



# La parità e il ritardo: analisi di due percorsi scolastici

Benassi F., Porciani L., Valentini A., Venturi S., Barsotti O.

Dipartimento di Statistica e Matematica Applicata all'Economia

Università di Pisa



# Sommario

- Obiettivo dello studio
- Definizioni e concetti operativi
- I dati utilizzati
  - potenzialità e limiti
  - coorti di studio
- Metodi
  - Il modello statistico applicato
  - Le caratteristiche dei modelli applicati alle due coorti di studio
- Risultati



# Obiettivo dello studio

- Individuare quali sono i fattori che influenzano le *performances* scolastiche degli individui in due momenti fondamentali della loro carriera scolastica:
  - 1) il superamento del primo biennio della scuola secondaria di II grado
  - 2) la conclusione della scuola secondaria di II grado



## Definizioni e concetti operativi

- Nel nostro caso le *performances* scolastiche sono state misurate in termini di successo ed insuccesso
- Un percorso scolastico si intende concluso con **successo**:
  - a) quando si compie nei tempi previsti, senza perdita di anni di studio
  - b) quando si conclude il percorso secondario di II grado con una votazione superiore a 80/100



## I dati utilizzati: potenzialità e limiti

- Dati di natura amministrativa: banca dati Osp Pisa
- Coprono l'intera popolazione scolastica presente nella provincia di Pisa per ogni ordine e grado del sistema scolastico (dalla scuola materna fino alla secondaria di II grado)
- Per ogni studente rilevano un insieme notevole di variabili: età, genere, cittadinanza, luogo di nascita e residenza, scuola frequentata etc. etc.
- Presentano un non trascurabile numero di missing specialmente per alcune variabili
- Le variabili raccolte riguardano naturalmente solo il singolo studente e non le dimensioni di contesto (variabili meso o ecologiche)



## I dati utilizzati: coorti di studio

- Due coorti di studio:
  1. popolazione studentesca iscritta nell'anno 2001 al primo anno della scuola secondaria di I grado
  2. popolazione studentesca iscritta nell'anno 2000 al primo anno della scuola secondaria di II grado
- Le due coorti di studio sono state seguite entrambe per 5 anni consecutivi attraverso il metodo delle **estrazioni successive**



## Metodi: scelta del modello statistico

- E' stato applicato un modello di analisi multivariata logistico
- La variabile dipendente è dicotomica: Successo / Insuccesso. Questa è stata posta in relazione ad alcune covariate:
  - socio demografiche (genere, cittadinanza)
  - esiti scolastici pregressi
  - zona di residenza



# Metodi: le caratteristiche dei modelli applicati alle due coorti

- Coorte 2001

*Parità / Ritardo = f (genere, cittadinanza, esiti scolastici pregressi, zona di residenza, e)*

- Coorte 2000

*Parità / Ritardo = f (genere, cittadinanza, esiti scolastici pregressi, zona di residenza, e)*

*Parità / Ritardo = f (genere, cittadinanza, esiti scolastici pregressi, zona di residenza, età, titolo studio dei genitori, tipologia di scuola, e)*

*>80/100 / <= 80/100 = f (genere, cittadinanza, esiti scolastici pregressi, zona di residenza, età, titolo studio dei genitori, tipologia di scuola, e)*

