

UNIVERSITÁ DEGLI STUDI DI PADOVA

FACOLTÀ DI SCIENZE STATISTICHE

CORSO DI LAUREA IN STATISTICA E FINANZA

TESI DI LAUREA

**LA PROBABILITÀ DI SOFFERENZA
NEI FINANZIAMENTI ALLE FAMIGLIE:
RELAZIONE DI UNO STAGE ALLA
BANCA POPOLARE DI VICENZA**

Relatore: Prof. Guglielmo Weber

Laureanda: Paola Brunello

Matricola n. 534409-SEF

Anno accademico 2007-2008

INDICE

Introduzione.....	3
1. L’approccio econometrico alle scelte dicotomiche.....	4
1.a I modelli probit e logit.....	8
1.b Il metodo di stima della massima verosimiglianza.....	12
1.c Gli effetti marginali nelle variabili esplicative.....	16
2. Lo stage alla Banca Popolare di Vicenza.....	18
2.a Storia e composizione della banca.....	18
2.b Esperienza nello stage.....	20
3. Analisi e stima del dataset.....	22
3.a Analisi descrittiva e incrociata delle variabili.....	23
3.b Analisi econometrica.....	41
Conclusioni.....	53
Bibliografia.....	54

INTRODUZIONE

I finanziamenti alle famiglie sono divenuti negli ultimi anni uno dei prodotti finanziari più diffusi. Banche ed intermediari creditizi hanno puntato molto su questa modalità di prestito monetario trasformandola in un vero e proprio boom.

Con la crisi dello scorso anno dei mutui subprime negli Stati Uniti, i finanziamenti si sono ridotti di molto ed anche se per le banche restano uno dei migliori prodotti, queste non applicano più la stessa spinta alla loro vendita.

Con finanziamenti includiamo i prestiti personali, i mutui chirografari e i mutui ipotecari che sono rivolti alle famiglie. Spesso i debitori, per la restituzione del prestito, non risultano regolari con il pagamento delle rate secondo il piano di ammortamento, entrando quindi in quella che noi chiamiamo sofferenza ovvero la probabilità che attiene al debitore e alla sua capacità di essere solvibile entro un anno.

L'obiettivo è scoprire quali variabili, in nostro possesso, influenzano la probabilità di sofferenza, quali in modo positivo e quali in modo negativo, e in che percentuale determinano questa probabilità. In questo modo potremo capire quali caratteristiche di un determinato tipo di finanziamento o di debitore sono più rischiose rispetto alla variabile di insolvenza e di conseguenza a percepire il comportamento delle controparti in certe situazioni.

Durante lo stage alla Banca Popolare di Vicenza alcuni colleghi mi hanno fornito un database composto da 22000 osservazioni e contenente circa venti variabili, con le quali costruiremo il modello per verificare l'obiettivo posto. Alla Banca lo studio di queste variabili viene effettuato attraverso l'uso di reti bayesiane, noi invece useremo un'applicazione più semplice, quella econometrica.

Essendo la probabilità di sofferenza una variabile dicotomica che assume valore 0 quando il debitore è insolvente ed 1 quando non lo è, useremo come approccio econometrico, per la stima del modello che specificheremo, i modelli probit e logit. Questi modelli infatti assumono che la variabile dipendente sia una variabile bivariata, come la nostra probabilità di sofferenza.

1. L'APPROCCIO ECONOMETRICO ALLE SCELTE DICOTOMICHE

Molti fenomeni economici, come la disoccupazione, il consumo aggregato, i profitti del settore industriale, ecc. possono essere spiegati attraverso dei modelli economici. Questi modelli prendono la forma di equazioni che relazionano il fenomeno che si intende spiegare ai valori assunti da una serie di variabili, con l'obiettivo di scoprire le cause del fenomeno oggetto di indagine. La verifica empirica (sulla base dei dati disponibili) della validità dei modelli economici, ossia della loro rispondenza alla realtà osservata, costituisce uno degli scopi fondamentali dell'analisi econometrica. In altre parole un'analisi empirica usa i dati disponibili per verificare la validità di una teoria economica attraverso la generazione di stime.

Come accennato, questi modelli, chiamati modelli di regressione lineare, studiano la relazione tra una variabile dipendente e una o più variabili indipendenti. La loro generica forma è:

$$y = \beta_1 + x_2\beta_2 + \dots + x_k\beta_k + \varepsilon$$

dove y è la variabile dipendente, $x_2, \dots, x_k$ le variabili indipendenti o esplicative, $\beta_1, \dots, \beta_k$ i parametri ed ε il termine di errore o di disturbo, per il quale si assume solitamente l'ipotesi di distribuzione normale.

Il modello è una semplificazione della realtà che include le caratteristiche principali della relazione di interesse, e ovviamente non può includere la miriade di aspetti essenziali che ci stanno intorno, così è necessario incorporare la variabile stocastica ε . Questo termine di disturbo rappresenta l'errore che si compie prendendo un certo numero di variabili e non tutte le possibili, ciò significa che esso cattura gli effetti delle variabili omesse, che possono essere positivi o negativi.

In genere, nel definire la distribuzione di ε condizionata al valore dato di x , si assume che la media del valore di ε non dipende dal valore di x , ossia:

$$E(\varepsilon|x) = E(\varepsilon) = 0$$

Come ci dice Gianni Amicano in *Lezioni di Econometria*, vi sono applicazioni econometriche in cui è necessario modellare fenomeni che si sostanziano in scelte dicotomiche (o comunque discrete):

$$y_t = 0$$

oppure

$$y_t = 1$$

dove $t = 1, 2, \dots, T$

La suddetta viene chiamata variabile dipendente binaria e può assumere solo i valori 0 e 1.

Il modello econometrico di riferimento cerca allora di spiegare la probabilità che l'agente sceglierà l'alternativa 1 in funzione di alcune variabili esplicative osservate, ovvero cerca di spiegare quali sono le circostanze che determinano i valori osservati su y_t :

$$y_t = f(1, x_{2t}, \dots, x_{kt}, \varepsilon_t),$$

$t = 1, 2, \dots, T$ indica la i -esima osservazione del campione.

Generalmente le diverse osservazioni si riferiscono a diverse unità (individui, imprese, famiglie) osservate ad un dato istante temporale. Ad esempio dato un campione T di individui, alcuni dei quali investono in borsa ($y_t = 1$) e altri no ($y_t = 0$), le variabili $x_{it}, \dots, x_{kt}$ costituiscono la misurazione dei fattori ritenuti rilevanti per determinare la scelta dell'individuo.

Come riportato in *Econometric Theory and Methods* di Davidson&Mackinnon, P_t denota la probabilità che $y_t = 1$ condizionatamente alle informazioni che ci danno i dati, indicati con Ω_t , i quali consistono delle variabili esplicative X_t :

$$P_t = \Pr(y_t = 1 | \Omega_t) = E(y_t | \Omega_t).$$

Ora qualsiasi modello binario deve assicurare che $E(y_t | \Omega_t)$ rientra nell'intervallo 0-1 e quindi che $0 < P_t < 1$, specificando:

$$P_t \equiv E(y_t | \Omega_t) = F(X_t \beta)$$

dove $X_t \beta$ è chiamata funzione indice ed è composta dal vettore X_t di variabili esplicative e dal vettore β dei parametri. $F(x)$ è la funzione di trasformazione che ha le seguenti proprietà:

$$F(-\infty) = 0, F(\infty) = 1, f(x) \equiv (dF(x) / dx) > 0.$$

“Queste proprietà di fatto corrispondono a quelle di una distribuzione di probabilità, e ci garantiscono che $F(x)$ è una funzione non lineare. Di conseguenza, cambiamenti nei valori di X_{ti} , che sono gli elementi di X_t , necessariamente influiscono su $E(y_t | \Omega_t)$ in modo non lineare. In modo più specifico, quando P_t è dato come specificato sopra, la sua derivata rispetto a X_{ti} è:

$$(\partial P_t) / (\partial X_{ti}) = (\partial F(X_t \beta)) / (\partial X_{ti}) = f(X_t \beta) \beta_i,$$

dove β_i è l'i-esimo elemento di β . Perciò, la grandezza della derivata è proporzionale a $f(X_t \beta)$ e ci dice che l'effetto su P_t di un cambiamento in una

delle variabili indipendenti è più grande quando $P_t = 0,5$ e molto più piccolo quando P_t è compreso tra 0 e 1.”

Per stimare un modello a scelta binaria, Amicano ci dice che è possibile utilizzare due approcci alternativi:

1) modello di probabilità lineare (MPL):

$$y_t = \sum_{i=1}^k \beta_i x_{it} + \varepsilon_t = X_t \beta + \varepsilon_t, t = 1, 2, \dots, T$$

$$E(\varepsilon_t) = 0, \forall t$$

in cui la variabile dipendente y_t (variabile casuale discreta) dipende in modo lineare da un insieme di k variabili esplicative;

2) approccio basato sull'esistenza di una variabile latente y_t^* che non è possibile osservare (o misurare) tale per cui:

$$y_t = \begin{cases} 1, & y_t^* > 0 \\ 0, & y_t^* \leq 0 \end{cases}$$

Questo approccio è alla base dei modelli probit e logit, che analizzeremo più precisamente ora.

1.a I MODELLI PROBIT E LOGIT

Questi modelli sono basati sull'assunzione che ci sia una variabile non direttamente osservabile y_t^* che viene determinata da un certo numero di regressori $x_1, \dots, x_k$ nel modo consueto:

$$y_t^* = X_t \beta + \varepsilon_t, \quad t = 1, 2, \dots, T$$
$$E(\varepsilon_t) = 0, \quad \forall t$$

Questa grandezza non è però osservabile, e si osserva invece la variabile dipendente y_t :

$$y_t = \begin{cases} 1, & y_t^* > 0 \\ 0, & y_t^* \leq 0 \end{cases}$$

Come riportato in *Lezioni di Econometria*: “La variabile latente y_t^* può essere interpretata come una propensione ad assumere valore $y_t = 1$ oppure come costo opportunità di $y_t = 0$. Si noti che se y_t^* viene moltiplicata per qualsiasi costante non nulla, il modello non cambia (problema di identificazione):”

$$\lambda y_t^* = \lambda X_t \beta + \lambda \varepsilon_t$$
$$y_t^{**} = X_t \beta^* + \varepsilon_t^*$$

Per questo motivo si è soliti “normalizzare” il modello in modo da avere un termine di disturbo ε_t che ha varianza unitaria:

$$E(\varepsilon_t^2) = \sigma^2 = 1.$$

Quindi:

$$y_t^* = X_t \beta + \varepsilon_t,$$

$$\varepsilon_t \sim \text{NID}(0,1)$$

è chiamato modello della variabile latente e determina il valore della variabile dipendente osservata secondo la relazione sopra descritta.

