A framework for estimating structural dynamic models with aggregate shocks with an application to mortgage default in Colombia between 1997 and 2004

Juan Esteban Carranza

Salvador Navarro*

Abstract

We propose a framework for estimating dynamic behavioral models accounting for the presence of unobserved state variables that are correlated across individuals and across time periods. We extend the standard literature on the structural estimation of dynamic models by incorporating an unobserved aggregate correlated shock that affects the individuals' static payoffs and the dynamic continuation payoffs associated with different decisions. Given a standard parametric specification the dynamic problem, we show that the aggregate shocks are identified from the variation in the observed aggregate behavior. The shocks and their transition are separately identified, provided there is enough cross-sectional variation of the observed states. We use our framework to estimate a model of mortgage default for a cohort of Colombian debtors between 1997 and 2004. Results indicate that the dynamic structure and the unobserved heterogeneity are crucial for identifying correctly the impact of different factors on default behavior.

1 Introduction

The estimation of discrete choice dynamic models is limited by the ability of standard microeconometric techniques to incorporate a rich pattern of unobserved heterogeneity affecting the

*University of Wisconsin-Madison

choices of individuals. Accounting for such unobserved heterogeneity is crucial in order to infer correctly the underlying relationship between individual choices and observed relevant variables. The basic techniques for estimating such behavioral models (e.g. Rust (1987), Wolpin (1984), Pakes (1986) and Hotz and Miller (1993); for one comprehensive review of the literature see, for example, Aguirregabiria and Mira (2002)) are based on the assumption that all the unobserved heterogeneity is *iid*. The more recent standard literature (e.g. Keane and Wolpin (1994)) can also accommodate the presence of unobserved states that vary systematically across individuals but stay constant over time. In this paper we develop a framework for estimating dynamic structural models under the presence of an unobserved state variable that is both correlated across individuals *and* over time.

An alternative approach for estimating dynamic choice models with serially correlated shocks has been proposed by Altug and Miller (1998). In their model the structure of the aggregate shocks is estimated separately and used as input into the dynamic model, which is then estimated using the technique developed by Hotz and Miller (1993). Such approach is practical when the aggregate shocks can be estimated from a separate model (e.g. a macroeconomic model). If that is not possible, they have to be estimated from the dynamic choice problem, as we do in this paper.

Our model is based on a Markovian decision problem with finite horizon in which the payoffs depend on observed and unobserved state variables that vary systematically across individuals. We use the model to compute the likelihood of a random panel of observed choices integrating over the parametric distribution of the unobserved states. Given any vector of model parameters, we integrate over the distribution of the random states that vary across individuals using standard numerical techniques.

The main contribution of the paper is the incorporation of an unobserved correlated shock that is common to all individuals and that is correlated over time. The estimation of these aggregate shocks exploits the variation in aggregate behavior, which is a piece of information that is not used directly by the existing literature. Moreover, we show conditions under which these aggregate unobserved shocks and their transition probability are separately identified in a standard specification of dynamic discrete choice model.

We use the proposed framework to estimate a dynamic model of mortgage default and estimate it using micro-level Colombian data spanning the years between 1998 and 2004. During this time, mortgage default rates in Colombia were unusually high due to an unprecedented economic downturn that was accompanied by a dramatic fall in home prices.

The extent to which the fall in household incomes and the fall in home prices contributed separately to the unprecedented rates of default is a relevant policy question that can be answered with the proposed estimation. Our results suggest that the factor driving the default decisions of debtors is home prices, not income. This is consistent with a model of rational dynamic behavior, but not with conventional wisdom. In fact, we show that estimates that don't account for the presence of unobserved correlated states would misleadingly indicate that prices are not a very important determinant of default.

In the second section of the paper we describe our methodological framework. We formulate an optimal stopping problem with correlated unobserved heterogeneity, describe our estimation approach and discuss the identification of the different components of the model. In the third section of the paper we present the application of the model to the Colombian mortgage market. We describe the data, the estimation and the results. The paper concludes with a discussion of the limitations of the proposed framework.

2 The empirical framework

2.1 An optimal stopping problem

Consider the standard optimal stopping problem of an individual i at time $t \leq T_i$, who has to choose action $j \in \{0, 1\}$ where $j = 0$ is an absorbing state over a finite horizon T_i which may be different across individuals. Each choice generates a static a payoff $\tilde{u}_{i,j,t} \equiv u(X_{i,j,t}; \gamma) + \varepsilon_{i,j,t}$ with an observed component $u(X_{i,j,t}; \gamma)$ that depends on a vector $X_{i,j,t}$ and is indexed by a vector of parameters γ. It also depends on an additive unobserved state variable $\varepsilon_{i,j,t}$ that is correlated across individuals and time periods.

At time t, the problem of the individual is to maximize the flow of payoffs from $\tau = t, ..., T_i$:

$$max_{\{d_{i,t},...d_{i,T_i}\}} E_t \sum_{\tau=t}^{T_i} \beta^{\tau-t} \tilde{u}_{i,d_\tau,\tau} \tag{1}$$

where $d_i = \{d_t, ...d_{T_i}\}$ is a sequence of feasible decisions such that once $d_{i,\tau} = 0$ is chosen, no other alternative can be chosen.

Normalize the payoff generated by the action $j = 0$ to zero and relabel $u_{i,1,t} \equiv u_{i,t}$. Let $\tilde{S}_{i,t} \equiv \{X_{i,t}, \varepsilon_{i,0,t}, \varepsilon_{i,1,t}\}$ be the set of relevant state variables for individual i at time t. The vector of observed states $X_{i,t}$ is assumed to follow a first order Markov process, indexed by the parameter vector $\rho_{X,i}$ which can be estimated directly from the data. The unobserved states $\{\varepsilon_{i,0,t}, \varepsilon_{i,1,t}\}$ are also assumed to be Markovian as described below.

We can use the Bellman representation to write recursively the problem for individual i who has not chosen $j = 0$ in the past at time $t \leq T$ as:

$$\tilde{V}_{i,1,t}(\tilde{S}_{i,t}) = max\{u(X_{i,t};\gamma) + \varepsilon_{i,1,t} - \varepsilon_{i,0,t} + \beta E_t\left[\tilde{V}_{i,1,t+1}(\tilde{S}_{i,j,t+1})|\tilde{S}_{i,j,t}\right], 0\} \tag{2}$$

where β is a known exogenous discount rate. At $t = T_i$ the continuation payoff of the problem is zero, so that:

$$E_{T_i}[\tilde{V}_{i,1,T_i+1}(\tilde{S}_{i,j,T_i+1})|\tilde{S}_{i,j,T_i}] = 0 \tag{3}$$

It has been shown before (Rust (1994)) that the model above is not identified non-parametrically. Therefore, the mapping of the model above into data is based on parametric assumptions on the distribution of the unobserved states ε. In order to allow a rich pattern of unobserved correlation, we will decompose this unobserved state as follows:

$$\varepsilon_{i,1,t} - \varepsilon_{i,0,t} \equiv \xi_t + \mu_i + \epsilon_{i,t} \tag{4}$$

where $\epsilon_{i,t}$ is an *iid* idiosyncratic logit disturbance, which is a standard and convenient assumption. The term μ_i is an individual-specific unobservable state that stays constant over time and is distributed among the population of individuals according to a known parametric distribution $\Phi(\mu_i; \sigma)$. The term ξ_t is a common aggregate unobserved shock that follows a first order Markov process. The parameters σ and ρ_ξ indexing the distribution of μ and ξ have to be estimated simultaneously with the whole model.

Notice that under this specification individual choices to be correlated over time and across debtors even after conditioning on the observed states; in addition, this unobserved heterogeneity can be allowed to depend on $X_{i,t}$ which would be equivalent to a model with heterogenous γ coefficients. The model is similar to the standard models except for the presence of the shock $\xi_t \neq 0$ which is allowed to be correlated over time. We will refer to these shocks as aggregate shocks that hist the economy over time, but they more generally can be understood as the common component of the unobserved heterogeneity.

This model nests the standard models mentioned above. Specifically, if we set $\mu_i = \xi_t = 0$, all the unobserved heterogeneity in the model is *iid* and the model is similar to the models in Rust (1987), Wolpin (1987), Hotz and Miller (1993) and Pakes (1986). If we assume away only the aggregate shocks so that $\xi_t = 0$, but account for a correlated individual shock $\mu_i \neq 0$ the model is similar to Keane and Wolpin (1994).

Let $S_{i,t} \equiv \{X_{i,t}, \mu_i, \xi_t\}$ be the the set of state variables, excluding the idiosyncratic *iid* error. Define the expected value function as the expectation of the value function in (2) with respect to the idiosyncratic *iid* shock, conditional on the current states:

$$\begin{aligned} V(S_{i,t}) &= E_\epsilon\left[\tilde{V}_{i,t}(S_{i,t},\epsilon_{i,t})|S_{i,t}\right] \\ &= log(1+e^{u(X_{i,t};\gamma)+\xi_t+\mu_i+\beta E_t[V_{i,1,t+1}(S_{i,t+1})|S_{i,t}]}) \end{aligned} \tag{5}$$

where the second equality is the standard "social surplus" equation which follows from the logit assumption.

For convenience, write the expectation of (5) as a function of the conditioning states as follows:

$$E_t[V(S_{i,t+1})|S_{i,t}] \equiv \Psi(S_{i,t};\rho_{X,i},\rho_\xi) \tag{6}$$

where the expectation is taken with respect to the dynamic states given their realization and their transition parameters. For given state variables and transition parameters, this value can be computed using standard numerical techniques starting at the terminal period.