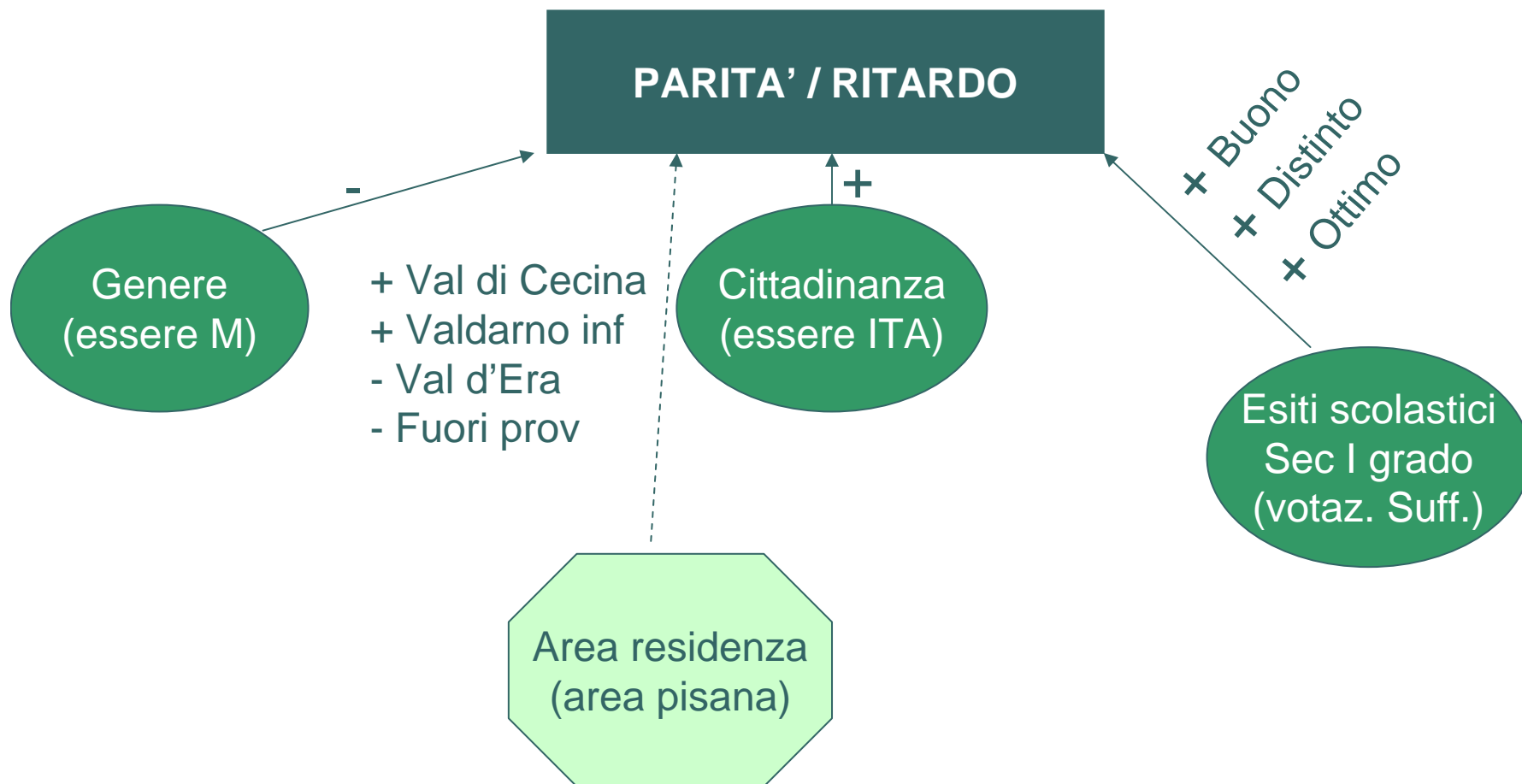




# Risultati

## Coorte 2001

*Parità / Ritardo = f (genere, cittadinanza, esiti scolastici pregressi, zona di residenza, e)*