Definiamo ora:

$$\begin{aligned} P_t &= \text{prob}(y_t = 1) = \text{prob}(y_t^* > 0) = \text{prob}(X_t \beta + \varepsilon_t > 0) = \\ &= \text{prob}(\varepsilon_t > -X_t \beta) = 1 - \text{prob}(\varepsilon_t < -X_t \beta) = 1 - F(-X_t \beta) = F(X_t \beta) \end{aligned}$$

dove con $F(\cdot)$ indichiamo la funzione di ripartizione della variabile casuale ε_t :

$$F(h) = \int_{-\infty}^h f(\varepsilon_t) d\varepsilon_t.$$

Si ipotizza anche che la funzione di densità di ε_t sia simmetrica rispetto allo zero. Si ha quindi che:

$$F(-X_t \beta) = 1 - F(X_t \beta).$$

Il **modello probit** assume che i termini disturbo del modello abbiano una distribuzione normale:

$$\varepsilon_t \sim \text{N}(0,1)$$

e

$$p(y_t = 1) = p(y_t^* > 0) = \Phi(X_t \beta)$$

in cui:

$$\Phi(x) = 1/\sqrt{2\pi} \int_{-\infty}^x \exp\{-X^2/2\} dX$$

dove con $\Phi(\cdot)$ indichiamo la funzione di ripartizione di una variabile normale standardizzata.

La funzione di verosimiglianza per un modello probit è:

$$L = \prod_{y_t=1} \Phi(X_t\beta) \prod_{y_t=0} [1 - \Phi(X_t\beta)],$$

di conseguenza la funzione di log-verosimiglianza assume la forma seguente:

$$\ln L = \sum_{y_t=1} \ln \Phi(X_t\beta) + \sum_{y_t=0} \ln [1 - \Phi(X_t\beta)].$$

E per ottenere la stima di massima verosimiglianza, come vedremo nel paragrafo successivo, è necessario usare un metodo di massimizzazione numerica.

Il **modello logit** invece assume che i termini di disturbo seguano una distribuzione logistica:

$$F(\varepsilon_t) = (\exp(\varepsilon_t)) / (1 + \exp(\varepsilon_t)) = 1 / (1 + \exp(-\varepsilon_t))$$

Una variabile casuale logistica ha le seguenti funzioni di densità e di ripartizione:

$$f(x) = e^x / (1 + e^x)^2 \quad x \in (-\infty, +\infty)$$

$$F(x) = e^x / (1 + e^x).$$

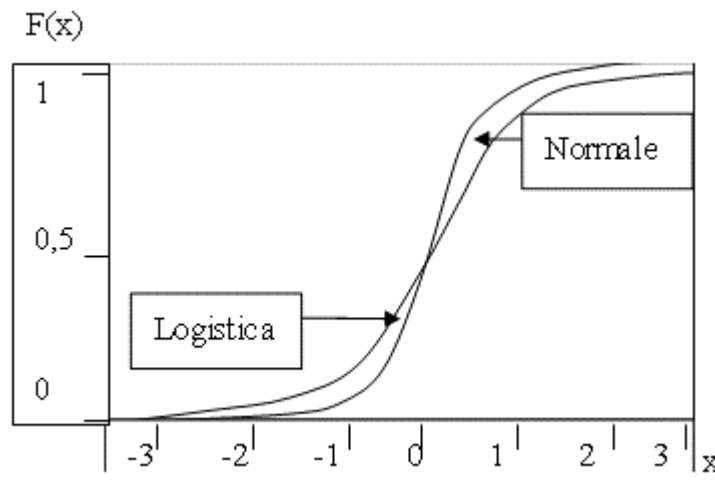
Il modello logit è più facile da derivare assumendo che:

$$\log(P_t / (1 - P_t)) = X_t \beta,$$

risolvendo per P_t troveremo quindi:

$$P_t = (\exp(X_t \beta) / (1 + \exp(X_t \beta))) = 1 / (1 + \exp(-X_t \beta)).$$

Qui sotto possiamo osservare il grafico della funzione logistica e quello della normale:



Dandone una breve descrizione, possiamo quindi affermare che la funzione logistica tende a 0 per X che tende a $-\infty$ e tende a 1 per X che tende a $+\infty$, vale 0,5 per $X = 0$ ed ha un andamento abbastanza lineare avvicinandosi a $X = 0$.

Una variabile casuale logistica è caratterizzata da maggiore dispersione rispetto alla normale standardizzata. Infatti:

$$E(x) = 0$$

$$V(x) = \pi^2 / 3 \approx 3.28 (> 1)$$

Il suo comportamento è abbastanza simile a quello di una normale standardizzata $N(0,1)$, tranne che nelle code.

Nella maggior parte dei casi, la sola reale differenza tra i due modelli sta nel modo in cui gli elementi di β sono misurati.

Infatti l'applicazione dei modelli logit e probit allo stesso insieme di dati conduce a risultati simili, ma dato che nel modello logit si ha:

$$V(\varepsilon_t) = E(\varepsilon_t^2) = \pi^2 / 3,$$

occorre dividere i risultati della stima logit per $\pi^2 / 3$ per poterli confrontare con quelli della stima probit (dove invece la varianza dei termini di disturbo è per ipotesi unitaria).

Inoltre, come ci dice Greene in *Econometric Analysis*: "la distribuzione logistica, piuttosto che quella normale, tende a dare una più ampia probabilità a $y_t=0$ quando $X_t\beta$ è estremamente piccolo."

1.b IL METODO DI STIMA DELLA MASSIMA VEROSIMIGLIANZA

Cominciamo innanzitutto definendo la funzione di verosimiglianza: siano $x_1, x_2, \dots, x_n$ un campione di n elementi estratti in modo indipendente ed identicamente dalla stessa popolazione, la cui densità la indichiamo con $f(x, \theta)$.

Se la popolazione fosse distribuita normalmente con valore atteso μ e varianza σ^2 , possiamo scrivere la funzione di verosimiglianza come:

$$L(x_1, x_2, \dots, x_n, \theta) = (2\pi)^{-n/2} \sigma^{-n} \exp \left\{ -1/2\sigma^2 \sum_{i=1}^n (x_i - \mu)^2 \right\}$$

Dato il campione e la forma della funzione di densità sopra descritti, che dipende da un vettore di parametri incogniti θ , come definisce Gianni Amicano in *Lezioni di econometria*: “la stima di massima verosimiglianza consiste nel cercare quei valori dei parametri del modello che rendono l'estrazione del campione osservato il più possibile verosimile”.

Si massimizza quindi la funzione di verosimiglianza rispetto ai parametri da stimare:

$$\text{Max}_{\theta} L(x_1, x_2, \dots, x_n, \theta).$$

La soluzione $\hat{\theta}$ viene indicata come stimatore di massima verosimiglianza di θ . Al fine di ottenere condizioni del primo ordine più semplici, si ricorre alla massimizzazione della cosiddetta funzione di log-verosimiglianza; infatti come ci ricorda Amicano, “se una funzione viene sottoposta ad una trasformazione monotonica conserva i punti di massimo e di minimo della funzione di partenza”. La funzione di log-verosimiglianza è:

$$\log L(x_1, x_2, \dots, x_n, \theta, \sigma^2) = -n/2 \log(2\pi) - n \log(\sigma) - 1/2\sigma^2 \sum_{i=1}^n (x_i - \mu)^2$$

La derivata logaritmica viene chiamata funzione score e possiamo identificarla nel modo seguente:

$$u = d \log f(x, \theta) / d\theta$$

e ha valore atteso nullo, infatti:

$$\begin{aligned} E(u) &= \int (d \log f(x, \theta) / d\theta) f(x, \theta) dx = \int (1/f(x, \theta)) (dF(x, \theta) / d\theta) f(x, \theta) dx \\ &= d/d\theta \int f(x, \theta) dx = d/d\theta * 1 = 0 \end{aligned}$$

La cosiddetta informazione attesa o di Fisher è invece la varianza dello score e può essere interpretata come l'ammontare di informazione contenuta da una variabile casuale osservabile x , concernente un parametro non osservabile θ , da cui dipende la distribuzione di probabilità di x . Solitamente indicata con $I(\theta)$, la sua varianza è pari al suo momento del secondo ordine:

$$I(\theta) = E\left[(\partial/\partial\theta \ln f(x, \theta))^2\right].$$

L'informazione di Fisher può essere letta come una misura della curvatura della verosimiglianza in corrispondenza della stima di massima verosimiglianza per θ . Una verosimiglianza piatta, con una derivata seconda modesta, comporterà minore informazione, mentre una maggiore curvatura apporterà una maggiore quantità di informazione.

Andiamo quindi a calcolarci le condizioni del primo ordine e a determinarci i parametri μ e σ^2 :

$$\begin{aligned} \partial \log L(x_1, x_2, \dots, x_n, \mu, \sigma^2) / \partial \mu &= 0 \Rightarrow 2 / 2\sigma^2 \sum_{i=1}^n (x_i - \mu) = 0 \\ \Rightarrow \hat{\mu} &= \bar{x}_n = 1/n \sum_{i=1}^n x_i \\ \partial \log L(x_1, x_2, \dots, x_n, \mu, \sigma^2) / \partial \sigma &= 0 \Rightarrow -n/\sigma + 1/\sigma^3 \sum_{i=1}^n (x_i - \mu)^2 = 0 \\ \Rightarrow \hat{\sigma}^2 &= 1/n \sum_{i=1}^n (x_i - \hat{\mu})^2 \end{aligned}$$

Dove $\hat{\mu}$ denota la media campionaria e $\hat{\sigma}^2$ la varianza campionaria.

Il metodo di stima dei modelli probit e logit è proprio quello della massima verosimiglianza, infatti come ci dice Wooldrige in *Introductory Econometrics*,

essendo i modelli a risposta binaria non lineari non possiamo applicare il metodo dei minimi quadrati.

Com'è riportato in *Econometric Theory and Methods*: “Quando la variabile dipendente può assumere valori discreti, la funzione di verosimiglianza per questi valori potrebbe essere definita come la probabilità che lo specifico valore si realizzi, piuttosto che come la densità di probabilità del valore”.

Se, per ogni osservazione t , il valore realizzato dalla variabile dipendente è y_t , così la verosimiglianza per questa osservazione se $y_t = 1$ è la probabilità che $y_t = 1$, e se $y_t = 0$ è la probabilità che $y_t = 0$. Il logaritmo della probabilità appropriata è il contributo alla verosimiglianza costituita dall'osservazione t .

“Dal momento che la probabilità che $y_t = 1$ è $F(X_t\beta)$, il contributo alla funzione di verosimiglianza per l'osservazione t quando $y_t = 1$ è $\log F(X_t\beta)$; così come il contributo alla funzione di verosimiglianza per l'osservazione t quando $y_t = 0$ è $\log (1 - F(X_t\beta))$.”

Quindi, se y è un n vettore con elementi y_t , la funzione di log-verosimiglianza per y può essere scritta così:

$$l(y, \beta) = \sum_{t=1}^n (y_t \log F(X_t\beta) + (1 - y_t) \log(1 - F(X_t\beta)))$$

Per ogni osservazione, uno dei termini all'interno dell'intera parentesi è sempre 0, e l'altro è sempre negativo. Il primo termine è 0 quando $y_t = 0$, e il secondo termine è 0 quando $y_t = 1$.

“Quando un termine non è 0, l'altro deve essere negativo, perchè diviene uguale al logaritmo di probabilità e questa deve essere minore di 1 ogniqualvolta $X_t\beta$ è finito. Per stimare il modello perfettamente $F(X_t\beta)$ dovrebbe essere uguale ad 1 quando $y_t = 1$ e uguale a 0 quando $y_t = 0$, e la completa espressione all'interno delle parentesi risultare uguale a 0. Questo potrebbe accadere solo se $X_t\beta = \infty$ quando $y_t = 1$, e $X_t\beta = -\infty$ quando $y_t = 0$.”

