

TESI DI LAUREA: UN PAIO DI SUGGERIMENTI¹

Efrem Castelnuovo, Università di Padova

26 febbraio 2007

La tesi di laurea è per molti una delle prime occasioni per scrivere una storia. Scrivere una storia non è facile, soprattutto se si ha l'ambizione di farlo in maniera chiara, convincente, magari appassionante. Scrivere è un'arte, e come per ogni arte alcuni sono naturalmente più dotati di altri; ma è anche scienza, ed in questo senso tutti possiamo imparare a scrivere meglio.

Per scrivere bene una tesi di laurea è importante curare l'impostazione della medesima; a mio parere, una tesi deve avere

1. una **prima pagina**, con tutte le informazioni rilevanti. Questa prima pagina deve apparire dalla prima versione della tesi! Un'esempio di prima pagina è il seguente:

UNIVERSITA' DEGLI STUDI DI PADOVA
FACOLTA' DI SCIENZE STATISTICHE
CORSO DI LAUREA IN
STATISTICA ECONOMIA E FINANZA

TESI:

"La regola di Taylor come modello descrittivo della politica monetaria
australiana: una verifica empirica"

Relatore: Dott. Efrem Castelnuovo

Laureanda: Nicole Kidman
Matricola: 123456 - SEF

Anno Accademico 2006/2007

¹A proposito di suggerimenti ... ogni vostro suggerimento sarà il benvenuto!
[efrem.castelnuovo@unipd.it]

2. un **introduzione**, con la quale lo scrittore introduce e motiva il lavoro di tesi (Qual è la domanda di ricerca a cui voglio rispondere con questa tesi? Perché la domanda è rilevante?). L'introduzione della tesi deve anche "ingolosire" il lettore con un accenno ai risultati raggiunti, al fine di convincerlo ad andare avanti con la lettura;
3. un **corpo centrale** del lavoro, che può prevedere l'analisi della letteratura di riferimento, del modello su cui ci si concentra, dei dati che si utilizza, eccetera. Il corpo centrale della tesi contiene ovviamente anche il contributo originale dell'autore (parte di analisi critica della letteratura se la tesi è compilativa; parte innovativa in termini modellistici se la tesi è teoria o applicata; risultati empirici se la tesi è applicata). Questo è il valore aggiunto della tesi, è qui che lo scrittore deve "fare il punto", cioè mettere bene in evidenza il valore aggiunto della tesi stessa;
4. una **conclusione**, con la quale lo scrittore riassume i risultati raggiunti, ribadendo ancora una volta sinteticamente ma chiaramente il suo contributo rispetto a quanto già noto in letteratura;
5. una **bibliografia**, che ha l'obiettivo di riassumere i contributi citati dallo scrittore della tesi. E.g., se citiamo i risultati presenti nel lavoro di Gregory N. Mankiw e Ricardo Reis e pubblicati nel 2002 sul *Quarterly Journal of Economics*, nel corpo della tesi scriveremo " ... come in Mankiw e Reis (2002) ... ", e nella Bibliografia metteremo

Mankiw, G.N., e R. Reis, 2002, Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve, *Quarterly Journal of Economics*, 117, 1295-1328, Novembre.

Io sono solito indicare in *Italics* le riviste, e con normale scrittura i libri (in alternativa, potete sottolineare il titolo del libro).

Attenzione: vanno citate anche le banche dati dalle quali si prendono le statistiche eventualmente citate nel testo della tesi, oppure i dati usati per il lavoro empirico; inoltre, se un lavoro scientifico citato nella tesi è stato scaricato da Internet, occorre mettere in bibliografia l'indirizzo web di tale lavoro (a meno che questo faccia parte di una collana di working papers, o sia già stato pubblicato: in questi casi, occorre invece indicare la collana di working papers / la rivista ove il lavoro è stato pubblicato);

6. un'**appendice tecnica**, nel caso in cui vi siano risultati di test che appesantirebbero la lettura se posizionati nel corpo centrale, oppure dimostrazioni troppo lunghe, analisi dei dati un po' "noiose", e così via.

Mi raccomando, siate chiari e concisi nello scrivere la tesi! A me personalmente il "bla bla bla" non piace, mi fa perder tempo, e mi irrita un po'. Scrivete la tesi senza tanti giri di parole, evidenziando in maniera precisa il vostro contributo; il resto non è interessante! Sia ben chiaro, però, che essere sintetici non significa essere superficiali: Una tesi scritta in maniera approssimativa, poco rigorosa, non completa, non è una buona tesi, e come tale verrà valutata!

In generale, consiglio di rileggere sempre quanto scritto, magari dopo aver lasciato da parte la tesi per un paio di giorni; se nel rileggere la vostra stessa tesi vi chiedete "Uhm, cosa volevo dire con questa frase?", beh, come potete pretendere che quella frase venga compresa dal lettore?!?