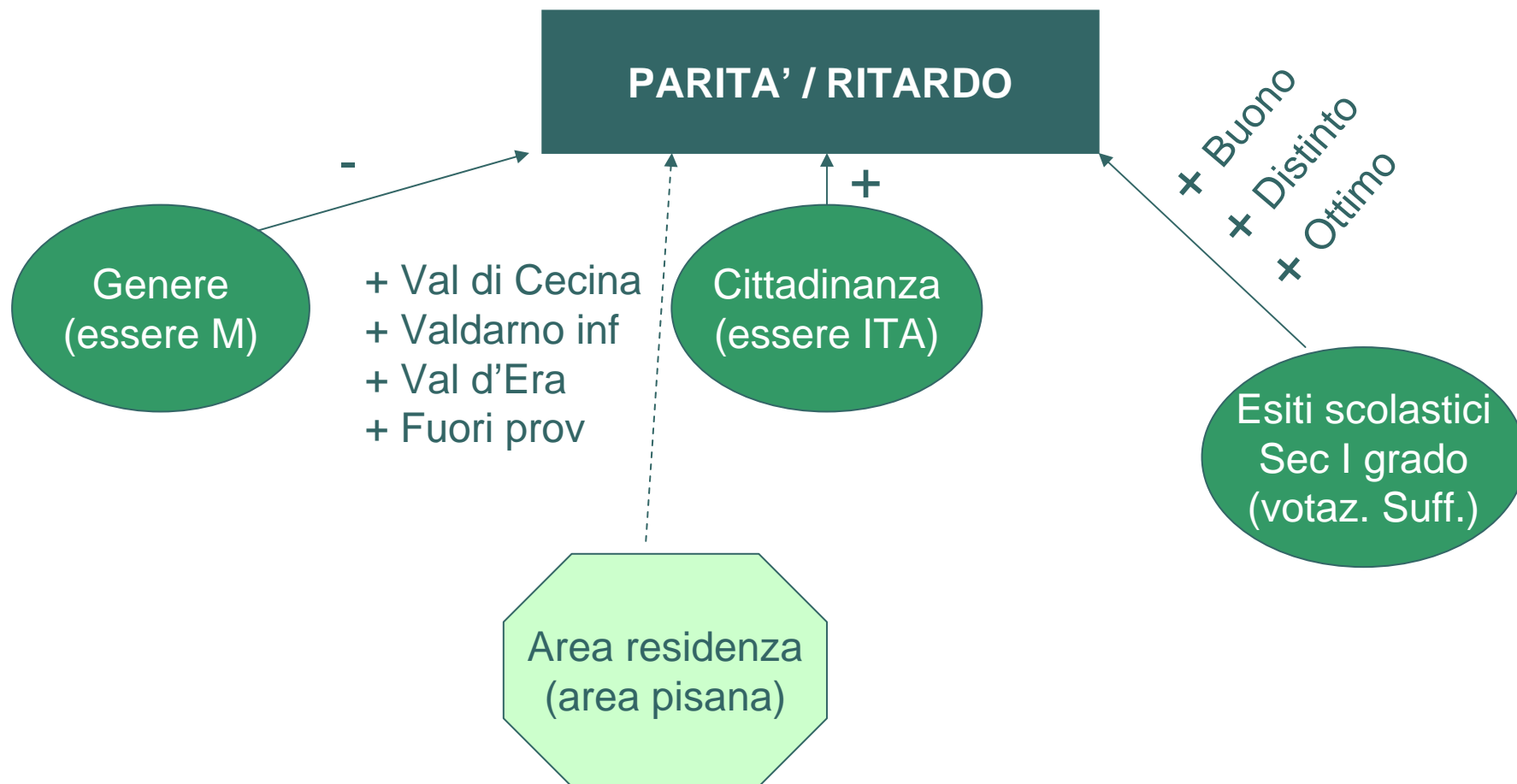




# Risultati

## Coorte 2000

*Parità / Ritardo = f (genere, cittadinanza, esiti scolastici pregressi, zona di residenza, e)*