Conditional on survival, the predicted probability that individual i chooses $j=1$ at time t is given by:

$$\begin{aligned} Pr_{i,1,t} &= Prob\left[u(X_{i,t};\gamma)+\mu_i+\xi_t+\epsilon_{i,t}+\beta E_t\left[V_{i,1,t+1}(S_{i,t+1})|S_{i,t}\right]>0\right] \\ &= \frac{e^{u(X_{i,t};\gamma)+\xi_t+\mu_i+\beta\Psi(S_{i,t};\rho_{X,i},\rho_\xi)}}{1+e^{u(X_{i,t};\gamma)+\xi_t+\mu_i+\beta\Psi(S_{i,t};\rho_{X,i},\rho_\xi)}} \end{aligned} \tag{7}$$

where the continuation payoffs correspond to the expectation of (6). Notice that this probability depends on the realization of the unobserved individual heterogeneity μ_i.

We define now the probability of any given sequence of choices which we will use below. Given (7), denote as $\tilde{Pr}_i$ the probability of an individual history which can be computed as the product of probabilities over the given sequence of choices, conditional on the realization of the individual heterogeneity and the aggregate shocks:

$$\tilde{Pr}_i = \prod_{t=1}^{\bar{T}_i} Pr_{i,1,t}^{d_{i,t}}(1-Pr_{i,1,t})^{(1-d_{i,t})}d\Phi \tag{8}$$

where $\bar{T}_i$ is the last time period at which the loan is observed to be outstanding either because it is defaulted on or because it reaches its maturity, i.e. the time when individual i first chooses $j=0$ or the final period T_i if it always chooses $j=1$.

2.2 Estimation

Consider estimating the model above using a random sample of $i = 1, ..., N$ individuals who are observed solving the described optimal stopping problem during a sequence of $\bar{T} = max\{\bar{T}_1, ..., \bar{T}_N\}$ time periods. For notational convenience assume that all individuals start to solve the problem simultaneously but then have potentially different problem horizons T_i. For each individual, a matrix of potentially time-varying exogenous state variables $X_i = \{X^0_{i,1}, ..., X^0_{i,\bar{T}_i}\}$ is observed, as well as a sequence of decisions $d^0_i = \{d^0_{i,1}, ..., d_{i,\bar{T}_i}\}$.

Given the observed states, its transition parameters $\rho_{X,i}$ which potentially vary across individuals, can be estimated directly before estimating the whole model if they are exogenous. The remaining parameters $\theta = \{\gamma, \sigma, \rho_\xi\}$ and $\xi = \{\xi_1, ..., \xi_{\bar{T}}\}$ have to be estimated. The sample likelihood is given by:

$$\begin{aligned} \ell(\theta, \xi) &= \prod_{i=1}^{N} \int \left[\prod_{\tau=1}^{\bar{T}_i} Pr_{i,1,\tau}^{d^0_{i,\tau}} (1 - Pr_{i,1,\tau})^{1-d^0_{i,\tau}} \right] d\Phi \\ &= \prod_{i=1}^{N} \int \tilde{Pr}_i(\gamma, \rho_\xi, \xi) d\Phi(\sigma) \end{aligned} \tag{9}$$

where the choice probabilities are integrated with respect to the initial distribution Φ of μ.

The model is estimated efficiently by maximizing the likelihood function over the space of parameters. Notice that the estimation of the model in principle is identical to the estimation of standard models, except for the presence of the aggregate shocks ξ and their transition ρ_ξ. Depending on the case, maximizing (9) can be difficult, specially if the number of periods $\bar{T}$ is large, because each shock ξ_t has to be estimated, for all t.

We show now that the model , if desired, the estimation of these aggregate shocks can be concentrated out from the wider estimation algorithm. In other words, we show that the estimation of the model is identical to the estimation of a standard model with the addition of a restriction that identifies the aggregate shocks. Specifically, take the derivative of (9) with respect to each ξ_t and set it equal to zero to obtain the following condition:

$$\begin{aligned} \frac{N_{d_{i,t}=1}}{N_t} \equiv \bar{s}_{1,t} = & \left[\frac{1}{N_t} \sum_{i=1}^{N} \int Pr_{i,1,t} \frac{\tilde{Pr}_i}{\int \tilde{Pr}_i d\Phi} d\Phi \right] \\ & + \left[\frac{1}{N_t} \sum_{i=1}^{N} \int \beta \frac{\partial \Psi_{i,t}(S_{i,t})}{\partial \xi_t} (Pr_{i,t}(S_{i,t}) - d_{i,t}) \frac{\tilde{Pr}_i}{\int \tilde{Pr}_i d\Phi} d\Phi \right] \end{aligned} \tag{10}$$

where $N_{d_{i,t}}$ is the number of individuals in the sample who choose action $j = 1$ at time t.

The first term on the right hand side of (10) is the expected aggregate choice probability conditional on the observed histories. The second term is the sample covariance of the prediction error and the derivative of the expected continuation payoff with respect to the aggregate shock, conditional on the observed histories. This implies that in finite samples, an efficient estimation of the model won't match the predicted and the observed aggregate choice probabilities exactly.

Equation (10) generates a set of $\bar{T}$ non-linear equations, which can be used to concentrate out the estimation of ξ from the problem of estimating θ. In other words, for any set of parameters θ_0, we can solve for the parameters ξ_0 that satisfy (**??**) as we look numerically for the estimator θ^* and its associated ξ^*.

Notice that in large samples (**??**) reduces to a set of intuitive average probabilities. Since the predicted choice probability and the expected continuation payoffs are conditioned on the same set $S_{i,t}$ of state variables, the covariance of the second term should converge to zero, since this covariance is zero in the population. It follows then, that when N_t is large the expression above can be approximated by the following expression:

$$\frac{N_{d_{i,t}=1}}{N_t} \equiv \bar{s}_{1,t} \approx \left[\frac{1}{N_t} \sum_{i=1}^{N} \int Pr_{i,1,t} \frac{\tilde{P}r_i}{\int \tilde{P}r_i d\Phi} d\Phi \right] \tag{11}$$

Which might be an easier expression to use when concentrating out the estimation of ξ.

Moreover, if we compute (10) in the population, we obtain a condition that we state as Lemma 1. This lemma can also be used to concentrate out the estimation of ξ when the population shares are observed and the sample is large. For this define the empirical distribution of the observed states as $F_t(x)$, which is by assumption independent of the distribution Φ of unobserved states. Let also $\mathbf{s}_{1,t}$ be the share of choice $j = 1$ at each time t among active agents.

Lemma 1 *Consider the estimation of the model described by the choice probabilities (7) and (8). At the true value of θ and ξ the following condition holds:*

$$\mathbf{s}_{1,t} = \int Pr_{i,1,t}(\theta,\xi) \frac{\tilde{P}r_i(\theta,\xi)}{\int \tilde{P}r_i(\theta,\xi) d\Phi(\sigma)} d\Phi(\sigma) dF_t(x) \equiv \tilde{s}_{1,t}(\theta,\xi) \tag{12}$$

This lemma states that at the true value of the parameters, the observed aggregate choice probability has to be equal to a weighted average of the predicted choice probabilities. The weighting is equivalent to conditioning the predicted choice probabilities on the observed choice history of each individual up until the terminal period $\bar{T}_i$.

As a corollary of this lemma, we point out below that if there is no persistent unobserved heterogeneity the condition (12) reduces to a simple average. This condition is similar to the standard BLP-style market-level condition that is used to concentrate out the estimation of choice-specific shocks from the estimation of discrete demand systems, except that it only holds in large samples. The proof follows trivially from Lemma 1, by noting that when there is no persistent unobserved heterogeneity, the integrals in the expressions above vanish.

Corollary 1 *Consider the estimation of the model described by the choice probabilities (7) and (8). Let $\mu_i = \mu\ \forall i$ so that the distribution Φ is degenerate. At the true value of θ and ξ the following condition holds:*

$$\mathbf{s}_{1,t} = \int Pr_{i,1,t}(\theta,\xi)dF_t(x) \tag{13}$$

An interesting feature of (10) and (11) is that the average choice probabilities at any period t are not conditioned on the survival until $t-1$ but on the whole history until $\bar{T}_i$. This property is not a consequence of the dynamic structure of the problem, but of the presence of unobserved correlated shocks. In fact, this condition extends to static models (as in Goolsbee and Petrin (2004)), in the sense that whenever there are unobserved correlated shocks, an efficient estimation with a finite sample would require that observed aggregate choice behavior matches the predicted behavior, conditional on the observed choices.

When the population shares $\mathbf{s}_{1,t}$ are known, so that the data set is a combination of micro-level and market-level information, Lemma 1 can be used to "concentrate out" the estimation of the aggregate shocks $\{\xi\}$ from the estimation algorithm using the aggregate choice probabilities.

Specifically, at each time t and given parameters θ_0 and ξ_t, the model generates a vector of aggregate predicted choice probability $\tilde{s}_{1,t}(\theta_0,\xi_0)$. If the model is correctly specified and the sample is large (12) must hold:

$$\mathbf{s}_{1,t} = \tilde{s}_{1,t}(\theta,\xi)\forall t \tag{14}$$

Given any value of θ^0, the expression in (14) generates a system of $\bar{T}$ non-linear equations, so that a unique value of $\xi(\theta^0)$ can be solved directly. If the population shares $\mathbf{s}_{1,t}$ are not observed, but only the shares $\bar{s}_{1,t}$ in the sample, then (10) or (11) can be used instead.

