

# DISCUSSION PAPER SERIES

No. 3302

**A STRUCTURAL ECONOMETRIC  
MODEL OF PRICE DISCRIMINATION  
IN THE MORTGAGE LENDING  
INDUSTRY**

Robert Gary-Bobo and Sophie Larribeau

*PUBLIC POLICY*



**Centre for Economic Policy Research**

[www.cepr.org](http://www.cepr.org)

Available online at:

[www.cepr.org/pubs/dps/DP3302.asp](http://www.cepr.org/pubs/dps/DP3302.asp)

# A STRUCTURAL ECONOMETRIC MODEL OF PRICE DISCRIMINATION IN THE MORTGAGE LENDING INDUSTRY

**Robert Gary-Bobo**, THEMA, Université de Cergy-Pontoise and CEPR  
**Sophie Larribeau**, THEMA, Université de Cergy-Pontoise

Discussion Paper No. 3302  
April 2002

Centre for Economic Policy Research  
90–98 Goswell Rd, London EC1V 7RR, UK  
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999  
Email: [cepr@cepr.org](mailto:cepr@cepr.org), Website: [www.cepr.org](http://www.cepr.org)

This Discussion Paper is issued under the auspices of the Centre's research programme in **PUBLIC POLICY**. Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as a private educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions. Institutional (core) finance for the Centre has been provided through major grants from the Economic and Social Research Council, under which an ESRC Resource Centre operates within CEPR; the Esmée Fairbairn Charitable Trust; and the Bank of England. These organizations do not give prior review to the Centre's publications, nor do they necessarily endorse the views expressed therein.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

Copyright: Robert Gary-Bobo and Sophie Larribeau

April 2002

## **ABSTRACT**

### **A Structural Econometric Model of Price Discrimination in the Mortgage Lending Industry\***

We propose a model of discrimination in the market for mortgages. The model explains accepted loan applications and determines loan sizes and interest rates simultaneously. A competitive and a discriminating monopoly version of the model are proposed. Offered interest rates and loan sizes are a function of observable borrower characteristics. The competitive model rests on a marginal condition, reflecting contract optimality, to which a zero-profit condition is added. In contrast, the discriminating monopoly maximizes profits under a borrower participation constraint, reflecting the availability of a rental market as an outside option. Each version of the model is a bivariate, non-linear model, and is estimated by standard maximum likelihood methods. The data used for estimation is a sample of clients of a French network of mortgage lenders. We show the presence of 'social discrimination' in the data, the loan conditions depending not only on the borrower's wage and down payment, but also on the borrower's occupational status. Abnormally high-risk premia in the competitive version of the model suggest the presence of market power, justifying an attempt at estimating its monopolistic version. The discriminating monopoly model estimates show that the borrowers' price-elasticity of demand for housing varies with occupational status, and is inversely related with the lender's interest rate mark-ups. This confirms that the lender exploits structural differences in the preferences to discriminate, as predicted by standard theories.

Keywords: discriminating monopoly, mortgage loans and price discrimination

Robert Gary-Bobo  
Université de Cergy-Pontoise  
THEMA  
33, boulevard du Port  
95011 Cergy-Pontoise Cedex  
FRANCE  
Tel: (33 1) 3425 6062  
Fax: (33 1) 3425 6233  
Email: robert.gary-bobo@eco.u-cergy.fr

Sophie Larribeau  
Université de Cergy-Pontoise  
THEMA  
33 boulevard du Port  
95011 Cergy-Pontoise Cedex  
FRANCE  
Email: sophie.larribeau@eco.u-cergy.fr

For further Discussion Papers by this author see:  
[www.cepr.org/pubs/new-dps/dplist.asp?authorid=125284](http://www.cepr.org/pubs/new-dps/dplist.asp?authorid=125284)

For further Discussion Papers by this author see:  
[www.cepr.org/pubs/new-dps/dplist.asp?authorid=157469](http://www.cepr.org/pubs/new-dps/dplist.asp?authorid=157469)

\* We thank Richard Arnott, Marc Ivaldi, and Daniel McFadden for their advice during a preliminary stage of the present research, the participants and organizers of the 12th CREST-NBER Franco-American Seminar on Economics and the 1st CEPR Workshop on Applied IO, held in Toulouse, October 1999, the participants and organizers of the CEPR Conference 'Discrimination and Unequal Outcomes', held in Le Mans, France, January 2002, Florian Heider, Carmen Matutes, Nathalie Picard, Jean-Charles Rochet and Nancy Wallace, for comments and helpful discussions. The first author thanks John Quigley for his help and hospitality in June 2000. We are responsible for remaining errors.

Submitted 13 March 2002

## 1. Introduction

To the best of our knowledge, a handful of contributions only have attempted to estimate a structural econometric model of price discrimination, and none have studied the related phenomenon of interest rate discrimination, which can be viewed as a particular case, with such methods<sup>1</sup>. In the present paper, we propose a structural econometric model of a market for mortgages, and estimate it on micro-data.

The model determines interest rates and loan sizes simultaneously as a function of observable borrower characteristics. It can thus be viewed as describing a case of first-degree price discrimination in which the lender imposes differing price-quantity pairs to different types of borrowers. In addition, our data set shows that the lender exerts a form of market power. We have therefore specified and estimated two versions of the model, a competitive, and a discriminating monopoly version.

The competitive variant of our model rests on the idea of "competition in contracts": in equilibrium, any additional entry of a lender with loan offers cannot at the same time attract borrowers and make strictly positive profits. Two equations then describe the equilibrium: the first is a zero-profit condition; the second is a necessary condition for contract optimality, expressing the efficiency of trading between lenders and borrowers. These two equations, once solved, give the amount lent and the interest rate charged as a function of borrower characteristics. The discriminating monopoly variant of the model rests on the idea of surplus-extraction by a monopolist: the lender maximizes profit subject to the participation constraint of borrowers. Each borrower has an outside option, which is to rent a house instead of buying one. The surplus-extraction or zero-surplus equation says that in monopolistic equilibrium, borrowers are indifferent between renting a house or accepting a mortgage contract offered by the lender. A second equation is a necessary condition for contract optimality, as in the competitive version. This is consistent with the classic result that a first-degree discriminating monopoly will propose efficient trades. These two equations determine again the loan size and interest rate as a function of borrower characteristics.

These assumptions give rise to nonlinear, bivariate econometric models which are estimated by standard maximum likelihood methods. The estimated risk-premia applied to borrowers in the competitive version are much too high to be reasonable. This indicates the presence of market power, because the estimated risk-premia are, in fact, interest-rate markups imposed on borrowers. The discriminating monopoly version of the model, in spite of its added complexity, gives much more reasonable estimated values of the risk-premium function.

We also find that the amount lent and the interest rate charged significantly vary with the borrower's income and downpayment. There is a clear indication that differences in treatment depend on the occupational status of borrowers: for instance, executives will pay less to reimburse their loans than blue collar workers, everything else being equal. More precisely, the mere fact of being indentified as an executive by the lender would

---

<sup>1</sup> In contrast, there is an important strand of empirical literature on racial discrimination in mortgage lending in the U.S., which is discussed below.

result in better loan conditions, even if the executive's income was in the range of a blue-collar worker's wage. Estimations also show that categories of borrowers have significantly different preference parameters, and these differences are exploited by the monopolist, as suggested by standard theory. For instance, workers have a smaller price-elasticity of demand for housing than executives. The borrower's price-elasticity of demand is inversely related with the interest-rate markup charged by the lender.

To the best of our knowledge, this type of nonlinear, structural approach had not been attempted before, and the estimations obtained show that the approach could be improved to exploit richer data sets, in other countries. We also think that our model could easily be adapted to test for the presence of racial discrimination in mortgage lending, in the U.S. (see the discussion in the next to last section of this paper).

Recent structural econometric approaches to nonlinear pricing or price discrimination include the pioneering work of Ivaldi and Martimort (1994), using data on energy provision to French dairy producers, and Bousquet and Ivaldi (1997), on telephone pricing. Third-degree price discrimination on the European car market has been studied by Goldberg and Verboven (2001). Clerydes (2001) uses data on book sales to study discriminatory pricing of paperbacks and hardcovers. Leslie (2001) studies price discrimination in the sales of tickets for a Broadway show. Cohen (2000) shows that in the U.S paper towel market, package sizes are used to price discriminate. McManus (2001) tests for the presence of second-degree discrimination, and therefore product design distortions, in the price-quality menus offered by coffee shops surrounding the University of Virginia. Miravete (2001) studies nonlinear tariffs and consumer choice in a menu of optional calling plans proposed by the Bell telephone company in Kentucky. Finally, Verboven (2002) uses differing driver preferences for gasoline and Diesel cars in Europe to estimate the extent of price discrimination by manufacturers. Many of the above quoted studies use a form of discrete choice model of product differentiation to represent the behaviour of heterogeneous consumers<sup>2</sup>.

#### *Relationships with the theoretical literature on price discrimination and credit rationing*

The theoretical literature on credit and banking has emphasized screening under adverse selection, proposing a theory which has the same formal structure as Rothschild and Stiglitz's (1976) model of competitive insurance markets. Among contributions to this topic, see Milde and Riley (1988), in which loan size is used as a screening instrument, and Bester (1985), establishing that collateral is a screening instrument, under asymmetric information. Calem and Stutzer (1995) have addressed the problem of racial discrimination with a theoretical screening model in which the probability of rejection of mortgage loan applications is used as a screening device: clients choose in a *menu of contracts* in which higher interest rates are traded off against higher probabilities of acceptance. This idea cannot be applied in the following, since we use data on accepted loan files only. Our theoretical model aims at modelling loans conditionally on acceptance. Brueckner (1994), and Stanton and Wallace (1998) address the delicate problem of informational asymmetries about lender mobility, and the associated risks of premature repayment and mortgage

---

<sup>2</sup> On the theory of discrete choice models and its application to oligopoly theory, see Anderson et al. (1992).

renegotiation. They construct a separating equilibrium in which borrowers with differing mobility select fixed rate mortgages with different combinations of rates and *points*<sup>3</sup>. This approach, which provides a good explanation for some of the observed mortgage "menu" in the United States, cannot be applied here, because the lender does not make use of points in our data set. It is also quite certain that the french provincial housing markets are characterized by much less mobility than the U.S. markets (this is an important cultural difference between the two countries which has not been studied, as far as we know). It follows that the interest-rate risk generated by prepayment on loans is apparently negligible in France, at least as a first approximation.

In the theoretical model of Brueckner (2000), borrowers self-select by choosing different initial loan-to-value ratios, high interest rates being associated with high LTV, and the unobservable borrower characteristic driving self-selection is the level of personal default costs. But Brueckner's model is a refinement of the typically American strategic default or default-option theory, which is not applicable to French data, because the French borrower's liability is not limited to the value of his (her) house. The initial LTV ratio (in fact, the related "downpayment ratio") plays a role in our empirical analysis; it has a statistically significant effect as an argument of our estimated risk-premium function. We find that the risk-premium is a decreasing function of the downpayment ratio. But this could simply be the mechanical result of a better collateralization of the loan. The model presented below maps the downpayment, the income, and other exogenous characteristics of borrowers into (interest-rate, loan size) pairs, and the LTV ratio is endogenous.

To sum up, our French lenders, which are certainly somewhat old-fashioned, as compared to US professionals, are also working in a different legal environment, and do not offer choices in a "menu of contracts" to their clients. This is an important reason for modelling our lender as a first-degree discriminating monopolist: given our data set, there are no banking practices, and no compelling dimensions of both the borrower characteristic and the credit contract spaces along which to construct a reasonable model of self-selection or second-degree discrimination.

At this point, it should be noted that none of the structural econometric approaches of price discrimination quoted above went as far as to use a condition expressing that product quality ranges maximize profit to help identify their model's parameters: this range is exogenously given. It is difficult to capture taste heterogeneity in a model of demand for differentiated products and to derive a measure of the extent of price discrimination or product-quality distortions, and the assumption that a producer's product line is optimal does not seem to have been tested or used as an identification restriction in the literature. Given this difficulty, a contribution of the present paper consists in the use of assumptions on the seller's profit-maximizing behaviour (i.e., our surplus-extraction equation) to help identify structural parameters in a price-discrimination problem.

Another difficulty, pointed out by recent theoretical work, is that a realistic model of second-degree discrimination would be likely to entail several dimensions of uncertainty about the borrower's characteristics. For instance, at least two dimensions: the borrower's

---

<sup>3</sup> The idea of separation by mobility, when borrowers are better informed than lenders about their probability of moving, has been first modelled by Chari and Jagannathan (1989), but in a model in which the interest rate is constant.

marginal willingness to pay for a larger loan *and* a parameter determining the utility level of his (her) outside option on the housing market. A good model would also probably involve several screening instruments: loan size, loan life, downpayment, points, prepayment penalties, are possible such instruments. Thus, a multidimensional discrimination model would be required, as studied in the work Rochet and Choné (1998). This is known to lead — apart from hard technical problems — to much less separation of consumer types, and thus much less discrimination power than in the classic, one-dimensional model. Rochet and Choné (1998) show that bunching is a robust feature of optimal solutions in the multidimensional screening problem. The optimal solution cannot be explicitly computed in general, and with the exception of Armstrong (1999), not much has been published on the approximation of the optimal discrimination policy by simple pricing rules. A consensus on the form of the appropriate model has not yet emerged. Our approach, which is to model data as if they reflected first-degree discrimination, based on observed client characteristics, is therefore justified, at least as a first step.