Un piccolo trucco per aumentare la vostra chiarezza espositiva è spezzare le frasi più lunghe in frasi più brevi. A proposito, ogni frase dovrebbe avere soggetto, verbo, e complementi ... se non li ha può essere che sia una frase speciale, magari poetica, che fa sognare ... e magari fa domandare al vostro relatore di tesi: "Che cosa stava sognando quando ha scritto questa frase?!?". Per favore, esprimetevi poeticamente in altri contesti, non quando scrivete la tesi!

IMPORTANTE!!! Ad un docente non si consegna una tesi scritta male, o in maniera disordinata, o senza introduzione e conclusioni, o senza bibliografia, o con le pagine non numerate, o senza aver prima controllato la correttezza ortografica dello scritto, "tanto queste son cose poco importanti che posso fare dopo ...". Una tesi scritta non al meglio dà un'impressione di trascuratezza che sicuramente irrita il docente e che vi può costare qualche punto-tesi in fase di valutazione del vostro lavoro: val davvero la pena correre questo rischio?

Passo ora a descrivere lo scottante tema del trattamento delle serie temporali macroeconomiche, serie che giocano spesso un ruolo importante nelle tesi che assegno. Nel farlo, non ho alcuna pretesa di essere esaustivo: infatti, il trattamento delle serie temporale macroeconomiche meriterebbe un intero corso di laurea! Vi suggerisco dunque di considerare le pagine che seguono come materiale complementare rispetto a quanto visto in classe nei corsi di statistica delle serie storiche/econometria che avete frequentato.

Il trattamento delle serie temporali macroeconomiche: una lista non esaustiva e non sistematica di elementi da valutare

Molte delle tesi che assegno sono lavori con i dati. I lavori empirici con serie temporali richiedono il trattamento delle serie macro, soprattutto delle variabili aggregate reali. E' buon comportamento osservare le serie (e.g. visualizzarle a schermo) prima di utilizzarle, così come è saggio conoscerle dal punto di vista del grado di integrazione (alcune proprietà statistiche valgono solo per serie stazionarie, ma spesso in macro le serie sono integrate di ordine 1 o 2 ...). Nel lavorare con i dati, occorre poi tenere in considerazione quanto segue:

- il passaggio da valori (values) $Z = PX$, i.e. serie nominali [quando non altrimenti specificato] a volumi (volumes), i.e. serie reali che rappresentano l'effettiva evoluzione temporale delle risorse in un sistema economico, si effettua così: $X = \frac{Z}{P}$, con P ad indicare l'indice dei prezzi dell'elemento X di riferimento (e.g. se X è PIL allora P sarà l'indice dei prezzi del PIL, se X è il consumo allora P sarà l'indice dei prezzi sui beni di consumo, e così via).
- Gli economisti scrivono spesso i loro modelli considerando le deviazioni delle serie macroeconomiche X_t rispetto al loro valore di stato stazionario \bar{X}_t , i.e. $\tilde{X}_t \equiv X_t - \bar{X}_t$ (tipicamente questo si fa con serie stazionarie, o che tendono ad esserlo); per alcune serie, invece, considerano log deviazioni, i.e. $\tilde{X}_t \equiv \log(X_t) - \log(\bar{X}_t)$ (serie che tendono ad essere integrate di ordine 1). In entrambi i casi, gli economisti mirano a computare dei *gaps* (e.g. come l'output gap, o il gap di disoccupazione), perchè i gaps sono spesso economicamente più interessanti delle serie "crude". Per calcolare un gap, però, occorre calcolare \bar{X}_t : come fare? \bar{X}_t a volte è calcolato da istituti di ricerca o enti come l'OCSE, il Congressional Budget Office, l'ISTAT, tramite sofisticate tecniche statistico-econometriche. Un'alternativa è quella di calcolare la misura di stato stazionario (o potenziale, o livello naturale) calcolando il *trend* della serie di interesse. Una scelta piuttosto diffusa tra i macroeconomisti è quella di calcolare il trend con un polinomio di grado n , il che può essere fatto stimando per minimi quadrati il seguente modello:

$$\underbrace{\log(X)}_{Tx1} = \underbrace{[1 \quad \dots \quad 1]'}_{Tx1} \underbrace{c}_{1x1} + \underbrace{t}_{Tx1} \underbrace{\beta_1}_{1x1} + t^2 \beta_2 + \dots + t^n \beta_n + \underbrace{\varepsilon}_{Tx1} \quad (1)$$

con $t = [1 \ 2 \ 3 \ 4 \ \dots \ T]'$, $T = \#$ osservazioni del campione. Una scelta molto spesso adottata è quella di stimare (1) con $n = 1$ o $n = 2$. Il fittato della (1) stimata ci dà la stima del potenziale \widehat{X}_t , mentre il gap (di consumo, reddito, investimenti, e così via) è rappresentato da $\widehat{\varepsilon}$, che viene tipicamente percentualizzato ottenendo così $\widehat{X}_t = 100\widehat{\varepsilon}$. Una delle possibili alternative al modello polinomiale (1) è rappresentata dal filtro Hodrick-Prescott, ben spiegato ed implementato dalla maggior parte dei *packages* econometrici per lo studio delle serie temporali. Si noti che il calcolo della deviazione [gap] di una serie rispetto al suo stato stazionario si riferisce solitamente alle serie reali come reddito, consumo, investimenti, e così via, mentre non si applica tipicamente ai tassi di crescita (e.g. inflazione) e nemmeno ai tassi di interesse.

