Semiparametric Estimation of a Finite Horizon Dynamic Discrete Choice Model with Application to Subprime Mortgage Default

Patrick Bajari, University of Washington, University of Minnesota and NBER
Sean Chu, Federal Reserve Board of Governors
Denis Nekipelov, University of California, Berkeley
Minjung Park, University of California, Berkeley
This version: April 11, 2013

Abstract

In this paper we develop and implement a computationally tractable semiparametric estimator for finite horizon dynamic discrete decision models with a terminal action. Our estimator is a two-step estimation method that does not rely on backward induction. Our methodology can be implemented using standard statistical packages without the need to write specialized computational routines, as it involves linear (or nonlinear) projections only and does not require simulation. Our methodology also offers a new approach to estimation of agents’ time preference parameter. We implement our estimator using a concrete empirical example of the optimal default behavior of subprime mortgage borrowers.

Keywords: Finite Horizon Optimal Stopping Problem, Time Preferences, Semiparametric Estimation, Inference with Estimated Regressors, Subprime Mortgage Default

JEL Classifications: C5, C14, G21

---

1We would like to thank seminar participants at Stanford GSB, Chicago Booth, Olin Business School, Berkeley ARE and IO fest for valuable comments. The views expressed are those of the authors and do not necessarily reflect the official positions of the Federal Reserve System. Correspondence: bajari@uw.edu; Sean.Chu@frb.gov; nekipelov@econ.berkeley.edu; mpark@haas.berkeley.edu.
1 Introduction

Dynamic discrete choice problems have traditionally been considered to be easier to study empirically when they have a finite horizon, as the value function can be found using backward induction. It is easy to characterize the choice probabilities in the final period, in which the agent’s decision problem becomes static, and the choice probabilities for all previous periods can be constructed iteratively by taking the value function one period ahead as given. However, this approach has two practical limitations. First, implementing the iterative procedure may require a prohibitively large amount of computer memory to store the value functions for all periods, which is especially problematic when the dimensionality of the state space is large and the decision horizon is long. Second, the iterative approach relies on being able to observe decisions in the final period. Such data are not available in many realistic settings. For example, panel data based on surveys that follow individuals over time may suffer from censoring if individuals drop out of the sample. A similar problem exists for program evaluation when the treatment has only been introduced recently. In such cases, the resulting estimator could be sensitive to assumptions regarding primitives of the final periods, which is undesirable because researchers cannot verify the plausibility of assumptions pertaining to time periods that are unobserved.

In this paper, we develop a novel multi-step procedure for estimating finite-horizon dynamic discrete choice problems in which one of the possible choices for an agent is a terminal action. Our estimation procedure does not rely on backward induction and provides consistent estimates for the structural parameters even if the available data do not extend to the last periods of the optimal decision problem. The procedure is conceptually different from multi-step estimation procedures that have been proposed for infinite horizon problems, such as Bajari, Benkard and Levin (2009) and Pesendorfer and Schmidt-Dengler (2010), as we must take into account the nonstationarity of agents’ optimal behavior due to the presence of a final period. Our method exploits Hotz and Miller’s (1993) intuition that, when there is a terminal action, one can represent the value function as a function of the one-period-ahead choice probabilities. We build on that intuition by proposing a nonparametric estimator of economic agents’ expectations about one-period-ahead value function and using the estimates to recover agents’ preferences within a regression framework. The resulting estimator involves linear (or nonlinear) projections only and does not require simulation. It is thus very easy to implement our estimation method using standard statistical software such as STATA, requiring a minimal computational burden. Our estimator is also attractive from an economic perspective, as we do not assume that agents form precise long-term forecasts for the transitions of the state variables. Rather, we only require that the agents have stable expectations of the transition process from the present state to the state one period ahead, and that they have stable
preferences over choices.

Our approach of nonparametrically estimating agents’ one-period-ahead expectations and using these estimates to recover their preferences is closely related to Ahn and Manski (1993), and can thus be regarded as a conceptual extension to dynamic settings of work by Manski (1991, 1993, and 2000), who examines the responses of economic agents to their expectations in models with uncertainty or endogenous social effects.

Furthermore, we demonstrate that our methodology offers a new approach to estimation of the time preference parameter that does not rely on exclusion restrictions, identification at infinity, or having observations from the final periods. We prove identification of the discount factor, which, unlike the prior literature on identification of time preferences (Magnac and Thesmar (2002); Fang and Wang (2012)), does not require the presence of a variable that affects the state transition but not the per-period utility, because in finite-horizon models, the time to maturity itself plays a role analogous to that of such variables.²

In order to demonstrate our estimation method in an empirically relevant setting, we examine the default behavior of subprime mortgage borrowers. Our empirical model has four key elements. First, the model is dynamic: borrowers are forward-looking and respond to both current and expected future shocks by adjusting their current behavior. Second, borrowers’ choice set is discrete—in each period a borrower chooses whether to default on a loan, prepay (or refinance) the loan, or make just the regularly scheduled monthly mortgage payment. We consider default to be a terminal action that ends the dynamic problem, with the borrower receiving a final, one-time “compensation” (or rather, utility loss) and no future utility flows. Third, the borrower’s decision problem has a finite horizon, reflecting the fact that mortgages have a fixed maturity—commonly 30 years. Fourth, our model is a single-agent model that abstracts from potential interactions among borrowers. Since our estimation method is designed for single-agent, finite-horizon dynamic discrete choice problems with a terminal action, these features make the empirical application perfectly suited for illustration of our estimation method. Furthermore, these features provide a realistic and tractable framework for borrowers’ default decisions.

Our empirical analysis employs panel data on a sample of subprime and “Alt-A” mortgage borrowers whose loans were securitized between 2000 and 2007.³ The unit of observation is an individual mortgage observed at a point in time. The data contain detailed borrower-level information from the loan

---

²Most of the prior literature that estimates time preferences relies on experimental data (see Frederick, Loewenstein and O’Donohue (2002) for a review of the literature). A few notable exceptions are Hausman (1979), Yao et al. (2012), Chung, Steenburgh, and Sudhir (2011), and Fang and Wang (2012).

³Alt-A’s are mortgages that are considered riskier than prime but less risky than subprime. In this paper, we casually use the term “subprime market” to refer to both subprime and Alt-A mortgages. The distinction between the two is in any case somewhat artificial.
application, including the terms of the contract, the appraised value of the property, the loan-to-value (LTV) ratio, the level of documentation, and the borrower’s credit score at the time of origination. We also observe the month-by-month stream of payments made by the borrower and whether the mortgage goes into default or is prepaid. To track movements in home prices, we merge the mortgage data with zip-code level home price indices.

Our structural estimates aid in understanding the relative importance of different theoretical drivers of borrower default, as well as the potential effects of several broad classes of policy interventions for mitigating default. Because the actual policy measures for mitigating default that have been introduced since the financial crisis have been complex and have contained many different programmatic components, we study an array of stylized interventions. We also use the structural estimates to determine the welfare consequences for borrowers under various counterfactuals.

Among other empirical findings, we obtain an estimate of 1.9% for the monthly discount rate for subprime borrowers. Our results also show that a uniform 10% reduction in outstanding mortgage balance for the pool of borrowers in our sample would reduce the overall default probability by 22%, and that borrowers’ average willingness to pay for the principal writedown would be $16,643.

The rest of this paper proceeds as follows. In Section 2, we present our model and discuss identification of the model primitives, including the discount factor. Section 3 discusses our estimation methodology and its sampling properties. In Section 4, we describe the data. Section 5 presents estimation results. In Section 6, we discuss our counterfactual exercises and welfare implications for borrowers. Section 7 concludes.

2 Model

In this section, we set up a single-agent, finite-horizon dynamic discrete choice model, in which one of the agent’s possible actions is terminal. For clarity of exposition, we write down a specific model of mortgage borrowers’ default decisions and prove identification results within the context of that model. However, our identification results hold more generally and do not rely on any specific feature of the considered model.

In our model of subprime mortgage borrowers, each agent enters a mortgage contract lasting $T$ time periods, where $T$ is common across all agents. The components of the model are as follows.
2.1 Actions

At each time period $t$ over the life of borrower $i$’s loan, the borrower chooses an action $a_{i,t}$ from the finite set $A = \{0, 1, 2\}$. The possible actions in $A$ are to default ($a_{i,t} = 0$), prepay the mortgage ($a_{i,t} = 1$), or make just the regularly scheduled monthly payment, which we refer to as “paying” ($a_{i,t} = 2$). We assume that there is no interaction among borrowers affecting their payoffs, so our setup is a single-agent model, not a game. We assume that default is a terminal action: once a borrower defaults, there is no further decision to be made and no further flow of utility starting from the next period.

2.2 Period Utility and State Transition

Each borrower observes a vector of state variables $s_{i,t} \in S$ in each period. The support $S$ is a product space that is a subset of $k$-dimensional Euclidean space. We allow the subspaces of this product space to be either continuous or discrete. The state vector $s_{i,t}$ includes borrower $i$’s characteristics, the current home value, monthly payments, etc. $s_{i,t}$ is fully observed by the econometrician. We assume that the borrower is also characterized by a time-invariant “type” $c_i \in C$ (observed by the econometrician) and a time-dependent vector of idiosyncratic shocks associated with each action $\varepsilon_{i,t} = (\varepsilon_{i,0,t}, \varepsilon_{i,1,t}, \varepsilon_{i,2,t})$ (unobserved by the econometrician). The set $C$ is assumed to be finite. Although certain elements of $s_{i,t}$ may also be time-invariant, the purpose of defining a separate type space $C$ will become apparent in Section 3, where we discuss our utility specification with random coefficients. Each element of $\varepsilon_{i,t}$ is assumed to have a continuous support on the real line. We make the following assumption regarding the marginal distributions of the random variables.