La funzione viene quindi massimizzata, e per i modelli logit e probit è totalmente concava rispetto a β ; le condizioni di primo ordine, o equazioni di verosimiglianza, definiscono unicamente lo stimatore $\hat{\beta}$. Queste equazioni di verosimiglianza possono essere così scritte:

$$\sum_{y_t=1}^n ((y_t - F(X_t\beta)) f(X_t\beta) X_{ti}) / (F(X_t\beta)(1 - F(X_t\beta))) = 0 \quad i=1, \dots, k$$

In particolare, per la distribuzione normale:

$$(\partial l(y, \beta)) / (\partial \beta) = \sum_{y_t=0} (-\phi_t / (1 - \Phi_t)) X_t + \sum_{y_t=1} (\phi_t / \Phi_t) X_t = 0$$

Per il modello logit, come suggerisce Greene, dopo un po' di manipolazione, le equazioni di verosimiglianza sono:

$$(\partial l(y, \beta)) / (\partial \beta) = \sum_{t=1}^n (y_t - (e^t / (1 + e^t))) X_t = 0$$

1.c GLI EFFETTI MARGINALI NELLE VARIABILI ESPLICATIVE

L'obiettivo di questo paragrafo è di mostrare l'effetto di una variazione marginale di una variabile esplicativa sulla probabilità che si verifichi l'evento.

Il modello logit stabilisce:

$$P(y_t = 1 | x_t) = (\exp(x_t)) / (1 + \exp(x_t)) = 1 / (1 + \exp(-x_t))$$

ed il modello probit stabilisce:

$$P(y_t = 1 | x_t) = \Phi(X_t \beta)$$

Quindi chiamando $P(y_t = 1 | x_t) = P_t$, possiamo verificare che variazioni delle variabili esplicative hanno diversi effetti su P_t a seconda del modello di riferimento:

$$(\partial P_t) / (\partial x_{it}) = (\exp(x_t)) / (1 + \exp(x_t)) * \beta_i = P_t (1 - P_t) * \beta_i$$

per il modello logit

$$(\partial P_t) / (\partial x_{it}) = \Phi(X_t \beta) * \beta_i$$

per il modello probit.

Come ci dice Greene: *“E’ ovvio che questi valori varieranno a seconda dei valori assunti dalle x_{it} . Nell’interpretare il modello stimato sarà utile calcolare l’effetto marginale associandolo alla significatività dei regressori.”*

Infatti i coefficienti che risultano dalla stima ci indicano solo la direzione della relazione, ma non la sua entità. L’entità della direzione viene appunto data dagli effetti marginali di cui abbiamo calcolato la formula sopra.

2. LO STAGE ALLA BANCA POPOLARE DI VICENZA

2.a STORIA E COMPOSIZIONE DELLA BANCA

La Banca Popolare di Vicenza è stata la prima banca ad essere costituita a Vicenza, nel 1866, e la prima banca popolare del Veneto. Da qui si è con il tempo estesa in tutto il Nord d'Italia, ed ampliata, inizialmente, attraverso l'acquisizione di piccole banche popolari di provincia, e più tardi acquisendo partecipazioni di maggioranza anche di banche popolari di regioni limitrofe. Inoltre, come dice il sito sulla storia della banca: *“Nel '98 la Banca Popolare di Vicenza è entrata a far parte, insieme al Banco Bilbao Vizcaya e all'INA, del nucleo degli azionisti di riferimento di BNL acquisendone una quota di capitale pari a circa l'8% (a inizio 2001 poco più della metà di questa quota è stata ceduta realizzando rilevanti plusvalenze).”* Nel 2006 la banca ha sottoscritto un contratto con BNP Paribas in forza del quale BPVi si è impegnata a vendere a BNP, che si è impegnata ad acquistare, parte delle Azioni BNL e segnatamente 75.000.000 azioni ordinarie. Dopo questo contratto, BPVi è rimasta proprietaria di 44.088.480 azioni ordinarie BNL.

Negli ultimi anni la Banca ha ampliato la sua rete anche nel Centro-Sud, realizzando un progetto con Banca Nuova e diventando operativa in Sicilia e Calabria con circa settanta sportelli che sono poi cresciuti, grazie all'acquisizione da Antonveneta MPS di altri 30 sportelli; in Toscana, invece, Cariprato conta circa 50 sportelli. Banca Popolare di Vicenza, Cariprato e Banca Nuova costituiscono il Gruppo Banca Popolare di Vicenza.

Il Gruppo BPVi possiede ora circa 600 filiali e circa 5000 dipendenti.

La Banca Popolare di Vicenza si trova, in territorio Veneto, delle concorrenti come Unicredit, IntesaSanPaolo, VenetoBanca, Banco Popolare di Verona, HipoAlpeAdriaBank, Sparkasse, Antonveneta MPS, Banca di Credito

Cooperativo, Cassa Rurale ed Artigiana di Brendola, ma resta comunque ben radicata nella sua terra d'origine, contando un buon numero clienti, tra cui privati ma anche piccole e grandi imprese, enti e associazioni, e di soci a cui vengono dati privilegi e servizi esclusivi.

Particolare attenzione è rivolta all'innovazione dei tradizionali prodotti/servizi inerenti all'Incasso/Pagamento, alla Gestione del Risparmio e alla Gestione del Credito, per renderli sempre più in grado di soddisfare le svariate esigenze della Clientela. Come ci dice anche il sito a riguardo: *“Numerosi i servizi offerti agli investitori, quali le contrattazioni di titoli, la custodia e l'amministrazione, la gestione dei valori mobiliari, tutti accompagnati da attività di consulenza personalizzata. Per rispondere alle tendenze del mercato immobiliare l'Istituto ha inoltre integrato la molteplicità di finanziamenti di medio-lungo termine con nuove formule (mutui trentennali, mutui con cap, mutui ad elevato loan to value).”*

Recentemente ad esempio BPVi e Camera di Commercio di Vicenza hanno rinnovato l'accordo che prevede significativi benefici per le imprese, attraverso un plafond per finanziamenti chirografari.

L'Istituto offre ai propri clienti il servizio di multicanalità Home Banking e Trading On-line, oltre che per tutti i normali servizi dello sportello bancario (operazioni sul c/c, bonifici, pagamenti vari, ecc.), anche per l'acquisto e la vendita di titoli azionari sia durante il giorno sia nella sessione after hours. Inoltre da poco esso offre un servizio di assistenza post vendita sui prodotti e servizi di investimento.

La Banca dispone anche di servizi per l'estero, infatti come riporta anche il sito: *“La Banca Popolare di Vicenza ha particolarmente specializzato la sua attività con l'estero, anche con servizi di assistenza e finanziamento offerti tramite la controllata irlandese BPV Finance International PLC - Dublino; assistenza alla costituzione ed amministrazione di entità giuridiche all'estero in collaborazione con Unione Fiduciaria; partecipazioni di minoranza nel capitale di banche operanti in Slovenia, Repubblica Ceca e Slovacca, Ungheria, Croazia, Bosnia e*

Romania, dove è operativo un Italian Desk. Dispone inoltre di un proprio ufficio di rappresentanza ad Hong-Kong. “

La Banca ha sempre dedicato un costante impegno alla valorizzazione della cultura veneta. Ne è la dimostrazione il palladiano Palazzo Tiene, sede storica dell'Istituto a Vicenza e patrimonio mondiale dell'Unesco, che ospita la collezione d'arte antica della Banca, nonché un piccolo museo della moneta. Qui si tengono inoltre mostre, concerti, conferenze e altre attività tese a valorizzare l'eredità culturale e l'immagine attuale del Veneto in Italia e all'estero.

2.b ESPERIENZA NELLO STAGE

Ho svolto lo stage presso la sede centrale della Banca Popolare di Vicenza all'U.O. Amministrazione Medio/Termine da febbraio 2008 a maggio 2008.

Qui mi sono occupata del controllo delle posizioni stipulate per l'erogazione o l'archiviazione di: Mutui Casa, Mutui Casa S.A.L., Mutui Ordinari, c/c ipotecari. Innanzitutto spieghiamo la differenza tra questi tipi di prodotti: i Mutui Casa riguardano esclusivamente l'acquisto, la costruzione o la ristrutturazione di una casa, quelli chiamati a S.A.L. hanno la caratteristica che l'importo richiesto viene concesso a seconda dello Stato di Avanzamento dei Lavori della casa che si sta per costruire o ristrutturare. I Mutui Ordinari possono avere scopi diversi, dalla semplice richiesta di liquidità all'acquisto di una licenza commerciale. Infine i c/c ipotecari dove l'importo finanziato viene versato su un apposito conto corrente aperto per l'occasione e in cui il debitore può andare a recuperare i soldi quando e quanto vuole.

Questo controllo prevedeva la verifica della presenza di alcuni documenti essenziali quali le copie autentiche del contratto di finanziamento, una delle quali in forma esecutiva, una copia della nota di iscrizione ipotecaria, la relazione legale definitiva che accertasse il grado ipotecario stabilito dalla delibera della

Banca e, se necessaria nel caso l'ipoteca gravasse su un immobile, una copia della polizza assicurazione incendio con vincolo a favore della Banca.

Dopo l'accertamento della presenza di questi documenti, si effettuava un controllo al contenuto di quest'ultimi confrontando nomi, cifre, scadenze, indirizzi con la bozza contrattuale e con altre attestazioni fornite prima della stipula. Altri documenti, quali quietanza degli oneri notarili, duplici di note di cancellazioni di eventuali precedenti ipoteche ed atto di compravendita, se non presenti, la Filiale di competenza veniva incaricata di raccogliarli tramite un'apposita mail.

Il controllo era affiancato da una scheda che doveva essere compilata per segnalare, oltre a ciò che era presente, eventuali mancanze di documenti che quindi erano stati richiesti o errori di scrittura negli stessi. La scheda era poi inserita all'interno del fascicolo che veniva in seguito registrato in un apposito file, assegnandogli uno specifico numero che corrispondeva alla fase di erogazione oppure all'archiviazione.

Nell'ultimo mese ho invece affiancato l'ufficio minute Mutui Casa, dove venivano predisposte le bozze contrattuali da inviare al Notaio.

La Filiale, innanzitutto, ci forniva una serie di documenti di cui dovevamo controllare l'esattezza: modelli di richiesta del finanziamento e delle condizioni da applicare, relazione legale preliminare del Notaio, perizia di stima dell'immobile di interesse e delibera di fido per accertarsi che la richiesta di mutuo fosse stata accolta dall'organo deliberante esatto, scelta che avveniva a seconda del rating del finanziato.

Sulla base di questi dunque, si svolgeva la personalizzazione delle condizioni a seconda del tipo di finanziamento scelto, come ad esempio mutuo a tasso variabile, tasso fisso, opzione sul tasso, rata fissa durata variabile e a seconda di eventuali convenzioni applicate, come quella inpdap. Dopodiché si passava alla compilazione della minuta contrattuale che veniva quindi inviata alla Filiale, la quale si occupava di fornirla alla controparte, per poi andare dal Notaio e procedere alla stipula del contratto di finanziamento.