In the following, section 2 is devoted to a description of the data and to the results of a preliminary linear econometric investigation. Section 3 presents the two versions of the model and section 4 their econometric estimations. Section 5 is devoted to a discussion of the empirical literature on racial discrimination in mortgage lending, and section 6 contains concluding remarks.

## 2. The data

### *2.1. Description*

We obtained a sample of observations on the clients of a French mortgage lender, the Crédit Hypothécaire de France (a nickname), hereafter CHF. The CHF is in fact a network of building societies, scattered on the French territory, the BSs. These local BSs have independent application screening and interest rate policies; they own in common a financial institution, which borrows money on national and international bond markets, and provides funds to the BSs. The CHF is a prudent and profitable institution, with a long history and a solid reputation. The BSs do not securitize their loans. Rating agencies have granted a very high note (AA+) to the CHF, so that the institution’s cost of funds is well approximated by, and closely parallels, the long-term rate on French state bonds (the "OAT" rate), with an almost constant difference of a few base points. Although the CHF has a special legal status, it is fair to describe the behavior of the local BSs as profit maximization. Until 1995, when the French government reformed its housing policy, the CHF had the privilege of distributing a particular kind of state subsidized home loans. This privilege has disappeared today, since all commercial banks can now initiate the same subsidized loans, but the CHF network has developed a strong expertise in mortgage lending to the working class, and a goodwill in accordance with this specialization. Its clientele is composed of a vast majority of modest income employees and workers. It is likely that many of the CHF clients would see their applications rejected elsewhere.

On top of distributing state subsidized loans, the characteristics of which are tightly regulated, the CHF also supplies the so-called "free loans", which are unregulated, ordinary mortgages. Until recently, the vast majority of these mortgages have been classic, fixed

rate, fixed repayment mortgages. The French mortgage law is in a sense simpler than the U.S. legal environment, since the borrowers' liability is *not* limited to the value of the house (lenders can pursue other borrower assets to mitigate default-related losses). In addition, house prices have not decreased very much in the provincial regions, which are the geographical origin of the sampled borrowers, during the observation period. It follows that strategic default (or the exercise of the default option) is not empirically relevant in the sample. In practice, mortgage defaults seem to be triggered by consumer insolvency, mostly due to loss of income. A form of unemployment insurance of mortgage loans does indeed exist, but it is not compulsory, it is expensive, and limited in scope. Informational asymmetries and moral hazard provide an explanation for the weakness of unemployment insurance of mortgages in France (on this topic, see Chiappori and Pinquet (1999)). These loans can in principle be renegotiated, the prepayment penalty being in all BSs around 3% of the principal's remaining value.

For the econometric investigations below, we have used a sample of 2610 observations on *accepted* free loans, originated from various BSs across France between 1989 and 1994. We have eliminated the subsidized loans. There is no information on rejected applications, and no observations of default or of repayment "incidents".

Each observation corresponds to a file, including, 1<sup>o</sup>) the amount of the loan, 2<sup>o</sup>) the loan interest rate (including insurance), 3<sup>o</sup>) the downpayment (savings used to buy the house by the borrower), 4<sup>o</sup>) the starting year, 5<sup>o</sup>) the loan term, 6<sup>o</sup>) the borrower's yearly wage, 7<sup>o</sup>) the age of the borrower, 8<sup>o</sup>) the family size, 9<sup>o</sup>) the borrower's occupational status, falling into 6 categories, and 10<sup>o</sup>) the geographical location of the BS granting the loan. We kept only four of the occupational status categories, the blue-collar workers (1180 observations), the white-collar employees (908 observations), the so-called intermediate professions (363 observations), and the executives (159 observations). We also do not use the geographical location variable in the estimations, and all regions have been pooled.

The model presented below has been constructed to be estimated with this limited set of information. It can of course easily be adapted to use more explanatory variables.

Table 1 provides descriptive statistics on the data. The amounts lent are not very high. The downpayment ratios, that is, (downpayment / loan + downpayment) are small, around 17%. The loan terms are distributed between 1 year and 20 years (the empirical distribution of loan terms is depicted on Figure 9). The nominal interest rates are very high with a mean value of 11,7%, but the real interest rates were also very high at the beginning of the nineties in France, the inflation rate being already quite low around 2% per year. The interesting aspect of the data is the substantial variance of the loan rates. This will allow the estimation of a risk-premium function, and of an interest-rate elasticity of the demand for housing, in spite of the fact that the observation period is short. Another striking fact is the markup on state bond rates, which is equal to 3% in the average. The lenders seem to exert a form of market power.

## *2.2. An exploratory linear model*

To gain some understanding of our data set's content, we have estimated a simple simultaneous equation model, explaining loan sizes and rates. More precisely, we started from the point of view that the relevant endogenous variable for this study is not the

interest rate itself, but the constant repayment annuity, denoted  $p$ , and defined as follows. A household who borrows for  $T$  years at continuous-time rate  $r$ , will repay the amount,  $p(r, T) = r/(1 - e^{-rT})$  at each instant of time to reimburse the loan completely, interest and principal, per euro borrowed (see below, subsection 3.1). The discrete-time equivalent of  $p$ , *i.e.*, the amount repaid each year, which is denoted  $P$ , is given by the expression

$$P(\bar{r}, T) = \frac{1 - \frac{1}{1+\bar{r}}}{1 - \left(\frac{1}{1+\bar{r}}\right)^T},$$

where  $\bar{r} = e^r - 1$  is the annual interest rate. If  $i$  denotes the lender's cost of funds, define  $C = P(i, T)$ . This is a measure of the lender's cost, which is commensurate with  $P$ . If borrowers are liquidity-constrained,  $P$  is a relevant endogenous variable, otherwise, the interest rate  $r$  would be the only important variable (on liquidity constraints, see Deaton (1992)). Let then  $m$  denote the amount lent and  $a$  the household's downpayment, so that  $H = m + a$  is the value of the house. Let finally  $w$  denote the household's yearly wage, and  $n$  the family size. Using the subset of 1180 unregulated loans granted to blue-collar workers to estimate a linear model explaining  $\log(H)$  and  $\log(P)$  simultaneously, we obtained the following results:

$$\begin{aligned} \log(H) = & 6.239 - 1.551 \log(P) \\ & (23.15) \quad (-20.75) \\ & + 0.102 \log(w) + 0.193 \log(a) + 0.028 n, \quad (A) \\ & (4.187) \quad (19.60) \quad (5.823) \end{aligned}$$

and

$$\log(P) = -0.4787 + 0.7239 \log(C) - 0.01862 \log\left(\frac{a}{H}\right). \quad (B)$$

(-19.39)
(62.82)
(-6.96)

Equations *A* and *B* have been estimated with the 3SLS method, to take care of simultaneity problems. The Student  $t$ -ratios are in parentheses. The adjusted  $R^2$  is .42 for equation *A* and .77 for equation *B*; the estimated correlation of *A* and *B*'s error terms is .244.

Note that  $a/H$ , the *downpayment ratio*, has a significant negative impact on  $P$ , and thus on the loan rate, as expected. But the striking fact is merely the significance of  $\log(P)$  in equation *A*, which determines house value. The coefficient on  $\log(P)$  can be interpreted (with precautions) as a form of price-elasticity of housing demand. We estimated the same model separately with the subsets of white collar employees, intermediate professions, and executives; the 3SLS estimated values of the coefficient on  $\log(P)$  are  $-1.347$  for white collars,  $-1.629$  for intermediates, and  $-1.648$  for executives, all highly significant. These findings suggest that the price-elasticity of house size varies with occupational status and even seems to increase (in absolute value) with the level in the social hierarchy. In the following, we try to provide an explanation for this phenomenon, and to relate it with observed interest-rate markups.

### 3. A model of mortgage lending

To describe the model, we first define the demand side, developing a continuous-time model of an expected utility maximizing, infinitely-lived and liquidity-constrained consumer-borrower. We then turn to the banking firm's profit function. This preliminary modelling work allows the presentation of two variants of the theory, corresponding to two assumptions relative to competition on mortgage lending markets: a competitive, and a monopolistic variant.

#### 3.1. The borrowers

Let  $\alpha$  denote a household's vector of characteristics. Each household of type  $\alpha$  is represented by an infinitely-lived, expected-utility maximizing consumer. The instantaneous utility of consumer  $\alpha$ , denoted  $u$ , is defined on the set of bundles  $(c, h)$ , where  $c$  is instantaneous aggregate consumption and  $h$  is housing, measured in constant-quality square meters. To sum up,  $u = u(c, h; \alpha)$ . Time  $t$  varies continuously, and it is assumed that consumer  $\alpha$ 's utility for a *certain* consumption path  $t \rightarrow (c(t), h(t))$  is

$$\delta \int_0^{+\infty} e^{-\delta t} u(c(t), h(t); \alpha) dt, \quad (1)$$

where  $\delta$  is a positive discount factor. Now, each consumer is subject to a liquidity constraint of an extreme form: she cannot borrow against future income, and can only obtain a home loan, with the entire house as a collateral. For simplicity, it is assumed that each consumer is a worker with a constant wage  $w$ . Yet, the worker is subject to a risk of default, which can be interpreted as loss of income, or as a layoff. To this, we add the rather simplifying assumption (with a European flavor), that once layed off, the worker remains unemployed forever. Default (or unemployment) randomly occurs at time  $t^*$ . The probability of defaulting between time 0 (the starting point of the loan here) and time  $t$  is assumed exponential and its cumulative distribution function is denoted  $F$ , that is,

$$F(t^* \leq t) = 1 - e^{-\theta t}, \quad (2)$$

where  $\theta$  is a nonnegative parameter. The density of  $F$  is  $f(t) = \theta e^{-\theta t}$ . The only event triggering mortgage defaults is unemployment, and it implies a complete loss of the house, due to mortgage foreclosure by the bank, for the sake of simplicity.

We consider classical, fixed interest, direct mortgage loans, with constant repayment (or self-amortizing annuities). Mortgage loans are characterized by an amount lent  $m$ , a fixed (continuous-time) interest rate  $r$ , a term  $T$ , and a starting date  $T^*$ ; they can be described by the array  $(r, m, T, T^*)$ . The borrower is endowed with a downpayment  $a$ ; this represents money that has been accumulated in the past, and is assumed entirely invested in the purchase of the house. The consumer rents a house if he or she does not borrow to buy one, and the accumulated sum of money  $a$  is then deposited in some bank and yields an interest  $i_0$ . To simplify the model, we assume that  $i_0 = 0$  in the following.

Now, if the house price per square meter is denoted  $\pi$ , then the house size (or housing consumption level) of a borrower is simply

$$h = \frac{m + a}{\pi}. \quad (3)$$

We define the *downpayment ratio* of a consumer as  $a/(m + a)$ .

The continuous-time constant amortizing repayment of a  $m$  euros loan of term  $T$ , originated at time  $T^* = 0$ , is  $p(r, T)m$ , and satisfies by definition,

$$me^{rT} = \int_0^T p(r, T)me^{rt} dt.$$

Straightforward integration then yields,

$$p(r, T) = \frac{r}{1 - e^{-rT}}. \quad (4)$$

For convenience, we call  $p(r, T)$  the *price* of the loan, and denote it simply by  $p$ . Note that all continuous-time variables can easily be transformed into their discrete-time counterparts, with the appropriate formulas.<sup>4</sup>

Then, because of the assumed liquidity constraint, the borrower faces a budget constraint at each time  $t$ . Taking consumption as a numeraire, the budget constraint has the simple form,

$$c = w - pm. \quad (5)$$

Definitions (3) and (5) show that, given  $w$ ,  $a$  and  $\pi$ , there is a one-to-one relationship between  $(p, m)$  and  $(c, h)$ . Given  $T$ , the interest rate  $r$  can be retrieved from  $p$  by inverting (4). It follows from this that a mortgage  $(r, m, T, T^*)$  is equivalently characterized by the vector  $(c, h, T, T^*)$ , given  $w$ ,  $a$  and  $\pi$ . In the following, it will be more convenient to reason in the  $(c, h)$  consumption-bundle plane, instead of the  $(r, m)$  plane.

#### *The expected utility of a borrower*

We define different levels of instantaneous utility as follows.

1°) When consumer  $\alpha$  repays a mortgage with price  $p$  and enjoys a house of size  $h$ , her utility level writes  $u = u(w - pm, h; \alpha)$ .

2°) If the consumer defaults, she loses her income and her house. In this case, her utility level is assumed to be independent of  $p$  and  $h$ ; it is denoted  $u = u_0(\alpha)$ .

3°) If the consumer never defaulted, and has completely repaid the loan, then  $u = u(w, h; \alpha)$ .

4°) If the consumer loses her income after the loan is completely repaid, she does not lose her house, and  $u = u(0, h; \alpha)$ .