The feasible computation of the model requires that for any set of feasible parameters θ_0, the vector ξ_0 that solves (10) be always defined. Moreover, the identification of the model will require that the vector ξ_0 be unique, at least around the true vector ξ^*. The following lemma establishes the sufficient conditions under which the solution to (14) exists and is unique. The proof of this lemma, shown in the appendix, relies on the monotonicity of the average predicted default rates (12) on the aggregate shock.

Lemma 2 *Let $E_t[\xi_{t+1}|\xi_t] = h(\xi_t;\theta)$, such that $h(.)$ is strictly monotone and $-1 < h'(\xi_t) < 1$. Then, for the system of T equations implied by $\bar{s}_{1,t} = s_{1,t}(\theta^0, \xi)$ for $t = 1, ..., T$ has a unique solution $\xi(\theta^0)$, if the sample size N is large.*

The sufficient conditions for the lemma to be true are very weak in the sense that they are far from necessary. Moreover, they imply restrictions that are usually natural in empirical environments. For example, if the aggregate shocks follow a linear autoregressive process, a sufficient condition for the lemma and the corollary to hold is that the process is stationary.

Lemmas 1 and 2 will be used to show our identification result below. For practical purposes, they imply that the model can be estimated using the standard techniques with the addition of (14) as a separate restriction, thereby reducing the computational dimension of the estimation algorithm. In other words, there is no need to maximize the likelihood over all the parameters of the model, which is useful specially when the number of periods is large. Specifically, the model can be estimated maximizing the likelihood (9) over the parameters θ, solving numerically for ξ from (14) along the estimation algorithm:

$$max_{\theta}\ell(\theta, \xi(\theta)) \tag{15}$$

Notice that the model is overidentified, in the sense that the levels of ξ and their transition parameters ρ_ξ are estimated separately. Therefore, additional moments can be added to (15) to guarantee the consistency of both, which might be desirable in long panels.

Before presenting an application of our methodology, in the following sections we discuss the identification of the components of the model and the applicability of the methodological framework to more general environments.

2.3 Identification of the model

We have already pointed out that in general the described model is nont identified non-parameterically as shown by Rust (1994). We discuss now the identification of the parametric model described above and show the conditions under which such identification is possible. The main problem lies in the separate identification of the aggregate shocks and their transition, which we show is possible only when micro level information is available. Importantly, the identification conditions are sufficient and necessary.

The choice probabilities in (7) are similar to the choice probabilities in standard empirical dynamic models with unobserved heterogeneity like in Keane and Wolpin (1994), except for the presence of the aggregate shocks ξ and their transition ρ_ξ. Therefore, the identification of the preference parameters γ and the parameters σ of the distribution of μ is based on similar arguments as in the standard literature. Therefore, we provide a brief discussion of their identification and then discuss in detail the identification of the aggregate shocks ξ and their transition parameters ρ_ξ.

The finite horizon of the problem facilitates the identification of much of the parameters of the model. Since at T_i the continuation payoffs of the problem are zero, the probability that individual i chooses $j = 1$,obtained from (7), doesn't contain the transition parameters:

$$Pr_{i,1,T_i} = \frac{e^{u(X_{i,T_i};\gamma)+\xi_{T_i}+\mu_{T_i}}}{1+e^{u(X_{i,T_i};\gamma)+\xi_{T_i}+\mu_i}} \tag{16}$$

Conditional on the aggregate and individual-level shocks ξ_{T_i} and μ_i, the identification of γ from (16) is obvious even with a single cross section of individuals who face the same terminal period[1].

The individual-level unobserved shocks are identified as long as we observe the same individuals making choices over time. Specifically, if an individual who according to his observed states $X_{i,t}$ has a high probability of choosing $j = 0$ is observed choosing $j = 1$ consistently over time, it can be inferred that he has a high μ_i. This identification argument is similar to the identification argument in random utility discrete choice models with heterogeneous preference parameters and relies partly on the functional form of the utility function.

The novel part of this paper is the separate identification of the aggregate shocks and their transition. Intuitively, the identification of the aggregate shocks comes from the variation in the data of the aggregate behavior, which is not fully exploited in the standard literature. Notice that, in practice, our estimation approach is equivalent to a standard estimation of a Markovian decision model, with the addition of an "aggregate" restriction (14), which directly identifies the aggregate shocks.

[1]Notice that, more generally, these preference parameters are identified non-paramaterically at T_i as shown by Heckman and Navarro (2007).

The separate identification of the levels ξ of the aggregate shocks and their transition parameters ρ_ξ has to be explained in detail. From inspecting (7) it can be seen that both the aggregate shocks ξ and their transition parameters ρ_ξ enter the continuation payoffs. Moreover, ξ enters additively the instant payoffs, so that potentially it can happen that changes in ξ that are offset by changes in their expected serial correlation generate identical predictions, so that they would not be separately identified.

We have two sources for the separate identification of the two set of unobservables. On one hand, notice from (16) that in the terminal periods $\{T_1, ..., T_N\}$ the parameters ρ_ξ don't enter the choice probabilities and therefore the aggregate shocks are identified up to the constant of the utility function. Therefore, if we observe individuals who face their terminal period at each time period of our sample, ξ will be identified separately from ρ_ξ.

The second and more general source of identification is the functional form of the choice probabilities, so that ξ and ρ_ξ will be separately identified even in a sample of individuals who face the same terminal period. To see this, notice that at the true value ρ_ξ^* of the transition parameters, our estimation algorithm looks for the unique vector ξ^* that satisfies (14) which we can rewrite as follows:

$$
\begin{aligned}
\mathbf{s}_{1,t} &= \int Pr_{i,1,t}(.;\hat{\gamma},\xi^*,\rho_\xi^*)\frac{\tilde{Pr}_i(.;\hat{\gamma},\xi^*,\rho_\xi^*)}{\int \tilde{Pr}_i(.;\hat{\gamma},\xi^*,\rho_\xi^*)d\Phi}d\Phi dF_t(x) \\
&= \int \tilde{Pr}_{i,t}(.;\hat{\gamma},\xi^*,\rho_\xi^*)dF_t(x) \qquad (17)
\end{aligned}
$$

where $\bar{s}_{1,t}$ is the observed proportion of individuals who choose $j = 1$ at time t and where $\tilde{Pr}_{i,t}$ is the choice probability integrated over the distribution of individual heterogeneity, conditional on each choice history:

$$
\tilde{Pr}_{i,t}(.;\hat{\gamma},\xi^*,\rho_\xi^*) = \int Pr_{i,1,t}(.;\hat{\gamma},\xi^*,\rho_\xi^*)\frac{\tilde{Pr}_i(.;\hat{\gamma},\xi^*,\rho_\xi^*)}{\int \tilde{Pr}_i(.;\hat{\gamma},\xi^*,\rho_\xi^*)d\Phi}d\Phi
$$

As we change ρ_ξ, the algorithm will find new vectors of ξ consistent with (17). The implicit function theorem implies that the variation of ξ as ρ_ξ changes is given by:

$$
\frac{\partial \xi_t}{\partial \rho_\xi} = -\frac{\int (\partial \tilde{Pr}_{i,1,t}/\partial \rho_\xi)dF_t(x)}{\int (\partial \tilde{Pr}_{i,1,t}/\partial \xi)dF_t(x)} \qquad (18)
$$

If such variation in ξ leads to the same choice probabilities as in (18), then the two sets of parameters are not separately identified. Notice, though, that at any given ρ_ξ and for every agent i, the implicit variation of ξ as ρ_ξ changes such that $\tilde{Pr}_{i,1,t}$ is constant is given by:

$$\frac{\partial \xi_t}{\partial \rho_\xi} = -\frac{(\partial \tilde{Pr}_{i,1,t}/\partial \rho_\xi)}{(\partial \tilde{Pr}_{i,1,t}/\partial \xi)} \tag{19}$$

which is in general different than (18), as long as the predicted choice probabilities vary across individuals. Consequently the predicted choice probabilities change as the transition parameters change.

In other words, if there is variation in the observed states across individuals, the derivative of the individual choice probabilities with respect to the ρ_ξ is different from zero. Consequently, the sample likelihood will necessarily fall around the estimated parameter ρ_ξ^* so that ξ and ρ_ξ are separately identified as formally established in the following proposition, which we prove in the appendix.

Proposition 1 *Consider the model with sample likelihood $\ell(\gamma^0, \sigma^0, \rho_\xi)$ given by (9) with known parameters γ^0 and σ^0. Assume that the conditions in Lemmas 1 and 2 hold. The parameter vector ρ_ξ is identified if and only if the states $X_{i,t}$ vary across individuals for at least one individual i for all t.*

The proposition establishes the identification of ρ_ξ, conditional on the known parameters γ and σ, whose identification was explained before. Moreover, the identification of ρ_ξ is formally independent from the identification of ξ. Therefore, if we estimate ρ_ξ using the estimated ξ we might find substantial discrepancies with the estimated ρ obtained from the estimation above, specially in short samples.

The identification of the parametric model is not surprising. The more important result is the *nonidentification* of the model when no micro level data are available. There is a growing literature on the estimation of structural dynamic models of demand using market-level data (e.g. Carranza (2007) and Gowrisankaran and Rysman (2006)). Our result highlights the limits of the identification of that general class of models.

2.4 Further remarks on the methodology

For illustrative purposes, we have described our methodological framework using a simple binomial optimal stopping problem. The general approach extends naturally to more general dynamic Markov decision problems with multiple repeated choices.

For example, if instead of an absorbing state, we let individuals choose $j = 0$ repeatedly, the only difference is that a continuation payoff has to be computed for both $j = 0$ and $j = 1$. This adds to the computational burden of the algorithm, but the fact that we would observe the same individuals making the same choices repeatedly over time would also strengthen the identification of the individual-level unobserved heterogeneity.