3. ANALISI E STIMA DEL DATASET

In questo dataset sono riportate 22000 osservazioni riferite a persone fisiche (e quindi non ad imprese) contenenti più variabili, tra cui la variabile dipendente che andremo a studiare.

La variabile dipendente è la probabilità di sofferenza (“stato_def”), ovvero la probabilità di default che attiene al debitore e alla sua capacità di essere solvibile entro un determinato arco temporale, in questo caso un anno. Tutte le informazioni sono state raccolte al 31/12/2006 e la probabilità di default stabilita al 31/12/2007. La determinazione dello stato di insolvenza è data dal confronto di due voci: “operativo” ed “utilizzato”, che sono rispettivamente il residuo da pagare del piano di ammortamento ed il residuo reale che attiene al debitore; infatti “stato_def” assume valore 0 quando la posizione non risulta in sofferenza, ovvero le variabili monetarie “operativo” ed “utilizzato” sono uguali, ed il valore 1 quando la posizione è in sofferenza, cioè il debitore è insolvente e l’utilizzato è maggiore dell’operativo.

La probabilità di insolvenza viene spiegata facendo ricorso a diverse variabili. Una variabile importante è il tipo di finanziamento monetario tra cui troviamo il prestito personale (“prespers”), il mutuo chirografario (“mchir”) ed il mutuo ipotecario (“mipot”). Per mutuo chirografario si intende quel prestito monetario dove la garanzia è data dal solo impegno scritto del debitore e non da un bene materiale, come invece risulta essere nel caso del mutuo ipotecario, dove la garanzia è un’ipoteca, solitamente su beni immobiliari o terreni; solitamente lo scopo dei mutui ha un fine specifico che può essere la costruzione, la ristrutturazione o l’acquisto di un immobile o di un terreno, oppure anche l’acquisto di una licenza commerciale ed in alcuni, ma più rari casi un semplice supporto di liquidità. Il prestito personale, invece, rientra nella categoria dei prestiti non finalizzati cioè in quelle operazioni di finanziamento non direttamente collegate all’acquisto di uno specifico bene o servizio, ma al

fornimento di liquidità da usare a proprio piacimento e privi di una garanzia materiale.

Altre variabili utili a spiegare la probabilità di insolvenza sono: le variabili monetarie tra cui il reddito annuo netto percepito dalla controparte (“fatturato”), l’importo del finanziamento stabilito alla stipula (“importo_stipula”) e l’importo della garanzia ipotecaria se sussiste (“impgrav”).

Il tipo di garanzia prestato può essere un’ipoteca interna quando l’immobile a garanzia è di proprietà del soggetto che chiede il prestito (“ipotint”), esterna quando l’immobile a garanzia è di proprietà di un terzo soggetto (“ipotest”) oppure non avere ipoteca (“noipot”); il debitore può essere un lavoratore autonomo (“auton”), dipendente (“dipen”) oppure non occupato (“inocc”), ad esempio una casalinga o un pensionato. L’età (“eta”) e il sesso (“cod_sesso”) del debitore, il numero di persone che ha a carico (“num_pers_acarico”), la durata del finanziamento (“durata”), le eventuali garanzie da parte di terzi (“gar_pers”), la nazionalità (mediante una variabile che vale 1 se il debitore è straniero) e la provincia di nascita (considerando delle variabili binarie corrispondenti ai 14 comuni con più debitori, escluse le isole) costituiscono, assieme a tutte le altre sopra citate, le variabili esplicative.

Tutte le analisi e le stime econometriche saranno condotte utilizzando il software statistico STaTa[®].

3.a ANALISI DESCRITTIVA E INCROCIATA DELLE VARIABILI

Cominceremo andando ad analizzare le principali statistiche descrittive delle variabili di interesse. Le tabelle che riporteremo ci indicheranno, infatti, oltre alle principali statistiche, i percentili della distribuzione e i cinque valori più grandi e più piccoli.

Passeremo poi invece ad un'analisi incrociata analizzando come le diverse variabili si distribuiscano in media tra i tre tipi di finanziamenti, e infine un'analisi più approfondita rispetto alla variabile dipendente stato_def, in cui studieremo la distribuzione delle percentuali di insolvenza nelle variabili esplicative.

Cominciamo l'**analisi descrittiva** partendo dalla variabile dipendente, la probabilità di sofferenza:

probabilità di default	Freq.	Percent	Cum.
-----+-----			
non sofferenza	21,056	95.71	95.71
in sofferenza	944	4.29	100.00
-----+-----			
Total	22,000	100.00	

Come possiamo osservare, sia dalla distribuzione dei percentili e sia dalla tabella delle frequenze, la maggior parte dei debitori non è in sofferenza. Infatti solo il 4,29% delle controparti risulta essere insolvente.

Passiamo ora all'analisi delle variabili esplicative iniziando dal rapporto utilizzato/operativo:

utop				

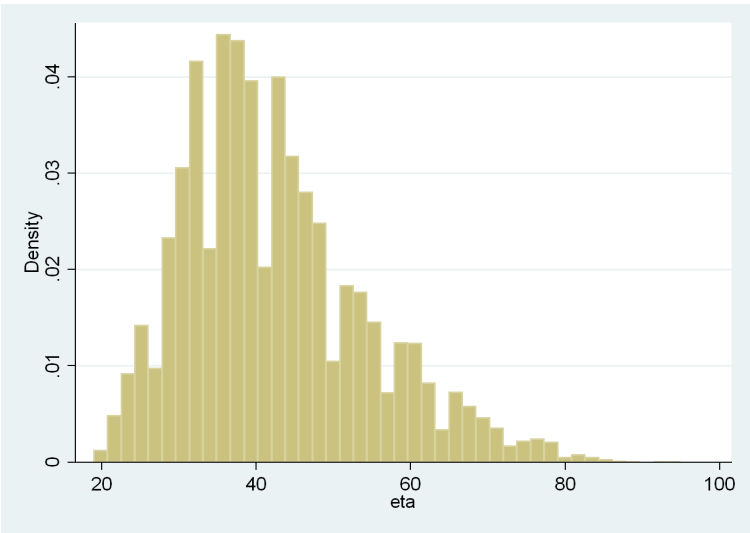
	Percentiles	Smallest		
1%	1	.5		
5%	1	.64		
10%	1	.9935484	Obs	21965
25%	1	1	Sum of Wgt.	21965
50%	1		Mean	1.001408
		Largest	Std. Dev.	.019522
75%	1	1.627103		
90%	1	1.781792	Variance	.0003811
95%	1	1.788224	Skewness	22.89868
99%	1.034864	2.018044	Kurtosis	845.607

La distribuzione, sopra riportata, di utilizzato/operativo ci mostra che questo rapporto è nella maggior parte uguale ad 1, e quindi che il debitore non è insolvente; mentre quando il rapporto è superiore ad 1 (“utilizzato”>”operativo”), vuol dire che la controparte non è regolare con il pagamento delle rate.

L’età dei finanziati:

eta					

	Percentiles	Smallest			
1%	22	19			
5%	26	19			
10%	29	19	Obs	22000	
25%	34	19	Sum of Wgt.	22000	
50%	40		Mean	42.42118	
		Largest	Std. Dev.	11.84653	
75%	49	88			
90%	59	89	Variance	140.3402	
95%	65	92	Skewness	.7942294	
99%	76.5	95	Kurtosis	3.408884	



Creando sette classi di età, a partire dai debitori con meno di 30 anni e andando di decennio in decennio fino agli ultraottantenni, vediamo come si distribuisce l'età:

classietà		Freq.	Percent	Cum.
< 30		2,978	13.54	13.54
30-40		8,087	36.76	50.30
40-50		6,038	27.45	77.74
50-60		2,950	13.41	91.15
60-70		1,393	6.33	97.48
70-80		484	2.20	99.68
>80		70	0.32	100.00
Total		22,000	100.00	

L'età media dei 22000 debitori è quindi intorno ai 42 anni. Lo si può dedurre dai percentili, dal grafico sopra riportato e dalla distribuzione delle classi di età descritta nella tabella, che ci indica che tra i 30 e i 50 anni troviamo la più alta concentrazione delle controparti.

Il sesso dei debitori:

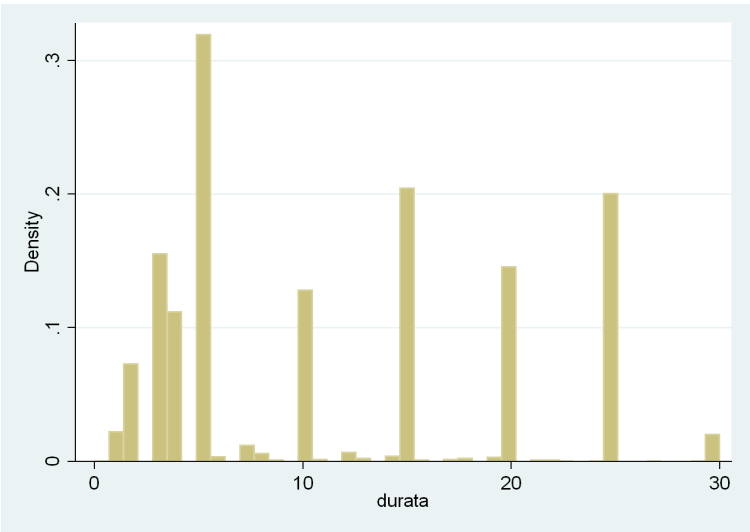
sesso		Freq.	Percent	Cum.
femmina		7,501	34.10	34.10
maschio		14,499	65.90	100.00
Total		22,000	100.00	

Osserviamo che la maggioranza dei debitori è di sesso maschile, 65.90% contro il 34.10% delle debitrice.

La durata dei finanziamenti, espressa in anni:

durata

Percentiles		Smallest	
1%	1	0	
5%	2	0	
10%	3	0	Obs 22000
25%	4	0	Sum of Wgt. 22000
50%	10		Mean 11.282
		Largest	Std. Dev. 8.276997
75%	20	30	
90%	25	30	Variance 68.50868
95%	25	30	Skewness .578831
99%	30	30	Kurtosis 1.919097



classidurata	Freq.	Percent	Cum.
-----+			
<5	10,480	47.64	47.64
5-10	2,330	10.59	58.23
10-15	3,375	15.34	73.57
15-20	2,365	10.75	84.32
20-25	3,136	14.25	98.57
25-30	314	1.43	100.00
-----+			
Total	22,000	100.00	

In media la durata dei diversi finanziamenti monetari è di circa 11 anni.

Osservando meglio dal grafico, troviamo dei picchi a 5, 10, 15, 20, 25, 30 che sovente sono le scadenze dei mutui e una concentrazione molto più alta tra i 0 e i 5 anni, come si vede anche dalla distribuzione suddivisa per classi di 5 anni, perché solitamente i prestiti personali hanno durata e importo di finanziamento inferiore ai mutui e quindi in più breve tempo restituibili.

In percentuale nettamente inferiore sono i finanziamenti con durata tra i 25 e i 30 anni.