5°) Finally, the consumer can rent a house. Let  $\rho$  denote the rent per square meter. If the consumer rents, she will choose  $(c, h) \geq 0$  so as to maximize  $u(c, h; \alpha)$  subject to the

---

<sup>4</sup> For instance, if  $T$  is expressed in years, and  $\hat{r}$  denotes the annual interest rate, then  $r = \log(1 + \hat{r})$ .

constraint  $c + \rho h = w$ . The solution of this standard maximization problem is the pair of demand functions  $(c^*(\rho, w; \alpha), h^*(\rho, w; \alpha))$ . The instantaneous indirect utility of a renter is thus defined as follows,

$$v^r(\rho, w; \alpha) = u[c^*(\rho, w; \alpha), h^*(\rho, w; \alpha); \alpha]. \quad (6)$$

Let now  $v(t, t^*, T^*, T; \alpha)$  denote the instantaneous utility of a consumer of type  $\alpha$  at time  $t$ , given that default occurs at time  $t^*$ , under a mortgage contract  $(r, m, T, T^*)$ . The expression for  $v$  can be computed with the utility levels defined above, since by assumption, the consumer rents her house from  $t = 0$  until  $T^*$ , loses her income at  $t^*$ , and loses her house at  $t^*$  if  $T^* < t^* < T + T^*$ . We can now define the discounted sum of utilities for the mortgage  $(r, m, T, T^*)$  as

$$U(t^*, T^*, T; \alpha) = \delta \int_0^{+\infty} v(t, t^*, T^*, T; \alpha) e^{-\delta t} dt. \quad (7)$$

Using the assumption that the date of default is exponentially distributed, we define the expected utility of a loan of term  $T$ , originated at  $T^*$ , as follows,

$$V(T^*, T; \alpha) = \int_0^{\infty} \theta e^{-\theta t^*} U(t^*, T^*, T; \alpha) dt^*. \quad (8)$$

It is now possible to compute  $V$ . After some cumbersome, but straightforward computations, we obtain the following result.

**Result 1**

$$\begin{aligned} V(T^*, T; \alpha) &= v^r(\rho, w; \alpha) [1 - e^{-(\theta+\delta)T^*}] \left( \frac{\delta}{\theta + \delta} \right) \\ &+ u_0(\alpha) [1 - e^{-(\theta+\delta)(T+T^*)}] \left( \frac{\theta}{\theta + \delta} \right) \\ &+ u(w - pm, h; \alpha) [e^{-(\theta+\delta)T^*} - e^{-(\theta+\delta)(T+T^*)}] \left( \frac{\delta}{\theta + \delta} \right) \\ &+ u(w, h; \alpha) \left( \frac{\delta}{\theta + \delta} \right) e^{-(\theta+\delta)(T+T^*)} \\ &+ u(0, h; \alpha) \left( \frac{\theta}{\theta + \delta} \right) e^{-(\theta+\delta)(T+T^*)}. \end{aligned} \quad (9)$$

From this result, one can derive the consumer's tenure choice conditions, with static expectations, that is, on the assumption that a mortgage characterized by  $(r, m, T)$  will always be available in the future. These tenure choice conditions are in fact equivalent to the borrower's *participation constraint*, from the point of view of the banker. Define  $V^r(\rho, w; \alpha)$  as the expected utility of a consumer who rents a house forever. Clearly, with the help of (9), one easily checks that

$$V^r(\rho, w; \alpha) = \lim_{T^* \rightarrow +\infty} V(T^*, T; \alpha) = v^r(\rho, w; \alpha) \left( \frac{\delta}{\theta + \delta} \right) + u_0(\alpha) \left( \frac{\theta}{\theta + \delta} \right). \quad (10)$$

Now, consumer  $\alpha$  will accept a loan  $(r, m, T, 0)$  at time 0 if and only if, for all  $T^* \geq 0$ ,

$$V(0, T; \alpha) \geq V(T^*, T; \alpha). \quad (11)$$

This condition has a remarkably simple equivalent formulation.

**Result 2**

Condition (11) is equivalent to

$$V(0, T; \alpha) \geq V^r(\rho, w; \alpha). \quad (12)$$

To prove this result, it is sufficient to show that  $V$  is decreasing with respect to  $T^*$ , if and only if condition (12) holds.

From now on, and to clarify notation, we denote

$$\bar{V}(c, h, T; \alpha) = V(0, T; \alpha), \quad (13)$$

letting  $c$  and  $h$  explicitly appear as arguments of the expected utility. Parameter  $\alpha$  must be viewed as a vector of characteristics including  $\theta$ ,  $w$ ,  $a$ ,  $\delta$ .

### 3.2. The lender

We now construct the bank's expected profit function from a mortgage loan to a borrower of type  $\alpha$ , starting at time  $T^* = 0$ . From now on, the reader must understand that all mortgages start at date  $T^* = 0$ .

We assume that the bank's liability structure can be modelled as if it sold bonds or certificates of deposit in exchange for an interest rate denoted  $i$  to finance its loans. The bank is also assumed to avoid any form of interest rate or refinancing risk. The timing of payments on its debt, interest and principal, matches the timing of revenues from its mortgage portfolio (amortization schedules are parallel). The loans are not securitized, and kept on the asset side of the bank's balance sheet (this assumption corresponds to common French practice). The interest rate  $i$  is used to discount the bank's future cash flows. If a client defaults at date  $t^*$ , for the sake of simplicity, we assume that the liquidation value of the house is zero. More realistic representations of mortgage foreclosure lead to very complex nonlinear expressions of expected profit, and happened to be intractable with the kind of data that we have.

With our simplifying assumptions, the date 0 net present value of a loan, conditional on  $t^*$ , denoted  $b(t^*)$ , writes

$$b(t^*) = -m + m \int_0^{t^*} p(r, T) e^{-it} dt. \quad (14)$$

The expected profit from a loan, can therefore be computed as,

$$\Pi(c, h, T; \alpha) = \int_0^T \theta e^{-\theta t^*} b(t^*) dt^* + e^{-\theta T} b(T). \quad (15)$$

Straightforward computations yield the following formula,

$$\Pi(c, h, T; \alpha) = -m + \frac{p(r, T)m}{p(\theta + i, T)}. \quad (16)$$

Given (16), the zero-profit condition  $\Pi(c, h, T; \alpha) = 0$  writes

$$p = p(\theta + i, T), \quad (17)$$

meaning that interest rates charged can be expressed as  $i$  plus a risk-premium  $\theta$ , in competitive equilibrium.

With this formulation, a simple change of variables in (16), using (3) and (5), yields the expression,

$$\Pi(c, h, T; \alpha) = \left( \frac{w}{p(\theta + i, T)} + a \right) - \pi h - \frac{c}{p(\theta + i, T)}, \quad (18)$$

which is linear with respect to  $(c, h)$ . The banker's marginal rate of substitution between  $c$  and  $h$  is therefore constant for each given type of borrower. More precisely,

$$\frac{\frac{\partial \Pi}{\partial c}}{\frac{\partial \Pi}{\partial h}} = \frac{1}{\pi p(\theta + i, T)}. \quad (19)$$

### 3.3. Competitive equilibrium in mortgage contracts

We assume here, (i), that a large number of banks compete to supply loans to consumers, and (ii), that a consumer's type  $\alpha$  is observable by every bank. Banks "compete in contracts". For given  $T$ , the banker who offers the most advantageous contract  $(c(T, \alpha), h(T, \alpha), T)$  will attract all the clients of type  $\alpha$ . It follows that competition will drive profits to zero for each pair  $(T, \alpha)$ : Formally, in equilibrium, if the contract  $(c(T, \alpha), h(T, \alpha), T)$  is traded,  $\Pi[c(T, \alpha), h(T, \alpha), T; \alpha] = 0$ . In addition, each contract  $(c(T, \alpha), h(T, \alpha), T)$  in the menu offered in equilibrium, will satisfy the familiar necessary condition for optimality,

$$\frac{\frac{\partial \Pi}{\partial c}}{\frac{\partial \Pi}{\partial h}} = \frac{\frac{\partial \bar{V}}{\partial c}}{\frac{\partial \bar{V}}{\partial h}}, \quad (20)$$

for otherwise, a competitor could offer a more advantageous contract to consumers of type  $\alpha$ , given  $T$ .

INSERT FIGURE 1 HERE

The competitive equilibrium is represented on figure 1(a). For any given  $T$ , the equilibrium contract  $(c^*, h^*) = (c(T, \alpha), h(T, \alpha))$  simply maximizes  $\bar{V}(c, h, T; \alpha)$  under the zero-profit constraint  $\Pi[c, h, T; \alpha] = 0$ , which is linear here.

### 3.4. The discriminating monopolist's menu of contracts

If banks are endowed with a form of market power, due to concentration in the mortgage lending sector, or due to the fact that competitors consistently reject the applications of some consumer types, then, a more appropriate model will be that of a first-degree discriminating monopolist. Perfect discrimination requires that, for each type pair  $(T, \alpha)$ , for which trade occurs, the contract  $(c(T, \alpha), h(T, \alpha), T)$  should maximize  $\Pi(c, h, T; \alpha)$  subject to the borrower's participation constraint (12), for given  $T$ . The necessary condition for optimality (20) will then again be satisfied for all  $(T, \alpha)$ . But the zero-profit condition should be replaced by the zero-surplus condition,

$$\bar{V}(c(T, \alpha), h(T, \alpha), T; \alpha) = V^r(\rho, w; \alpha), \quad (21)$$

which says that the entire consumer surplus is extracted by the monopolist. This situation is represented on figure 1(b): for each  $(T, \alpha)$ , the monopolist maximizes profit  $\Pi$ , with respect to  $(c, h)$ , subject to the constraint that  $\bar{V}(c, h, T; \alpha) \geq V^r(\rho, w; \alpha)$ . Expected profit is decreasing with respect to  $c$  and  $h$ : the bankers always prefer to reduce house size  $h$  (lend less) for a fixed reimbursement flow  $pm$ , and they always prefer to increase the rate, that is, reduce consumption  $c$ , for a given loan size  $m$ . The optimal contract is therefore located on the lowest iso-profit line compatible with borrower participation.

## 4. Estimation of the models

We will first propose a way of estimating the competitive model described above, and discuss the possibilities of testing for the presence of discrimination. We then present some estimation results. The same job will be done with the discriminating monopolist model.

### 4.1. Specification and estimation of the competitive model

In order to obtain a reasonably tractable formulation we assume that the instantaneous utility  $u$  is quasi-linear, with the particular parametric form,

$$u(c, h) = u_0 + c + \gamma \frac{h^\epsilon}{\epsilon}, \quad (22)$$

where  $\gamma$  is a function of individual characteristics and  $\epsilon$ , an individual preference parameter. The parameter  $\beta$  defined as

$$\beta = \frac{1}{1 - \epsilon},$$

with  $0 < \epsilon < 1$ , can be interpreted as the "price-elasticity of the demand for housing", as will be shown below. With the above specification of  $u$ , using (9) with  $T^* = 0$ , one gets,

$$\begin{aligned} \bar{V}(c, h, T; \alpha) = u_0 + \frac{w\delta}{\delta + \theta} e^{-(\delta + \theta)T} + \frac{c\delta}{\delta + \theta} (1 - e^{-(\delta + \theta)T}) \\ + \frac{\gamma}{\delta + \theta} \left[ \delta + \theta e^{-(\delta + \theta)T} \right] \frac{h^\epsilon}{\epsilon}. \end{aligned} \quad (23)$$

From (23), it is easy to compute the marginal rate of substitution of a borrower,

$$\frac{\partial \bar{V} / \partial c}{\partial \bar{V} / \partial h} = \frac{h^{\frac{1}{\beta}}}{\gamma K(\theta, \delta, T)}, \quad (24)$$

where by definition,

$$K(\theta, \delta, T) = \frac{1 + (\theta/\delta)e^{-(\theta+\delta)T}}{1 - e^{-(\theta+\delta)T}}. \quad (25)$$

Remark that if the borrower becomes very impatient, then  $\lim_{\delta \rightarrow +\infty} K(\theta, \delta, T) = 1$  and if the borrower becomes extremely patient,  $\lim_{\delta \rightarrow 0} K(\theta, \delta, T) = +\infty$ .

The contract optimality condition (20) can then be rewritten,

$$h = \left( \frac{\gamma K(\theta, \delta, T)}{\pi p(\theta + i, T)} \right)^{\beta}. \quad (26)$$

Using then the change of variable  $h = (m + a)/\pi$  and taking logs yields the expression,

$$\log(m + a) = \beta \log(\gamma) + (1 - \beta) \log(\pi) + \beta \log[K] - \beta \log[p(\theta + i, T)]. \quad (27)$$

Remark that (27) is very close to being a simple "demand for housing", in the expression of which  $p(\theta + i, T)$  plays the role of a housing price.

Recall that the zero-profit condition (17) writes  $p = p(\theta + i, T)$ . If we assume that both equations (27) and (17) are satisfied up to a random error term, (27) and (17) become a bivariate nonlinear system.