ASSUMPTION 1

(i) **Conditional independence of idiosyncratic payoff shocks:** $s_{i,t} \perp \varepsilon_{i,t} \mid c_i$.

(ii) **Conditional independence over time of idiosyncratic payoff shocks:** $\varepsilon_{i,t} \mid \varepsilon_{i,t-1}, a_{i,t-1}, c_i \sim \varepsilon_{i,t} \mid a_{i,t-1}, c_i$.

(iii) **Exclusion restriction** ($c_i$ cannot be represented as a linear combination of the elements of $s_{i,t}$): $C$ does not belong to any proper linear subspace of $S$.

(iv) **Markov transition of state variables:** $s_{i,t}$ follows a reversible Markov process, conditional on $a_{i,t}$.

In our empirical analysis we will use a conventional specification of the distribution of the idiosyncratic shocks, assuming that components of $\varepsilon_{i,t}$ are mutually independent, have a type I extreme value distribution, and are i.i.d. across borrowers and over time. However, this assumption is not essential,

---

4Note that we use $t$ to denote the loan’s age, not calendar time. Since $T$ is assumed to be common across all agents, loans with the same $t$ have the same number of months remaining until maturity.
and we establish the existence of an optimal strategy and prove our identification results for an arbitrary continuous distribution of random shocks that satisfy Assumption 1.

The transition of some of the state variables may be influenced by the current action taken, \( a_{i,t} \). We also allow the state variables potentially to follow a high-order Markov process by including the lagged state variables from the previous periods in the vector \( s_{i,t} \). This structure allows for greater realism, as certain important state variables may exhibit lag dependence. For instance, in our empirical application, the evolution of housing prices depends on the prices in previous periods.

We assume that the per-period utility of the borrower is separable in the idiosyncratic shock component. We can characterize the borrower’s utility as:

\[
U(a_{i,t}, s_{i,t}; c_i) = u(a_{i,t}, s_{i,t}; c_i) + \varepsilon_{i,a_{i,t},t}, \quad \text{for } t < T,
\]

\[
U(a_{i,t}, s_{i,t}; c_i) = u_T(a_{i,t}, s_{i,t}; c_i) + \varepsilon_{i,a_{i,t},t}, \quad \text{for } t = T.
\]

As specified, the per-period utility has a deterministic component, \( u(\cdot; \cdot; \cdot) \), which is a time-invariant function of the action, state, and the borrower’s type. The payoff function in the final period can in general differ from that of earlier periods. For example, in our empirical application, the borrower obtains full ownership of the house once the mortgage is fully paid off at maturity, which we can think of as adding a lump-sum boost to the period utility in the final period.

### 2.3 Decision Rule and Value Function

We consider the borrower’s problem as an optimal stopping problem, and assume that the default decision is irreversible and that the borrower cannot “re-start” borrowing after default. This assumption is realistic because default usually damages a borrower’s credit to such an extent that borrowing for another house is impossible for a long time. Provided that the default (“stopping”) decision is irreversible, the choice of the default option is equivalent to taking a one-time “compensation” (more specifically, a utility loss) without future utility flows. If the borrower pays or refinances his mortgage, he receives the corresponding period payoff (which is a combination of utility from consumption of housing services and disutility from payments for the mortgage) plus the expected discounted stream of future utility. Parameter \( \beta \) is the one-period-ahead discount factor that characterizes the time impatience of the borrower.

The borrower’s decision rule \( D_t \) for each period \( t \) is a mapping from the vector of payoff-relevant variables into actions, \( D_t : S \times \mathbb{R} \to A \). We denote the borrower’s decision probabilities by \( \sigma_t(\cdot|s_{i,t}, c_i) = E[\{D_t(s_{i,t}, c_i, s_{i,t}) = k\} | s_{i,t}, c_i] \) for \( k \in A \). We collect \( \sigma_t(\cdot|s_{i,t}, c_i) \) for all \( k \) and \( t \) such that \( \sigma_t(s_{i,t}, c_i) = [\sigma_t(k = 0|s_{i,t}, c_i), \sigma_t(k = 1|s_{i,t}, c_i), \sigma_t(k = 2|s_{i,t}, c_i)] \) and \( \sigma = (\sigma_1(s_{i,1}, c_i), \ldots, \sigma_T(s_{i,T}, c_i)) \), and refer to \( \sigma \)
as the policy function.

Considering the expected discounted sum of utility of the borrower who has not defaulted prior to period \( t < T \), we introduce the ex ante value function:

\[
V_t(s_{i,t}; c_i) = E_{\sigma,f(s)} \left[ \sum_{\tau = t}^{T} \left( \beta^{T-\tau} U(a_{i,\tau}, s_{i,\tau}; c_i) \prod_{\tau_1=1}^{\tau-1} 1(a_{i,\tau_1} > 0) \right) | s_{i,t}, c_i \right],
\]

where \( f(s) \) represents the state transitions. The term \( \prod_{\tau_1=1}^{\tau-1} 1(a_{i,\tau_1} > 0) \) reflects that once a borrower defaults, there is no further flow of utility starting from the next period.

The choice-specific value function, denoted by \( V_t(a_{i,t} = k; s_{i,t}; c_i) \), corresponds to the deterministic component of the discounted sum of payoffs that the borrower receives when choosing action \( k \) in period \( t \):

\[
V_t(a_{i,t} = k; s_{i,t}; c_i) = u(a_{i,t} = k; s_{i,t}; c_i) + \beta E[V_{t+1}(s_{i,t+1}; c_i) | s_{i,t}, c_i, a_{i,t} = k] \text{ for } t < T,
\]

\[
V_t(a_{i,t} = k; s_{i,t}; c_i) = u_T(a_{i,t} = k; s_{i,t}; c_i) \text{ for } t = T.
\]

In particular, the choice-specific value of default is equal to the period utility of default, i.e., \( V_t(a_{i,t} = 0; s_{i,t}; c_i) = u(a_{i,t} = 0; s_{i,t}; c_i) \) for \( t < T \), because default is a terminal action, which makes the future value term \( E[V_{t+1}(s_{i,t+1}; c_i) | s_{i,t}, c_i, a_{i,t} = 0] \) zero.

### 2.4 Optimal Policy Functions

The following theorem establishes a formal existence and uniqueness result characterizing the borrower’s optimal decision.

**Theorem 1** Under Assumption 1 there exists a unique decision rule \( D_t^*(s_{i,t}; c_i, \varepsilon_{i,t}) \) supported on \( A \) for \( t = 1, 2, \ldots, T \) that solves the maximization problem

\[
\sup_{(D_1, D_2, \ldots, D_T) \in A^T} V_1(s_{i,1}; c_i).
\]

**Proof.** Our argument uses backward induction. In the last period (at mortgage maturity) the borrower faces a static optimization problem of choosing among \( V_T(0, s_{i,T}; c_i) + \varepsilon_{i,0,T}, V_T(1, s_{i,T}; c_i) + \varepsilon_{i,1,T}, \) and \( V_T(2, s_{i,T}; c_i) + \varepsilon_{i,2,T} \). The optimal decision delivers the highest payoff, yielding the decision rule \( D_T^*(s_{i,T}; c_i, \varepsilon_{i,T}) = \arg \max_{k \in A} \{ V_T(k, s_{i,T}; c_i) + \varepsilon_{i,k,T} \} \). Provided that the payoff shocks are idiosyncratic and have a continuous distribution, the optimal choice probabilities are characterized by continuous
functions of \((V_T(k, s_{i,T}; c_i), k \in A)\). Knowing the optimal decision rule in period \(T\), we can obtain the choice-specific value function in period \(T - 1\) as

\[
V_{T-1}(k, s_{i,T-1}; c_i) = u(k, s_{i,T-1}; c_i) + \beta \mathbb{E}\left[\sum_{k' \in A} 1\{D_T = k'\} \left(V_T(k', s_{i,T}; c_i) + \varepsilon_{i,k',T}\right) \mid s_{i,T-1}, c_i, a_{i:T-1} = k\right].
\]

Provided that the \(T^{th}\) period optimal decision has already been derived, the optimal decision problem in \(T - 1\) becomes a static choice among three alternatives. Its solution, again, trivially exists and is (almost surely) unique because the distribution of \(\varepsilon_{i,T-1}\) is continuous. We iterate this procedure back to \(t = 1\).

If we specify that the distribution of idiosyncratic shocks has the standard type I extreme value distribution, then the probabilities of default, prepayment, and payment—for a given state, borrower type, and period \(t \leq T\)—can be expressed in closed form. These expressions have the standard multinomial logit form and are in terms of the choice-specific value functions. Using the Hotz-Miller inversion, we can also obtain explicit expressions for the differences between the choice-specific values for different actions, which are in terms of (the logarithms of) the optimal choice probabilities. In our model, we can identify the levels of the choice-specific values themselves, and not just the differences, because default is a terminal action, pegging the choice-specific value of default to a fixed function \(u(a_{i,t} = 0, s_{i,t}; c_i)\). The easiest example through which to illustrate this is when the payoff from the default option is trivially normalized to zero. In this case we can recover the choice-specific value functions and the ex ante value function directly from the data. As we show below, our model is identified from the data as long as the payoff from default is set to a fixed function.