In addition, we can allow for multiple choices each with its associated continuation payoff. The computation of multiple continuation payoffs along the estimation algorithm is feasible but computationally costly. In addition, the data requirements are stronger, as the identification of the aggregate shocks relies on the computation of choice-specific aggregate probabilities. Otherwise, the estimation approach is the same.

3 An application to the Colombian mortgage market between 1998 and 2004

3.1 Description of the data

We will use the described empirical model to estimate a dynamic model of optimal default using two separate data sets with information on the behavior of Colombian mortgage debtors between 1997 and 2004. The first (or "main") data set contains information on a set of random mortgages that were outstanding between 1997 and 2002. The monthly payment history of each mortgage, its original and current value and term of the mortgaged home are included. A "secondary" data set contains non-matching individual-level demographic data, including income and real estate holdings.

The total number of loans contained in the main data set is 16000. Nevertheless, this set of mortgages includes loans that started at different points in time, most of them before 1997. From this subset of loans that started before 1997 we only observe those that survived until 1997. Since our model predicts that loan survival is endogenous, for the estimation below we selected the cohort of loans started during the year 1997 and assumed that the distribution of unobserved attributes of new debtors is the same throughout that year. After eliminating from our sample those loans with incomplete or inconsistent payment histories, we ended with a total of 925 loans which are observed from the time they start in 1997 until 2004[2]. Despite the reduction in the number of mortgages in our data set, we end up with a

[2]For a detailed study of the default behavior observed in the whole sample using a simpler empirical

panel with 14250 observations.

The data set contains only the price of each home at the time the loan started as reported by the bank. These prices are very reliable because banks are very serious about the value of the collateral. The expected prices of individual homes at any point in time $\bar{P}_{it}$ are updated using housing price indices constructed by the Colombian Central Bank. In addition, all data is aggregated into quarters, so that default observations are not confounded with missed payments or coding errors. All variables are expressed in constant 1997 real Colombian pesos.

Since this main data set contains no information on the income of debtors over the span of the sample, survey data from the secondary data set was used to control for the changing distribution of income. This data set is part of an annual survey conducted by DANE that contains demographic information of large samples of individual household. We selected households in the sample who reported having a home loan. We use the reported income and matching housing payments to simulate the joint distribution of income and the other state variables.

In the data it is observed that some debtors stop making their payments, sometimes only temporarily and sometimes definitively. Therefore what 'default' means and its timing has to be defined. Specifically, in the estimation below, loans that accumulate past due payments of more than 3 months are assumed to be defaulted and are dropped off from the data set. Therefore, 'default' is defined as the event in which the number of past due payments in a loan history changes from 3 or less to more than 3 between two quarters. After a loan is defined to be defaulted, it is dropped off the sample[3].

Table 1 contains some summary statistics of the main data set, which goes from the first quarter of 1997 to the second quarter of 2004[4]. The number of loans in the data set increases during the first four quarters of 1997 as new loans are initiated until reaching 925 which is the total number of loans in the cohort. Notice from column (3) that the number of number of non-defaulted loans decreases gradually over time which is a reflection of the high numbers of defaults observed in the sample. The default rate, defined as the number of defaults over

model see Carranza and Estrada (2007).

[3]The default rate based on this definition is highly correlated with default rates based on longer default periods. The 3-month threshold was chosen in order to observe as much default as possible and in order to capture *all* defaulted loans, including those that are terminated soon after default.

[4]Since default is inferred from the change in the number of past due mortgage payments, no default is reported during the first period of the sample.

the total number of outstanding loans in column (4), reaches a level higher than 7% during the fourth quarter of 1999, which is indicative of the severity of the market collapse. By the end of the sample more than half of the loans in the sample were defaulted.

To give a sense of the characteristics of the defaulted loans we computed the average price of homes with outstanding loans (column (5)) and the average price of all homes in the sample (column (6)). Notice that up until the middle of 1999, the average price of homes with outstanding mortgages was higher than the average price of the homes of all the loans in the sample which implied that defaults tended to occur among the mortgages of the least expensive homes. After 1999 the price of homes with outstanding loans was lower than the average price of all homes in the sample, which implied that it was among mortgages of the more expensive homes where defaults were concentrated.

Besides the rich modelling of the structural error in our model, we use the secondary data set to account for the unobserved variation in individual incomes. The data correspond to the quarterly household survey collected by the Colombian national statistics agency (DANE). The survey collects demographic and economic information of a random sample of households. All households are asked their household income. In addition, once a year they are asked whether they have a mortgage or not and the corresponding monthly payments.

In order to control for the unobserved variation in income we use the distribution of income that we observe in this data set, conditional on whether the household has a mortgage or not and on their monthly payments. Specifically, for each household we simulate several income draws from the data to integrate out part of the unobserved heterogeneity. The draws are taken from the corresponding quintile of the distribution of income ordered according to the monthly mortgage payments which is assumed to match the distribution of income conditional on the ratio of balance to remaining term.

To understand the roots of the extraordinarily high observed default rates in Colombia in these years, it is important that we describe the history and institutional details of the Colombian mortgage financing system. The centerpiece of the system, established in the 1970's, were the mortgage banks whose only purpose was to fund construction projects. In order to guarantee enough funding, these banks were the only institutions allowed to issue interest-bearing savings accounts[5].

In addition, mortgage loans were denominated in a constant value unit called "UPAC"[6],

[5] Regular commercial banks had exclusive rights to issue checking accounts bearing no interest.

[6] UPAC stands for Unidad de Poder Adquisitivo Constante: Constant Purchasing Power Unit

whose value changed over time according to a rate (called the “monetary correction") determined by the Central Bank which was supposed to reflect the inflation rate. The UPAC protected institutions and debtors against inflationary risks and facilitated the long-run financing of housing projects, which in turn gave a boost to the economy during the following decades.

Each month, debtors had to pay a proportion of the outstanding balance of their debt. In addition, each month debtors made an interest payment on the balance. This additional interest rate was fixed for the lifetime of the loan and was not set on a debtor-by-debtor basis, but was rather negotiated between the mortgage bank and the developer in charge of the construction of any type of housing project, before individual homes were sold. The following month the remaining balance was updated according to the “monetary correction".

Until the early 1990's the monetary correction pretty much tracked the inflation rate. But then the government decided to liberalize the financial sector and allowed commercial banks to offer savings accounts, which until then could only be offered by the mortgage banks. The government also decided to tie the “monetary correction" to a market interest rate, which meant that interest was added over time to the balance of the debts.

During these years the Colombian exchange rate was fixed and positive flows of capital kept the interest rate low. Then in the 1990's the flows of capital reversed, as happened in virtually all emerging economies. The Central Bank decided to defend the exchange rate at any cost, as did most countries in the area, which meant letting the interest rates increase to unprecedented levels which had itself a devastating effect on the real sector, including the housing industry. In addition, as home prices and household incomes started to fall, mortgage balances, that were now tied to the interest rate, ballooned. At the end of the decade and due to the default rates observed in the data, mortgage financing in Colombia came to a halt and was only reestablished several years later under a different regulatory framework.

One of the important policy questions raised by the crisis is the extent to which the observed default rates were caused by the government policies and what extent was caused by the fall in income, which was probably inevitable. We will be able first to measure the effect of changes on each variable on the default probabilities. Moreover, we will be able to simulate the effect of counterfactual policies.

3.2 The model of default

We study the behavior of mortgage holders ("debtors") who live in the mortgaged piece of real estate ("home"). Let the utility that a debtor i gets from the home each period t be given by the following linear function:

$$\tilde{u}(q_{i,t}, y_{i,t} - m_{i,t}, \varepsilon_{i,t}) = \theta_0 + \gamma q_i + \alpha(y_{i,t} - m_{i,t}) + \varepsilon^u_{i,t}. \tag{20}$$

where $q_{i,t}$ is a measure of subjective home quality, $y_{i,t} - m_{i,t}$ is the difference between household income and mortgage payments and $\varepsilon^u_{i,t}$ is an additive unobserved state variable, which incorporates unobserved (to the econometrician) variables that may affect default, e.g. home attributes that are only valued by its owner and other preference shocks that vary across consumers and time.

Since no home attributes are observed in our application, we further assume that the unobserved "quality" of homes $q_{i,t}$ is random:

$$q_{i,t} \equiv \kappa + \varepsilon^q_{i,t}, \tag{21}$$

where $\varepsilon^q_{i,t}$ is a random variable that is potentially correlated over time and across debtors. Any systematic difference in the subjective home quality across debtors will be captured by the correlation structure of the error which will be described in the estimation section below.

In our data set we have no information on the required payments $m_{i,t}$ of each debtor. However, it is known that the required payments are linear functions of mortgage balances $b_{i,t}$ and remaining term $L_{i,t}$, with some random variation across debtors:

$$m_{i,t} = \rho_0 + \rho_1 b_{i,t} + \rho_2 L_{i,t} + \varepsilon^m_{i,t}. \tag{22}$$

where $\varepsilon^m_{i,t}$ is an error term.

We assume that "default" leads to an absorbing state. Let $W_{i,t}$ denote the value for individual i of defaulting on her mortgage at time t. This value is the result of a complex scenario. Specifically, the individual may be waiting to see whether the following period she can pay back her dues; she may try to sell the home and cash the difference between price and loan balance; she may let the bank take over the property to cover her obligation; finally, she could also just stop making payments indefinitely and face forfeiture or a renegotiation with the bank.