### *Econometric specification*

For the purpose of estimation, our bivariate model (17)-(27) should be fully specified. We consider  $(a, w, T)$  as exogenous variables. The discrete variable  $T$  could in principle be treated as endogenous, but this would lead us to a much more complicated model with three endogenous variables, and the equation determining  $T$  would be very difficult to estimate. In addition the empirical distribution of  $T$  is very much concentrated on 10, 15 and 20 years (see figure 9). This is why we treat  $T$  as exogenous. In addition, there are other exogenous observations on the households, such as family size, age, and occupational status, that can be used to explain differences in preferences for housing and default risk. Let  $X$  denote the vector of all exogenous variables, including,  $a, w, \pi, T$  and  $i$ . The vector  $(\beta, \gamma, \delta, \theta)$  fully characterizes a consumer. A possibility is then to assume that  $\gamma, \beta$  and  $\theta$  are themselves functions of  $X$ , and to specify functional forms  $\gamma(X), \theta(X)$  and  $\beta(X)$ , with parameters to be estimated ( $\delta$  could also, in principle, be specified that way).

We have been able to estimate a somewhat simplified version of (17)-(27). To limit the number of parameters, we have set  $K = 1$ : this is tantamount to assuming that consumers are very impatient<sup>5</sup>. We obtain the following model.

$$\log(m + a) = \beta(X) \log(\gamma(X)) + (1 - \beta(X)) \log(\pi) - \beta(X) \log[p(\theta(X) + i, T)] + e_1, \quad (28)$$

$$\log(p) = \log[p(\theta(X) + i, T)] + e_2. \quad (29)$$

---

<sup>5</sup> An estimate of  $\delta$  has been obtained in a variant of the monopoly model (see the appendix).

The perturbation vector  $e = (e_1, e_2)$  is assumed normally distributed with mean zero and covariance matrix  $\Omega$ . Of course, in the above equation, it must be understood that  $p = p(r, T)$  (Recall that  $r$  and  $T$  are observed). The endogenous variables are simply  $\log(m+a)$ , the logarithm of the total house price (in euros), and  $\log(p)$ , the logarithm of the continuous-time constant reimbursement annuity. The function  $\log(\gamma(X))$  is simply taken to be a linear function of the constant,  $\log(w)$ , the age of the borrower (in logarithms), and family size (in logarithms). Parameter  $\beta$  varies with occupational status dummies only (white collar, intermediate professions and executives), blue-collar workers being the reference group. Let *Exec*, *Interm*, *Whitecol*, *Bluecol* respectively denote the executive, intermediate profession, white collar and blue collar dummies. Finally,  $\theta$  is specified as follows:

$$\begin{aligned} \log(\theta(X)) = & \theta_0 + \theta_1 \log(w) + \theta_2 \log(w) \cdot (\textit{Exec}) \\ & + \theta_3 \log(w) \cdot (\textit{Interm}) + \theta_4 \log(w) \cdot (\textit{Whitecol}) + \theta_5 \log\left(\frac{m+a}{a}\right), \end{aligned} \quad (30)$$

a specification in which the wage (in logs) interacts with occupational status. The model has been estimated by the maximum likelihood method. It is less simple than it seems, since  $(m+a)$  appears in  $\theta(X)$  and is endogenous; this gives rise to complicated Jacobian terms in the expression of likelihood. The price of houses  $\pi$  has been calibrated: we set its value at 428 euros per square meter (this corresponds roughly to \$45 per square foot), a reasonable figure for the French provinces.

### *Estimation results*

The results are presented in Table 2. All parameters are significant, except the white-collar dummy in  $\beta(X)$ , the intermediate-professions-wage interaction variable in  $\theta(X)$ , and family size in  $\gamma(X)$ . This means that the  $\beta$  parameter of white collars is not significantly different from that of blue collars, and that the wage's effect on the intermediate professions' risk premium is not significantly different from the same effect on the blue-collar's risk premium. All parameter estimates also have the expected sign. We will concentrate our comments on  $\beta$  and on the estimated risk premium function.

Remark first that the "price-elasticity of demand for housing"  $\beta$  varies significantly with occupational status: For executives,  $\beta = 2.1766$ ; for intermediate professions  $\beta = 2.1475$ ; for white collars  $\beta = 2.1231$ ; and for blue collars, we find  $\beta = 2.1166$ . That is, the lower in the professional hierarchy, the less elastic<sup>6</sup>. Second, the estimated  $\theta$  exhibits "social discrimination". Workers are clearly discriminated against, just because they are workers, not simply because their wage is low. A negative coefficient on the wage in the  $\theta$  function, *which varies with the occupational status*, means that the way the wage is taken into account to assess default risk depends on the status. According to the banker, the richest are the less risky. Yet, the extent of the interest rate reduction which is granted for a given increase in the wage is higher, the higher the status. At the same time, the inverse of the downpayment ratio has a positive coefficient, as expected. The number of variables

---

<sup>6</sup> The reader could have expected the reverse ranking: we provide an explanation below for this result.

introduced in the risk-premium function is too small to guarantee that the estimates reflect prejudice against the working class (or a favorable prejudice for executives). The observed discrimination could be a form of statistical discrimination, that is, the occupational status variables are likely to act as proxy variables for unobserved factors correlated with default risk (see the discussion in section 5 below).

#### *The presence of market power*

The mean value of  $\theta$  in the blue-collar category is  $\theta(\text{Bluecol}) = 0.0398$ ; it is  $\theta(\text{Whitecol}) = 0.0425$  for white-collar;  $\theta(\text{Interm}) = 0.0367$  for intermediate professions, and  $\theta(\text{Exec}) = 0.0330$  for executives, the differences being significant, except for the intermediates. But the order of magnitude of these risk-premia is too high. Since the probability of default before term is  $F(T) = 1 - e^{-\theta T}$ , we can estimate  $F(15)$  and the figures are 0.464, 0.469, 0.380 and 0.350 for the blue collars, white collars, intermediates and executives, respectively. These probabilities of default do not correspond to what is known *a priori* on mortgage default rates in France. A reasonable figure would be something like 3%, not 35% or 46%! The chief executives of the CHF claim that their network does a very good screening job and that  $F(15) \approx 0.01$ . This is a sufficient indication for the presence of market power: the values of  $\theta$  should be interpreted as markups, not as risk-premia. The fact that these markups are inversely related with the corresponding price-elasticities of demand for housing, the  $\beta$ s, is an indication that price discrimination is taking place. Workers are less elastic than executives, they therefore pay more in (monopolistic) equilibrium. This is why we estimate a monopolistic version of the model in the next subsection.

Figures 2 to 4 represent numerical simulations of the model with the estimated parameters. Figures 2 and 3 depict the interest rate charged on borrowers as a function of the wage and of the downpayment ratio, respectively. To draw these figures, the values of exogenous variables are set equal to their overall sample mean, with the exception of the variable appearing on the x-axis. The schedules show striking differences of treatment between social categories, everything else being equal. The mere fact of being an executive leads to a reduction of the interest rate of approximately half a percentage point, relative to white collars. Figure 4 shows the house size schedule as a function of the wage. The same type of social discrimination appears clearly.

#### *4.2. The Monopoly Model*

To establish the equations of the monopoly model, it is sufficient to replace the zero-profit condition (17) with the zero-surplus condition (21). In order to obtain the analytical expression of (21), we first compute  $V^r$ , the indirect utility of a household on the rental market. The rental demand for housing is obtained by maximizing  $u(c, h)$ , as specified by (22) above, with respect to  $(c, h)$ , under the budget constraint  $c + \rho h = w$ , where  $\rho$  is the rent per square meter. This easily yields the instantaneous indirect utility of renting,

$$v^r(\rho, w) = u_0 + w + \frac{1}{\beta - 1} \gamma^\beta \rho^{1-\beta}, \quad (31)$$

from which we derive the expected utility of renting forever, defined by (10) above:

$$V^r(\rho, w) = v^r(\rho, w) \frac{\delta}{\theta + \delta} + u_0 \frac{\theta}{\delta + \theta}. \quad (32)$$

Now, we equate  $\bar{V}$ , given by (23), with  $V^r$ , to obtain the analytical form of (21), and simple computations yield the following expression.

$$-pm + \frac{\beta\gamma}{\beta - 1} K(\theta, \delta, T) h^{(\beta-1)/\beta} = \frac{\gamma^\beta \rho^{1-\beta}}{(\beta - 1)(1 - e^{-(\theta+\delta)T})}. \quad (33)$$

Note that  $u_0$  has disappeared from (33). The model to be estimated is (27) and (33). As before, we set  $\delta = +\infty$ , and thus  $K = 1$  (see the appendix, for an estimation of  $\delta$ ), and get the following bivariate model.

$$\log(m + a) = \beta(X)\log(\gamma(X)) + (1 - \beta(X))\log(\pi) - \beta(X)\log[p(\theta(X) + i, T)] + e_1, \quad (34)$$

$$pm = \frac{\beta(X)\gamma(X)}{\beta(X) - 1} \left( \frac{m + a}{\pi} \right)^{\frac{\beta(X)-1}{\beta(X)}} - \frac{\gamma(X)^{\beta(X)} \rho^{1-\beta(X)}}{\beta(X) - 1} + e_2, \quad (35)$$

where  $e_1$  and  $e_2$  are zero-mean random error terms, with a normal distribution,  $y_1 = \log(m + a)$  and  $y_2 = pm$  are the endogenous variables, and  $X$  is the vector of exogenous variables. Of course,  $p = p(r, T)$ . The specifications of  $\gamma(X)$  and  $\theta(X)$  are exactly the same as before ( $\theta$  being defined by (30);  $\gamma$  being a loglinear function of the constant,  $w$ , age and family size). For this estimation, we set  $\beta(X) = \exp[b_0 + b_1(Exec) + b_2(Interm) + b_3(Whitcol)]$ . The price of a square meter and the yearly rent per square meter are calibrated at reasonable values, respectively  $\pi = 428$  and  $\rho = 46$ , in euros.

The model is again estimated by standard maximum likelihood. Estimation results are summarized by Table 3. All coefficients are significant, except the coefficients of the intermediate professions dummy in  $\beta$ , of family size and age in  $\gamma$ , and of the (white-collar)(wage) interaction variable in  $\theta$ . It follows that "social discrimination" is present, since wage-status interaction terms are significant for executives and intermediate professions in the risk-premium function  $\theta$ .

Again, the values of  $\beta$  can be ranked according to occupational status, estimated values being  $\beta = 1.6117$  for blue collars,  $\beta = 1.6002$  for white collars, and  $\beta = 1.6308$  for executives. The estimated mean values of  $\theta$  for each occupational status reveal the expected ranking of risk premia, from the lowest, the executives, to the highest, the white and blue collars (see bottom of Table 3). These estimates of  $\theta$  correspond to much more reasonable probabilities of default than the competitive estimates obtained above. Computing  $F(15) = 1 - e^{-15\bar{\theta}}$  with the estimated average values of the risk premium,  $\bar{\theta}$ , in each category, yields  $F = .169$  for blue collars,  $F = .119$  for white collars,  $F = .055$  for intermediates and  $F = .018$  for executives: the expected ranking.

We conclude that housing demand elasticities (*i.e.*, the  $\beta$ s) are overestimated by the competitive model, because risk-premia are also overestimated. Since equation (29) (*i.e.*, the competitive pricing equation) is likely to be an incorrect representation, upward biased

risk-premia are transformed into upward biased housing demand elasticities while estimating equation (28) (*i.e.*, the house size equation) simultaneously with (29). We probably get better estimates of structural parameters with the monopoly version of the model, in spite of the added complexity of the zero-surplus equation (35).

We conclude that workers seem to be discriminated against by the bank, the origin of the discrimination being mostly due to differing elasticities of demand, and secondarily to differences in perceived default risks. These differences in elasticities  $\beta$  could simply capture the fact that blue-collar workers are more likely to see their loan applications rejected by other commercial banks, since the CHF is specialized in "social loans." Nevertheless, the results show that consumer heterogeneity is exploited by the banker to make more profit, as illustrated by the simulations. Figures 5 to 8 represent numerical simulations of the model with estimated values of the parameters. Figures 5 to 7 illustrate the same kind of phenomena as Figures 2 to 4, but under the assumption of monopolistic behavior. Figure 8 represents the estimated reservation utility or participation levels of the social categories, as defined by (21) above, in the form of indifference curves. It is easy to see that executives consume more than the workers for any given size of the house. To compute each of these indifference curves, we used the sample mean values of the exogenous variables, except for the wage, the mean of which is evaluated in each occupational status sub-sample. In the appendix, we propose a variant of the monopoly model, in which the impatience parameter  $\delta$ , and function  $K$ , are estimated.

## 5. Further remarks on racial discrimination in credit markets

Discrimination in credit markets has recently attracted considerable attention, and the question of deciding whether or not — and why — lenders discriminate against minority groups is a hotly debated topic among economists. The importance of the question is amplified by the fact that racial discrimination in mortgage lending has been made illegal by the Equal Credit Opportunity Act of 1974, in the United States, and by the availability of new data sources, allowing for new econometric tests. The recent literature on this question is mostly empirical, and concentrated on racial or sexual discrimination problems. Empirical studies of discrimination in mortgage lending have developed with the debate triggered by the contributions of Shafer and Ladd (1981) and Munnell *et al.* (1996). Several other important contributions to this literature are commented in Ladd's (1998) survey article.