### 2.5 Semiparametric Identification

In this section we demonstrate that our model is identified from objects observed in the data, namely, the choice probability of each option, conditional on the current state and the borrower’s observable type \((P_t(a_{i,t} = k \mid s_{i,t}, c_i));\) and the transition distribution for the state variables, characterized by the conditional cdf \(F(s_{i,t} \mid s_{i,t-1}, c_i, a_{i,t-1})\).

The model’s three structural elements are: (1) the deterministic component of the per-period payoff function, \(u(\cdot, \cdot, \cdot)^5\); (2) the time preference parameter \(\beta\); and (3) the conditional distribution of the idiosyncratic utility shocks to the borrowers, which have the type-specific joint cdf \(F_{\varepsilon}(\cdot | c)\). We shall argue that \(u(\cdot, \cdot, \cdot)\) is nonparametrically identified and that the time preference parameter \(\beta\) is identified,

---

5We omit discussion regarding identification of \(u_T(\cdot, \cdot, \cdot)\), as it is obvious that \(u_T(\cdot, \cdot, \cdot)\), in case \(u_T \neq u\), is identified if and only if decisions from the final period are observed.
for a given distribution of the idiosyncratic payoff shocks satisfying Assumption 1. We emphasize that our identification results do not rely on the extreme value assumption for the distribution of the idiosyncratic shocks.

We show the model is identified by demonstrating that there exists a unique mapping from the observable distribution of the data to the structural parameters. We start with the case in which the payoff from the default option is known, and then consider relaxing this assumption.

Theorem 2 (Identification with known default utility) Suppose that the payoff from the default option is a known function $u(0, \cdot, \cdot)$, and that the distribution of idiosyncratic shocks conditional on the borrower-specific heterogeneity variables $c, F_z(\cdot|c)$, has a full support with the density strictly positive on $\mathbb{R}^3$. Also, suppose that for at least two consecutive periods $t$ and $t'$, $\sigma_{k,t}(\cdot; \cdot) \neq \sigma_{k,t'}(\cdot; \cdot)$ for $k \in A$.

(i) If the data distribution contains information on at least two consecutive periods and the discount factor $\beta$ is fixed, the per-period utility $u(\cdot; \cdot; \cdot)$ is nonparametrically identified. Moreover, if $u(\cdot; \cdot; \cdot) = u_T(\cdot; \cdot; \cdot)$ and one of the observed periods is the period of mortgage maturity $T$, then the discount factor is also identified.

(ii) If the data distribution contains information on at least three consecutive periods, then both the discount factor and the per-period utility function $u(\cdot; \cdot; \cdot)$ are identified.

This theorem, proved in the Appendix, establishes the general result that the considered model is identified (including identification of the time preference parameter) if the payoff from default is a known function. The argument requires observing two consecutive time periods over which there is variation in the optimal decision probability, conditional on the state variables and the type. The theoretical justification for why two such periods exist stems from the finite horizon: Borrowers’ tendency to default should depend upon the time remaining until the mortgage maturity date. More generally, the optimal decision rules will depend on time $t$ in finite-horizon models, even conditional on the state variables, satisfying the assumption of the proposition. Unlike the prior literature on identification of time preferences (Magnac and Thesmar (2002); Fang and Wang (2012)), our identification of the discount factor does not require the availability of variables that affect the state transition but are excluded from the expressions for the per-period utility—rather, the presence of a finite horizon implies that the time to maturity itself plays a role analogous to that of such variables.

For cases in which the default utility is not a known function, we would need to normalize it. In the empirical literature on dynamic discrete optimization problems, it has been noted (e.g., see Bajari, Hong and Nekipelov (2012)) that different normalizations of the per-period utility are not innocuous and
can have an impact on the structural parameter estimates. Moreover, we show that in the finite-horizon optimal decision problem, the structural model is, under certain conditions, overidentified for a given normalization, such that no normalization is necessary. These two insights can be used to explore the identification of the elements of the structural model, including the payoff from default. We formally show that in the finite-horizon optimal stopping problem, the normalization of the per-period payoff from the default choice to zero is not innocuous. Moreover, we show that under stronger requirements on the data, one can identify the payoff from the default option without the need to normalize any payoff.

Theorem 3 (Identification with normalized default utility) Suppose that the distribution of idiosyncratic shocks conditional on the borrower-specific heterogeneity variables $c$, $F_x(\cdot | c)$, has a full support with the density strictly positive on $\mathbb{R}^3$. Also, suppose that for at least two consecutive periods $t$ and $t'$, $\sigma_{k,t}(\cdot; \cdot) \neq \sigma_{k,t'}(\cdot; \cdot)$ for $k \in A$.

(i) If $u(\cdot, \cdot; \cdot) \neq u_T(\cdot, \cdot; \cdot)$ or the choices of the borrowers in the final period are not observed, the default utility $u(0, s; c)$ is not identified. If in this case $u(0, s; c)$ is normalized to a fixed function, the recovered discount factor does not depend on the choice of normalization for the default utility. However, the recovered differences between the per-period payoffs from payment or prepayment and the per-period payoff from default depend upon the choice of normalization for the default utility.

(ii) Suppose that $u(\cdot, \cdot; \cdot) = u_T(\cdot, \cdot; \cdot)$ and that the choices of the borrowers in the period of mortgage maturity $T$ are observed along with the choices from earlier periods. Then, the utilities from all choices, $u(0, s; c)$, $u(1, s; c)$, and $u(2, s; c)$, are identified along with the discount factor $\beta$.

This theorem, proved in the Appendix, has a clear interpretation. In the last period, the decision of the borrower is static and thus there is no option of “delayed default.” As a result, the last-period decision depends only on the differences between the utilities from the payment and prepayment options and the utility from the default option. However, in any period before the last, the borrower has an option of defaulting in the following period if he pays or prepaies in the current period, but not if he defaults. This asymmetry implies that the normalization has a disproportional effect on the discounted payoffs from different options. Part (i) of the theorem holds because, while the utility from default in the current period is shifted by the normalization (as the future discounted payoff is zero), the payoffs from payment or prepayment are additionally shifted by the amount equal to the discounted expected payoff from defaulting in the next period. Nevertheless, $\beta$ is invariant to the normalization because the tradeoff between current payoffs and future option values is unaffected by the normalization. Part (ii) of

---

6We are grateful to Günter Hitsch, who encouraged us to present the formal argument supporting this statement.
the theorem implies that observing the final-period choices would allow us to pin down the actual levels of the payoffs and not just the differences in level, provided the final-period utilities are the same as in the previous periods.

In our empirical application, we do not have data on the behavior of the borrowers whose mortgages are close to maturity, simply because subprime mortgages were only introduced in recent years. Such truncation of data is common in other empirical settings as well. Furthermore, the per-period payoff function in the final period may differ from the per-period payoff function of earlier periods in our empirical setup. Thus, in order to identify the model using our data, we must normalize the utility from default. We shall discuss our choice of normalization in more detail when we discuss the estimation results.

To illustrate the logic behind the identification of the discount factor, we derive closed-form expressions for components of the model after normalizing the default utility and making the type I extreme value assumption for the idiosyncratic shocks. In this case, the ex ante value function takes the form

$$V_t(s_i,t; c_i) = \log \left( \sum_{k=0}^{2} \exp \left( V_t(a_i,t = k, s_i,t; c_i) \right) \right) = \log \left( \frac{1}{\sigma_t(0, s_i,t; c_i)} \right) + u(0, s_i,t; c_i),$$

where $u(0, s; c)$ is the normalized utility from default. We can then combine the expression for the ex ante value function with the expression for the Bellman equation for each of the non-default choices to obtain the following system of equations

$$\log \left( \frac{\sigma_t(k, s_i,t; c_i)}{\sigma_t(0, s_i,t; c_i)} \right) = u(k, s_i,t; c_i) - u(0, s_i,t; c_i) + \beta E \left[ \log \left( \frac{1}{\sigma_{t+1}(0, s_{i,t+1}; c_i)} \right) + u(0, s_{i,t+1}; c_i) | s_i,t, c_i, a_{i,t} = k \right], \quad k = 1, 2.$$ (2)

This system of equations can be written for each instant in time $t$. In particular, if the data contain at least three consecutive periods, the following system of equations can be used to find the discount factor:

$$\beta = \frac{\log(\sigma_{t+1}(k, s_{i,t+1}; c_i) \sigma_t(0, s_i,t; c_i))}{E [\log(\sigma_{t+1}(0, s_{i,t+1}; c_i)) | s_{i,t} = s_i, c_i, a_{i,t} = k] - E [\log(\sigma_{t+2}(0, s_{i,t+2}; c_i)) | s_{i,t+1} = s_i, c_i, a_{i,t+1} = k]}, \quad k = 1, 2.$$ (2)

By the assumption of Theorem 1, the denominator of this expression is not equal to zero. Thus, the discount factor is identified.