The resulting value of default $W_{i,t}$ is the weighted sum of payoffs across the random scenarios just described. We assume that $W_{i,t}$ has the following linear reduced form:

$$W_{i,t} = \omega_0 + \omega_1 y_{i,t} + \omega_2 \bar{\pi}_{i,t} + \omega_3 b_{i,t} + \varepsilon^w_{i,t}. \tag{23}$$

where $\bar{\pi}_{i,t}$ is the expected price of the home at time t, $b_{i,t}$ is the balance of the debt, $y_{i,t}$ is the debtor's income and $\varepsilon^w_{i,t}$ are other unobserved (to the econometrician) attributes. These are variables that enter directly the payoffs of the individual scenarios arising after a default decision as discussed above.

Group the unobserved components into one error term $\bar{\varepsilon}_{i,t} \equiv \gamma\varepsilon^q_{i,t} - \alpha\varepsilon^m_{i,t} + \varepsilon^u_{i,t} - \varepsilon^w_{i,t}$ and let $\tilde{S}_{i,t} = \{\bar{\pi}_{i,t}, y_{i,t}, b_{i,t}, L_{i,t}, \bar{\varepsilon}_{i,t}\}$ be the vector of observed and unobserved states and assume that it follows a first order Markov process. We can obtain the value of the debtor's problem at each point in time as function of variables that can be mapped to the data and of unobserved random variables:

$$\tilde{V}_{l_i,t}(\tilde{S}_{i,t}) = \max\left\{0, \zeta_0 + \zeta_1\bar{\pi}_{i,t} + \zeta_2 y_{i,t} + \zeta_3 b_{i,t} + \zeta_4 L_{i,t} + \bar{\varepsilon}_{i,t} + \beta E\left[\tilde{V}_{l_i,t+1}(\tilde{S}_{i,t+1})|\tilde{S}_{i,t}\right]\right\} \tag{24}$$

where it is assumed that at the last period of the mortgage T_i the continuation payoff of non-default is zero:

$$E\left[\tilde{V}_{l_i,T_i+1}(\tilde{S}_{i,T_i+1})|\tilde{S}_{i,T_i}\right] = 0 \tag{25}$$

The parameters to be estimated $\zeta = \{\zeta_0, \zeta_1, \zeta_2, \zeta_3, \zeta_4\}$ are linear combinations of the underlying structural parameters. Notice that this function can be computed recursively starting from the last period if all the state variables and their transition are known.

3.3 Estimation

In order to estimate the model we decompose the unobserved state $\bar{\varepsilon}_{i,t}$ as follows:

$$\bar{\varepsilon}_{i,t} = \xi_t + \mu_i + \epsilon_{i,t} \tag{26}$$

where μ_i is an individual-specific unobservable state and $\epsilon_{i,t}$ is an *iid* idiosyncratic logit disturbance. The term ξ_t is a common aggregate unobserved shock with a transition indexed by the vector $\rho^\xi = \{\rho^\xi_0, \rho^\xi_1, \rho^\xi_2\}$ as follows:

$$\xi_{t+1} = \rho^\xi_0 + \rho^\xi_1 \xi_t + \omega^\xi_t \tag{27}$$

where ω^ξ_t is an error with variance ρ^ξ_2.

We estimate the model above using debtor-level data on mortgage balances, mortgage terms and home prices over a set of $t = 1, ...T$ time periods. Since the Colombian mortgage data we use does not contain matching income data tracing the evolution of income for individual debtors, we use household survey data containing information on debtors' income and mortgage payments as described in the data section. Therefore, we treat income $y_{i,t}$ as an unobserved state with distribution given by $G_t^y(y|b/L)$, which is the empirical distribution of income conditional on the mortgage payments.

We also assume that μ correlated with the initial loan-to-value ratio (LTV) of each loan, which as said before is regarded as a good predictor of the risk attitude of debtors. We assume that this underlying correlation is determined by the following loading equation:

$$LTV_i = \alpha_0 + \alpha_1 \mu_i + \nu_i \tag{28}$$

where $\nu_i \sim N(0, \alpha_2^2)$, and μ is distributed according to the mixture of three normal distributions with parameters $\sigma = \{\bar{\mu}, \sigma_\mu^2, w_\mu\}$ such that $\bar{\mu}$, σ_μ^2 and w_μ are 3×1 vectors containing the means, the variances and the probabilities of each distribution, respectively, such that the mean of the mixture is zero and its variance is one. We denote this distribution as $\Phi(\mu; \sigma)$. The vector σ of distribution parameters and the coefficients of (28) above are estimated jointly with the other parameters of the model.

Let $X_t \equiv \{X_{1,t}, ..., X_{N_t,t}\}$ where $X_{i,t} = (\bar{\pi}_{i,t}, b_{i,t}, L_{i,t})$ contains the observed states. We estimate the transition $X_{i,t}$ directly from the data according to:

$$log(b_{it+1}) = \rho_0^b + \rho_1^b log(b_{it}) + \prod_{l=1}^{3} \rho_l^b L_{i,t}^l + \omega_{it}^b \tag{29}$$

$$\log(pi_{it+1}) = \rho_0^\pi + \rho_1^\pi log(\pi_{it}) + \omega_{it}^\pi$$

$$log(y_{it+1}) = \rho_0^y + \rho_1^y log(y_{it}) + \omega_{it}^y$$

where $\{\omega_{i,t}^b, \omega_{i,t}^\pi, \omega_{i,t}^y\}$ are *iid* errors and $\rho_X = \{\rho^y, \rho^b, \rho^\pi\}$ are parameters to be estimated. The transition of the balance is assumed to depend on both the balance and the remaining term of the mortgage. It is estimated using only non-defaulted mortgages so that it reflects the expected evolution of the balance for household that have not defaulted yet. Since house prices are updated using a price index, the transition is basically the same for everyone. The transition of income is common for households within the same quintile of the income distribution.

Under the given assumptions, the model above generates the following *non*-default probability for debtor i at time t conditional on not having defaulted on the mortgage up to $t-1$ and conditional on the realization of the random states:

$$Pr_{i,t}(\bar{\pi}_{i,t}, b_{i,t}, L_{i,t}, y_{i,t}, \mu_i, \xi_t, l_i) = \frac{e^{\zeta_0+\zeta_1\bar{\pi}_{i,t}+\zeta_2 y_{i,t}+\zeta_3 b_{i,t}+\zeta_4 L_{i,t}+\xi_t+\mu_i+\beta\Psi_{l_i,t+1}}}{1+e^{\zeta_0+\zeta_1\bar{\pi}_{i,t}+\zeta_2 y_{i,t}+\zeta_3 b_{i,t}+\zeta_4 L_{i,t}+\xi_t+\mu_i+\beta\Psi_{l_i,t+1}}}. \tag{30}$$

where $\Psi_{l_i,t} = E_\epsilon \tilde{V}_{l_i,t}$ is the expected value function as defined in (5) and (6) which is computed separately for each l_i using the specified transition probabilities.

For any realization of the aggregate shocks and any choice of parameters $\theta^0 = \{\zeta^0, \sigma^0, \alpha^0, \rho^{\xi 0}\}$ we can obtain the aggregate *non*-default probability for each time period as defined in (12):

$$s_t(\xi_t, X_t; \theta^0) = \frac{1}{N_t} \frac{\int \prod_{\tau=1}^{\bar{T}_i} Pr^0_{i,\tau} Pr^0_{i,t} dG^Y_t(Y|K) d\Phi(\mu; \sigma^0)}{\int \prod_{\tau=1}^{t-1} Pr^0_{i,\tau} dG^Y_t(Y|K) d\Phi(\mu; \sigma^0)} \tag{31}$$

where Φ is the normal distribution function. Given (31), the implied vector of aggregate shocks $\xi(\theta^0)$ can be solved from (14).

Let $d_{i,t} \in \{0,1\}$ be the observed choice of individual i at time $t \le T_i^*$, where T_i^* is the the time when i defaults, the last period of the mortgage or the last period at which she is observed. With the values of the aggregate shocks, $\xi(\theta^0)$, at hand we can compute the likelihood of the sample for any choice of parameters θ^0, which is the product across debtors of individual default/non-default histories, integrated over the distribution of the unobservables:

$$\ell(\theta^0) = \prod_{i\in N} \int \left[\prod_t P_{i,t}^{d_{i,t}} \left(1 - P_{i,t}^{(1-d_{i,t})}\right)\right] dG^Y_t(Y|K) d\Phi^\mu(\mu; \sigma^0) d\Phi(\nu) \tag{32}$$

where the likelihood accounts also for the the distribution of the errors ω of the LTV loading equation. Estimates of θ are obtained by finding the vector that maximizes (32).

3.4 Computation and results.

For any value of θ along the estimation algorithm, the computation of (32) requires the use of numerical techniques to integrate out the distribution of income and μ. We proceed as follows: For each mortgage i at time t, a set of S_i income draws $\{Y_{st}\}_{s=1,\ldots S_i}$ is simulated from the corresponding quintile of the empirical distribution of income conditional on the monthly mortgage payments, contained in the “secondary" data set. In addition, for each income draw and for any vector σ of mixture parameters, the distribution of μ is used to integrate them out using a quadrature method.

The computation of the likelihood of individual default/non-default observations requires in addition the computation of the expected value functions (5), which is done recursively starting from the last period for each mortgage term length l_i. There are three types of term length in the data: 5 years, 10 years and 15 years. For each term length, the expected value functions are computed backwards using interpolation as in Keane and Wolpin (1994), given the transition of the observed states and the assumed transition of the aggregate shocks. The interpolation is made using a multi-linear approximation.

Instead of concentrating out the estimation of ξ, the maximization of the likelihood function (32) is done over the whole parameter space, checking afterwards that at the estimated values the predicted default probabilities match the observed shares as in (14). We estimated eight versions of the model: four duration models with myopic debtors and four fully dynamic models. Each type of model was estimated with and without persistent unobserved heterogeneity and with and without income heterogeneity. The quarterly discount rate was set to $\beta = 0.97$.

