complessivo. Per le nuove realizzazioni e le altre tipologie di intervento i parametri stimati hanno valori di segno opposto (negativo nel primo caso, positivo nel secondo), tuttavia i valori non risultano significativi.

TAVOLA V.6 MODELLO DI DURATA RELATIVO ALLA DURATA DEL RITARDO ALL'AVVIO

Variabile	Modalità	β	Livello di significatività
Intercetta		5,76	***
Area	Centro	0,10	+
	Nord	-0,41	***
	<i>Sud</i>	0,00	
Log costo		0,03	+
Risorse FAS	No	-0,01	+
	<i>Si</i>	0,00	
Tipo di intervento	Altro	1,15	+
	Nuova realizzazione	-1,56	+
	Recupero-restauro	-2,90	+
	<i>Servizi</i>	0,00	
Log costo * Tipo di intervento	Altro	-0,06	+
	Nuova realizzazione	0,13	+
	Restauro – recupero	0,24	**
	<i>Servizi</i>	0,00	
<i>Weibull shape (γ)</i>		1,17	

*** : $p < 0,001$

** : $0,001 \leq p < 0,01$

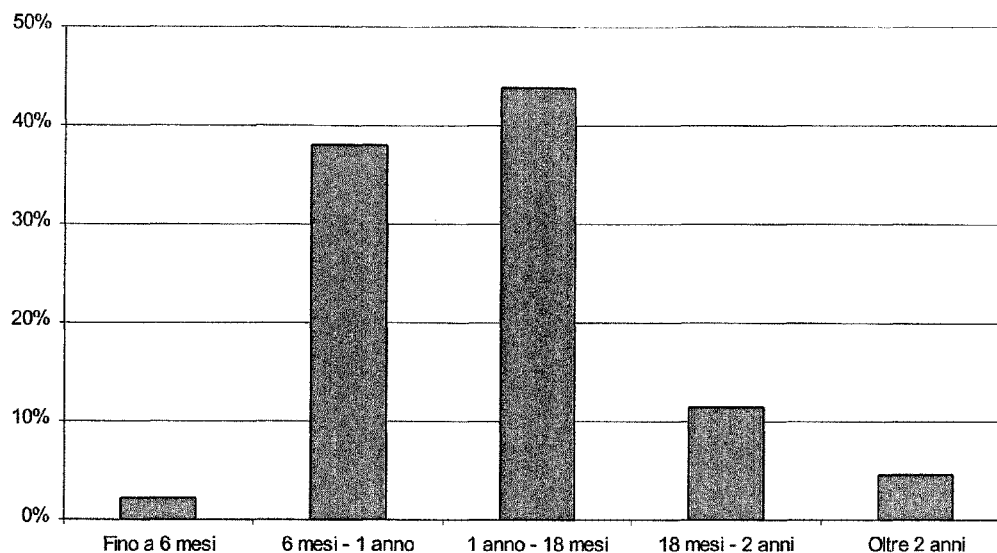
* : $0,01 \leq p < 0,05$

+ : $p \geq 0,05$

Il parametro γ pari a 1,17 indica che la funzione di rischio $h(t)$ e dunque la probabilità di partire in ritardo è crescente nel tempo, con incrementi progressivamente minori.

La durata stimata del ritardo viene quindi utilizzata solo per gli interventi che secondo il modello precedente hanno una probabilità di ritardo maggiore di 0,5. Nella Figura V.2 viene riportato il grafico della percentuale di interventi per classe di ritardo.

FIGURA V.2 PERCENTUALE DI INTERVENTI PER CLASSE DI RITARDO STIMATO PER GLI INTERVENTI NON ANCORA AVVIATI



Si prevede quindi che circa il 40 per cento degli interventi ritardatari partirà con un ritardo fino ad 1 anno, oltre il 40 per cento tra 1 anno ed 1 anno e mezzo e un ulteriore 10 per cento circa non supererà i due anni. La distribuzione dei ritardi ha un valore medio di 417 giorni, con una deviazione standard di 160 giorni, superiore a quella dei ritardi effettivi già registrati: si prevede cioè che, in media, gli interventi non avviati subiranno un ritardo di avvio maggiore rispetto a quello degli interventi già avviati.