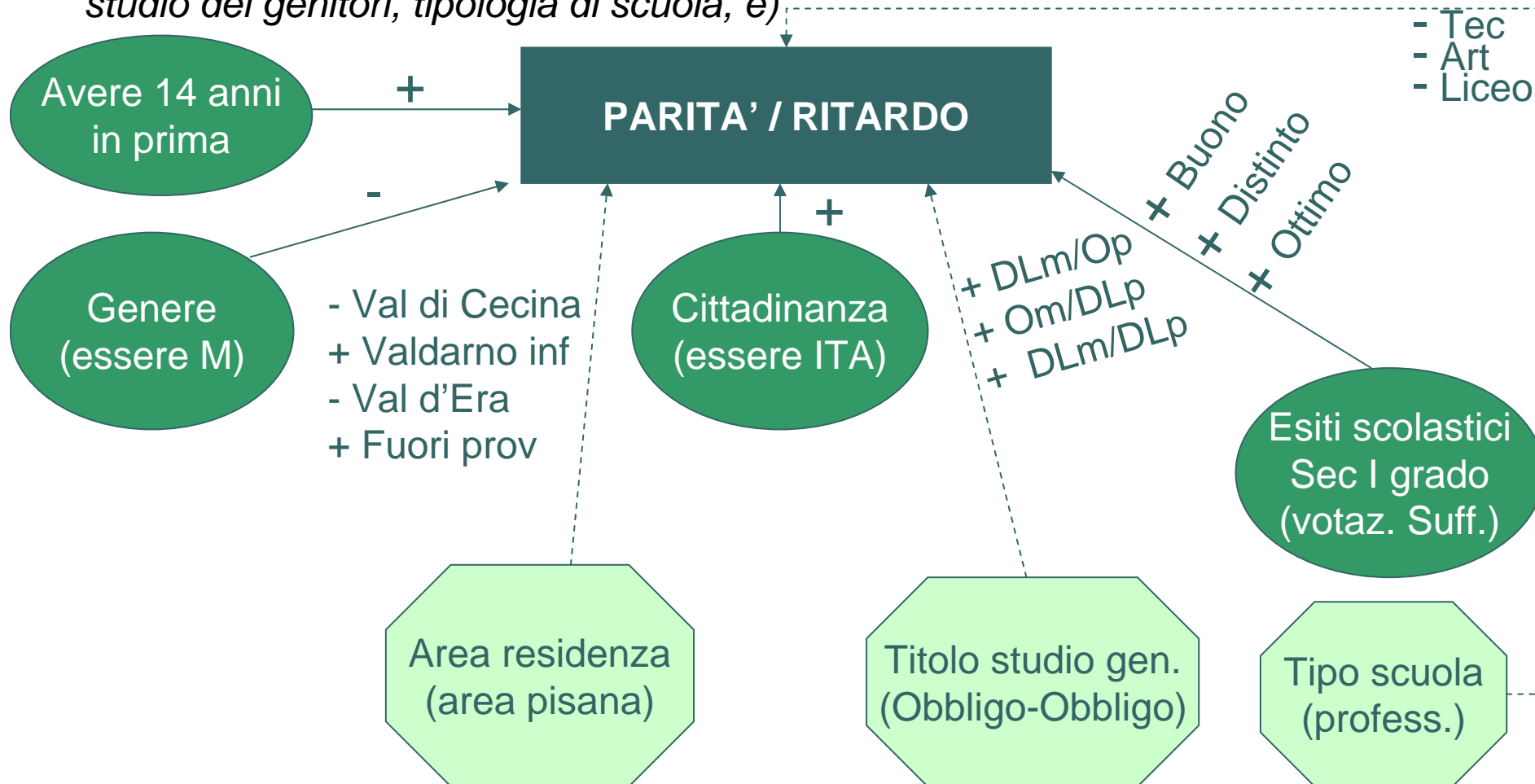




# Risultati

## Coorte 2000

*Parità / Ritardo = f (genere, cittadinanza, esiti scolastici, zona di residenza, età, titolo studio dei genitori, tipologia di scuola, e)*