One can also see the logic behind the identification of the discount factor using expression (2). Because the discount factor $\beta$ is the parameter attached to the expected one-period-ahead value function conditional on choosing action $k$ in $t$, $\beta$ can be identified if there is variation in the one-period-ahead value function that is uncorrelated with variation in the per-period utility function $u(\cdot, \cdot; \cdot)$. In finite-horizon
models, such variation exists with respect to \( t \), because the per-period utility function is time-invariant by assumption, whereas the optimal choice probability \( \sigma \) depends on \( t \).

3 Econometric Methodology

3.1 Semiparametric Estimator

Our specification of borrowers' per-period payoffs is a version of the random coefficients model, where the distribution of coefficients characterizes the borrower-level heterogeneity \( c \). For notational simplicity, from now on we drop the index \( i \) for borrowers, except where necessary for disambiguation. The per-period utility is parameterized by the random coefficients, and is defined as

\[
u(a, s; c) = u(s; \theta(a, c)),\]

where \( a \in A \) and \( \theta : A \times C \mapsto \Theta \), with \( \Theta \) denoting the parameter space. We allow the utility to be nonparametric, and the coefficient vector \( \theta(a, c) \) may be considered the vector of coefficients for the sieve representation of the per-period payoff function. Such a representation of the per-period utility gives us the flexibility to choose either a parametric or a fully nonparametric specification for the utility associated with each realization of the state variables, action, and borrower type. It also places our model in the class of dynamic discrete choice models in which unobserved heterogeneity is modeled using mixture distributions (e.g., Kasahara and Shimotsu (2009) and Arcidiacono and Miller (2011)).

To estimate the model, we use a plug-in semiparametric estimator. Our proposed estimator does not rely on backward induction, and provides consistent estimates even when data on the final periods are not available. Parallel to our identification argument, we nonparametrically estimate the borrowers’ policy functions and use them to recover the choice-specific value functions of the borrowers. We then use the latter to recover the distribution of random coefficients in the utility function and the time preference parameter. Below, we first characterize the general form of the estimator corresponding to an arbitrary distribution of idiosyncratic payoff shocks satisfying Assumption 1. Then, we discuss our specific implementation, with idiosyncratic payoff shocks that are distributed type I extreme value. For this case, estimation reduces to evaluating several linear projections, which does not require costly computations and can be implemented using any software that is capable of estimating a linear regression. Estimation for more general shock distributions may require solving nonlinear equations and thus necessitate using more advanced computational tools.
Step 1 First, we nonparametrically estimate the conditional choice probabilities of the borrowers. Our data represent a panel of \(i = 1, \ldots, J\) loans observed in periods \(\tau = 1, \ldots, T^*\) indexing calendar time. The panel is unbalanced due to defaults and issuance of new loans. We use \(T_{i,\tau}\) to denote the time elapsed (in months) from the period of mortgage origination for a loan \(i\) observed in period \(\tau\). We estimate the policy functions by evaluating the conditional distribution of observed actions for each observed loan age \(t\). It is important to recover policy functions separately for each \(t\) in order to take into account the nonstationarity of agents’ optimal behavior due to the presence of a final period. Our estimation procedure uses a projection onto the orthogonal series \(q^L(\cdot) = (q^1(\cdot), \ldots, q^L(\cdot))^\prime\), where \(L\) is the number of series terms. We consider the orthogonal representation for the choice probability as

\[
\log \frac{\sigma_t(k, s; c)}{\sigma_t(0, s; c)} = \sum_{l=1}^{\infty} r_l(t, k, c) q_l(s), \quad \text{for } k = 1, 2,
\]

where \(r_l(t, k, c)\) are coefficients of the series representation. This representation will provide a uniform approximation for the decision rule if the choice probability ratio is continuous and the state space \(S\) is a compact set, by the Weierstrass theorem. We construct our estimator by replacing the infinite sum with a finite sum for some (large) number \(L\). The parameters are estimated by forming a quasi-likelihood:

\[
\hat{Q}(r^L(t, 1, c), r^L(t, 2, c); c) = \frac{1}{J} \sum_{\tau=1}^{T^*} \sum_{i=1}^{J} \{T_{i,\tau} = t\} \{c_i = c\} \left[ \sum_{l=1}^{L} r_l(t, 1, c) q_l(s_{i,t}) \right. \\
+ \left. \sum_{l=1}^{L} r_l(t, 2, c) q_l(s_{i,t}) \right] \\
- \log \left( 1 + \exp \left( \sum_{l=1}^{L} r_l(t, 1, c) q_l(s_{i,t}) \right) + \exp \left( \sum_{l=1}^{L} r_l(t, 2, c) q_l(s_{i,t}) \right) \right),
\]

where \(r^L(t, k, c) = (r_1(t, k, c), \ldots, r_L(t, k, c))\). Then, we obtain the estimator as

\[
(\hat{\sigma}^L(t, 1, c), \hat{\sigma}^L(t, 2, c)) = \arg\max_{c \in S} \hat{Q}(r^L(t, 1, c), r^L(t, 2, c); c).
\]

(3)

The estimated policy functions correspond to the fitted values based on the estimated parameters:

\[
\hat{\sigma}_t(k, s; c) = \frac{\exp \left( \sum_{l=1}^{L} \hat{r}_l(t, k, c) q_l(s) \right)}{1 + \exp \left( \sum_{l=1}^{L} \hat{r}_l(t, 1, c) q_l(s) \right) + \exp \left( \sum_{l=1}^{L} \hat{r}_l(t, 2, c) q_l(s) \right)} \quad \text{for } k = 1, 2,
\]

\[
\hat{\sigma}_t(0, s; c) = 1 - \hat{\sigma}_t(1, s; c) - \hat{\sigma}_t(2, s; c)
\]

13
The number of series terms is a function of the total sample size, with \( L \to \infty \) as \( J \to \infty \).\(^7\) As we show below, for our asymptotic distribution results to be valid (and, thus, for the first-stage estimation error to have no impact on the convergence rate for the estimated structural parameters), it is sufficient to find an estimator for the choice probabilities with a uniform convergence rate of at least \( (J)^{1/4} \). Such estimators exist if the choice probability is a sufficiently smooth function of the state.

Unlike other multi-step estimators such as Bajari, Benkard and Levin (2009), Pesendorfer and Schmidt-Dengler (2010), and Aguirregabiria and Mira (2002), we do not need to separately estimate transition functions for the state variables. Instead of simulating future state variables using the estimated state transition functions, we propose to nonparametrically estimate borrowers’ expectations about the one-period-ahead value function, and use these estimates to recover the structural parameters within a regression framework. We discuss this approach in detail below.

**Step 2** In the case of an arbitrary (known) distribution of idiosyncratic shocks, we must perform a functional inversion to recover the choice-specific and ex ante value functions from the estimated policy functions. Suppose that \( F_c(\cdot | c) \) is the distribution of payoff shocks (conditional on the “type” of the borrower) with density \( f_c(\cdot | c) \). Consider the functions

\[
\begin{align*}
\sigma_0(z_1, z_2; c) &= \int_{z_0 \geq z_1 + \varepsilon_1, z_0 \geq z_2 + \varepsilon_2} f_c(\varepsilon | c) d\varepsilon_0 d\varepsilon_1 d\varepsilon_2, \\
\sigma_1(z_1, z_2; c) &= \int_{z_1 + \varepsilon_1 \geq z_0, z_1 + \varepsilon_1 \geq z_2 + \varepsilon_2} f_c(\varepsilon | c) d\varepsilon_0 d\varepsilon_1 d\varepsilon_2, \\
\sigma_2(z_1, z_2; c) &= \int_{z_2 + \varepsilon_2 \geq z_0, z_2 + \varepsilon_2 \geq z_1 + \varepsilon_1} f_c(\varepsilon | c) d\varepsilon_0 d\varepsilon_1 d\varepsilon_2.
\end{align*}
\]

\(^{(4)}\)

In Lemma 1 in the Appendix, we show that these functions are well-defined and that if the default probability is strictly between zero and one for almost all values of the state variables, the system determining the choice probabilities is everywhere invertible for \( z_1, z_2 \in \mathbb{R} \), assuming a large support for the idiosyncratic payoff shocks.

We can use the above system of equations to recover the difference between the choice-specific value of each non-default choice and the utility from default, for each period \( t \), each value of the state variables \( s \), and each point in the support of borrower-specific heterogeneity \( c \). Specifically, we reexpress the choice probabilities as functions of these differences, equate them to their empirical analogues (which were recovered in the first step), and drop the expression for action 2 (because there are only two degrees

\(^7\)We provide precise prescription for the choice of the number of series terms in the theoretical part of this section.
of freedom). This yields

\[
\hat{\sigma}_{0,t}(s; c) = \sigma_0 \left( V_t(1, s; c) - u(s; \theta(0, c)), V_t(2, s; c) - u(s; \theta(0, c)); c) \right), \\
\hat{\sigma}_{1,t}(s; c) = \sigma_1 \left( V_t(1, s; c) - u(s; \theta(0, c)), V_t(2, s; c) - u(s; \theta(0, c)); c) \right).
\]

The solution to the above system can be expressed as:

\[
\hat{V}_t(k, s; c) = u(s; \theta(0, c)) + F_k \left( \hat{\sigma}_{0,t}(s; c), \hat{\sigma}_{1,t}(s; c) \right), \quad k = 1, 2,
\]

where function \( F_k \) is the solution of the inversion problem (4) for the argument \( z_k \). We can characterize the ex ante value function as

\[
\hat{V}_t(s; c) = u(s; \theta(0, c)) + F \left( \hat{\sigma}_{0,t}(s; c), \hat{\sigma}_{1,t}(s; c) \right),
\]

where the function \( F \) is determined by the solutions \( F_1 \) and \( F_2 \) and the distribution of the idiosyncratic payoff shocks.