A difficult problem in most empirical studies is to detect the presence of discrimination in the sense of G. Becker (1957). More precisely, a lender, or seller, is said to discriminate in the sense of Becker, if she is ready to forego profits just because of her prejudices. This form of discrimination is based on a particular "taste for discrimination" of the sellers, and is not usually considered by standard I.O. theories of price discrimination.

In contrast, a lender might treat a minority group differently, because racial or ethnic characteristics are correlated with some variables, important in the determination of creditworthiness and default risk, and which remain unobserved. This latter form of behavior is called *statistical discrimination* in the sense of Arrow (1973) and Phelps (1972). An important difficulty stems from the fact that econometricians can never be sure of having introduced enough explanatory variables to control for possible risk differences in

their estimation of default probabilities. It follows that a significant coefficient on race in regressions could only indicate that statistical discrimination is taking place.

Another difficulty is the need to separate the effect of risk from that of market power in the formation of interest rates. On the role of concentration in the explanation of interest rates, see Cavaluzzo and Cavaluzzo (1998). On the empirical relationship between concentration and discrimination, see Berkovec *et al.* (1998). The risk premium charged on some borrowers can also be interpreted as a standard monopolistic markup. To be more precise, it might be that the borrowers' preference parameters are correlated with their social, racial or ethnic group, because individual preferences depend on a group's particular economic conditions. Then, if market power is present, standard price discrimination can in turn become the explanation for differential treatment, without necessarily reflecting the presence of prejudice. This type of approach, however, is not without its dangers, which would be to attribute the bulk of observed differences in treatment to taste differences correlated with race. However, competition should tend to eliminate discrimination in the sense of Becker, since prejudiced lenders would lose business in favor of the unprejudiced. It follows that market power and discrimination in this sense must be closely interrelated.

In the results presented above, it is very likely that the preference differences between occupational status groups, *i.e.*, the differences in parameter  $\beta$ , reflect the differing bargaining strengths of the borrowers, or the differing probabilities of obtaining a loan elsewhere, blue collars being more or less a captive clientele of the CHF.

Finally, the most difficult problem in detecting the presence of prejudice is that discrimination in the market for mortgages might reflect the existence of discrimination in other markets, such as the labour and housing markets, and thus be purely statistical in nature. Some minority groups would pay higher mortgage rates because they have higher probabilities of losing their jobs, and this, in turn, could simply be a consequence of their employer's behavior.

Much of the published work on discrimination in the mortgage market, to the best of our knowledge, has been devoted to the study of default and of credit denial rates (again see H. Ladd (1998)). In contrast, the model presented above aims at explaining the structure of *accepted* loan applications and determines loan sizes and interest rates simultaneously. This model could be used to test for the presence of discrimination in the sense of Becker, paying attention to the role of local market conditions, of preference heterogeneity and of differing default risks. In theory, the model allows one to separate, in the interest rate and in loan size differences, what can be attributed to prejudice, from the impact of differences in preferences and default risks. Variations in  $\beta$ ,  $\gamma$  and  $\delta$  can reflect differing preferences while  $\theta$  reflects default risk differences as perceived by the banker. An observable characteristic can significantly change  $(\beta, \gamma, \delta)$ , and thus lead to changes in treatment by the banker.

If  $X$  contains enough information (enough control variables) to estimate a risk premium reasonably, the function  $\theta(X)$  should not significantly depend on race. If it indeed does depend on race significantly, then, discrimination in the sense of Becker is taking place. Of course, if the information contained in  $X$  is not sufficient to control for differences in riskiness, then, statistical discrimination in the sense of Arrow-Phelps can be the explanation for a significant coefficient on race in function  $\theta$ .

The model presented above allows for separation of these effects from plain discrimination effects based on observable differences in preferences. There is a danger here, however, which would be to attribute the bulk of differences in treatment to differences in preferences: minority consumers would have relatively smaller or bad quality houses just because they happen not to like nice housing! To avoid ambiguities of this type, it seems to us that race should not be introduced as a variable in the specifications of  $\beta$ ,  $\gamma$  and  $\delta$ .

## 6. Conclusion

To conclude, in the present contribution, we proposed a model of the mortgage lending market. The model can be used to test for the presence of discrimination, using information on accepted loan applications only. It rests on the idea of discrimination by the lender, based on observable attributes of the borrower. It explains the interest rate and the loan size of accepted loan applications simultaneously. We study a competitive equilibrium variant and a discriminating monopolist variant of the model, in order to take phenomena related to market power into account.

The model has been estimated with a sample of loan files originating from branches of a French mortgage lender. We reject the competitive model because estimated interest rate markups are too high to reflect default risks only. The monopolistic model gives a better account of the discrimination phenomena at work in the data. We conclude that "social discrimination" is present, in the sense that *ceteris paribus*, a member of the working class would pay higher interest rates than an (equally rich) executive, just because he or she is identified as a blue collar. Part of the differences in interest rates must be attributed to observable differences in preferences, since blue-collar workers have a significantly smaller price-elasticity of demand for housing than executives. The model also shows how borrower characteristics and interest rates affect the size of granted loans.

## References

- Anderson, S. de Palma, A. and J. Thisse (1992), *Discrete Choice Theory of Product Differentiation*, MIT Press, Cambridge.
- Armstrong, M. (1999), "Price Discrimination by a Many-Product Firm", *Review of Economic Studies*, 66, 151-168.
- Arrow, J. K. (1973), "The Theory of Discrimination", in A.O. Ashenfelter and A. Ries, eds. *Discrimination in Labour Markets*, Princeton University Press, Princeton, New Jersey.
- Becker, G. (1957), *The Economics of Discrimination*, University of Chicago Press, Chicago, Illinois.
- Berkovec, J. A., Canner, G. B., Gabriel, S. A. and T. H. Hannan (1998), "Discrimination, Competition, and Loan Performance in FHA Mortgage Lending", *Review of Economics and Statistics*, 80, 241-250.
- Bester, H. (1985), "Screening vs. Rationing in Credit Markets with Imperfect Information," *American Economic Review*, 75, 850-855.
- Bousquet, A. and M. Ivaldi (1997), "Optimal Pricing of Telephone Usage: an Econometric Implementation", *Information Economics and Policy*, 9, 219-239.
- Brueckner, J. K. (1994), "Borrower Mobility, Adverse Selection, and Mortgage Points", *Journal of Financial Intermediation*, 3, 416-441.
- Brueckner, J. K. (2000), "Mortgage Default with Asymmetric Information", *Journal of Real Estate Finance and Economics*, 20, 251-274.
- Calem, P. and M. Stutzer (1995), "The Simple Analytics of Observed Discrimination in Credit Markets", *Journal of Financial Intermediation*, 4, 189-212.
- Cavaluzzo, K.S. and L.C. Cavaluzzo (1998), "Market Structure and Discrimination: The Case of Small Businesses", *Journal of Money, Credit and Banking*, 30, 771-792.
- Chari, V. V., and R. Jagannathan (1989), "Adverse Selection in a Model of Real Estate Lending", *Journal of Finance*, 36, 471-483.
- Chiappori, P. A. and J. Pinquet, (1999), "L'assurance-chômage des emprunteurs", *Revue française d'économie*, 14, 91-115.
- Clerydes, S. K., (2001), "Product Selection with Multi-Peaked Preferences in Book Publishing", University of Cyprus, Nicosia, Cyprus, mimeo.
- Cohen, A. (2000), "Package Size and Price Discrimination in Paper Towels", University of Virginia, mimeo.
- Deaton, A. (1992), *Understanding Consumption*, Clarendon Lectures in Economics, Oxford University Press, Oxford.
- Goldberg, P.K., and F. Verboven (2001), "The Evolution of Price Dispersion in the European Car Market", *Review of Economic Studies*, 68, 811-848.
- Ivaldi, M. and D. Martimort (1994), "Competition under Nonlinear Pricing", *Annales d'Economie et de Statistique*, 34, 72-114.
- Ladd H. F. (1998), "Evidence on Discrimination in Mortgage Lending", *Journal of Economic Perspectives*, 12, 41-62.

- Leslie, P.J.(2001), "Price Discrimination in Broadway Theatre", UCLA, mimeo.
- McManus, B. (2001), "Nonlinear Pricing in an Oligopoly Market: The Case of Specialty Coffee", Washington University-St Louis, Olin School of Business, mimeo.
- Milde H. and J. Riley (1988), "Signaling in Credit Markets", *Quarterly Journal of Economics*, 103, 101-130.
- Miravete, E. (2001), "Estimating Demand for Local Telephone Service with Asymmetric Information and Optional Calling Plans", University of Pennsylvania, mimeo.
- Munnell, A.H., G.M.B. Tootell, L.E. Browne and J. McEneaney (1996), "Mortgage Lending in Boston: Interpreting HMDA Data", *American Economic Review*, 86, 25-53.
- Phelps, E. S. (1972), "The Statistical Theory of Racism and Sexism", *American Economic Review*, 62, 659-661.
- Rochet, J-C. and P. Choné (1998), "Ironing, Sweeping, and Multidimensional Screening", *Econometrica*, 66, 783-826.
- Rothschild M. and J. Stiglitz (1976), "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information", *Quarterly Journal of Economics*, 90, 630-649.
- Shafer, R. and H.F. Ladd (1981), *Discrimination in Mortgage Lending*, MIT Press, Cambridge, Mass.
- Stanton, R. and N. Wallace (1997), "Mortgage Choice: What's the Point?", *Real Estate Economics*, 26, 173-205.
- Verboven, F. (2002), "Quality-based Price Discrimination and Tax Incidence: Evidence from Gasoline and Diesel Cars", Catholic University of Leuven, Belgium, mimeo.

## Appendix. A Variant of the Monopoly Model

We also estimated a variant of the monopoly model, with a more parsimonious specification of utility, that is,

$$u(c, h) = u_0 + c + \gamma \log(h), \quad (36)$$

but we tried to estimate the impatience parameter  $\delta$ . With this specification, we find that the instantaneous reservation utility of borrowers writes,

$$v^r(\rho, w) = u_0 + w - \gamma + \gamma \log\left(\frac{\gamma}{\rho}\right). \quad (37)$$

The necessary condition for optimality (20) now writes,

$$h = \frac{\gamma K(\theta, \delta, T)}{\pi p(\theta + i, T)}, \quad (38)$$

where  $K$  is still defined by (25). With the above specification of  $u$ , using (9) with  $T^* = 0$ , one gets,

$$\begin{aligned} \bar{V}(c, h, T; \alpha) = \frac{\delta c}{\delta + \theta}(1 - e^{-(\delta + \theta)T}) + \frac{\gamma \log(h)}{\delta + \theta} [\delta + \theta e^{-(\delta + \theta)T}] \\ + \frac{w\delta}{\delta + \theta} e^{-(\delta + \theta)T} + u_0. \end{aligned} \quad (39)$$

Then, equating (39) with  $V^r(\rho, w; \alpha)$ , using (37), yields the specific form of the zero-surplus condition (21), that is, after some computations,

$$(c - w)(1 - e^{-(\delta + \theta)T}) + \gamma \log(h)(1 + (\theta/\delta)e^{-(\delta + \theta)T}) = \gamma[\log(\gamma/\rho) - 1]. \quad (40)$$

The system to be estimated is (38) and (40), with  $c - w = -pm$  and  $h = (m + a)/\pi$ . Taking logs on both sides of (38), we find the system,

$$\log(m + a) = \log(\gamma(X)) + \log[K(\theta, \delta, T)] - \log[p(\theta + i, T)] + e_1, \quad (41)$$

$$pm = -\gamma(X) \frac{[1 + \log(\gamma(X)/\rho)]}{[1 - e^{-(\delta + \theta)T}]} + \gamma(X) \log\left(\frac{m + a}{\pi}\right) K(\theta, \delta, T) + e_2, \quad (42)$$

where  $e_1$  and  $e_2$  are zero-mean random error terms, with a normal distribution,  $y_1 = \log(m + a)$  and  $y_2 = pm$  are the endogenous variables, and  $X$  are exogenous variables.

We assume that for each individual  $j$ ,  $\theta$  can be expressed as  $\theta_j = -(1/T_j)\log(1 - F)$  where  $F$  is a common probability of default. The parameters to estimate are  $D = \log(\delta)$ ,  $F$ ,  $\rho$  and the parameters of function  $\gamma$ , which is specified as follows,

$$\begin{aligned} \log(\gamma(X)) = \gamma_0 + \gamma_1 \log(w) + \gamma_2 \log(\text{age}) + \gamma_3 \log(T) \\ + \gamma_4(\text{Exec}) + \gamma_5(\text{Interm}) + \gamma_6(\text{Whitecol}). \end{aligned} \quad (43)$$

This model was difficult to estimate. We estimated  $\delta = e^D$  and the  $\gamma$ s by standard maximum likelihood for fixed  $\rho$  and  $F$ , and then estimated  $\rho$  and  $F$  by a grid-search procedure. The results are presented in Table 4. All the  $\gamma$  parameters are significantly different from zero, except the coefficient on the white-collar dummy, meaning that white collars can be merged with the blue collars. The estimated value of  $\delta$  is 0.158, a reasonable figure. Given that we assume liquidity constrained borrowers, they must be impatient enough to be willing to borrow, so that it is reassuring to find  $\delta$  above the maximal interest rates in the sample. The estimated rent  $\rho$ , which is FFrancs 294 per square meter and per year is also reasonable, given that our observations are outside the very expensive Paris area (this makes a rent of \$4.41 per square foot and per year, or an approximate monthly rent of \$202 for a two-rooms flat of 550 square feet). Finally, the estimated value of  $F$  is 0.011, very close to the expected 1% value. This, again, corresponds to a much more realistic average estimated value,  $\bar{\theta} = 0.00073$ , than those obtained with the competitive model. In fact  $\theta$  seems to be of the order of magnitude of 1 to 10 base points.