Il modello di spesa

Il modello si basa sulle informazioni relative alla spesa realizzata annualmente dai singoli interventi (Modello di spesa in Figura V.1).

La variabile dipendente del modello è la percentuale cumulata di spesa realizzata⁸⁵ fino al tempo t . Nel modello il tempo t rappresenta una variabile discreta che consente di esprimere la spesa per un intervento in funzione della corrispondente spesa realizzata fino al periodo precedente. In sostanza, si ha:

$$p_t = f(p_{t-1}, \bar{X}) \quad (9)$$

⁸⁵ Espressa in termini relativi, ovvero compresa tra 0 ed 1.

dove p_t è la percentuale cumulata di spesa realizzata fino al tempo t e \bar{X} un vettore di variabili esplicative.

L'utilizzo della cumulata di spesa al tempo $t-1$ come variabile esplicativa è un aspetto centrale del modello e consente, in fase previsiva, di ottenere per ogni periodo valori più affidabili e realistici, perché in linea con i valori precedenti.

L'ipotesi alla base del modello è che la spesa cumulata abbia una forma che è ben approssimata dalla logistica; dunque la relazione assunta tra le variabili esplicative X_i e la variabile dipendente del modello è del tipo:

$$p_t = \frac{\exp(\beta_0 + \beta_1 p_{t-1,i} + \beta_2 x_{2,i} + \dots + \beta_n x_{ni})}{1 + \exp(\beta_0 + \beta_1 p_{t-1,i} + \beta_2 x_{2,i} + \dots + \beta_n x_{ni})} \quad (10)$$

Si assume dunque che la relazione tra la variabile di spesa e le variabili esplicative sia di tipo non lineare.

Per la stima dei parametri, la variabile dipendente, che assume valori compresi nell'intervallo $[0,1]$ è stata opportunamente trasformata attraverso la funzione *logit*, che assume valori nell'intervallo $(-\infty, +\infty)$.

$$y_t = \log \frac{p_t}{1-p_t} = \beta_0 + \beta_1 y_{t-1,i} + \beta_2 x_{2,i} + \dots + \beta_n x_{ni} \quad (11)$$

Il modello finale è dunque un modello di regressione lineare sulla trasformata *logit* della percentuale cumulata di spesa.

Stima della spesa degli APQ: trasformazione dei dati su base trimestrale ed applicazione del modello

L'introduzione nel modello (11) della spesa realizzata al tempo $t-1$ come variabile esplicativa, oltre a ridurre il numero di osservazioni valide per la stima, potrebbe fornire risultati distorti in quanto, nel caso di spesa annuale, sarebbero esclusi tutti gli interventi un solo anno di spesa. Pertanto si è ritenuto opportuno espandere la base informativa ricostruendo, attraverso un'interpolazione lineare, la serie di dati di spesa su base trimestrale.

Per gli interventi APQ, dunque, la stima del modello si riferisce alla cumulata di spesa realizzata fino ad un determinato trimestre t . Le previsioni della spesa sono state successivamente ricondotte ad intervalli di tempo annuali.