We then substitute the obtained expressions into the Bellman equation for the borrower, obtaining a system of nonlinear simultaneous equations

\[
E \left[ F_k \left( \sigma_{0,t}(s_t; c), \sigma_{1,t}(s_t; c) \right) - u(s_t; \theta(k, c)) + u(s_t; \theta(0, c)) - \beta u(s_{t+1}; \theta(0, c)) - \beta F \left( \sigma_{0,t+1}(s_{t+1}; c), \sigma_{1,t+1}(s_{t+1}; c) \right) \right] = 0, \quad k = 1, 2
\]

where the current state \( s_t \), the current action, and the borrower type serve as instruments.\(^8\) This system can be estimated directly by substituting in the state at time \( t, s_t \), and treating \( F_1 \) and \( F_2 \) as the “outcome variables,” with the right-hand side containing the parameters to be estimated (i.e., the unknown utilities parameterized by \( \theta(a, c) \) and the time preference parameter \( \beta \)). Alternatively, we can estimate this system using a nonlinear IV methodology. For estimation, it is convenient to replace this system of conditional moment equations with a system of unconditional moment equations, using the set of instruments \( Z_t = \{W_m(s_t; c), m \in \mathcal{M}\} \). \( Z_t \) is constructed from the current state variables, the current action, and borrower-level heterogeneity elements \( c \) using functions from some set \( \mathcal{M} \), which could be a finite list of orthogonal

\(^8\)In our empirical application, we choose a normalization of default utility such that it is a function of some time-invariant state variables that affect default utility only and not the per-period utility of prepayment or payment. This allows us to identify \( \theta(0, c) \) separately from \( \theta(1, c) \) and \( \theta(2, c) \). One might choose a different normalization in other settings, but as discussed in the previous section, we need a normalization of some kind in order to identify the model when data on agents’ choices in the final period are unavailable.
polynomials of the state variables. Denote

\[ \epsilon_{kt} = F_k(\sigma_{0,t}(s_i; c), \sigma_{1,t}(s_i; c)) - u(s_i; \theta(k, c)) + u(s_i; \theta(0, c)) - \beta u(s_{t+1}; \theta(0, c)) \]

\[ - \beta F(\sigma_{0,t+1}(s_{t+1}; c), \sigma_{1,t+1}(s_{t+1}; c)), \text{ } k = 1, 2. \] (6)

Define \( \epsilon_t = (\epsilon_{1t}, \epsilon_{2t})' \) and \( Z_t = (Z_{1t}, Z_{2t}) \). Then the system of unconditional moment equations takes the familiar form

\[ E[\epsilon_t Z_t] = 0. \]

Provided that the model that we analyze is smooth with respect to the nonparametrically estimated components, such as choice probabilities, we can apply the results in Chen, Linton, and van Keilegom (2003) and Mammen, Rothe, and Schienle (2012) to establish the impact of the first-stage estimation error on the standard errors of the structural estimates. In the next section, we will analyze the properties of this two-step estimator for the case of a general distribution of the idiosyncratic payoff shocks.

When the idiosyncratic payoff shocks have an i.i.d. type I extreme value distribution, the elements of the derived system of conditional moment equations can be expressed in closed form:

\[ F(\sigma_0(s; c), \sigma_1(s; c)) = \log (\sigma_0(s; c)^{-1}). \]

And the functions \( F_1(\cdot) \) and \( F_2(\cdot) \) in this case take the form:

\[ F_k(\sigma_0(s; c), \sigma_1(s; c)) = \log (\sigma_k(s; c)/\sigma_0(s; c)), \text{ for } k = 1, 2. \]

The choice-specific and ex ante value functions in this case can be recovered directly from the estimated choice probabilities (up to normalization \( u(s; \theta(0, c)) \)), without requiring any iterations or complicated inversions. The per-period payoffs are then recovered from the second-stage plug-in estimator, using the recovered choice-specific and ex ante value functions as inputs.

Thus, when the choice utilities are linearly parameterized and we assume, in addition, that the idiosyncratic payoff shocks are distributed type I extreme value, the second stage estimates can be obtained using standard least squares, which does not even require the construction of the GMM objective function. And our entire estimation procedure reduces to performing three linear regressions:

1. We nonparametrically estimate the borrower’s choice probabilities \( \sigma_{k,i,t}(\cdot; c) \). Using these choice probabilities we recover the ex ante value function for each borrower in each time period, constructing the variable \( F_{i,t} = F(\tilde{\sigma}_0(s_{i,t}; c_i), \tilde{\sigma}_1(s_{i,t}; c_i)) = \log (\tilde{\sigma}_0(s_{i,t}; c_i)^{-1}) \).
2. We estimate the one-period-ahead expected ex ante value $E[F_0(s_{i,t+1}; c_i), \sigma_{1,t+1}(s_{i,t+1}; c_i)] | a_{i,t}, s_{i,t}, c_i]$ by estimating a nonparametric regression of variable $F_{i,t+1}$ (constructed in the previous regression) on $a_{i,t}, s_{i,t}$ and $c_i$ for each borrower. The fitted values from this regression form the variable

$$\hat{F}_{k,i,t+1} = \hat{E}[F_0(s_{i,t+1}; c_i), \sigma_{1,t+1}(s_{i,t+1}; c_i)] | a_{i,t} = k, s_{i,t}, c_i]$$

for $k = 1, 2$ which we construct for each borrower in each time period.

3. We construct the “outcome” variables $Y_{k,i,t} = \log(\tilde{\sigma}_{k,t}(s_{i,t}; c_i)/\tilde{\sigma}_{0,t}(s_{i,t}; c_i))$ for $k = 1, 2$. Assuming that the per-period utility is represented by a linear index of the state variables, we also construct a vector of “regressors”

$$X_{k,i,t} = \left(s_{i,t}, -s_{i,t} + \beta \hat{E}[s_{i,t+1}|s_{i,t}, a_t = k], \hat{F}_{k,i,t+1}\right)'$$

The components of this vector correspond to $u(s_{i,t}, \theta(k, c_i))$, $-u(s_{i,t}, \theta(0, c_i)) + \beta u(s_{i,t+1}, \theta(0, c_i))$ and $F(\sigma_{0,t+1}(s_{i,t+1}; c_i), \sigma_{1,t+1}(s_{i,t+1}; c_i))$, respectively. Then, estimating the coefficient $\delta_k$ in the linear regression

$$Y_{k,i,t} = \delta_k' X_{k,i,t} + \epsilon_{k,i,t}$$

yields the structural parameters of interest. In fact, by construction of $Y$ and $X$, we get $\delta_k = (\theta(k, c), \theta(0, c), \beta)'$. We can further improve inference by simultaneously estimating the equations for $k = 1$ and 2 as a system.

In the first and second regressions, the estimation is performed separately for each $t$ in order to account for the nonstationarity of the problem. It is also worth noting that we recover agents’ expectations nonparametrically in the second regression. As discussed in Ahn and Manski (1993), this approach relies on the assumption that agents’ expectations are realized and are conditioned only on variables observed by the researcher. Our approach is more robust than imposing parametric state transition functions and then simulating future state variables using the estimated transition functions in order to construct $F_{k,i,t+1}$. Our methodology is also computationally more attractive than many existing techniques for estimating dynamic decision models, because performing a linear projection is significantly faster than simulating over a distribution.

Our estimator exploits Hotz and Miller’s (1993) intuition that, in the presence of a terminating action, one can represent the future value term as a function of agents’ choice probabilities one period ahead. In effect, we do not require agents to form precise long-term forecasts for the transition of state variables.
We only require that borrowers have rational expectations regarding the state variables one period ahead, which we believe to be a more tenable assumption in many empirical settings. This property is also useful in contexts such as ours that feature heavy data truncation.

3.2 Asymptotic Theory for the Plug-In Estimator

The previous section outlined the structure of the two-step plug-in estimator for the structural parameters, which include the per-period payoffs and the discount factor. This section provides the asymptotic theory for the constructed estimator. We assume a parametric specification for the per-period utility, although our theory allows for an immediate extension to a nonparametric specification of the per-period utility. A key requirement of the plug-in semiparametric procedure is that the first-stage nonparametric estimator of the policy functions converge at a sufficiently fast rate. Our results for the consistency and the convergence rate of the first-stage estimator rely on the results in Wong and Shen (1995), Andrews (1991), and Newey (1997).

To assure consistency and a fast convergence rate for the first-stage estimator, we need the following assumption.

**ASSUMPTION 2**

1. In addition to the Markov assumption (Assumption 1.iv), for each period $t$ the distribution of states $s_{i,t} \mid (s_{i,t-1}, c_i)$ is identical across borrowers and over time conditional on borrower heterogeneity $c_i$, and the choice probabilities $\sigma_{k,t}(s; c)$ are uniformly bounded from 0 and 1 for each $k = 0, 1, 2$ and $c \in C$. The state space $S$ is compact.

2. The eigenvalues of $\mathbb{E} \left[ q^{L^t}(s_{i,t})q^{L^t}(s_{i,t}) \bigg| a_{i,t}, c_i \right]$ are bounded away from zero uniformly over $L$, and $|\eta_i(\cdot)| \leq C$ for all $l$.