**Table 1 . Descriptive Statistics**

	Overall Sample	Executives	Intermediate Professions	White Collars	Blue Collars
Number of Observations	2610	159	363	908	1180
Loan Size in euros					
- average	33540	44055	35188	35599	31569
- stand. dev.	18284	35705	23550	16779	12745
Downpayment Ratio					
- average	0.172	0.297	0.247	0.162	0.139
- stand. dev.	0.182	0.271	0.232	0.174	0.137
Loan Term in years					
- average	14.89	13.08	13.04	14.91	15.68
- stand. dev.	4.08	4.05	3.88	4.02	3.94
Loan Interest Rate					
- average	11.658	10.316	10.847	11.791	11.986
- stand. dev.	1.739	2.076	1.951	1.608	1.545
Cost of Funds					
- average	8.330	7.888	8.028	8.257	8.538
- stand. dev.	1.238	1.038	1.033	1.244	1.275
Annual Wage in euros					
- average	20011	30671	22900	21018	16911
- stand. dev.	10936	20743	12976	10760	6160

**Table 2 . Competitive Model**

Parameter	Estimation	T-Ratio	P-Value
<b>Beta</b>			
- Constant	2.1166	29.908	0.0000
- Executive	0.0600	3.146	0.0008
- Intermediate	0.0309	2.223	0.0131
- White Collar	0.0065	0.645	0.2594
<b>Gamma</b>			
- Constant	5.2850	39.677	0.0000
- Wage	0.2061	15.588	0.0000
- Age	0.1017	5.662	0.0000
- Household Size	0.0067	0.848	0.1981
<b>Theta</b>			
- Constant	-2.8077	-9.330	0.0000
- Wage	-0.0749	-2.858	0.0021
- Wage*Executive	-0.0049	-1.065	0.1434
- Wage*Intermediate	0.0011	0.386	0.3498
- Wage*White Collar	0.0072	3.980	0.0000
- Inverse of Downpayment Ratio	0.1945	19.081	0.0000

Mean of Theta for	Beta for
- Executive ..... 0.0330	- Executive ..... 2.1766
- Intermediate ..... 0.0367	- Intermediate ..... 2.1475
- White Collar ..... 0.0425	- White Collar ..... 2.1231
- Blue Collar ..... 0.0398	- Blue Collar ..... 2.1166

Mean Log-Likelihood ..... L = 0.305248	Number of observations ..... N=2610
Estimated Covariance ... $\Omega = \begin{pmatrix} 0.2083 & -0.0087 \\ -0.0087 & 0.0110 \end{pmatrix}$	Empirical variance of
matrix of errors	- Log(m+a) ..... Var=0.2958
Correlation Coefficient ..... = -0.1814	- Log(p) ..... Var=0.0286

**Table 3 . Monopoly Model**

Parameter	Estimation	T-Ratio	P-Value
<b>Beta</b>			
- Constant	-0.4915	-7.826	0.0000
- Executive	0.0191	1.577	0.0574
- Intermediate	-0.0069	-0.710	0.2388
- White Collar	-0.0115	-1.363	0.0865
<b>Gamma</b>			
- Constant	5.1841	31.915	0.0000
- Wage	0.2896	19.273	0.0000
- Age	0.0163	0.666	0.2527
- Household Size	0.0007	0.068	0.4729
<b>Theta</b>			
- Constant	3.3770	2.050	0.0202
- Wage	-0.7824	-5.430	0.0000
- Wage*Executive	-0.1097	-2.747	0.0030
- Wage*Intermediate	-0.0436	-2.415	0.0079
- Wage*White Collar	-0.0059	-0.600	0.2741
- Inverse of Downpayment Ratio	0.7096	15.905	0.0000

Mean of Theta for	Beta for
- Executive ..... 0.00126	- Executive ..... 1.6308
- Intermediate ..... 0.00382	- Intermediate ..... 1.6048
- White Collar ..... 0.00852	- White Collar ..... 1.6002
- Blue Collar ..... 0.01236	- Blue Collar ..... 1.6117

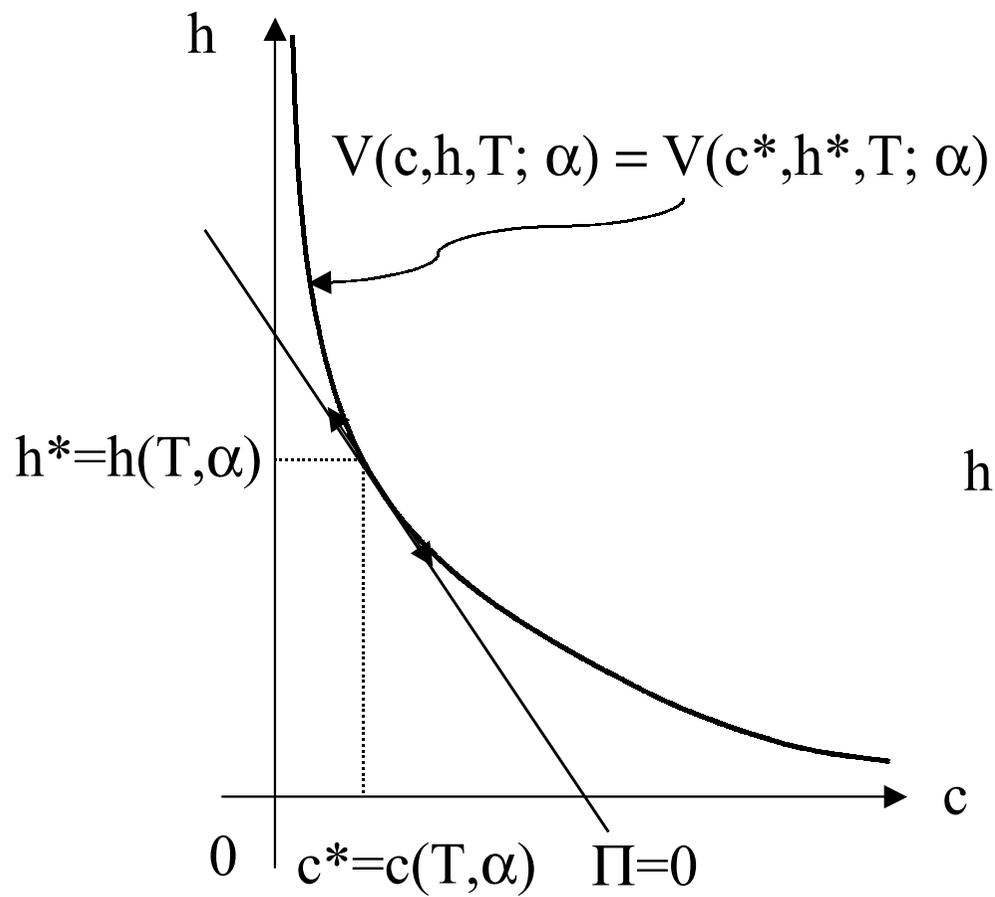
Mean Log-Likelihood ..... L = -1.90688	Number of observations ..... N=2610
Estimated Covariance ... $\Omega = \begin{pmatrix} 0.2125 & -0.1477 \\ -0.1477 & 1.0899 \end{pmatrix}$	Empirical variance of
matrix of errors	- Log(m+a) ..... Var=0.2958
Correlation Coefficient ..... = -0.3070	- pm ..... Var=2.5141

**Table 4 . Variant of Monopoly Model**

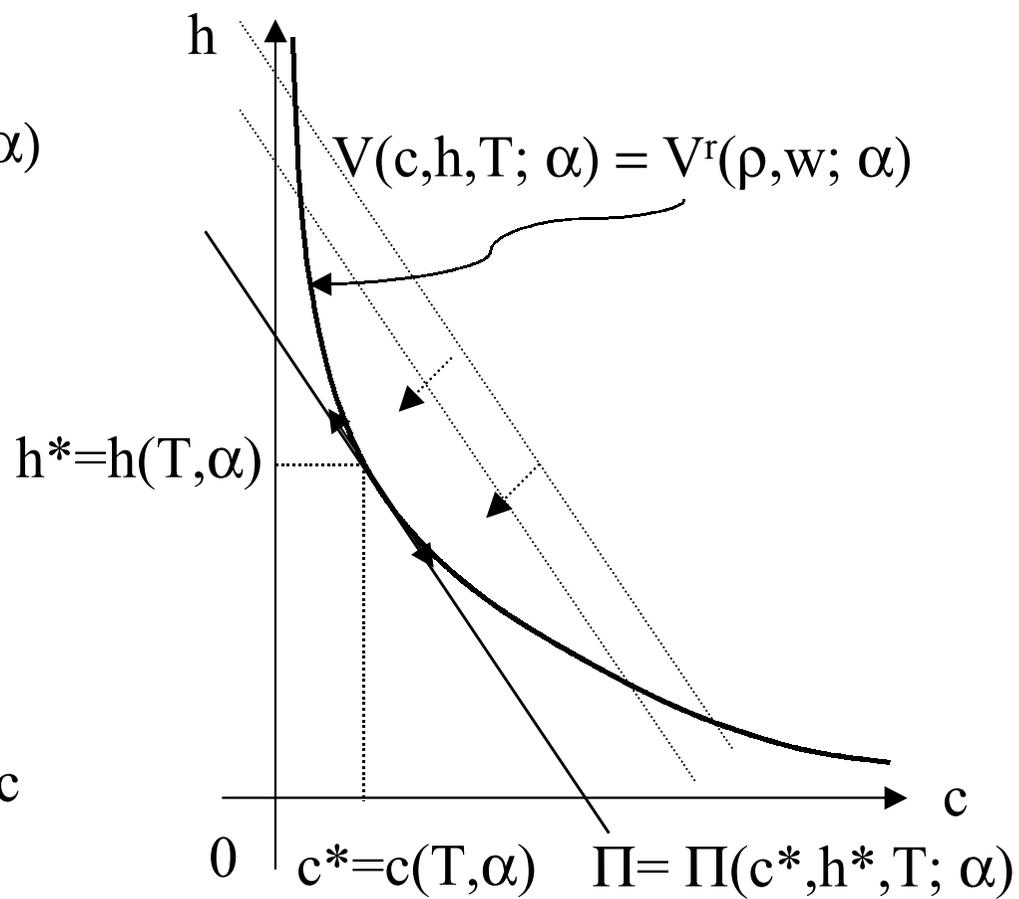
Parameter	Estimation	T-Ratio	P-Value
Constant	1.5411	7.923	0.0000
Wage	0.5662	40.083	0.0000
Age	0.0826	2.850	0.0022
Loan Term	0.6231	28.647	0.0000
Executive	0.0874	3.473	0.0003
Intermediate	0.0260	1.438	0.0752
White Collar	-0.0072	-0.604	0.2728
Ln(Delta)	-1.8443	-86.394	0.0000

Prob. of Default = 0.011 Mean of Theta = 0.00073 Delta = 0.1581	Rent = $\rho$ = 294.4 FF (44.95 euros) Price of square meter = $\pi$ = 2800 FF (427.48 euros)
---	---

Mean Log-Likelihood..... L = -1.89102  Estimated Covariance ... $\Omega = \begin{pmatrix} 0.2080 & -0.2276 \\ -0.2276 & 0.9725 \end{pmatrix}$ matrix of errors  Correlation Coefficient. = -0.5059	Number of observations ..... N=2610  Empirical variance of - Log(m+a) ..... Var=0.2958 - pm ..... Var=2.5141
---	--



1(a)



1(b)