La distribuzione del reddito annuo netto:

fatturato				

	Percentiles	Smallest		
1%	0	0		
5%	6177	0		
10%	9358.5	0	Obs	21220
25%	13000	0	Sum of Wgt.	21220
50%	16900		Mean	21849.39
		Largest	Std. Dev.	22395.35
75%	23500	450000		
90%	37000	547625	Variance	5.02e+08
95%	51532	549555	Skewness	7.878627
99%	104000	596670	Kurtosis	115.0513

Innanzitutto possiamo osservare che 780 debitori non possiedono un reddito, e che la distribuzione è decisamente spostata verso l'estrema destra. Infatti fino al cinquantesimo percentile abbiamo un reddito pari a € 16.900,00 e da lì fino al novantanovesimo percentile arriviamo fino a € 104.000,00, quindi la differenza tra le due metà della distribuzione è significativa. La media del reddito annuo netto delle controparti è intorno ai € 22.000,00.

Il rapporto importo_stipula/fatturato (cioè finanziamento su reddito):

impfat				

	Percentiles	Smallest		
1%	.1230769	3.33e-06		
5%	.2212389	3.89e-06		
10%	.3111111	.0040393	Obs	21011
25%	.5882353	.006779	Sum of Wgt.	21011
50%	1.684636		Mean	1477.359
		Largest	Std. Dev.	105995.7
75%	4.705883	857747		
90%	7.5	5500000	Variance	1.12e+10
95%	10	8500000	Skewness	91.97128
99%	50	1.15e+07	Kurtosis	8903.844

Il rapporto finanziamento su reddito ci indica che più alto è il rapporto più basso è il reddito del debitore rispetto all'importo del finanziamento. Vediamo che la distribuzione è decisamente spostata verso l'estrema destra, dove molto velocemente si alzano i valori. Ciò sta a significare che molte persone possiedono un reddito genericamente basso rispetto all'importo del finanziamento che richiedono.

Il tipo di garanzia prestato e di mutuo sottoscritto:

garanzia	Freq.	Percent	Cum.
-----+-----			
pegno interno	130	0.59	0.59
pegno esterno	64	0.29	0.88
ipoteca interna	9,073	41.24	42.12
ipoteca esterna	1,859	8.45	50.57
pluralità garanzie interne	13	0.06	50.63
pluralità garanzie esterne	1	0.00	50.64
pluralità garanzie	58	0.26	50.90
assenza garanzia	10,802	49.10	100.00
-----+-----			
Total	22,000	100.00	

Relativamente ai tipi di garanzia osserviamo che quasi la metà dei debitori non hanno un'ipoteca mentre l'altra metà è suddivisa tra chi ne possiede una interna (41,24%) e chi una esterna (8,45%); in ciò troviamo conferma osservando, come sotto descritto, che esattamente il 50% dei finanziamenti sono mutui ipotecari ("MI") e l'altra metà è suddivisa tra mutui chirografari ("MC") col 21% circa e prestiti personali ("PP") col quasi 29%, che non necessitano di ipoteca.

Una minoranza ha come garanzia pegni (solitamente su azioni/obbligazioni o su altri beni come oggetti di valore) oppure una pluralità di garanzie reali come beni materiali, che possono essere dell'affidato o di terzi.

tipo_fin		Freq.	Percent	Cum.
MC		4,647	21.12	21.12
MI		11,000	50.00	71.12
PP		6,353	28.88	100.00
Total		22,000	100.00	

L'impiego dei debitori:

impiego		Freq.	Percent	Cum.
dipendente		14,251	64.78	64.78
autonomo		5,790	26.32	91.10
non occupato		1,959	8.90	100.00
Total		22,000	100.00	

Vediamo che quasi il 65% dei debitori è lavoratore dipendente, il 26% è lavoratore autonomo, e il 9% è non occupato.

Passiamo ora all'**analisi incrociata** cominciando dall'età media dei debitori a seconda del tipo di finanziamento:

Summary of eta			
tipo_fin	Mean	Std. Dev.	Freq.
-----+-----			
MC	45.650742	12.792444	4647
MI	41.421091	10.095217	11000
PP	41.790493	13.420738	6353
-----+-----			
Total	42.421182	11.846526	22000

Vediamo dunque che per i mutui chirografari l'età media è leggermente superiore (45 anni) rispetto a quella dei mutui ipotecari e prestiti personali che risulta circa tra i 41 e 42 anni.

La durata media suddivisa per finanziamento:

Summary of durata			
tipo_fin	Mean	Std. Dev.	Freq.
-----+-----			
MC	5.0365827	3.100268	4647
MI	18.238818	5.9134868	11000
PP	3.8048166	1.1920426	6353
-----+-----			
Total	11.282	8.2769971	22000

La tabella ci conferma che in media i mutui ipotecari hanno durata maggiore, qui pari a 18 anni, mentre mutui chirografari e prestiti personali rispettivamente pari a 5 e 3,8.

I maschi e le femmine debitori:

tipo_fin	sesso		Total
	femmina	maschio	
MC	1,434	3,213	4,647
	19.12	22.16	21.12
MI	3,902	7,098	11,000
	52.02	48.96	50.00
PP	2,165	4,188	6,353
	28.86	28.88	28.88
Total	7,501	14,499	22,000
	100.00	100.00	100.

Osservando la tabella si può dire che la percentuale tra maschi e femmine nei prestiti personali è pressoché uguale, mentre nei mutui ipotecari le femmine sono in leggera maggioranza e nei chirografari, al contrario, sono i maschi ad avere una percentuale di poco superiore.

La media del fatturato condizionata al finanziamento:

tipo_fin	Summary of fatturato		
	Mean	Std. Dev.	Freq.
MC	26061.211	29621.169	4550
MI	22814.082	22340.903	10468
PP	17131.194	13901.201	6202
Total	21849.387	22395.349	21220

Osserviamo che in media sembra abbiano reddito più basso i debitori aventi un prestito personale, circa € 17.000,00, rispetto a chi ha un mutuo ipotecario, quasi € 23.000,00 e il reddito più alto chi ne ha uno chirografario, con circa € 26.000,00.

Per quanto riguarda invece l'importo del finanziamento alla stipula, come si vede dalla tabella sottostante, il mutuo ipotecario è intorno ai € 100.000,00, di importi nettamente inferiori il chirografario con circa € 25.000,00 e il prestito personale con quasi € 10.000,00.

Summary of importo_stipula			
tipo_fin	Mean	Std. Dev.	Freq.
-----+-----			
MC	25517.765	69037.728	4550
MI	97606.803	64333.774	10468
PP	9497.9239	5901.8185	6202
-----+-----			
Total	56397.736	68983.689	21220

In media il residuo da pagare secondo il piano di ammortamento:

Summary of operativo			
tipo_fin	Mean	Std. Dev.	Freq.
-----+-----			
MC	16944.687	64978.153	4647
MI	81743.578	63351.441	11000
PP	5633.3319	4820.0143	6353
-----+-----			
Total	46077.722	64751.247	22000

I residui da pagare seguono le medie degli importi alla stipula, ovvero per i mutui ipotecari il residuo più alto con circa € 82.000,00, poi i chirografari con quasi € 17.000,00 e i prestiti personali con € 5.600,00.

I tipi di garanzia:

Summary of tipo_gar			
tipo_fin	Mean	Std. Dev.	Freq.
-----+-----			
MC	124.25694	4.0328829	4647
MI	112.23455	.90159886	11000
PP	124.84999	1.8393458	6353
-----+-----			
Total	118.417	6.5639564	22000

Riguardo i mutui chirografari e i prestiti personali vediamo che la media riportata è di circa 125 che corrisponde, ovviamente, all'assenza di garanzia. Mentre per i mutui ipotecari la media è circa 112 corrispondente all'ipoteca interna, infatti solo in misura molto inferiore compare l'ipoteca esterna pari a 113.

Dalla tabella sottostante vediamo la distribuzione dei lavoratori nei vari finanziamenti: i lavoratori dipendenti (1) hanno una percentuale molto bassa nei mutui chirografari (15,77%), seguita dai prestiti personali (32,95%) e infine i mutui ipotecari (51,28%). Analizzando invece i lavoratori autonomi (2), scopriamo che risultano essere in buona percentuale (51,23%) nei mutui ipotecari e la situazione si ribalta, rispetto ai dipendenti, tra mutui chirografari (32,49%) e prestiti personali (16,29%). Infine i non occupati (3) sembrano distribuiti in maniera più uniforme, leggermente minore la percentuale dei chirografari (26,44%) rispetto agli ipotecari (37,06%) e prestiti personali (36,50%).

tipo_fin	cod_impiego			Total
	1	2	3	
MC	2,248	1,881	518	4,647
	15.77	32.49	26.44	21.12
MI	7,308	2,966	726	11,000
	51.28	51.23	37.06	50.00
PP	4,695	943	715	6,353
	32.95	16.29	36.50	28.88
Total	14,251	5,790	1,959	22,000
	100.00	100.00	100.00	100.00

Quindi vediamo in che percentuale troviamo gli stranieri:

```
count if tipo_fin=="MC" & comune=="
524
```

```
count if tipo_fin=="PP" & comune==" "
1632
```

```
count if tipo_fin=="MI" & comune==" "
1465
```

Sulle nostre 22000 osservazioni la percentuale di stranieri è del 16,45%. E' del 45,07% nei prestiti personali, del 40,45% nei mutui ipotecari e del 14,47% in quelli chirografari.

Analizziamo anche la distribuzione dell'eventuale garanzia da parte di persone terze:

	gar_pers		
tipo_fin	0	1	Total
-----+-----+-----			
MC	4,451	196	4,647
	20.69	40.08	21.12
-----+-----+-----			
MI	11,000	0	11,000
	51.14	0.00	50.00
-----+-----+-----			
PP	6,060	293	6,353
	28.17	59.92	28.88
-----+-----+-----			
Total	21,511	489	22,000
	100.00	100.00	100.00

Come possiamo vedere le garanzie personali riguardano solamente i finanziamenti che non hanno ipoteca, ovvero i mutui chirografari e i prestiti personali; in particolare con il quasi 60% i debitori possedenti un prestito personale ricorrono a garanzie da parte di persone terze.

Ed infine l'**analisi incrociata variabile dipendente-variabili esplicative**.

I tipi di finanziamento:

tipo_fin	stato_def		Total
	0	1	
MC	4,410	237	4,647
	20.94	25.11	21.12
MI	10,673	327	11,000
	50.69	34.64	50.00
PP	5,973	380	6,353
	28.37	40.25	28.88
Total	21,056	944	22,000
	100.00	100.00	100.00

Risultano essere in sofferenza circa il 40,25% dei debitori possedenti un prestito personale, il 34,64% dei possedenti un mutuo ipotecario e il 25,11% di chi ne ha uno ipotecario.

La durata dei finanziamenti, espressa in anni:

Classidurata	stato_def		Total
	0	1	
<5	9,889	591	10,480
	46.97	62.61	47.64
5-10	2,262	68	2,330
	10.74	7.20	10.59
10-15	3,296	79	3,375
	15.65	8.37	15.34
15-20	2,305	60	2,365
	10.95	6.36	10.75

20-25		3,005	131		3,136
		14.27	13.88		14.25
-----+-----+-----					
25-30		299	15		314
		1.42	1.59		1.43
-----+-----+-----					
Total		21,056	944		22,000
		100.00	100.00		100.00

Possiamo osservare che la percentuale maggiore di sofferenti si trova al di sotto dei 5 anni di durata (62,61%), dove peraltro abbiamo più finanziamenti. Tra i 20-25 anni troviamo il 13,88% di insolventi, e poi in scalare tra 10-15 l'8,37%, tra 15-20 il 6,36% e infine, la percentuale minore, tra i 25-30 anni, col 1,59%.