3. $\frac{\sigma_{k,t}(s;c)}{\sigma_{0,t}(s;c)}$ belongs to a separable functional space with basis $\{q_l(\cdot)\}_{l=1}^\infty$. For each $t \leq T$ and $k \in \{1, 2\}$ the selected series terms provide a uniformly good approximation for the probability ratio

$$\sup_{s \in S} \left\| \log \frac{\sigma_{k,t}(s;c)}{\sigma_{0,t}(s;c)} - \text{proj} \left( \log \frac{\sigma_{k,t}(s;c)}{\sigma_{0,t}(s;c)} \right) \right\| = O(L^{-\alpha})$$

for some $\alpha \geq \frac{1}{2}$.

4. Conditional choice probabilities are twice differentiable uniformly in the observed heterogeneity component $c$. 

18
Assumption 2 can be verified for particular classes of polynomials and sieves (see Chen, 2007). We also impose an additional assumption restricting the complexity of the class of functions that is associated with our “nonparametric multinomial logit” estimator, which we discuss in the Appendix. Assumption 2 implies the following result establishing the consistency and convergence rate of the first-stage estimator for the policy functions.\footnote{Proof is in the Appendix.}

**Theorem 4** Under Assumptions 1 and 2, the estimator (3) is consistent uniformly over $s$:

$$\sup_{s \in S} ||\hat{\sigma}_{k,t}(s|c) - \sigma_{k,t}(s|c)|| = o_P(J^{-1/4})$$

provided that $L \to \infty$ with $J \log(J) \to \infty$ as $J \to \infty$.

The asymptotics in this theorem is in terms of the number of loans $J$, reflecting the fact that each loan is observed only once for a given $t$. We use the estimated first-stage policy functions as inputs for the estimation of the second-stage structural parameters. Our approach is based on applying existing plug-in implementations for estimating the system of equations (5). These techniques involve constructing nonparametric elements based on a statistical model (in our case, the policy functions) that are then plugged into a fully parametric second step. Estimation in the second step is commonly performed by means of a weighted minimum distance procedure, with weights that are chosen optimally to maximize the efficiency of the resulting estimator. We use the vector of state variables and the nonparametric estimates for the one-period-ahead ex ante value function (which can be easily constructed using the first-stage policy function estimates) as instruments, using the identity matrix as the weighting matrix. Although this estimator is not semiparametrically efficient, it has the advantage of being estimable using least squares rather than requiring GMM.

To establish the asymptotic properties of the designed procedure we impose the following assumptions.

**ASSUMPTION 3**

1. Parameter space $\Theta$ is a compact subset of $\mathbb{R}^p$.

2. The per-period payoff is Lipschitz-continuous in parameters.

3. The variance of the one-period-ahead policy function is bounded ($\sup_{s \in S} E \left[ \sigma_{k,t+1}(s_{t+1}; c)^2 \right] s_t = s < \infty$ and strictly positive ($\inf_{s \in S} E \left[ \sigma_{k,t+1}(s_{t+1}; c)^2 \right] s_t = s > 0$) for any $t < T$.

Under this assumption and the technical assumption described in the Appendix, we can use the results
regarding semiparametric plug-in estimators in Ai and Chen (2003) and Chen, Linton and van Keilegom (2003), and establish the following result for the estimator for the second-stage structural parameters.

**Theorem 5** Under Assumptions 1, 2 and 3, the estimator (5) is consistent and has asymptotic normal distribution:
\[
\sqrt{T} \left( \left( \hat{\theta}(a, c), \hat{\beta} \right) - \left( \theta_0(a, c), \beta \right) \right) \xrightarrow{d} N(0, V),
\]
where variance \(V\) is determined by the functional structure of the model.

The result of this theorem follows from Theorem 4.1 in Ai and Chen (2003). A significant difference between equations (5) used for our estimation and the conditional moment equations implied by infinite-horizon Markov dynamic decision processes is that the one-period-ahead values in our moment equations are estimated separately. As a result, the estimated choice-specific value function and the one-period-ahead ex ante value function can be considered to be unrelated nonparametric objects (in contrast to infinite-horizon dynamics, in which the two are connected via a fixed point). This feature facilitates the evaluation of the asymptotic variance. We give an explicit expression for the variance in the Appendix.

### 4 Data

We illustrate our proposed estimation methodology using data from CoreLogic on subprime and Alt-A mortgages that were originated between January 2000 and September 2007 and securitized in the private-label market. For each loan, we observe the loan terms and borrower characteristics reported at the time of origination, such as the type of mortgage (fixed rate, adjustable rate, etc.), the initial contract interest rate, the level of documentation (full, low, or none\(^{10}\)), the appraised value of the property, the LTV ratio, the location of the property by zip code, and the borrower’s FICO score.

To simplify the modeling, we focus on 30-year fixed-rate mortgages, the most common mortgage type. We further restrict our sample to loans that are first liens and that are for properties located in 20 major Metropolitan Statistical Areas (MSAs).\(^{11}\) We also exclude “cash-out” refinance loans\(^{12}\) and focus on loans that are for home purchases or refinances with no cash out. Many homeowners use cash-out

---

\(^{10}\)Full documentation indicates that the borrower’s income and assets have been verified. For low documentation loans, only a limited amount of information about assets has been verified. No documentation indicates there has been no verification of information about either income or assets.

\(^{11}\)Atlanta, Boston, Charlotte, Chicago, Cleveland, Dallas, Denver, Detroit, Las Vegas, Los Angeles, Miami, Minneapolis, New York, Phoenix, Portland, San Diego, San Francisco, Seattle, Tampa, and Washington D.C.

\(^{12}\)That is, when the borrower takes out a larger loan than needed to pay off the old loan. Our data identify individual loans and not the identities of borrowers, so we cannot identify the previous loan taken out by a particular borrower. However, the data identify whether the purpose of each loan was for a home purchase or a refinance, and whether cash was taken out on a refinance.
refinance loans for debt consolidation or home improvement, which would significantly change their non-
housing-debts or home value in a way we do not observe in the data. It makes our model less suitable
for explaining the default behavior of borrowers with cash-out refinance loans, so we drop them from the
estimation sample.

A challenge for empirical research on mortgages in general (and especially for subprime mortgages) is
the paucity of information on borrower income, an important state variable.\textsuperscript{13} Although it is impossible
to observe a borrower’s realized income stream over time, we can proxy for it using the borrower’s initial
income as of the time of loan origination. A key difficulty stems from the fact that the database does
not directly report the borrower’s income. Instead, it reports the back-end debt-to-income (DTI) ratio
and the front-end DTI ratio. The former is defined as the ratio of scheduled monthly payments on the
borrowers’ total debt (including credit cards, student loans, etc.) to their monthly income; the latter is
defined as the ratio of the scheduled monthly payment on only the mortgage to monthly income. The
back-end DTI ratio is much more frequently reported than the front-end DTI ratio, because lenders
perceive the former as being a more important determinant of the borrower’s credit risk.\textsuperscript{14} However,
we cannot reliably infer income from the back-end DTI ratio, because we do not observe borrowers’
non-mortgage debt. Thus, we use a small share (3.5\%) of our loans for which the front-end DTI ratio
is available in order to obtain a reliable measure of initial income. With this restriction, we have about
12,000 borrowers in the sample. We shall use this selected sample in order to illustrate our estimation
methodology, but we do not claim that our sample is representative of the entire subprime population.

The data track each loan over the course of its life, reporting the outstanding balance, delinquency
status, and scheduled payment in each month. We define default as occurring if the bank takes possession
of the home or if the loan has been delinquent for 90 days or more, a commonly used definition of default
in the mortgage literature. Default is a terminal event, so if a loan defaults in month \( t \), it drops from
the sample starting at \( t + 1 \). We define prepayment as occurring if the loan balance is observed going to
zero before maturity, presumably because the borrower has paid off the loan in full. We do not observe
the new loan used to refinance the original loan, because our data identify individual loans but not the
identities of specific borrowers. When the borrower prepays, we assume that the borrower refinances into
a new loan.\textsuperscript{15} We assume that the new loan matures at the same maturity date as the old loan and has

\textsuperscript{13}Borrower income is reported to the government’s Home Mortgage Disclosure Act (HMDA) database for loan applications
in urban areas. However, the HMDA data do not contain information on the property value or the borrower’s behavior
over time, and attempted matches between the HMDA data and commercial databases such as CoreLogic typically yield a
very low percentage of unique matches for each loan.

\textsuperscript{14}The CoreLogic database is primarily marketed toward lenders and institutional investors rather than academic re-
searchers.

\textsuperscript{15}In practice, some borrowers may prepay because they have sold the house. From the data, we cannot distinguish
between prepayments due to refinancing and prepayments due to home sales. It is our understanding that most subprime
borrowers prepay in order to refinance into a new loan with lower interest rates.
an interest rate equal to the current market interest rate.\textsuperscript{16,17} If the borrower does not default or prepay in month \( t \), the borrower continues to make just the regularly scheduled payment, and the loan survives into the next month. We track the status of each loan in our sample through December 2009. Thus, we have data on default behavior for only up through the first ten years of each loan, although the loans have a maturity of 30 years.