Figure 1

Figure 1. Competitive Model: Interest Rate on Loans

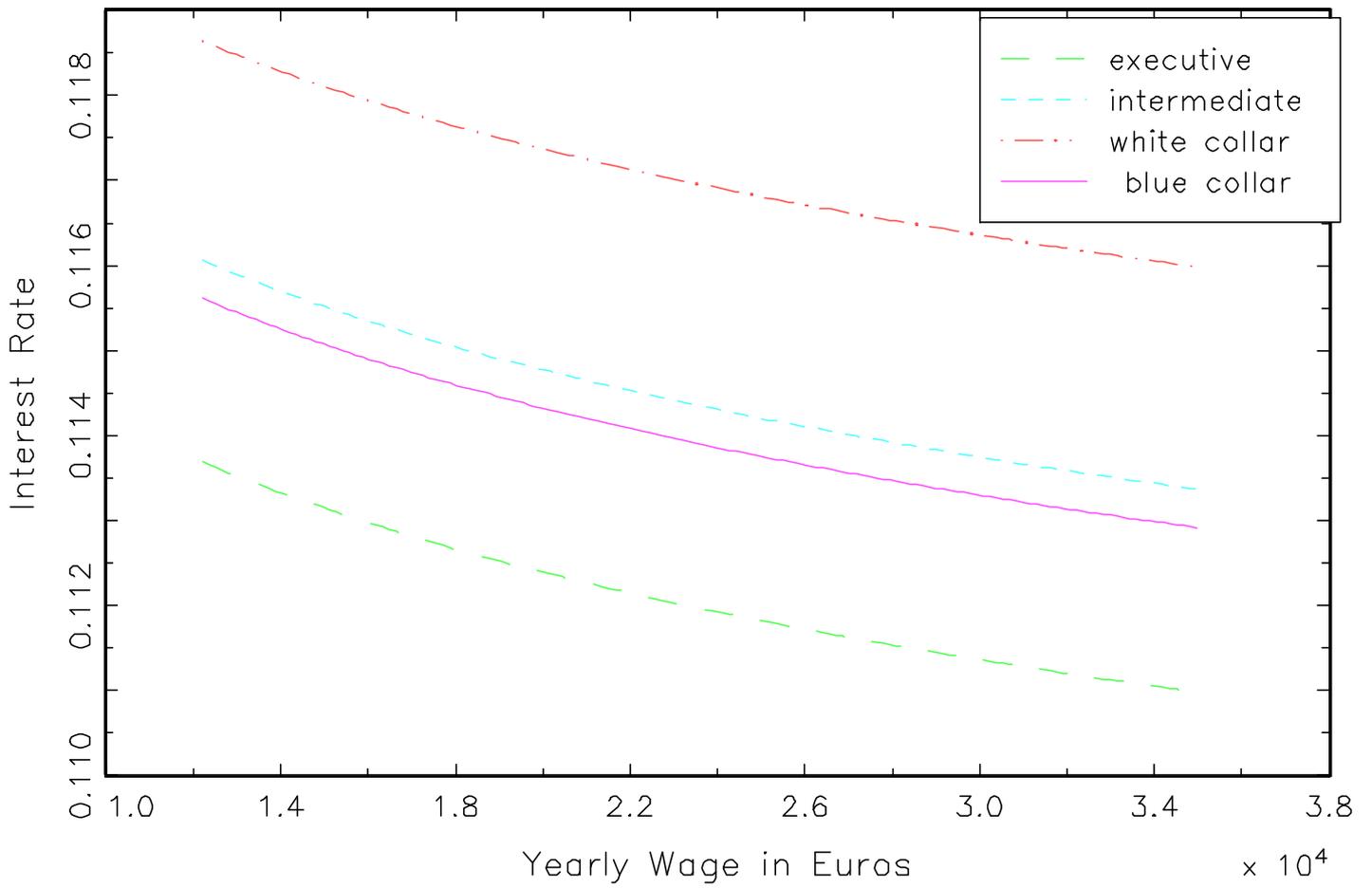


Figure 2. Competitive Model: Interest Rate on Loans

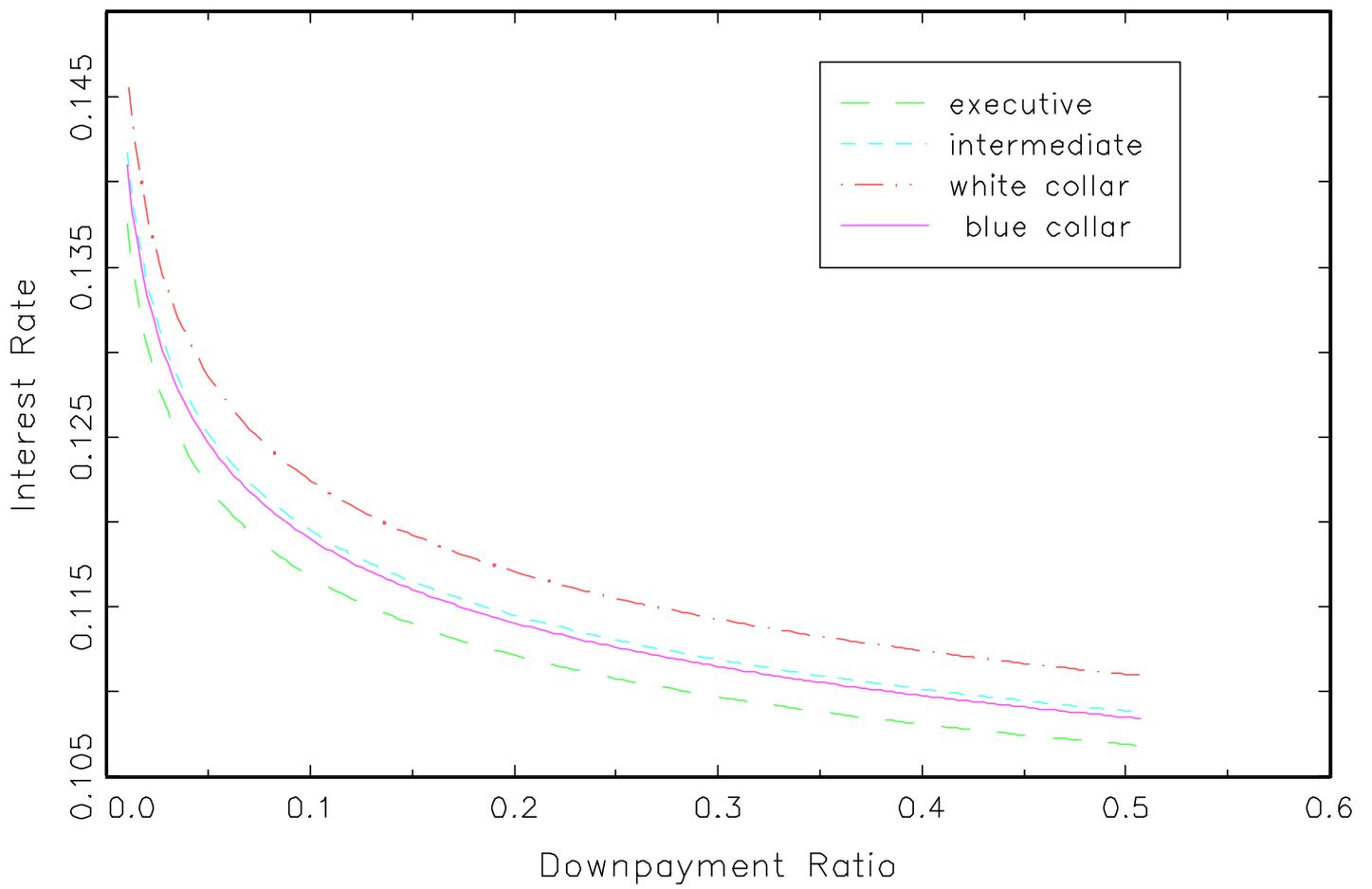


Figure 3. Competitive Model: House Size

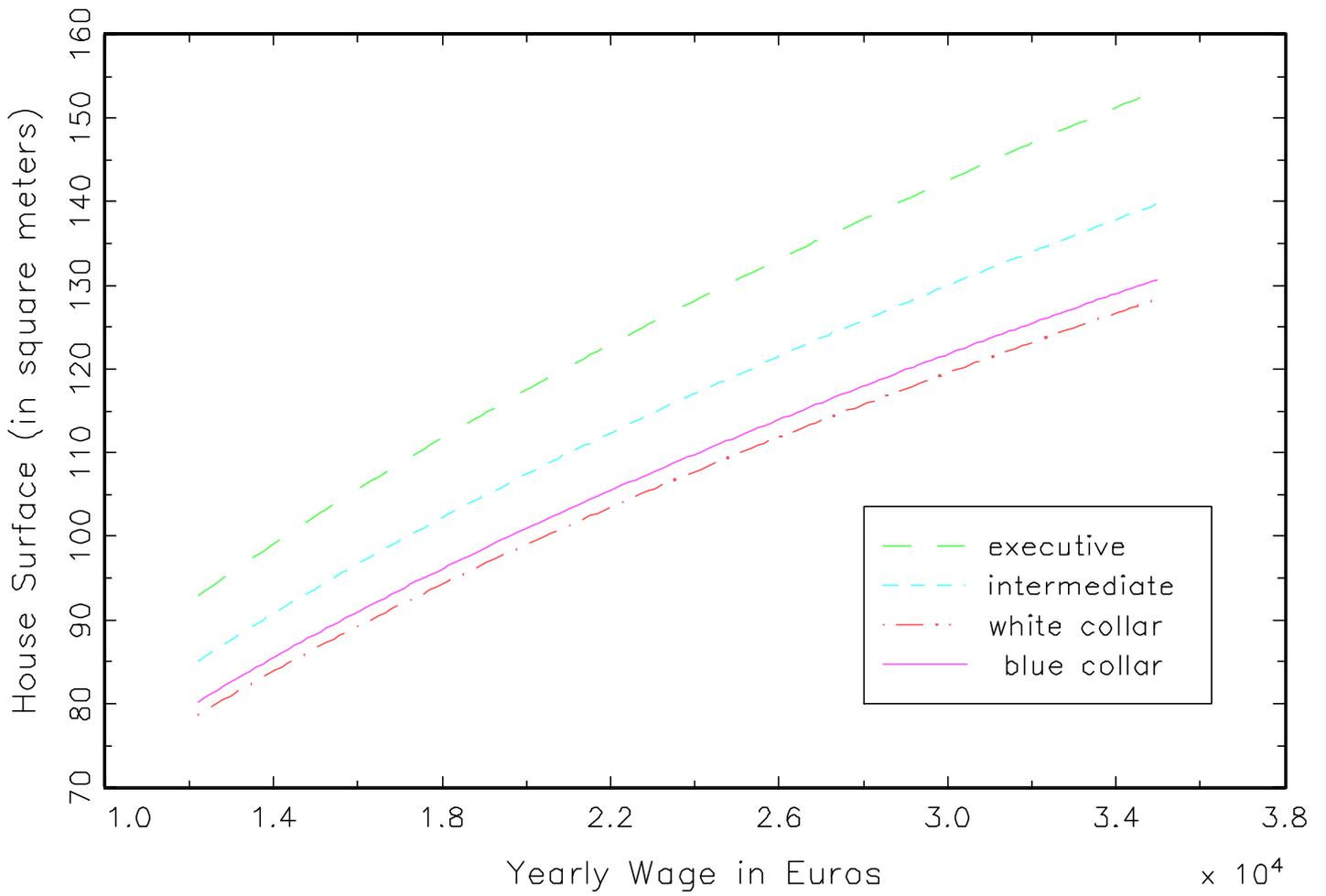


Figure 6. Monopoly Model: Interest Rate on Loans