We show on table 2 the estimated parameters of the duration models, which are equivalent to the model described above, except that we set the discount rate equal to zero $\beta = 0$. In these models the aggregate shocks are just time-changing constants. Model I contains no dynamics, no persistent unobserved heterogeneity and no income heterogeneity. Model II adds only income heterogeneity to model I, whereas model III adds persistent unobserved heterogeneity to model I. Model IV is a duration model with both persistent unobserved heterogeneity and heterogeneous income.

On table 3 we display the estimated parameters of the fully dynamic model. Model V contains no persistent unobserved heterogeneity and no income heterogeneity. Model VI is a dynamic model with income heterogeneity, whereas model VII has persistent unobserved heterogeneity but no income. Model VIII has full dynamics, persistent unobserved heterogeneity and heterogeneous income.

For each model, we show the estimated coefficients and the estimated marginal effects integrated over the distribution of debtors, with corresponding standard errors. The marginal effects are computed with respect to a 10% change in price, balance and income and a one quarter change in term length. In the case of the dynamic models (table 3), the marginal effects are computed accounting for the effects of changes in the state variables on the continuation payoffs.

We discuss first the results of the estimation of the duration models displayed in table 2, where the dependent variable is the probability of *not* making default, as indicated above. The results imply that, conditional on all other variables, home price has a negative effect on default probability, while the value of the mortgage balance and the remaining number of quarters left in the mortgage have a positive effect on the default probability, which was expected.

The first salient feature of the estimates of the duration model is the effect of accounting for the persistent unobserved heterogeneity on the estimated price and balance coefficients. Comparing the estimates in models I and II with the results of models III and IV, we can see that the price and balance coefficients are in absolute value much bigger in the models that included explicitly the persistent unobserved heterogeneity. The estimated marginal effects, which are precisely estimated, are literally doubled. These effects imply that a 10% increase in balance or home price changes in average the quarterly default probability in one percent point, which economically is a very significant figure.

The second salient feature of the results is the economic irrelevance of income on the default rates. Statistically, models I and II seem to indicate that income is positively correlated with default, which does not make much sense. After controlling for the persistent heterogeneity such correlation becomes insignificant. In any case, the magnitude of the estimated effects is very small.

We also report on the lower part of the table the estimated coefficients of the loading equation that correlates the persistent unobserved heterogeneity with the initial loan-to-value LTV of the loans. The estimates suggest that higher initial LTV is associated with a higher “taste" for default, which simply means that riskier debtors select themselves into more leveraged mortgages. Unfortunately the estimates are statistically insignificant. We also report the variance of the persistent heterogeneity which is computed over the mixture of estimated normal distributions and its respective probabilities (not shown).

The estimates corresponding to the fully dynamic models are presented on table 3. The upper part of the table contains the estimates of the dynamic models without persistent unobserved heterogeneity (models V and VI), while the lower part contains the estimates of the models with persistent heterogeneity (models VII and VIII). The first thing to notice is that the inclusion of persistent unobserved heterogeneity has as significant an effect on the price and balance coefficients as it did in the duration models. In models VII and VIII the estimated marginal effects of 10% changes in price and balance are higher in absolute value than in the duration models, even though the difference is not statistically significant.

It should be remembered that in the dynamic model a change in a variable has an effect on the current default probability through its effect on the current payoffs via the parameter γ, which is the same as ion the duration models. In addition, such change has on effect through its effect on the expected evolution of the variable in the future which affects the continuation payoffs associated with any choice. The marginal effects reported for the dynamic model account for these two effects.

As a consequence the effects of a purely transitory shock to the states that does not affect its transition will be in general smaller in magnitude than the reported marginal effects. In general we cannot compare coefficients across specifications. Nevertheless, the coefficients are more or less comparable across specifications that have no persistent unobserved heterogeneity. To see this, denote the estimated marginal effect as $\hat{me}$ and let $\hat{ce}$ be the estimated coefficient of interest. Abusing notation a little bit, the estimated marginal effect is given then by:

$$\hat{me} = \int \hat{ce}\hat{Pr}_i^2 dF_i$$

where $\hat{Pr}_i$ is the predicted choice probability of debtor i and F_i is the distribution of observed and unobserved debtor characteristics. In the case without unobserved persistent heterogeneity (models I, II, V and VI) the distribution F_i is the same across specifications. We know that at the estimated parameters and for all specifications $\int \hat{ce}\hat{Pr}_i dF_i \approx s$, where s is the observed market share. If the estimated probabilities are more or less similar across specifications along the distribution F, then we know that the estimated coefficients have more or less a similar scaling and are therefore more or less comparable.

If we compare the estimated coefficients in the duration models I and II in table with the estimated coefficients from the dynamic models V and VI in table 3, we can see the estimated coefficients are much larger in the duration models than in the dynamic models. Since these models have no unobserved persistent heterogeneity, we can roughly compare the magnitude of their coefficients. The reason for this difference is that the coefficients of the duration models capture the entire effect of the variables, whereas in the dynamic models, the coefficient captures the purely static effect. This highlights the fact that the dynamic model makes it possible to distinguish between transitory shocks and shocks that spill over time periods.

The estimates on table 3 of the parameters ρ^{ξ} are not very precise. The displayed results correspond to an estimation of the model in which no restriction was imposed to restrict the estimates of ρ^{ξ} to be consistent with the estimates of ξ. As we have pointed out, in our model both sets of parameters are *separately* identified, so that we can actually estimate the implicit beliefs of debtors about the evolution of ξ separately from its realization.

We chose to impose no restrictions, because the span of our sample was relatively short and its timing was extraordinary. One drawback of our decision is that the estimated ρ^{ξ} coefficients have too large standard deviations. We found that persistence coefficient ρ_1^{ξ} of the autoregressive process that drives the expected evolution of ξ is negative. If we estimated the coefficient directly on the estimated ξ, such persistence coefficient is positive. This difference would imply that debtors were too pessimistic about the evolution of the aggregate shocks, and were therefore *anticipating* their default decisions. The lack of statistical significance, though, does not allow us to draw any strong conclusion.

We do not include measures of the fit of the model in the tables of results because the fit of all models at the market level is virtually perfect. We have already shown that the unrestricted maximization of the model likelihood implies that at the estimated parameters (14) holds. In other words for every set of estimated parameters and for every specification of the model, the observed default rate is equal to the average default rate across surviving debtors, weighted by the corresponding history probability.

We finalize our discussion of the results with a counterfactual policy simulation that illustrates the usefulness of the model. As we indicated when describing the data, the observed default rates were driven both by an economic slowdown and an exogenous policy decision that drove up the mortgage balances. We will compute the counterfactual default behavior of debtors under a natural policy alternative. Specifically, we will assume that the "monetary correction" rate which was set by the Central Bank was tied to the inflation rate instead of the market interest rate.

Under the counterfactual policy assumption, each debtor pays a proportion of its real balance each period depending on the number of periods left in the mortgage. Therefore the evolution of *real* balances was perfectly anticipated by debtors. Under the counterfactual assumption, the transition of real balances is given by:

$$b_{i,t+1} = b_{i,t} - b_{i,t}/L_{i,t} = b_{i,t}(1 - 1/L_{i,t}) \tag{33}$$

This transition approximates the initial spirit of the UPAC system, as an institutional arrangement to protect banks and debtors against inflationary risks.

We perform our counterfactual analysis using the estimates of model VIII. Given that our sample size falls rapidly over time as debtors default on their loans, we compute first a baseline simulation with the given transitions. We take all debtors in or sample and have them start their mortgages simulataneously on the first quarter of 1997. For each debtor we simulate ten histories of observed states and unobserved heterogeneity using the estimated distribution of states. The analog of the default rate in the simulation is the hazard rate, which we can average across simulated debtors as we follow their survival and default probability over time. We obtain the counterfactual default rates performing the same computation on the simulated sample using the counterfactual transition of balances (33) instead of the one we estimated from the data.

We show the results of the baseline simulation and the counterfactual computation in Figure 1 over the 30 periods in our sample. As can be seen, the counterfactual default rates are consistently lower than the baseline simulation. Moreover, since as debtors default they can not start again, these differences accumulate over time. At the end of the sample around 70% of debtors have defaulted under the baseline simulation. Under the counterfactual simulation around 50% of debtors default. In other words, the policy of tying the balances to a market interest rate was the cause of at least 2/7 of the observed defaults.

This difference is substantial and is only a lower bound estimate of the impact of the counterfactual policy, because we have kept all other variables at their observed levels. Specifically, we would expect that home prices were affected negatively by the observed default rates. If we allowed for general equilibrium effects, the home prices would be higher in the counterfactual simulation and the equilibrium counterfactual default rates would be even lower.

4 Final remarks

The dynamic model of default described above was estimated with a methodology that accounts for a very rich structure of unobserved heterogeneity. Specifically, it incorporates individual-level heterogeneity using both survey and simulated data. Our main contribution is the addition of aggregate time-varying heterogeneity, allowing for a rich pattern of unobserved heterogeneity.

The standard techniques for estimating dynamic structural models have limited applicability due to difficulties associated with incorporating correlated unobserved states. In that sense, the applicability of our methodology goes beyond the estimation of default models. It can be used to estimate dynamic structural models in environments with both micro-level and aggregate data.

The proposed framework identifies the aggregate heterogeneity exploiting the aggregate variation of choices over time. We showed that the aggregate shocks are separately identified from their transition, as long as there is micro-level variation in the observed states. This result is important because it highlights the limitations of identification of dynamic models when only market-level information is available.