As described above, we only directly observe the borrower’s income as of the time of loan underwriting. To proxy for potential income fluctuations for a given borrower over time, we include the monthly county-level unemployment rate as an additional state variable. Data on monthly county-level unemployment rates come from the Bureau of Labor Statistics. If we observed individuals’ employment status in every month, we would be able to identify defaults that occur due to income shocks. However, we are limited in our ability to do so due to the noisiness of our county-level proxy. This is an issue that researchers commonly face in the empirical literature on the subprime mortgage crisis, and we are no exception. We acknowledge that the lack of information on changes in income and job status of individual borrowers could affect our structural estimates and counterfactual results. Unfortunately, we cannot predict the direction or size of the resulting bias.

We track movements in home prices using housing price indices (HPI) at the zip code level, also from CoreLogic. CoreLogic reports the home price indices at a monthly frequency, and constructs them using the transaction prices of properties that undergo repeat sales at different points in time in a given zip code area. We impute the current value of a home by adjusting its appraised value at the time of origination by the index. Because home price declines are thought to be one of the main drivers behind the recent surge in mortgage defaults, and because there is a high degree of variation in home price movements across locations even within the same MSA, it is important to have home price data at a fine geographic level. Hence, we believe that the use of the zip code level HPI from CoreLogic enhances the robustness of our results. See Table 1 for variable definitions.

Table 2 reports summary statistics, both for the entire estimation sample and separately according to the mode by which loans come to an end in the sample—by prepayment, by default, or by censoring at the end of the sample. Maturation is not a relevant category for our sample. As shown, default is

\textsuperscript{16}For example, if the borrower refines when the loan is 50 months old, we assume that the new loan will mature in 310 months. In practice, borrowers typically refinance into a mortgage with a standardized maturity such as 30 years or 15 years. Since some people refinance into 15-year mortgages while others refinance into 30-year mortgages, our assumption is a reasonable approximation of the average outcome.

\textsuperscript{17}We assume that the market interest rate available to borrower \( i \) at time \( t \) is equal to \( r_{i,t} = r_t(z_{i,t}) + \xi_t \), where \( r_t(z_{i,t}) \) is the prevailing rate available at time \( t \) for loans with observable characteristics \( z_{i,t} \), and \( \xi_t \) is a borrower-specific spread that is constant over time. For a given borrower, we can identify \( \xi_t \) as the residual from regressing the observed interest rate on the observed characteristics of the borrower’s original loan.
associated with lower-income borrowers and lower credit scores. For instance, the average FICO score among all loans is 672.6, compared with an average of 626.4 conditional on default.

The second panel of Table 2 presents summary statistics for time-varying variables as of the last period in which we observe each loan. Relative to the overall average across all borrowers, borrowers who default tend to have less net equity and lower housing value at the point of default. For instance, the amount of net equity in the last observed period is on average $103,000 among all borrowers, but only $38,000 for loans that default.\(^\text{18}\)

Table 2 also reports the share of loans for home purchases as opposed to no-cash-out refinances, as well as the share of loans to borrowers who intend to live in the house (owner-occupiers) versus people who buy the house for investment purposes (investors). These partitions of borrowers will become important later when we discuss borrower heterogeneity.

Table 3 shows that loan characteristics differ significantly across origination years. If loans of different vintages do not systematically differ from each other and home price transitions are uniform over time, we would expect that the fraction of loans that default by December 2009 (the end of the sample period) should be monotonically decreasing as we move to the right in Table 3, because loans from earlier vintages have had more time over which to default. Instead, we see that the fraction of loans that default by December 2009 is actually much higher for loans from the 2005–2007 vintages compared with loans originated in 2000–2004, suggesting that the later loans were riskier. This difference can partly be explained by differences in observable loan characteristics and by the decline in home prices in later years. However, loan characteristics and home price transitions might not fully explain the differences across origination years, as researchers have found that loans originated in the later years have a higher propensity to default even after controlling for these factors (Demyanyk and van Hemert (2011)). To account for any systematic differences in unobservable characteristics, we include origination year dummies in the state vector.

Table 4 reports the empirical distribution of the time to default, conditional on a loan eventually defaulting (upper panel), and the empirical distribution of the time to prepayment for loans that eventually prepay (lower panel). Both distributions have a humped shape. A similar humped shape in the

---

\(^{18}\)Table 2 shows that the fraction of loans that have an active prepayment penalty as of the last observed period is significantly lower for censored loans. The reason for this difference is that the prepayment-penalty period expires before the censoring time for most loans that don’t default or prepay sooner.
default hazard and prepayment hazard is also well-documented in the mortgage literature\textsuperscript{19} (Gerardi et al. (2008); von Furstenberg (1969)), but there is no agreement on what economic forces lead to the initial increase in default hazard. For example, it could be generated by borrowers’ stronger determination to make payments on loans that they have just obtained (due to some behavioral biases, for instance) or by greater uncertainty about income or employment shocks further into the future (e.g., conditional on a lender approving a loan, presumably the borrower has a steady income in most cases. An unexpected income shock is thus likelier 18 months down the road as opposed to just after origination). In one of the alternative specifications that we examine below, we modify our analysis to address the impact of such time-varying unobserved factors on our results.

\[ \text{Table 4 about here} \]

While many of the state variables listed in Table 1 are time-invariant, three of them—home value, market interest rate and unemployment rate—evolve stochastically over time.\textsuperscript{20} These state variables are nonstationary (based on an independent analysis of their time series behavior), but we can estimate their transition paths in terms of the growth rate in home value and the first differences in market interest rate and unemployment rate, which are stationary.\textsuperscript{21}

Our analysis assumes that the regulatory regime remained constant over the sample period. In particular, we assume that borrowers do not experience or expect changes to their mortgages due to various foreclosure mitigation programs that the United States government implemented in the later stages of the housing crisis. Two major programs would conceivably be relevant to loans resembling those in our sample. Hope for Homeowners (based on legislation in Spring 2008) is a principal writedown policy. Due to lenders’ apparent unwillingness to reduce principal, there has been virtually no participation in the program. The Home Affordable Modification Program (HAMP, based on legislation in 2009) is a payment-reduction program. However, most institutions started their HAMP trials in late 2009, largely after our sample period. Furthermore, since these programs have had extremely low take-up rates, it is a reasonable approximation to assume that borrowers did not expect to benefit from participation in such

\textsuperscript{19}This effect is also confirmed by unreported regressions using our data. All unreported results are available upon request.

\textsuperscript{20}Net equity also stochastically evolves over time. Its evolution is determined by the evolution of home value and the evolution of the outstanding loan principal. Because the outstanding principal follows a deterministic evolution fully specified by the contractual interest rate (fixed over time) and loan maturity (also fixed over time), estimating the evolution of home value is sufficient to infer the evolution of net equity.

\textsuperscript{21}As discussed in the earlier section, our estimation method does not require estimation of state transition functions. However, estimated state transition functions will be required for counterfactual analyses. Thus, we estimate state transitions (we discuss the chosen specifications in the following section on empirical results) and use them in our counterfactual analyses. Consequently, our structural parameter estimates do not hinge on parametric assumptions regarding state transition functions, while the counterfactual results do.
5 Estimation Results

5.1 Main Results

Our baseline specification has a fixed coefficient design and assumes that heterogeneity across borrowers is fully captured by the state vector $s$ (equivalent to assuming that the set of observable borrower types $C$ is a singleton). In our analysis of alternative specifications later in this section, we will relax this assumption by considering a random coefficient design in which borrower $i$’s type, $c_i$, is defined by some observable time-invariant characteristics that are not part of the state vector. In our analysis of alternative specifications, we also consider the possibility of time-varying unobserved heterogeneity. In this subsection, we focus on the baseline specification.

5.1.1 Results on Policy Function Estimates

We start by discussing the first-step estimates of the policy functions. Because the policy function estimates are reduced-form in nature, the estimates themselves do not have well-defined economic interpretations. Thus, we focus on the goodness of fit of the policy function estimates instead of discussing the coefficients. Having policy function estimates that do a reasonable job of matching empirical probabilities is crucial for the reliability of the structural parameter estimates and the plausibility of the counterfactual results. We investigate the performance of our policy function estimates in three ways. First, we report the within-sample fit of our estimates using the full sample, comparing the predicted probabilities of default, prepayment and payment in each period to the empirical counterparts. Second, we report the out-of-sample fit: we use half of the sample for estimation, use these estimates to compute the fitted probabilities in each period for the other half of the sample (the validation sample) and compare the fitted values to the empirical counterparts in the validation sample. The comparison addresses potential concerns about overfitting. Third, for each loan in the data, we start with the observed state as of the period in which the loan is first observed, and forward-simulate the borrower’s decisions until the end of 2009 (the censoring date) using the estimated policy functions and state transitions. We then compare the predicted probability of eventual default or prepayment by the censoring date to the empirical counterparts. The fit implied by the first two methods depends on the precision of policy function estimates.

\footnote{It is possible that loans in our dataset participated in private modifications without any government involvement. We find that less than 3 percent of the mortgages in our sample are modified. This lessens potential concerns over the impact of private modifications on borrower behavior in our dataset.}
only (since we use the realized values of the state variables in each period in computing the predicted probabilities), whereas the fit from the third method depends on the precision of both the policy function estimates and the state transition estimates. More noise is introduced in the third method, so the fit is necessarily worse.