I lavoratori:

	stato_def		
cod_impiego	0	1	Total
-----+-----+-----			
1	13,670	581	14,251
	64.92	61.55	64.78
-----+-----+-----			
2	5,498	292	5,790
	26.11	30.93	26.32
-----+-----+-----			
3	1,888	71	1,959
	8.97	7.52	8.90
-----+-----+-----			
Total	21,056	944	22,000
	100.00	100.00	100.00

Quelli, in percentuale, maggiormente in sofferenza risultano essere i lavoratori dipendenti (1) con il 61,55%, poi gli autonomi (2) col 30,93% e infine i non occupati (3) con il 7,52%.

Delle 944 osservazioni che si trovano in sofferenza, 254 sono debitori stranieri, ovvero il 27%.

Le garanzie:

gar	stato_def		Total
	0	1	
112	8,816	257	9,073
	42.36	27.93	41.75
113	1,791	68	1,859
	8.60	7.39	8.55
125	10,207	595	10,802
	49.04	64.67	49.70
Total	20,814	920	21,734
	100.00	100.00	100.00

In netta maggioranza, con il 64,67% risultano più insolventi i debitori che non hanno garanzia (125), seguono coloro che hanno un'ipoteca interna (112) con il 27,93% e infine chi possiede un'ipoteca esterna (113) col 7,39%.

I maschi e le femmine:

sesso	stato_def		Total
	0	1	
femmina	7,176	325	7,501
	34.08	34.43	34.10
maschio	13,880	619	14,499
	65.92	65.57	65.90
Total	21,056	944	22,000
	100.00	100.00	100.00

Seguendo, diciamo, le corrispondenti percentuali di debitori, risultano maggiormente in sofferenza i maschi con 65,57% contro il 34,43% delle femmine.

Per quanto concerne l'età, invece, nella tabella sottostante vediamo che l'età più insolvente è quella tra i 30-40 anni con il 34,75%, segue tra i 40-50 col 27,12%, in queste fasce d'età dopotutto riscontriamo anche la più alta concentrazione di debitori. Nelle fasce <30 e tra i 50-60 abbiamo una percentuale media del circa 15%, mentre dai 60 anni in su le percentuali scendono e si aggirano in media attorno al 3-4%.

classietà	stato_def		Total
	0	1	
< 30	2,812	166	2,978
	13.35	17.58	13.54
30-40	7,759	328	8,087
	36.85	34.75	36.76
40-50	5,782	256	6,038
	27.46	27.12	27.45
50-60	2,825	125	2,950
	13.42	13.24	13.41
60-70	1,339	54	1,393
	6.36	5.72	6.33
70-80	472	12	484
	2.24	1.27	2.20
>80	67	3	70
	0.32	0.32	0.32
Total	21,056	944	22,000
	100.00	100.00	100.00

Relativamente alla garanzia di persone:

gar_pers	stato_def		Total
	0	1	
0	20,615	896	21,511
	97.91	94.92	97.78
1	441	48	489
	2.09	5.08	2.22
Total	21,056	944	22,000
	100.00	100.00	100.00

Circa il 95% delle controparti senza garanzie personali risulta in sofferenza, dopotutto quasi il 98% di tutti i debitori non possiede questo tipo di garanzia, così solo il 5% di coloro che la possiedono sono insolventi.

L'analisi approfondita di queste variabili ci ha meglio illustrato la composizione del nostro database, in particolare cogliendo quali sono le caratteristiche principali dei finanziamenti che vengono fatti alle famiglie. Ovvero che tipo di lavoratore, con quale fatturato e con quale età richiede un certo tipo di finanziamento con specifico importo e durata.

Poi, parte più importante, abbiamo voluto analizzare quali di queste caratteristiche risultino legate alla sofferenza, ovvero al debitore insolvente che non sta pagando regolarmente il residuo del finanziamento secondo il piano di ammortamento. Nella maggior parte dei casi queste variabili non ci rivelano qualcosa di particolare, in quanto seguono all'incirca le loro percentuali di presenza nel database, ma ci danno lo stesso un'idea generale di dove sia più concentrata la probabilità di default.

Variabile significativa è la garanzia, alla cui assenza è legata una percentuale abbastanza alta di debitori insolventi e superiore alla sua presenza nel dataset, così come i prestiti personali che rispetto alla loro presenza risultano però essere quelli dov'è concentrata quasi la metà delle posizioni in sofferenza. Da questa

analisi quindi le variabili che più corrispondono alla probabilità di sofferenza sono l'assenza di garanzia e il prestito personale.

3.b ANALISI ECONOMETRICA

Passiamo ora alla stima econometrica del modello in STaTa[®], specificando però prima come avverrà l'interpretazione dei risultati.

Per quanto riguarda l'interpretazione dei coefficienti il segno ci indica la direzione della relazione, ma il numero non è l'effetto marginale come nella stima OLS. Quindi se il segno del coefficiente è positivo la variabile determina una propensione alla sofferenza (valore 1), se è negativo il contrario.

Le variabili che più influiscono sulla probabilità di sofferenza, per cui cioè si rifiuta l'ipotesi nulla che il coefficiente sia pari a 0, sono quelle che hanno $P > |z|$ minore di 0.05 (al 5%) o 0.01 (all'1%).

Per migliorare la bontà di adattamento del nostro modello trasformiamo in logaritmi le variabili monetarie come “fatturato”, “operativo”, “utilizzato”, “importo_stipula”, “impgrav”, in modo da ridurre la variabilità e renderne più interpretabili i coefficienti.

Siccome la variabile dipendente è costruita sulla relazione tra operativo ed utilizzato non è il caso di tenerle contemporaneamente in considerazione, utilizzeremo quindi nella stima, tra le variabili indipendenti, o $\log(\text{operativo})$ oppure $\log(\text{utilizzato})$.

Più tardi, quando avremo scelto il nostro modello di riferimento, con il comando `dprobit` andremo ad analizzarci gli effetti marginali e proveremo la stima `logit`.

Inseriamo i logaritmi delle variabili monetarie, evidenziando che “lf” è $\log(\text{fatturato})$, “lu” è $\log(\text{utilizzato})$, “lo” è $\log(\text{operativo})$, “limps” è $\log(\text{importo_stipula})$ e “limpgrav” è $\log(\text{impgrav})$.

Cominciamo la specificazione omettendo log operativo) e vediamo la stima che ne esce:

Probit regression	Number of obs	=	21011
	LR chi2(32)	=	352.29
	Prob > chi2	=	0.0000
Log likelihood = -3555.1836	Pseudo R2	=	0.0472

stato_def	Coef.	Std. Err.	z	P> z	DEFINIZIONE
ipoint	.2425613	.400411	0.61	0.545	"IPOTECA INTERNA"
ipotest	.4244812	.4035456	1.05	0.293	"IPOTECA ESTERNA"
noipot	-.5065008	.1337714	-3.79	0.000	"ASSENZA DI IPOTECA"
mipot	-3.058547	.7618107	-4.01	0.000	"MUTUO IPOTECARIO"
prespers	.0217123	.045509	0.48	0.633	"PRESTITO PERSONALE"
maschio	-.0300499	.0348492	-0.86	0.389	
dipend	.0768199	.0844265	0.91	0.363	"LAVORAT. DIPENDENTE"
auton	.3311025	.0854047	3.88	0.000	"LAVORAT. AUTONOMO"
gar_pers	.2926844	.0872081	3.36	0.001	"GARANZIA DI PERSONE"
num_pers_a~o	-.1770152	.0402708	-4.40	0.000	"N.PERSONE A CARICO"
eta	.0023486	.0087845	0.27	0.789	"ETA'"
eta2	-.000031	.000097	-0.32	0.749	"ETA' AL QUADRATO"
durata	.010084	.0045275	2.23	0.026	"DURATA DEL FINANZ."
limpgrav	.1791273	.0543975	3.29	0.001	"LN(IMPORTO IPOTECA)"
lf	-.0350892	.0165482	-2.12	0.034	"LN(FATTURATO)"
lu	.2094998	.0326684	6.41	0.000	"LN(UTILIZZATO)"
lims	-.3641637	.0444106	-8.20	0.000	"LN(IMPORTO STIPULA)"
straniero	.2560921	.0419396	6.11	0.000	
belluno	-.1552489	.1484301	-1.05	0.296	
firenze	-.0560105	.1639383	-0.34	0.733	
milano	.0648871	.1309222	0.50	0.620	
napoli	.3294013	.1750735	1.88	0.060	
padova	-.0910385	.1338461	-0.68	0.496	
prato	-.0857803	.1005713	-0.85	0.394	
roma	.121265	.2380284	0.51	0.610	
torino	.3198153	.1847368	1.73	0.083	
treviso	-.4237876	.1821245	-2.33	0.020	
trieste	.1064759	.0992507	1.07	0.283	
udine	-.0758808	.0837652	-0.91	0.365	
venezia	.0403363	.1092428	0.37	0.712	
verona	-.6966596	.2606638	-2.67	0.008	
vicenza	-.1728745	.0733625	-2.36	0.018	

```

      _cons |      .7161285      .3742123      1.91      0.056
-----+-----

```

Proviamo invece ora a togliere il logaritmo di utilizzato e ad inserire log operativo): analizzeremo quindi questo ultimo modello, in quanto la variabile che più ci interessa è proprio operativo ovvero il residuo che un debitore dovrebbe avere per non essere insolvente.

```

Probit regression                                Number of obs   =      20977
                                                LR chi2(32)    =      341.21
                                                Prob > chi2    =      0.0000
Log likelihood = -3540.6106                    Pseudo R2      =      0.0460

```

stato_def	Coef.	Std. Err.	z	P> z	DEFINIZIONE
ipoint	.2428332	.4002012	0.61	0.544	"IPOTECA INTERNA"
ipotes	.4245899	.4033377	1.05	0.292	"IPOTECA ESTERNA"
noipot	-.5261348	.1343383	-3.92	0.000	"ASSENZA DI IPOTECA"
mipot	-3.111399	.7617793	-4.08	0.000	"MUTUO IPOTECARIO"
prespers	.0215781	.0456307	0.47	0.636	"PRESTITO PERSONALE"
maschio	-.0308911	.0349072	-0.88	0.376	
dipend	.1016296	.0851845	1.19	0.233	"LAVORAT. DIPENDENTE"
auton	.3519629	.0861765	4.08	0.000	"LAVORAT. AUTONOMO"
gar_pers	.2851478	.087915	3.24	0.001	"GARANZIA DI PERSONE"
num_pers_a-o	-.1800526	.0406238	-4.43	0.000	"N.PERSONE A CARICO"
eta	.0007925	.0087882	0.09	0.928	"ETA'"
eta2	-.0000118	.000097	-0.12	0.903	"ETA' AL QUADRATO"
durata	.0100384	.0045389	2.21	0.027	"DURATA DEL FINANZ."
limpgrav	.1829126	.0544217	3.36	0.001	"LN(IMPORTO IPOTECA)"
lf	-.0364349	.0165754	-2.20	0.028	"LN(FATTURATO)"
lo	.1827498	.0322776	5.66	0.000	"LN(OPERATIVO)"
lims	-.3365071	.0441941	-7.61	0.000	"LN(IMPORTO STIPULA)"
straniero	.2609107	.041946	6.22	0.000	
belluno	-.1984407	.1542698	-1.29	0.198	
firenze	-.1016036	.1705989	-0.60	0.551	
milano	.0672689	.1308755	0.51	0.607	
napoli	.3322911	.1747695	1.90	0.057	
padova	-.0888391	.1338167	-0.66	0.507	
prato	-.0957185	.1018433	-0.94	0.347	
roma	.1209069	.2376034	0.51	0.611	

torino	.3227184	.1846205	1.75	0.080
treviso	-.4196148	.1818697	-2.31	0.021
trieste	.1106451	.0992017	1.12	0.265
udine	-.0724611	.0837168	-0.87	0.387
venezia	.0437276	.1091799	0.40	0.689
verona	-.6897917	.2607614	-2.65	0.008
vicenza	-.1690653	.0733279	-2.31	0.021
_cons	.7323431	.3752384	1.95	0.051

Il soggetto di riferimento è femmina, non occupato, ricorre al mutuo chirografario e ad una pluralità di garanzie o pegni.