We applied the methodology to address the factors that determined the mortgage default rates in Colombia during the economic crisis that it faced during the late 1990's and the early years of the current decade. We showed that the policy of tying the variation of mortgage balances to the interest rate, instead of the inflation rate, was the cause of a substantial part (but presumably not all) of the observed defaults.

The use of dynamic structural model to study mortgage default highlights the often overlooked fact that default behavior does not only depend on the difference between home price and mortgage balance. As we showed, default depends also on the expected evolution of these variables, which affects the option value of defaulting in the future. For example, it is possible to design policies that increase the value of *not* defaulting, while keeping the current states (including balance) constant. The extent to which this is possible is an empirical issue which can only be addressed with the specific data and an empirical dynamic model, like ours.

References

Aguirregabiria, V., and P. Mira (2002): "Swapping the Nested Fixed Point Algorithm: A Class of Estimators for Discrete Markov Decision Models," *Econometrica*, 70(4), 1519–1543.

Altug, S., and R. A. Miller (1998): "The Effect of Work Experience on Female Wages and Labour Supply," *Review of Economic Studies*, 65(1), 45–85.

Carranza, J. E. (2007): "Product innovation and adoption in market equilibrium: The case of digital cameras," Unpublished manuscript, University of Wisconsin-Madison, Department of Economics.

Carranza, J. E., and D. Estrada (2007): "An empirical characterization of mortgage default in Colombia between 1997 and 2004," Unpublished manuscript, University of Wisconsin-Madison, Department of Economics.

Goolsbee, A., and A. Petrin (2004): "The Consumer Gains from Direct Broadcast Satellites and the Competition with Cable TV," *Econometrica*, 72(2), 351.

Gowrisankaran, G., and M. Rysman (2006): "Dynamics of Consumer Demand for New Durable Goods," wp, John M. Olin School of Business.

Heckman, J. J., and S. Navarro (2007): "Dynamic Discrete Choice and Dynamic Treatment Effects," *Journal of Econometrics*, 136(2), 341–396.

Hotz, V. J., and R. A. Miller (1993): "Conditional Choice Probabilities and the Estimation of Dynamic Models," *Review of Economic Studies*, 60(3), 497–529.

Keane, M. P., and K. I. Wolpin (1994): "The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence," *The Review of Economics and Statistics*, 76(4), 648–672.

Pakes, A. (1986): "Patents as Options: Some Estimates of the Value of Holding European Patent Stocks," *Econometrica*, 54(4), 755–784.

Rust, J. (1987): "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 55(5), 999–1033.

——— (1994): "Structural Estimation of Markov Decision Processes," in *Handbook of Econometrics, Volume*, ed. by R. Engle, and D. McFadden, pp. 3081–3143. North-Holland, New York.

WOLPIN, K. I. (1984): "An Estimable Dynamic Stochastic Model of Fertility and Child Mortality," *Journal of Political Economy*, 92(5), 852–874.

——— (1987): "Estimating a Structural Search Model: The Transition from School to Work," *Econometrica*, 55(4), 801–817.

Appendix

Proof of Lemma 2

Assume: (i) the aggregate shocks follow an autoregressive process such that $\xi_{t+1} = h(\xi_t) + \upsilon_{t+1}$, where υ is an *iid* error with cdf F_υ, such that $E_t[\xi_{t+1}|\xi_t] = h(\xi_t)$;(ii) $-1 < \frac{\partial h(\xi_t}{\partial \xi_t} < 1$; (iii) the sample size N is large. We need to show that for any parameters θ^0 such that the assumptions (i), (ii) and (iii) hold, equation (14) has a unique solution $\xi(\theta^0)$.

First, assume that $\theta = \theta^0$ and rewrite the mapping as follows:

$$\mathbf{s}_{1,t} \equiv s_{1,t}(S_{i,t};\theta^0,\xi_t) = \int Pr_{i,1,t}(S_{i,t};\theta^0,\xi_t)\frac{\tilde{Pr}_i}{\int \tilde{Pr}_i d\Phi} d\Phi dF_t(x) \tag{A1}$$

where $0 < \bar{s}_{1,t} < 1$ and $s_{1,t}$ are the observed and predicted proportions of individuals who choose $j = 1$ at time t, respectively. The integral is computed with respect to the distribution Φ_t, conditioned on the observed history. The expected continuation payoffs can be computed recursively starting at T, when $E_T V(S_{i,T+1}) = 0$. For $t < T$, $E_t V(S_{i,t+1}) = E_t \left[log(1 + e^{u(X_{i,t+1};\gamma)+\xi_{t+1}+\mu_i+\beta E_t[V_{i,1,t+2}(S_{i,t+2})|S_{i,t+2}]}) \right]$.

We prove existence and uniqueness by showing that under the given conditions the mapping $s_{1,t}(.,\xi_t)$ shown above is bounded by zero and one and is strictly monotone in ξ_t. The derivative of $s_{1,t}$ with respect to ξ_t is given by:

$$\begin{aligned}\frac{\partial s_{1,t}(.,\xi_t)}{\partial \xi_t} &= \int \left[Pr_{i,1,t}(S_{i,t})(1 - Pr_{i,1,t}(S_{i,t}))\left(1 + \beta \frac{\partial E_t[V(S_{i,t+1})|S_{i,t}]}{\partial \xi_{t+1}}\right)\right] \frac{\tilde{Pr}_i}{\int \tilde{Pr}_i d\Phi} d\Phi dF_t(x) \\ &+ \frac{1}{N_t}\sum_{i=1}^{N}\int \left[Pr_{i,1,t}(S_{i,t})\left((\kappa(S_{i,t}) - \int \kappa(S_{i,t})\frac{\tilde{Pr}_i}{\int \tilde{Pr}_i d\Phi}\right)\right] \frac{\tilde{Pr}_i}{\int \tilde{Pr}_i d\Phi} d\Phi dF_t(x)\end{aligned} \tag{A2}$$

whereas the derivative of $S_{1,t}$ with respect to $\xi_{t'}$ for $t \neq t'$ is:

$$\frac{\partial s_{1,t}(.,\xi_t)}{\partial \xi_t'} = \int \left[Pr_{i,1,t}(S_{i,t})\left((\kappa(S_{i,t'}) - \int \kappa(S_{i,t'})\frac{\tilde{Pr}_i}{\int \tilde{Pr}_i d\Phi}\right)\right] \frac{\tilde{Pr}_i}{\int \tilde{Pr}_i d\Phi} d\Phi dF_t(x) \tag{A3}$$

where the function $\kappa(.)$ is given by:

$$\kappa(S_{i,t}) = (-Pr_{i,1,t}(S_{i,t}))^{1-d_{i,t}}(1 - Pr_{i,1,t}(S_{i,t}))^{-d_{i,t}}\left(1 + \beta\frac{\partial E_t[V(S_{i,t+1})|S_{i,t}]}{\partial \xi_{t+1}}\right)$$

The first thing to notice is that the second term in (A2) and (A3) are the average of an expectation error. Therefore, as the sample size goes to infinity, these terms become zero.

Therefore, all we need to do to show that in large samples the mapping (A1) is monotone is show that the first term in (A2) is either positive or negative.

We will show now that the (A2) is always positive. Notice first that the derivative of the continuation payoffs with respect to ξ_t is given by:

$$\frac{\partial E_t V(S_{i,t+1})}{\partial \xi_t} = \int \left[Pr_{i,t+1} \frac{\partial h(\xi_t)}{\partial \xi_t} \left(1 + \beta \frac{\partial E_t V_{i,1,t+2}(S_{i,t+2})}{\partial \xi_t} \right) \right] dF_v \tag{A4}$$

for $t < T$. At $t = T$, this derivative is $\frac{\partial E_T V(S_{i,T+1})}{\partial \xi_T} = 0$.

Assumptions (i) and (ii) imply that $-1 < \frac{\partial E_t V(S_{i,t+1})}{\partial \xi_t} < 1$. To see this, start computing (A4) at $t = T - 1$ and then solve backwards. This, in turn implies that (A2) is strictly positive. Therefore, $s_{1,t}(., \xi_t)$ is strictly monotone (increasing) $\forall \xi_t$.

Another implication of $-1 < \frac{\partial E_t V(S_{i,t+1})}{\partial \xi_t} < 1$ is that as $\xi_t \to \infty$, in (A1) $s_{1,t} \to 1$. Conversely, as $\xi_t \to -\infty$, $s_{1,t} \to 0$, which completes the proof.

Proof of Proposition 1

The probability that a particular history $\{d_{i,1}, ..., d_{i,\bar{T}_i}\}$ is observed is given by (8):

$$\int \tilde{Pr}_i d\Phi = \int \prod_{t=1}^{\bar{T}_i} Pr_{i,1,t}^{d_{i,t}} (1 - Pr_{i,1,t})^{(1-d_{i,t})} d\Phi \tag{A5}$$

where, given Lemma 1 and Lemma 2, the vector $\xi(\gamma, \rho_\xi)$ is uniquely obtained from (14):

$$\mathbf{s}_{1,t} = \int Pr_{i,1,t}(.;\gamma, \rho_\xi, \xi_t(\gamma, \rho_\xi)) \frac{\tilde{Pr}_i(.;\gamma, \rho_\xi, \xi(\gamma, \rho_\xi))}{\int \tilde{Pr}_i(.;\gamma, \rho_\xi, \xi(\gamma, \rho_\xi)) d\Phi} d\Phi dF_t(x) \tag{A6}$$

The implicit function theorem implies that as ρ_ξ changes, ξ changes in (A6) according to the following derivative:

$$\frac{d\xi_t}{d\rho_\xi} = -\frac{\partial \tilde{s}_{1,t} / \partial \rho_\xi}{\partial \tilde{s}_{1,t} / \partial \xi} \tag{A7}$$

Given Lemma 2, this derivative is well defined, provided that its conditions are met.