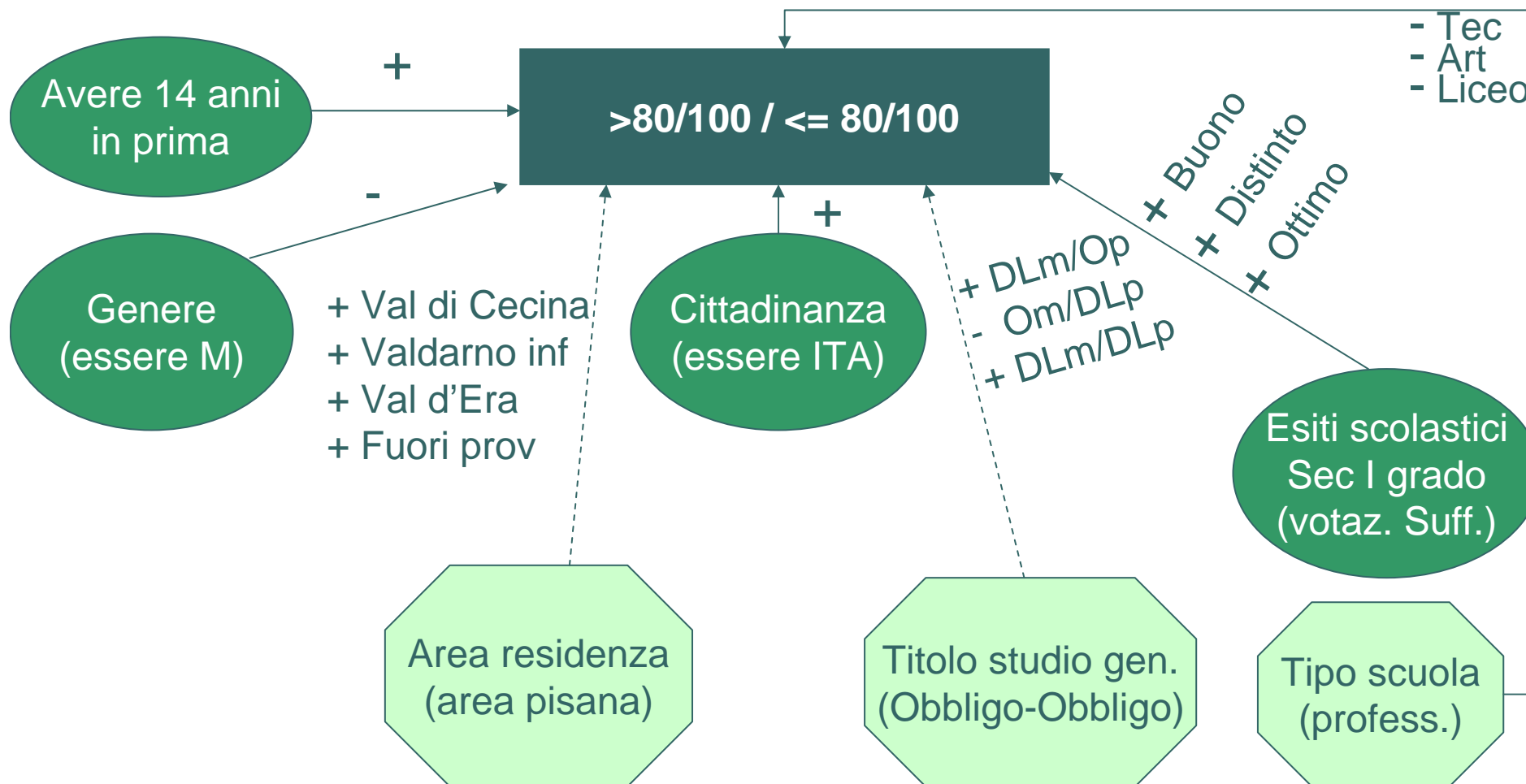




# Risultati

## Coorte 2000

$>80/100 / \leq 80/100 = f(\text{genere, cittadinanza, esiti scolastici, zona di residenza, età, titolo studio dei genitori, tipologia di scuola, e})$





# Risultati

- La capacità esplicativa dei modelli (pseudo Rquadro) varia tra un minimo di 0,26 ad un massimo di 0,32. Si riesce a spiegare quindi nel peggiore dei casi un quarto della variabilità della variabile dipendente e nel migliore dei casi circa un terzo. Dati i limiti dei dati a disposizione e l'alto numero di record ai quali i modelli sono stati applicati il risultato pare soddisfacente
- I modelli hanno valenza descrittiva e non propriamente esplicativo-causale