Innanzitutto possiamo dire quali variabili sono risultate significative e che quindi influiscono maggiormente sulla probabilità di sofferenza: l'assenza di garanzia ipotecaria, che sia un mutuo ipotecario, che il debitore sia un lavoratore autonomo, che ci siano o meno eventuali garanzie da parte di terzi, il numero di persone che ha a carico il finanziato, la durata del finanziamento e il logaritmo di fatturato (queste ultime due significative solo al 5%), l'importo dell'ipoteca se c'è (logaritmo), l'operativo (logaritmo), l'importo stabilito alla stipula (logaritmo), che il debitore sia straniero o meno e che sia nato nelle province di Treviso, Vicenza (queste ultime due significative solamente al 5%) e Verona.

L'analisi dei coefficienti ci indica soltanto il segno della direzione nella relazione, ad esempio le variabili "noipot", "mipot", "num_pers_acarico", "lf", "lims", "Treviso", "Verona" e "Vicenza" sembra riducano la probabilità mentre "auton", "gar_pers", "durata", "limpgrav", "lo", "straniero" sembrano aumentarla.

Per capire in quale percentuale le variabili che sono risultate significative influiscano sulla probabilità di sofferenza, andiamo ora a calcolare gli effetti marginali di questo modello:

Probit regression, reporting marginal effects

Number of obs = 20977

LR chi2(32) = 341.21

Prob > chi2 = 0.0000

Log likelihood = -3540.6106

Pseudo R2 = 0.0460

stato_~f	dF/dx	Std. Err.	z	P> z	x-bar	DEFINIZIONE
-----+-----						
ipotint*	.0203436	.035179	0.61	0.544	.408495	"IPOTECA INTERNA"
ipotest*	.0465601	.0577176	1.05	0.292	.082233	"IPOTECA ESTERNA"
noipot*	-.0431387	.0115371	-3.92	0.000	.498165	"ASSENZA DI IP."
mipot*	-.3985948	.1451853	-4.08	0.000	.493636	"MUTUO IPOTECAR."
prespers*	.0017425	.0037142	0.47	0.636	.292845	"PREST. PERS."
maschio*	-.0024972	.0028473	-0.88	0.376	.663155	
dipend*	.007917	.0064549	1.19	0.233	.658006	"LAVOR. DIPEND."
auton*	.0330472	.0093786	4.08	0.000	.257949	"LAVOR. AUTONOMO"
gar_pers*	.0290692	.0110734	3.24	0.001	.021452	"GARANZIA PERS."
num_pe-o	-.0144234	.0032297	-4.43	0.000	1.16108	"N.PER. A CARICO"
eta	.0000635	.000704	0.09	0.928	42.3903	"ETA'"
eta2	-9.46e-07	7.77e-06	-0.12	0.903	1937.08	"ETA'AL QUADRATO"
durata	.0008041	.0003628	2.21	0.027	11.2956	"DURATA DEL FIN."
limpgrav	.0146526	.0043439	3.36	0.001	5.91231	"LN(IMP.IPOTECA)"
lf	-.0029187	.0013271	-2.20	0.028	9.73845	"LN(FATTURATO)"
lo	.0146395	.0025675	5.66	0.000	9.80737	"LN(OPERATIVO)"
lims	-.0269565	.0035052	-7.61	0.000	10.2737	"LN(IMP.STIPULA)"
strani-o*	.0244708	.0045377	6.22	0.000		
belluno*	-.0133719	.0085957	-1.29	0.198		
firenze*	-.0074409	.0113707	-0.60	0.551		
milano*	.0057125	.0117601	0.51	0.607		
napoli*	.0355143	.0238938	1.90	0.057		
padova*	-.0065893	.0091599	-0.66	0.507		
prato*	-.0070651	.0069005	-0.94	0.347		
roma*	.0107773	.0234254	0.51	0.611		
torino*	.0342285	.0249126	1.75	0.080		
treviso*	-.0234057	.0065093	-2.31	0.021		
trieste*	.0097361	.0095414	1.12	0.265		
udine*	-.0054717	.0059474	-0.87	0.387		
venezia*	.0036364	.0094182	0.40	0.689		

verona*		-.0305804	.0050639	-2.65	0.008
vicenza*		-.0118944	.0044847	-2.31	0.021

```
-----+-----
obs. P | .0428565
pred. P | .0365744 (at x-bar)
-----
```

(*) dF/dx is for discrete change of dummy variable from 0 to 1
z and P>|z| correspond to the test of the underlying coefficient being 0

L'effetto marginale è dato dal prodotto tra il beta e la densità della normale nella specificazione media dei dati.

Le variabili con l'asterisco sono tutte variabili dummy, il comando dprobit riporta anche il loro cambiamento discreto.

L'assenza di ipoteca riduce del 4% la probabilità di essere insolvente, l'avere un mutuo ipotecario piuttosto che uno chirografario la riduce del circa 39%. Essere un lavoratore autonomo piuttosto che non occupato e l'eventuale garanzia da parte di terzi l'aumentano ciascuna del 3% circa, così come la durata del finanziamento del 0,08%, dell'1,4% il log(importo dell'ipoteca) e il log operativo) e del 2,5% l'essere straniero.

La riducono invece dell'1,44 % il numero di persone a carico, dell'0,2% il log(fatturato), del 2,7% il log(importo alla stipula), del 2% l'essere nato nella provincia di Treviso, del 3% Verona e dell'1% Vicenza rispetto all'essere nati nelle isole.

L'avere come finanziamento un mutuo ipotecario riduce, come visto sopra, addirittura del circa 39% la probabilità di default; questo risultato può essere riferito al fatto che avere un mutuo ipotecario significa avere un'ipoteca e quindi una garanzia materiale per chi concede il finanziamento e un impegno maggiore per chi lo deve restituire.

Di seguito riportiamo la stima logit del modello preso in considerazione:

Logistic regression	Number of obs	=	20977
	LR chi2(32)	=	339.02
	Prob > chi2	=	0.0000
Log likelihood = -3541.7025	Pseudo R2	=	0.0457

stato_def	Coef.	Std. Err.	z	P> z	DEFINIZIONE
ipotint	.6404907	1.014007	0.63	0.528	"IPOTECA INTERNA"
ipotest	1.029249	1.020221	1.01	0.313	"IPOTECA ESTERNA"
noipot	-1.043677	.2604007	-4.01	0.000	"ASSENZA DI IPOTECA"
mipot	-6.873516	1.791805	-3.84	0.000	"MUTUO IPOTECARIO"
prespers	.0423536	.0970611	0.44	0.663	"PRESTITO PERSONALE"
maschio	-.0599098	.0759856	-0.79	0.430	
dipend	.2594167	.1944547	1.33	0.182	"LAVORAT. DIPENDENTE"
auton	.7900644	.1958509	4.03	0.000	"LAVORAT. AUTONOMO"
gar_pers	.5709035	.1752226	3.26	0.001	"GARANZIA DI PERSONE"
num_pers_a~o	-.4089687	.0948082	-4.31	0.000	"N. PERSONE A CARICO"
eta	-.000655	.019196	-0.03	0.973	"ETA'"
eta2	-3.77e-06	.0002133	-0.02	0.986	"ETA' AL QUADRATO"
durata	.0247002	.0101507	2.43	0.015	"DURATA DEL FINANZ."
limpgrav	.4002564	.1238398	3.23	0.001	"LN(IMPORTO IPOTECA)"
lf	-.0716117	.0337508	-2.12	0.034	"LN(FATTURATO)"
lo	.3929721	.0720236	5.46	0.000	"LN(OPERATIVO)"
lims	-.7228335	.0958282	-7.54	0.000	"LN(IMPORTO STIPULA)"
straniero	.5455405	.0884233	6.17	0.000	
belluno	-.456086	.3630429	-1.26	0.209	
firenze	-.2217932	.3712706	-0.60	0.550	
milano	.1299433	.2896081	0.45	0.654	
napoli	.6999177	.3541338	1.98	0.048	
padova	-.239991	.3118129	-0.77	0.441	
prato	-.211498	.2192561	-0.96	0.335	
roma	.2882718	.5163203	0.56	0.577	
torino	.6587541	.3757875	1.75	0.080	
treviso	-.9686738	.4542795	-2.13	0.033	
trieste	.2357317	.2162706	1.09	0.276	
udine	-.165427	.1903537	-0.87	0.385	
venezia	.0894481	.2410139	0.37	0.711	
verona	-1.714711	.7128054	-2.41	0.016	
vicenza	-.3849155	.170599	-2.26	0.024	

Possiamo osservare, come già sappiamo dalla teoria, che le stime logit e probit danno sostanzialmente gli stessi risultati (la differenza sta nel fatto che il termine di disturbo di probit ha varianza unitaria mentre quello di logit segue una distribuzione logistica con varianza pari a $(\pi^2/3)$).

Le variabili che risultano significative sono le stesse risultate dalla stima probit, con l'aggiunta di Napoli, significativa però, al limite, al 5% e che ci dice che l'essere nato a Napoli, piuttosto che nelle isole, aumenta la probabilità di sofferenza e che l'essere nati a Verona è significativo al 5% e non più all'1%.

Con queste analisi siamo giunti alla conclusione che la presenza o meno di un'ipoteca in un finanziamento, ma in modo particolare il mutuo ipotecario, è una variabile rilevante ed influisce sulla probabilità di insolvenza riguardante il debitore. Questa variabile può dividere il nostro database in due parti, finanziamenti con garanzia ipotecaria o meno; vogliamo quindi verificare l'ipotesi nulla di stabilità strutturale della relazione, contro l'ipotesi alternativa che le stime differiscano tra finanziamenti con presenza di ipoteca o meno. Una differenza significativa nelle stime starebbe ad indicare un cambiamento strutturale della relazione tra la variabile dipendente e le esplicative.

Costruiamo allora una regressione ausiliaria in cui la variabile "noipot" (ovvero assenza di garanzia ipotecaria) moltiplichiamo tutte le altre esplicative, escluse le dummy di città, prestito personale e garanzia di persone terze in quanto creerebbero problemi di collinearità. Specificiamo quindi un modello di regressione in cui tutti i parametri differiscono per i finanziamenti con ipoteca e senza ipoteca.