- serie temporali: percentualizzate, aggiustate per la stagionalità ... occorre sempre controllare che il dataset contenga serie omogenee in termini di caratteristiche, e.g. occorre ad esempio evitare che il dataset contenga l'inflazione non percentualizzata - ad esempio un valore come 0.02, che indica 2% - e un tasso di interesse a breve percentualizzato - e.g. 2, che indica 2%. Inoltre, occorre verificare che le serie che possono soffrire di effetti stagionali siano di fatto destagionalizzate; nel caso in cui non lo fossero, occorre destagionalizzarle [differenti pacchetti econometrici possiedono procedure per farlo con allegata spiegazione].
- inflazione annuale vs. trimestrale vs. annualizzata. Supponiamo di realizzazioni trimestrali dell'indice di prezzo P: l'inflazione annuale si calcola come $\pi^{annuale} = \frac{P_t - P_{t-4}}{P_{t-4}}$, mentre quella trimestrale si calcola come $\pi^{trimestrale} = \frac{P_t - P_{t-1}}{P_{t-1}}$, e quella trimestrale annualizzata si calcola sulla trimestrale, i.e. $\pi^{annualizzata} = 4\pi^{trimestrale}$. Per percentualizzare i tassi di inflazione qui elencati occorre moltiplicarli per 100. Invece, con osservazioni mensili, quanto abbiamo detto cambia come segue: $\pi^{annuale} = \frac{P_t - P_{t-12}}{P_{t-12}}$, $\pi^{mensile} = \frac{P_t - P_{t-1}}{P_{t-1}}$, $\pi^{annualizzata} = 12\pi^{mensile}$. Ancora una volta, per percentualizzare le serie occorre moltiplicare per 100. Si noti che il calcolo dell'inflazione annuale (così come del tasso di crescita annuale di qualsiasi variabile macroeconomica) attribuisce persistenza spuria a tale serie, quindi non è una buona idea usare l'inflazione annuale se lo studio che si vuole condurre mira a misurare la persistenza della serie stessa (i.e. è meglio usare l'alternativa, cioè

l'inflazione trimestrale annualizzata).

- per diversi stimatori è possibile calcolare matrici di varianza-covarianza robuste alla presenza di eteroschedasticità/correlazione seriale, come la matrice di VCV Newey-West: Quando il pacchetto econometrico che utilizzate per fare le stime ve lo consente, scegliete l'opzione Newey-West (e.g. E-views ve lo consente).

- ma da dove si scaricano le serie temporali? Esistono diverse fonti:

1. Federal Reserve Bank of St. Louis [per serie sugli Stati Uniti]:

<http://research.stlouisfed.org/fred2/>

2. Federal Reserve Bank of Philadelphia [per Survey data sugli Stati Uniti]: <http://www.phil.frb.org/econ/spf/>

3. Banca Centrale Europea [dati su Euro area]: <http://www.ecb.int> [lato sinistro dello schermo, voce: "key indicators"]

4. Economagic [per serie su molti stati]: <http://www.economagic.com>

5. The Center for International Data: <http://www.internationaldata.org/>

6. OCSE [OECD, per serie su molti stati, compresi aggregati come "7 major OECD countries", "Euro 12", "OECD total"]: trovate il sito tramite <http://www.google.com> con le parole chiave "economic outlook statistics", e cercate il link che sta sotto "SourceOECD: Statistics: Economic Outlook"

7. Diverse banche dati (Eurostat, OECD, IMF, etc.) su serie macroeconomiche si trovano partendo da <http://195.145.59.167/>, cliccando "here", "International Statistical Yearbook"

8. Dati su serie finanziarie: <http://finance.yahoo.com/>

9. Una raccolta di banche dati gratuitamente disponibili sul web: <http://www.economicnetw>

Attenzione! Per i dati OCSE dovete usare il computer in sala "Dati Statistici" [biblioteca di Statistica]: occorre avvisare il bibliotecario, che vi permetterà di utilizzare quella postazione dopo averla "sbloccata" tramite password.

- Importante! Non tutte le serie sono a disposizione per tutte le frequenze! Ad esempio, il PIL statunitense è a disposizione a livello trimestrale ma non a livello mensile ... occorre in questi casi approssimare le serie mancanti con altre (e.g. al posto del PIL si può usare la disoccupazione, oppure l'indice della produzione industriale, etc ...).