Assume (i) that the preference parameter γ is known; and (ii) that for some $i, j \in N_t$ it is true that $X_i \neq X_j$. Assumption (ii) implies that for at least two agents $i, j \in N_t$, the predicted choice probabilities are different, $\int Pr_{i,1,t} d\Phi \neq \int Pr_{j,1,t} d\Phi$. Given (A5), (ii) also implies that for at least two agents $i, j \in N_t$, $(d \int Pr_{i,1,t} d\Phi / d\xi_t) \neq (d \int Pr_{j,1,t} d\Phi / d\xi_t)$.

A sufficient condition for the identification of ρ_ξ is that, for some $i \in N_t \forall t$, the derivative of the predicted probabilities with respect to ρ_ξ is different from zero:

$$\frac{d \int Pr_{i,1,t}(.; \gamma, \rho_\xi, \xi_t(\gamma, \rho_\xi)) d\Phi}{d\rho_\xi} \neq 0 \tag{A8}$$

In other words, we need to show that for at least one agent the predicted choice probability changes as ρ_ξ changes. We prove that this is true by contradiction. Suppose that for all $i \in N_t$, the derivative of the predicted choice probabilities with respect to ρ_ξ are zero. Using the chain rule and replacing (A7), we obtain:

$$\begin{aligned} \frac{d \int Pr_{i,1,t}(.) d\Phi}{d\rho_\xi} &= \frac{\partial \int Pr_{i,1,t}(.) d\Phi}{\partial \rho_\xi} + \frac{\partial \int Pr_{i,1,t}(.) d\Phi}{\partial \xi} \frac{d\xi}{d\rho_\xi} = 0 \\ &= \frac{\partial \int Pr_{i,1,t}(.) d\Phi}{\partial \rho_\xi} - \frac{\partial \int Pr_{i,1,t}(.) d\Phi}{\partial \xi} \left(\frac{\partial \tilde{s}_{1,t} / \partial \rho_\xi}{\partial \tilde{s}_{1,t} / \partial \xi} \right) = 0 \end{aligned} \tag{A9}$$

which would imply that for all $i \in N_t$:

$$\frac{\partial \int Pr_{i,1,t}(.) d\Phi / \partial \rho_\xi}{\partial \int Pr_{i,1,t}(.) d\Phi / \partial \xi} = \frac{\partial \tilde{s}_{1,t} / \partial \rho_\xi}{\partial \tilde{s}_{1,t} / \partial \xi} \tag{A10}$$

But this is impossible because we have already argued that (A5), (ii) imply that for at least two agents $i, j \in N_t$, $(dPr_{i,1,t}/d\xi_t) \neq (dPr_{j,1,t}/d\xi_t)$. Therefore (A10) is false and the proposition is proved.

Table 1: Summary statistics (main data set)

(1) Quarter	(2) Number of loans	(3) Outstanding loans	(4) Default rate	(5) Mean Price 1	(6) Mean Price 2	(7) Price/ Balance
1997 : 1	93	93	0.00 %	167.98	167.9828	53.23 %
1997 : 2	355	351	1.14 %	85.69	85.3543	47.17 %
1997 : 3	591	575	2.09 %	87.28	86.3226	47.25 %
1997 : 4	925	892	1.91 %	85.12	84.0201	46.01 %
1998 : 1	925	856	4.21 %	91.18	88.4451	44.96 %
1998 : 2	925	831	3.01 %	95.70	91.6633	43.60 %
1998 : 3	925	810	2.59 %	95.11	90.2188	45.65 %
1998 : 4	925	788	2.79 %	95.70	89.3045	47.57 %
1999 : 1	925	750	5.07 %	100.14	91.3159	48.65 %
1999 : 2	925	704	6.53 %	94.14	91.0233	49.66 %
1999 : 3	925	680	3.53 %	77.67	85.8669	51.58 %
1999 : 4	925	634	7.26 %	61.55	92.029	48.44 %
2000 : 1	925	598	6.02 %	59.76	87.9514	49.14 %
2000 : 2	925	586	2.05 %	65.43	96.0334	43.04 %
2000 : 3	925	555	5.59 %	58.92	94.8815	44.00 %
2000 : 4	925	539	2.97 %	59.74	95.4666	42.74 %

Continues in next page

Prices and balances are in 1997 COL$

Mean Price 1 and Mean Price 2 are computed over outstanding and all loans, respectively.

Table 1, continued

(1)	(2)	(3)	(4)	(5)	(6)	(7)
2001 : 1	925	526	2.47 %	67.15	107.0776	37.51 %
2001 : 2	925	513	2.53 %	61.34	97.2037	42.04 %
2001 : 3	925	502	2.19 %	66.04	104.0606	39.06 %
2001 : 4	925	491	2.24 %	69.75	108.6502	36.62 %
2002 : 1	925	489	0.41 %	63.46	98.703	39.29 %
2002 : 2	925	483	1.24 %	71.43	110.7895	34.48 %
2002 : 3	925	473	2.11 %	66.99	103.5303	35.78 %
2002 : 4	925	462	2.38 %	76.25	117.7744	30.71 %
2003 : 1	925	456	1.32 %	70.26	108.3027	32.00 %
2003 : 2	925	453	0.66 %	73.77	113.6786	30.25 %
2003 : 3	925	450	0.67 %	72.92	112.0695	29.47 %
2003 : 4	925	448	0.45 %	73.87	114.9951	27.67 %
2004 : 1	925	444	0.90 %	72.45	113.0203	27.33 %
2004 : 2	925	439	1.14 %	80.93	125.7102	23.91 %

Prices and balances are in 1997 COL$

Mean Price 1 and Mean Price 2 are computed over outstanding and all loans, respectively.

Table 2: Estimation results: Duration Models

	Model I		Model II	
Coefficient	Est. (s.e.)	Marginal effect (s.e.)	Est. (s.e.)	Marginal effect (s.e.)
γ_1 (Price)	0.072 (0.016)	0.004 (0.001)	0.073 (0.013)	0.004 (0.001)
γ_2 (Balance)	-0.185 (0.028)	-0.005 (0.001)	-0.174 (0.025)	-0.005 (0.001)
γ_3 (Term)	-0.016 (0.005)	-0.002 (0.001)	-0.016 (0.002)	-0.002 (0.000)
γ_4 (Income)			-0.001 (0.000)	-0.001 (0.000)
	Model III		Model IV	
Coefficient	Est. (s.e.)	Marginal effect (s.e.)	Est. (s.e.)	Marginal effect (s.e.)
γ_1 (Price)	0.120 (0.043)	0.008 (0.002)	0.126 (0.045)	0.008 (0.003)
γ_2 (Balance)	-0.422 (0.133)	-0.012 (0.003)	-0.417 (0.136)	-0.012 (0.004)
γ_3 (Term)	-0.023 (0.011)	-0.003 (0.001)	-0.025 (0.012)	-0.003 (0.002)
γ_4 (Income)			-0.001 (0.001)	-0.001 (0.001)
α_0	0.483 (0.007)		0.483 (0.007)	
α_1	-0.004 (0.005)		-0.003 (0.005)	
α_2	0.036 (0.003)		0.036 (0.003)	
$var(\mu)$	5.800		5.750	

In models I and II $\mu_i = 0$; all models include aggregate shocks (not shown)

Table 3: Estimation results: Dynamic Models

	Model V		Model VI	
Coefficient	Est. (s.e.)	Marginal effect (s.e.)	Est. (s.e.)	Marginal effect (s.e.)
γ_1 (Price)	0.008 (0.003)	0.003 (0.001)	0.008 (0.002)	0.003 (0.001)
γ_2 (Balance)	-0.049 (0.019)	-0.004 (0.001)	-0.047 (0.013)	-0.004 (0.002)
γ_3 (Term)	-0.003 (0.001)	-0.002 (0.000)	-0.003 (0.001)	-0.002 (0.001)
γ_4 (Income)			0.000 (0.000)	0.000 (0.000)
ρ_0^ξ	0.005 (0.156)		0.034 (0.069)	
ρ_1^ξ	-0.429 (3.693)		-0.389 (0.632)	
ρ_2^ξ	0.000 (0.004)		0.000 (0.000)	

	Model VII		Model VIII	
Coefficient	Est. (s.e.)	Marginal effect (s.e.)	Est. (s.e.)	Marginal effect (s.e.)
γ_1 (Price)	0.094 (0.029)	0.012 (0.006)	0.092 (0.029)	0.012 (0.006)
γ_2 (Balance)	-0.478 (0.099)	-0.016 (0.006)	-0.463 (0.087)	-0.015 (0.006)
γ_3 (Term)	-0.021 (0.009)	-0.004 (0.002)	-0.021 (0.008)	-0.004 (0.002)
γ_4 (Income)			0.000 (0.001)	0.000 (0.000)
ρ_0^ξ	-2.271 (1.938)		-2.271 (2.314)	
ρ_1^ξ	-0.480 (1.146)		-0.506 (1.525)	
ρ_2^ξ	0.238 (0.978)		0.237 (1.428)	
α_0	0.483 (0.007)		0.483 (0.007)	
α_1	-0.001 (0.003)		-0.002 (0.004)	
α_2	0.036 (0.003)		0.036 (0.003)	
$var(\mu)$	3.030		3.006	

In models V and VI $\mu_i = 0$

Figure 1: Simulated and counterfactual default