Usiamo il **test di Chow** per verificare l'ipotesi:

H0: stabilità strutturale della relazione

H1: stime differenti tra finanziamenti con ipoteca oppure no

Regressione del modello ristretto:

Probit regression	Number of obs	=	20977
	LR chi2(32)	=	341.21
	Prob > chi2	=	0.0000
Log likelihood = -3540.6106	Pseudo R2	=	0.0460

stato_def	Coef.	Std. Err.	z	P> z	DEFINIZIONE
ipotint	.2428332	.4002012	0.61	0.544	"IPOTECA INTERNA"
ipotest	.4245899	.4033377	1.05	0.292	"IPOTECA ESTERNA"
noipot	-.5261348	.1343383	-3.92	0.000	"ASSENZA DI IPOTECA"
mipot	-3.111399	.7617793	-4.08	0.000	"MUTUO IPOTECARIO"
prespers	.0215781	.0456307	0.47	0.636	"PRESTITO PERSONALE"
maschio	-.0308911	.0349072	-0.88	0.376	
dipend	.1016296	.0851845	1.19	0.233	"LAVORAT. DIPENDENTE"
auton	.3519629	.0861765	4.08	0.000	"LAVORAT. AUTONOMO"
gar_pers	.2851478	.087915	3.24	0.001	"GARANZIA DI PERSONE"
num_pers_a~o	-.1800526	.0406238	-4.43	0.000	"N. PERSONE A CARICO"
eta	.0007925	.0087882	0.09	0.928	"ETA'"
eta2	-.0000118	.000097	-0.12	0.903	"ETA' AL QUADRATO"
durata	.0100384	.0045389	2.21	0.027	"DURATA DEL FINANZ."
limpgrav	.1829126	.0544217	3.36	0.001	"LN(IMPORTO IPOTECA)"
lf	-.0364349	.0165754	-2.20	0.028	"LN(FATTURATO)"
lo	.1827498	.0322776	5.66	0.000	"LN(OPERATIVO)"
lims	-.3365071	.0441941	-7.61	0.000	"LN(IMPORTO STIPULA)"
straniero	.2609107	.041946	6.22	0.000	
belluno	-.1984407	.1542698	-1.29	0.198	
firenze	-.1016036	.1705989	-0.60	0.551	
milano	.0672689	.1308755	0.51	0.607	
napoli	.3322911	.1747695	1.90	0.057	
padova	-.0888391	.1338167	-0.66	0.507	
prato	-.0957185	.1018433	-0.94	0.347	
roma	.1209069	.2376034	0.51	0.611	
torino	.3227184	.1846205	1.75	0.080	
treviso	-.4196148	.1818697	-2.31	0.021	

trieste	.1106451	.0992017	1.12	0.265
udine	-.0724611	.0837168	-0.87	0.387
venezia	.0437276	.1091799	0.40	0.689
verona	-.6897917	.2607614	-2.65	0.008
vicenza	-.1690653	.0733279	-2.31	0.021
_cons	.7323431	.3752384	1.95	0.051

Regressione del modello non ristretto, composto anche dalle nuove variabili costituite dall'interazione di "noipot" con le altre esplicative:

Probit regression	Number of obs	=	20977
	LR chi2(43)	=	364.86
	Prob > chi2	=	0.0000
Log likelihood = -3528.7834	Pseudo R2	=	0.0492

stato_def	Coef.	Std. Err.	z	P> z	DEFINIZIONE
ipotint	.2449756	.4018684	0.61	0.542	"IPOTECA INTERNA"
ipotest	.4014465	.405095	0.99	0.322	"IPOTECA ESTERNA"
noipot	1.315626	1.145486	1.15	0.251	"ASSENZA DI IPOTECA"
mipot	-1.388763	1.237558	-1.12	0.262	"MUTUO IPOTECARIO"
prespers	.0114744	.0463201	0.25	0.804	"PRESTITO PERSONALE"
maschio	-.004069	.0549454	-0.07	0.941	
dipend	-.0285614	.1356995	-0.21	0.833	"LAVORAT. DIPENDENTE"
auton	.2893042	.1366102	2.12	0.034	"LAVORAT. AUTONOMO"
gar_pers	.2791759	.0885296	3.15	0.002	"GARANZIA DI PERSONE"
num_pers_a-o	-.1064854	.0554622	-1.92	0.055	"N. PERSONE A CARICO"
eta	-.0079653	.0167886	-0.47	0.635	"ETA'"
eta2	.0001395	.0001845	0.76	0.450	"ETA' AL QUADRATO"
durata	.0155509	.0057011	2.73	0.006	"DURATA DEL FINANZ."
limpgrav	.027634	.1045662	0.26	0.792	"LN(IMPORTO IPOTECA)"
lf	-.0175085	.0273619	-0.64	0.522	"LN(FATTURATO)"
lo	.2094229	.0990653	2.11	0.035	"LN(OPERATIVO)"
lims	-.2275882	.1345808	-1.69	0.091	"LN(IMPORTO STIPULA)"
straniero	.2820253	.0665392	4.24	0.000	
belluno	-.20225	.1548012	-1.31	0.191	
firenze	-.1088274	.1738331	-0.63	0.531	
milano	.074787	.1311412	0.57	0.568	
napoli	.3167479	.1751651	1.81	0.071	
padova	-.0799477	.1342562	-0.60	0.552	

prato		-.1004628	.1046454	-0.96	0.337	
roma		.1277212	.2377438	0.54	0.591	
torino		.3279809	.185528	1.77	0.077	
treviso		-.4172428	.1825156	-2.29	0.022	
trieste		.1128754	.0994036	1.14	0.256	
udine		-.0705669	.0839177	-0.84	0.400	
venezia		.0441206	.1091287	0.40	0.686	
verona		-.6905861	.2627169	-2.63	0.009	
vicenza		-.1670539	.0736061	-2.27	0.023	
nomas		-.0446618	.0712602	-0.63	0.531	"NOIPOT*MASCHIO"
nodip		.2027405	.1736025	1.17	0.243	"NOIPOT*DIPEND"
noaut		.0736088	.1755393	0.42	0.675	"NOIPOT*AUTON"
nopers		-.1495237	.0818063	-1.83	0.068	"NOIPOT*N_PERS_ACAR"
noeta		.0097921	.0197362	0.50	0.620	"NOIPOT*ETA"
noeta2		-.0001825	.0002173	-0.84	0.401	"NOIPOT*ETA2"
nodur		.0004985	.0116169	0.04	0.966	"NOIPOT*DURATA"
nolf		-.0299332	.0345019	-0.87	0.386	"NOIPOT*LF"
nolo		-.0241539	.1048606	-0.23	0.818	"NOIPOT*LO"
nolims		-.1209736	.1425922	-0.85	0.396	"NOIPOT*LIMS"
nostran		-.0574473	.0831091	-0.69	0.489	"NOIPOT*STRANIERO"
_cons		-.8635534	1.089499	-0.79	0.428	

Nella stima le variabili "noint", "noest", "nomip", "nograv" (corrispondenti alla moltiplicazione di "noipot" rispettivamente con "ipotint", "ipotest", "mipot", "impgrav") non appaiono in quanto il loro prodotto è uguale a 0; inoltre nessuna delle variabili incrociate è risultata significativa nella regressione probit del modello non ristretto, al contrario di "auton", "gar_pers", "durata", "lo" e "straniero" che aumentano la probabilità di sofferenza e l'essere nati nelle province di Treviso, Verona, Vicenza che la diminuiscono.

Il test di Chow è il test di significatività congiunta delle restrizioni:

```
test nomas nodip noaut nopers noeta noeta2 nodur nolf nolo nolims nostran
```

- (1) nomas = 0
- (2) nodip = 0
- (3) noaut = 0
- (4) nopers = 0

```
( 5) noeta = 0
( 6) noeta2 = 0
( 7) nodur = 0
( 8) nol1f = 0
( 9) nolo = 0
(10) nolims = 0
(11) nostran = 0
```

```
      chi2( 11) =    22.98
Prob > chi2 =    0.0178
```

Osservando il p-value possiamo dire di non rifiutare l'ipotesi nulla all'1%. Così come, osservando il valore del test 22,98 che è minore di 24,72, valore di una χ^2 con 11 gradi di libertà al 99% di probabilità, non rifiuto H_0 . Al 5% dovrei invece rifiutarla.

In conclusione potremmo dire che, ad un livello di significatività dell'1%, le nostre nuove variabili sono uguali a 0, e quindi che la variabile assenza di ipoteca ("noipot") non è significativa, al di fuori dell'effetto sull'intercetta.

CONCLUSIONI

Grazie al dataset fornitomi dalla Banca Popolare di Vicenza e all'analisi econometrica, basata sull'approccio probit e logit, abbiamo constatato che alcune variabili connesse al prestito monetario ed altre, caratterizzanti la controparte, influenzano la probabilità di essere insolvente.

Avere un mutuo ipotecario piuttosto che uno chirografario riduce la probabilità di sofferenza di circa il 40%, ciò significa che il tipo di finanziamento influisce sul comportamento di un debitore. Una percentuale così alta si può spiegare collegandola al fatto che, avendo quest'ultimo come ipoteca un bene materiale, è più probabile che un debitore si attenga alle scadenze dei pagamenti piuttosto che andare incontro all'eventuale pignoramento del bene. Al contrario si sentirà meno vincolato possedendo un mutuo chirografario oppure un prestito personale dove spesso la garanzia è data semplicemente da altre persone.

Attraverso la stima e l'analisi dei coefficienti, abbiamo quindi stabilito che altre variabili influiscono sulla nostra variabile dipendente, come in parte ci aspettavamo; in particolare hanno un impatto negativo e diminuiscono la probabilità l'assenza di garanzia ipotecaria, il numero di persone a carico, il reddito annuo, l'importo del finanziamento alla stipula e l'essere nati nelle città di Treviso, Verona, Vicenza. L'aumentano invece la garanzia di persone terze, che il debitore sia un lavoratore autonomo piuttosto che non occupato, la durata del prestito monetario, l'importo dell'ipoteca se sussiste, il residuo da pagare del piano di ammortamento, l'essere straniero. L'entità di questi impatti (positivi e negativi) sulla probabilità di sofferenza oscilla tra 0,08% e il 4% circa; solo il mutuo ipotecario ha una percentuale molto alta, come sopra riportato.

Inoltre il test di Chow ci ha rivelato, ad un livello di significatività dell'1%, che le stime non differiscono tra finanziamenti con ipoteca o meno, e quindi che la presenza o meno di ipoteca non è significativa relativamente alle caratteristiche di cui abbiamo calcolato le interazioni, tranne che per l'effetto sull'intercetta.

BIBLIOGRAFIA

GIANNI AMISANO, *Lezioni di Econometria*, 1999.

WILLIAM H. GREENE, *Econometric Analysis*, fifth edition, Prentice Hall International, 2003.

JEFFREY M. WOOLDRIGE, *Introductory Econometrics*, A modern approach, Thomson, South- Western, 2^a ed, 2002.

DAVIDSON R. & MACKINNON J.G., *Econometric Theory and Methods*, Oxford University Press, 2004.

MARNO VERBEEK, *Econometria*, Zanichelli, 2005.

J. JOHNSTON, *Econometric Methods*, New York, McGraw Hill, 1971.