Come variabili esplicative, oltre alla percentuale di spesa realizzata fino al trimestre $t - 1$, sono state considerate:

- **la percentuale di spesa cumulata fino ai trimestri t e $t - 1$ prevista nella dichiarazione di spesa fornita alla data di stipula dell'APQ.** Il piano di spesa presentato alla stipula rappresenta la previsione iniziale, degli stessi responsabili dell'intervento, della spesa annuale fino alla conclusione dell'intervento⁸⁶. Come per la variabile dipendente, per tutte le variabili esplicative che rappresentano una percentuale cumulata di spesa viene utilizzata la trasformata *logit*⁸⁷;

- **l'inizio dei lavori in ritardo rispetto a quanto dichiarato alla stipula.** Alla base della scelta di considerare questa variabile dicotomica come esplicativa del comportamento di spesa degli interventi vi è l'ipotesi che problemi nella fase di progettazione e/o affidamento all'origine del ritardo nell'avvio possano determinare una maggiore lentezza e dunque ulteriori ritardi anche nella realizzazione vera e propria dell'intervento. Se questa ipotesi fosse verificata già dalle fasi che precedono l'esecuzione dei lavori sarebbe possibile individuare in anticipo eventuali criticità in fase realizzativa;

- **variabili relative a caratteristiche strutturali dell'intervento.** Si tratta di variabili categoriche, quali:

- Area geografica di localizzazione dell'intervento (Nord, Centro, *Sud*);
- Asse comunitario dell'accordo (Città, Reti e nodi di servizio, Risorse culturali, Risorse naturali, *Sistemi locali e Risorse umane*);
- Tipo di intervento (Nuova realizzazione, Recupero - restauro, Altro tipo di realizzazione, *Servizi*);
- Stato dell'intervento (Attivo⁸⁸, *Concluso*);
- Durata prevista alla stipula (Fino a 2 anni, 3 anni, 4 anni, *5 o più*);
- Classe di costo (*Bassa, Media, Alta*).

⁸⁶ Questa informazione è da ritenersi tanto più rilevante a fini previsivi quanto minore è il tempo trascorso dalla stipula. Le dichiarazioni possono infatti essere modificate ed aggiornate nelle successive versioni.

⁸⁷ Poiché la funzione *logit* non risulta definita nei valori estremi della cumulata di spesa, valori di spesa nulla sono stati considerati pari a 0,00001; sono invece stati considerati pari a 0,99 tutti i valori superiori.

⁸⁸ Gli interventi attivi includono anche quelli sospesi. Questi interventi non sono stati considerati per la stima dei parametri del modello, in quanto rappresentano casi anomali; sono stati però considerati nella fase di previsione della spesa.

Le modalità di riferimento delle variabili categoriche sono quelle indicate in corsivo. Sono stati inoltre considerati gli effetti dell'interazione tra le variabili di spesa con la durata prevista.

Risultati

I dati utilizzati per la stima del modello si riferiscono a 6.171 osservazioni relative a 3.012 interventi avviati⁸⁹ all'ultima data di monitoraggio. Sono stati esclusi i dati di spesa ricostruiti per i trimestri precedenti alla data di avvio. Per verificare la bontà del modello di stima è stata applicata la procedura di *simple validation*⁹⁰.

I risultati del modello stimato su un campione casuale (*training set*) pari all'80 per cento del totale degli interventi avviati (2.410 interventi su 3.012) sono stati applicati al campione di *test*, costituito dal restante 20 per cento degli interventi (602) non considerati per la stima. L'operazione è stata ripetuta su 10 coppie di campioni. Come appare dalla Tavola V.7, il confronto tra la correlazione tra valore osservato e valore stimato calcolata sui due sottoinsiemi di osservazioni (*training set* e *test set*) è superiore a 0,9 in entrambi i casi e conferma così la bontà del modello.

TAVOLA V.7 APPLICAZIONE DELLA TECNICA DI *SIMPLE VALIDATION*. Correlazione tra valori stimati e previsti per i campioni training e test

	Interventi	Correlazione	σ
Traning set	2.410	0,92	0,003
Test set	602	0,91	0,001

L'effetto di tutte le variabili considerate è risultato significativo⁹¹, ad eccezione del ritardo all'avvio.

⁸⁹ Sono considerati avviati tutti gli interventi con data di inizio dei lavori precedente a quella di monitoraggio, anche se indicata nel cronoprogramma come prevista.