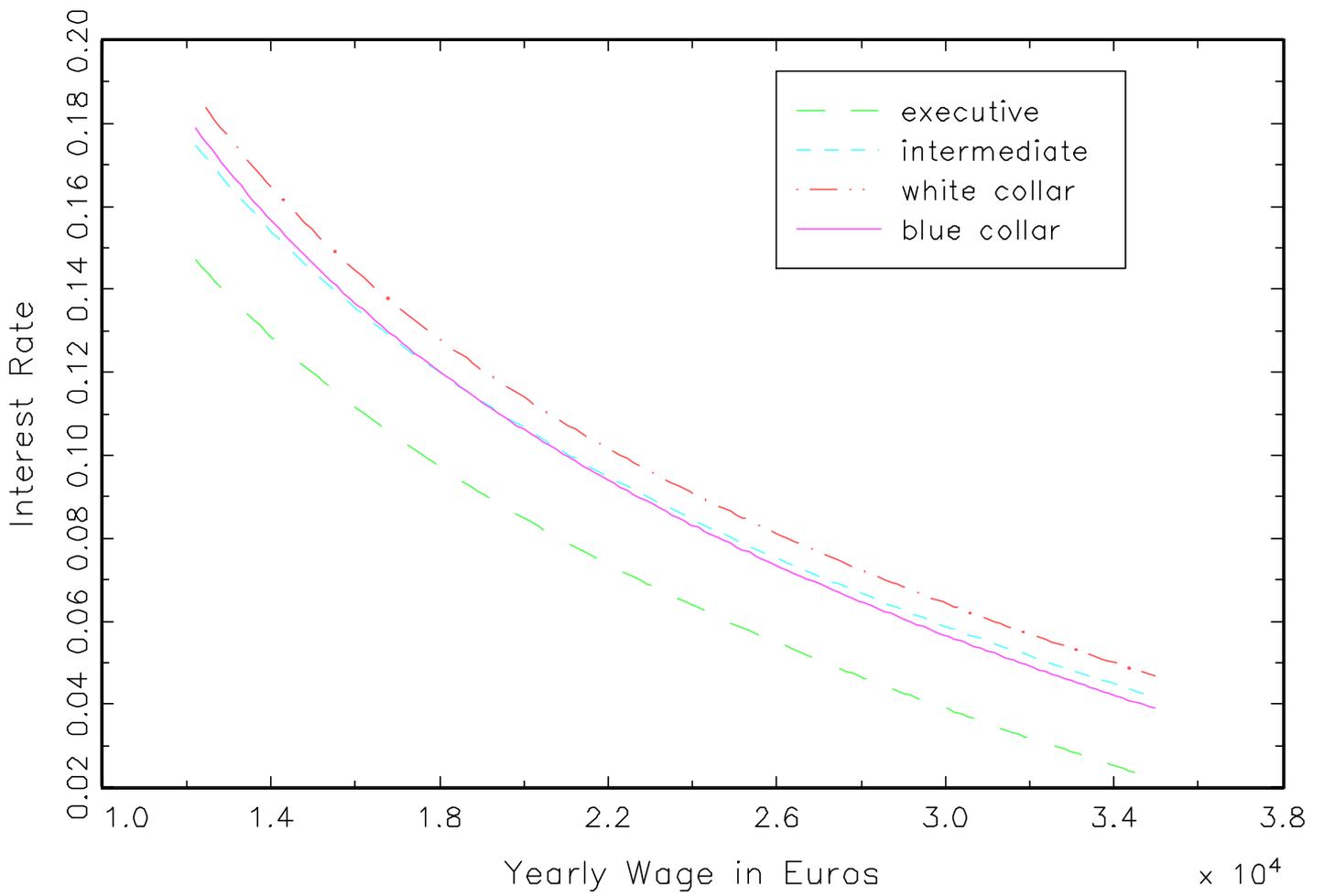


Figure 8. Monopoly Model: Interest Rate on Loans

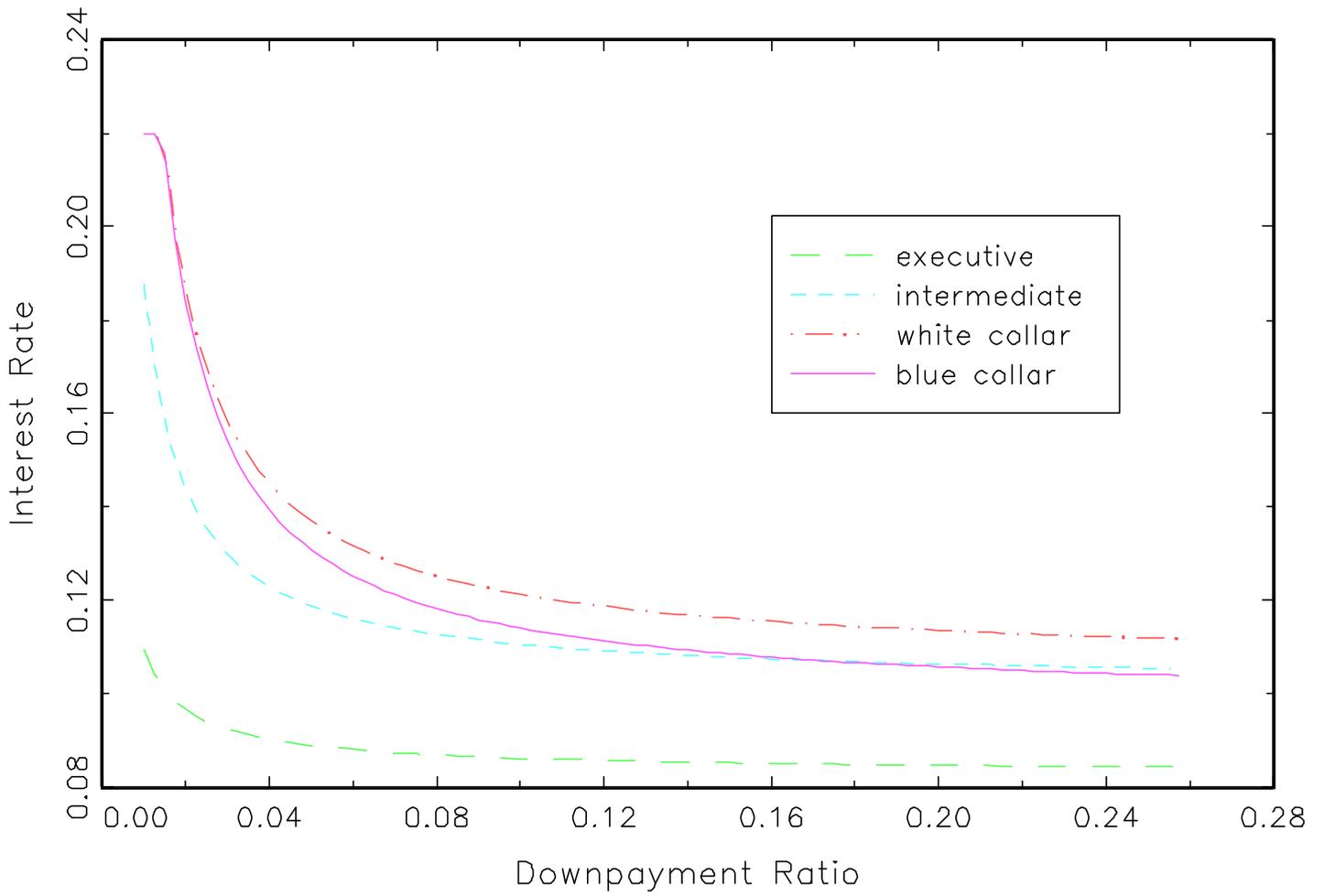


Figure 9. Monopoly Model: House Size

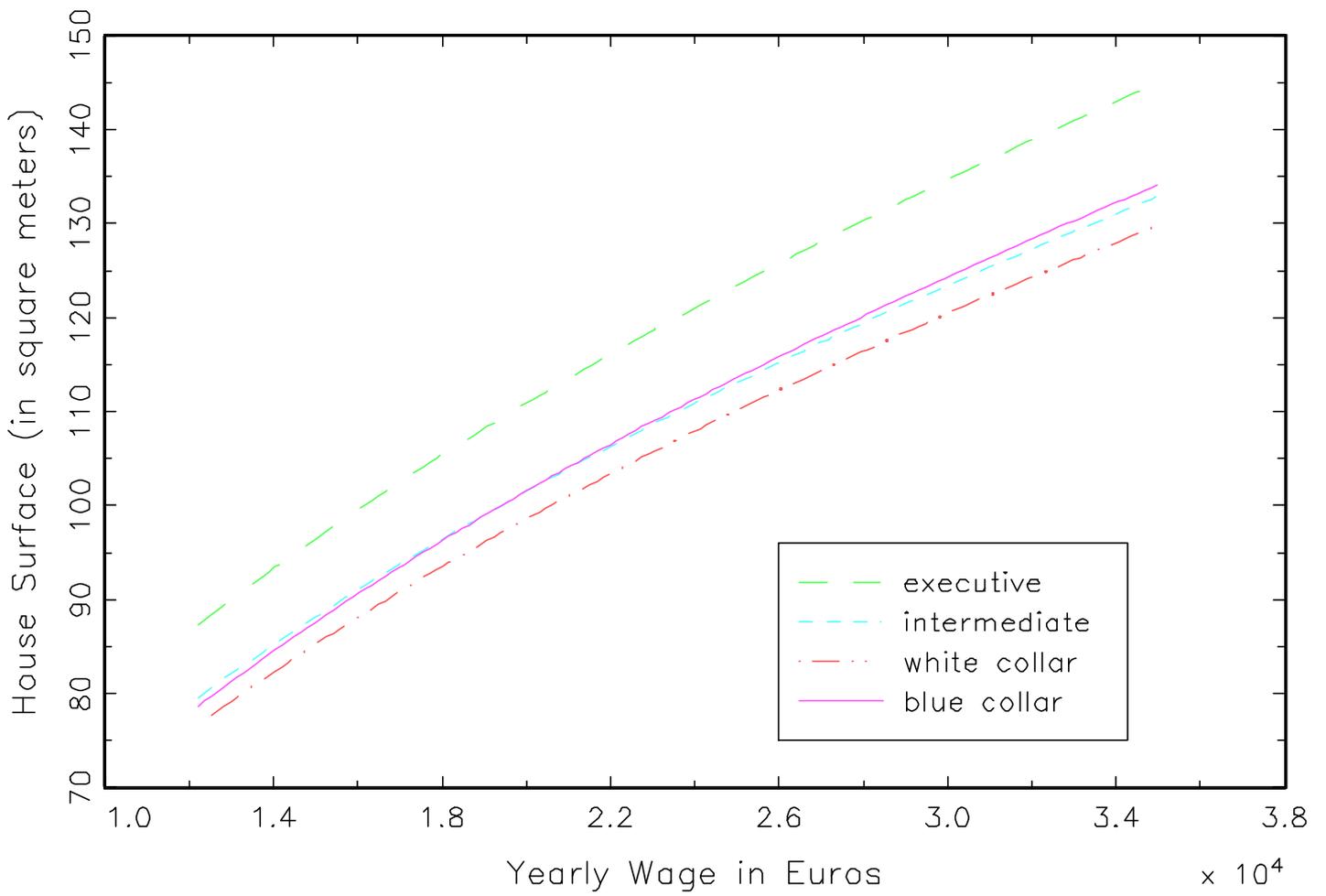


Figure 12. Monopoly Model: Indifference Curves

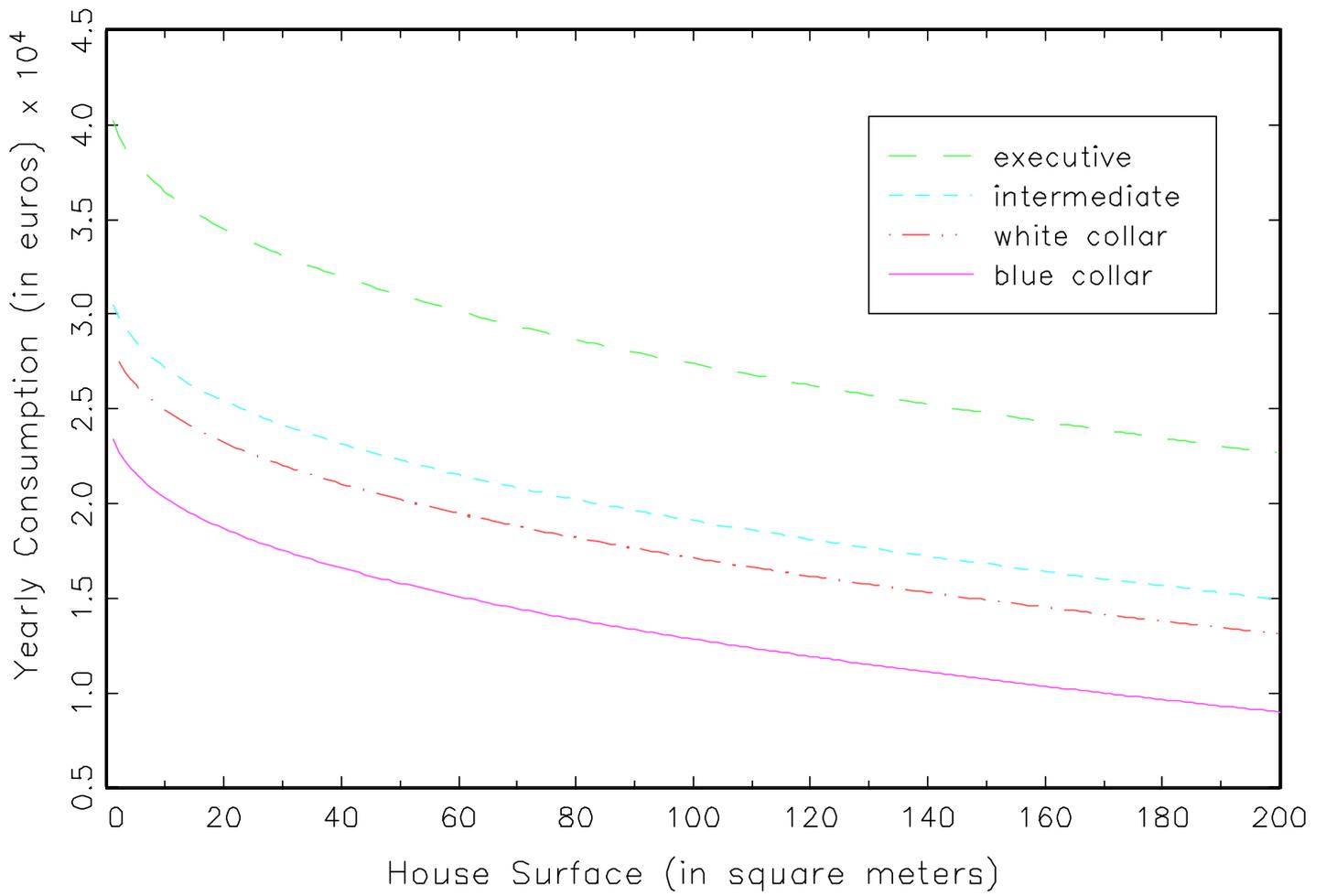


Figure 13. Loan Term Distribution