⁹⁰ cfr. *Two Crows Corporation* (1999).

⁹¹ Per quanto riguarda l'interpretazione dei parametri, a parità di altri fattori, per le variabili categoriche a valori più elevati corrisponde una maggiore velocità di spesa, viceversa a valori più alti del parametro relativo alla cumulata di spesa al tempo t-1 corrisponde una maggiore lentezza.

TAVOLA V.8 VARIABILI ESPLICATIVE INCLUSE NEL MODELLO DI SPESA

Variabile	Livello di significatività
Area	***
Asse	***
Stato dell'intervento	***
Tipo di intervento	***
Ritardo all'avvio	+
Classe di costo	***
Durata prevista	***
Y_{t-1}	***
Spesa dichiarata $_t$	***
Spesa dichiarata $_{t-1}$	***
Y_{t-1} * Durata prevista	***
Spesa dichiarata $_t$ * Durata prevista	***
Spesa dichiarata $_{t-1}$ * Durata prevista	***

*** : $p < 0,001$
** : $0,001 \leq p < 0,01$
* : $0,01 \leq p < 0,05$
+ : $p \geq 0,05$

In particolare dal modello emerge che l'effetto complessivo della cumulata di spesa realizzata fino al tempo t-1 (che è dato dalla combinazione dell'effetto semplice e dell'interazione con la durata prevista) è fortemente significativo ed evidenzia una maggiore velocità di avanzamento percentuale per i progetti più brevi.

TAVOLA V.9 MODELLO DI SPESA

Variabile	Modalità	β	Livello di significatività
Intercetta		0,48	***
Area	Centro	-0,06	***
	Nord	0,09	***
	<i>Sud</i>	0,00	
Asse	Città	0,00	+
	Reti e nodi di servizio	0,01	+
	Risorse culturali	-0,13	***
	Risorse naturali	-0,05	+
	<i>Sistemi locali e ris. Umane</i>	0,00	
Stato dell'intervento	Attivo	-0,38	***
	<i>Concluso</i>	0,00	
Tipo di intervento	Altro	0,01	+
	Nuova realizzazione	0,13	***
	Restauro – recupero	0,15	***
	<i>Servizi</i>	0,00	
Ritardo all'avvio	No	-0,03	+
	<i>Si</i>	0,00	
Classe di costo	Bassa	0,19	***
	Media	0,08	***
	<i>Alta</i>	0,00	
Durata prevista	Fino a 2 anni	0,85	***
	3 anni	0,23	***
	4 anni	0,07	*
	<i>5 anni e più</i>	0,00	
Y_{t-1}		0,97	***
Spesa dichiarata $_t$		0,05	***
Spesa dichiarata $_{t-1}$		-0,06	***
Y_{t-1} * Durata prevista	Fino a 2 anni	-0,34	***
	3 anni	-0,05	***
	4 anni	-0,04	***
	<i>5 anni e più</i>	0,00	
Spesa dichiarata $_t$ * Durata prevista	Fino a 2 anni	0,03	*
	3 anni	-0,01	+
	4 anni	-0,04	*
	<i>5 anni e più</i>	0,00	
Spesa dichiarata $_{t-1}$ * Durata prevista	Fino a 2 anni	-0,04	*
	3 anni	0,02	+
	4 anni	0,05	*
	<i>5 anni e più</i>	0,00	

*** : $p < 0,001$ ** : $0,001 \leq p < 0,01$ * : $0,01 \leq p < 0,05$ + : $p \geq 0,05$

Per quanto riguarda i parametri legati al piano di spesa alla stipula per i trimestri t e $t-1$, essi sono di segno opposto e di valore, nel complesso, quasi equivalente. Tale risultato suggerisce che il contributo complessivo di questa variabile alla previsione di spesa si a marginale. Inoltre, analizzando i risultati per le variabili categoriche, possono essere fatte le seguenti osservazioni:

- **Area:** rispetto agli interventi nel Mezzogiorno, quelli del Nord fanno registrare un avanzamento di spesa più veloce, viceversa gli interventi localizzati nel Centro risultano più 'lenti';
- **Asse:** gli interventi relativi alle Risorse culturali e alle Risorse naturali, cui sono associati parametri negativi, hanno una velocità di avanzamento della spesa più contenuta in confronto agli interventi per i Sistemi locali e le Risorse umane, considerate come modalità di riferimento;
- **Tipo di intervento:** le nuove realizzazioni e le opere di recupero e restauro sono associate a valori positivi (e significativi) dei parametri, che indicano una maggiore rapidità nella realizzazione della spesa rispetto ai servizi;
- **Stato dell'intervento:** come era lecito attendersi, gli interventi ancora attivi hanno mediamente una più lenta progressione della spesa rispetto agli interventi già conclusi;
- **Classe di costo:** come atteso, gli interventi di dimensioni più piccole sono più 'veloci' rispetto a quelli di dimensioni più elevate;
- **durata prevista:** in linea con i risultati relativi alla dimensione economica, per quanto riguarda la durata dei lavori inizialmente prevista, gli interventi più 'brevis' sono i più rapidi: il valore del parametro associato a questa variabile diminuisce all'aumentare della classe di durata.

V.1.2 Previsioni di spesa

Le previsioni finali sono il risultato della combinazione dei tre modelli presentati. Il valore di spesa previsto anno per anno è stato ottenuto, per ogni singolo intervento, a partire dalla previsione della cumulata della spesa trimestrale ottenuta dalla (11).

Per i trimestri per cui esiste spesa realizzata è stato considerato il valore osservato:

$$\hat{y}_t = y_t$$

Inoltre, poiché le previsioni dipendono dalle condizioni iniziali degli interventi, per quelli non avviati è stato necessario operare le seguenti scelte relativamente alla data di avvio e al valore di spesa per il primo trimestre:

- **data di avvio:** per gli interventi in ritardo è stato fissato come trimestre di avvio quello corrispondente alla data di avvio stimata con il modello di durata. Se tale data è antecedente a quella di monitoraggio, allora viene considerata quest'ultima. Per tutti gli altri interventi è stato considerato il trimestre della data di avvio indicata alla stipula.

- **valore di spesa per il primo trimestre p_0 :** l'inclusione nel modello della variabile y_{t-1} richiede, in fase previsiva, di fissare un valore per il trimestre $t=0$ da cui avviare la previsione. Per tutti gli interventi con spesa realizzata nulla⁹² si è fissato arbitrariamente come valore di spesa iniziale $p_t=0,01$. Per gli interventi con spesa realizzata prima dell'avvio dei lavori si è invece considerata come valore iniziale la percentuale cumulata di spesa già realizzata.

Dalla percentuale cumulata per trimestre, ottenuta come trasformata della \hat{y}_t , dalla (10), viene infine ricostruita la percentuale di spesa realizzata su base annuale:

$$\hat{p}_a = \hat{p}_{t+4} - \hat{p}_t$$

e da questa la spesa annuale prevista in valore assoluto:

$$\hat{s}_a = \hat{p}_a * \text{Costo dell'intervento}$$

V.1.3 Confronto tra la precedente versione del modello e il nuovo sistema di previsioni

Il sistema di previsione presentato in questo lavoro rappresenta una evoluzione del precedente modello sviluppato dall'UVER per la previsione della spesa degli interventi in APQ. Innanzitutto si è passati da un modello di spesa ad un sistema di modelli interconnessi, inoltre lo stesso modello di spesa ha subito significativi cambiamenti.

Infatti, nel modello precedente la spesa cumulata al tempo t , è considerata su base annuale e viene espressa in funzione di una serie di variabili esplicative

⁹² In questo caso la funzione *logit* non è definita