We use a sieve logit with splines of the state variables in order to flexibly model borrowers’ choice probabilities. The state variables are the set of variables listed in Table 1 (excluding “owner-occupiers” and “purchase loans,” which will be used to define borrower types later in the analysis of alternative specifications) as well as dummies for origination year and MSA. The sieve basis includes restricted cubic splines for continuous state variables, interpolating between 3 to 5 equally spaced percentiles of each state variable’s marginal distribution. It also includes interactions among the state variables. To capture the dependence of the optimal choice probabilities on time to maturity \((T - t)\) where \(T\) is fixed at 360 in our estimation sample, we also include 5 splines of \(t\) and interact them with the other state variables.

We assume that borrowers condition their expectations about the home price and unemployment rate in the next period on the current values of these variables and on their first and second lags. Accordingly, the state vector at time \(t\) for a borrower \(i\) includes the lagged and current values of \(i\)’s home price and the unemployment rate in \(i\)’s county. Similarly, we assume that borrowers condition their expectations about the market interest rate in the next period on the current value and the first lag of the market interest rate, which we similarly include in the state vector.

For the purpose of the counterfactual analysis, we must specify parametric assumptions about the state transition functions. To model the home price and unemployment rate transitions, we assume that the home price growth rate (in percentage terms) and the first difference of the unemployment rate follow AR(2) processes. We model the first difference of the market interest rate as an AR(1) process. These parametric choices are consistent with our assumptions about what the borrowers condition their future expectations on and with how we specify the state vector. Assuming that home price growth rates follow an AR(2) process and that borrowers have rational expectations is admittedly somewhat restrictive and may affect our counterfactual results—in reality, subprime borrowers’ beliefs may have evolved far less predictably.\(^{23}\) This explains why the model’s fit in Table 7 is worse than its fit in Table 5 or 6. Importantly, however, these assumptions would only affect our counterfactuals, but not our structural estimates. Our approach is thus more robust than standard estimation methods, in which particular assumptions about the state transitions would affect both the structural estimates and any counterfactual results.

Table 5 shows the within-sample fit, both overall and for various subgroups. The table shows that

\(^{23}\)See Foote, Gerardi and Willen (2012) for a discussion of borrowers’ expectations about home price movements before and during the subprime mortgage crisis.
the within-sample fit of the first-step policy function estimates is excellent. The likelihood ratio index, also called McFadden’s pseudo $R^2$, is 0.153, a significant improvement over a constant-only model.

[Table 5 about here]

Because we included very flexible splines of the state variables in estimation of the policy functions, one might worry about overfitting, leading to poor out-of-sample predictions. To check for this possibility, we randomly split our sample into two halves, and use the first half for estimation and second for validation. The fit for the validation sample is reported in Table 6.

[Table 6 about here]

Table 6 shows that the fit is excellent even in the validation sample, though it is unsurprisingly slightly worse than the within-sample fit in Table 5. Tables 5 and 6 demonstrate that the first-step estimates predict borrowers’ observed behavior accurately, but another critical piece that will play an important role in the counterfactual simulations is the accuracy of the estimated state transitions. To evaluate the combined fit of the estimated policy functions and transition functions, we start with the earliest observation for each borrower, simulate the sequence of state transitions and borrower’s actions using the estimated policy functions and transition functions, and then compare the simulated path with the actual data. Table 7 compares the observed paths with the predicted paths in terms of the probability of eventual default or prepayment by the end of 2009 and the duration until default or prepayment.

[Table 7 about here]

The table again shows comparisons for the overall sample as well as various slices of the sample. Clearly, the fit is not as good as in the previous tables due to the additional noise introduced by possible misspecification of the state transitions. However, we still find that the first-step estimates explain the data very well.

5.1.2 Results on Projection of One-Period-Ahead Ex Ante Value Function

Once we recover the policy functions in the first stage, we can compute the ex ante value up to the normalization $u(0, s; c)$ for each observation in the data, which is expressed by (1) under the assumption of type I extreme value payoff shocks. Similarly to how we obtained our nonparametric estimator for the policy function, we use a sieve estimator to fit the ex ante value one period ahead (time $t + 1$), conditional on taking action $k$ in the current period, onto splines and interactions of the state variables.
for the current period (time $t$). We specify the sieve in a manner that allows the predicted one-period-ahead ex ante value to depend on the time to maturity, in order to reflect the nonstationarity of the problem. To construct the borrower’s expectation of the next-period ex ante value conditional on taking action $k$ today, which appears in the structural equation (7), we use the fitted value from the series estimator. Note that the one-period-ahead ex ante value is zero conditional on defaulting in the current period, which holds trivially because default is a terminal action.

Our projection estimates (not reported but available upon request) explain an extremely high fraction of variation for the observed ex ante values, with an $R^2$ value of approximately 0.92. Our projection method is conceptually similar to an alternative approach that computes the next period’s ex ante value using forward simulation based on the estimated state transition functions. However, the projection method is computationally much faster than the forward-simulation approach. Furthermore, the projection method does not require us to specify parametric state transition functions in order to recover the structural parameters, as borrowers’ expectations are implicitly embedded in the nonparametric prediction of the one-period-ahead ex ante value function.

Once we construct the expected ex ante value functions using the projection method, we have all the components of the borrower’s Bellman equation, and are ready to estimate the structural parameters.

5.1.3 Results on Structural Estimates

Following the estimation procedure outlined in Section 3, we estimate the structural parameters of the per-period utility as well as the discount factor. As discussed in our model section, we must normalize the utility of default. Because the normalization is not innocuous (Theorem 3), we must choose a normalization that reasonably reflects borrowers’ economic conditions. We assume that a borrower’s utility from default—or equivalently, the cost of default—depends on the borrower’s credit score, because the cost of damaged credit resulting from default conceivably depends on the borrower’s existing credit quality. Furthermore, we assume that the location (MSA) of the borrower’s residence also influences the default utility, because mortgages are recourse loans in some states but non-recourse loans in others.\footnote{Under non-recourse, lenders cannot go after a defaulter’s assets other than the mortgage collateral (i.e., the house), which lowers the perceived cost of default to the borrower.}

Because the FICO score and MSA of residence plausibly do not have a direct impact on the utility from payment or prepayment, we include these variables only in the utility function of default.

We estimate the structural equations, (7) for $k = 1, 2$, using a seemingly unrelated regression (SUR). We impose the cross-equation restriction that the parameters of the default payoff—corresponding to FICO score and the MSA dummies—should be the same in both the payment equation and the prepay-
ment equation. We also impose two additional cross-equation restrictions constraining the discount factor and the degree of disutility from the monthly payments to be the same in the prepayment and payment equations. We parameterize the utility functions to be linear in the state variables (including the monthly payments). In effect, we assume that borrowers are risk-neutral. Table 8 reports our estimates of the structural parameters.

[Table 8 about here]

Our results for the per-period payoff from default and the per-period payoff from paying are sensible. Borrowers with higher FICO scores have a lower utility of default, consistent with default resulting in a higher cost of damaged credit for borrowers with good existing credit. Higher home value and lower monthly payments increase the utility of continuing to pay. Higher local unemployment, a proxy for the likelihood of job loss by the borrower, reduces the utility of continuing to pay. Borrowers with low-documentation loans, who are presumably riskier, have lower utility from paying. These indicators of greater borrower risk may proxy for a higher probability of binding liquidity constraints, which effectively raises the cost of making the monthly payments and thus decreases the probability of choosing the “pay” option.

The intuition for the prepayment equation estimates is somewhat more nuanced. Importantly, the probability of prepayment depends on both the borrowers’ willingness to refinance and their ability to do so. In particular, borrowers may be unable to refinance if lenders deem them to be too risky. Thus, it makes sense that the coefficient estimates imply that higher house value and lower local unemployment increase the propensity to prepay.

Moreover, the coefficients of the prepayment equation may be determined in part by unobserved heterogeneity in borrower or loan characteristics. While this is certainly true as well of the “pay” equation, intuitively, unobserved heterogeneity would seem to be even more important in the case of prepayment, because households have many idiosyncratic reasons for wanting to prepay. For example, the positive coefficient on “Multiple Liens” makes sense if borrowers generally take out multiple liens only as a temporary measure and intend to consolidate their loans through refinancing relatively quickly. Finally, although we limit our attention to 30-year fixed-rate mortgages, we cannot rule out unobserved loan characteristics (e.g., degree of stringency in the prepayment conditions) that may be correlated with some of the state variables. For example, the coefficient on the prepayment penalty is positive, which seems

\[\text{We do not separately include the outstanding loan balance in the utility function, because we believe that controlling for the monthly payment is sufficient to capture agents’ (dis)utility from having a loan obligation. That is, having debt does not directly result in disutility; only having to make payments on the debt does. Furthermore, the monthly payment is highly correlated with the outstanding loan balance, with correlation of over 0.97.}\]
counterintuitive as one would expect prepayment penalties to decrease the utility of refinancing. However, consumers who were unsophisticated enough to obtain loans with a prepayment penalty may likely have obtained a loan that was also unattractive along other dimensions (unobserved to econometricians), and accordingly might be more eager to refinance. The result is also consistent with the theoretical model and empirical evidence presented by Mayer, Piskorski, and Tchistyi (2013), which indicate that mortgages with prepayment penalties tend to go to the least creditworthy borrowers, who are also the most likely to refinance in response to positive credit shocks.

Finally, our estimate of 0.9815 for the monthly discount factor is just below one, as one might expect, and is statistically significantly different from 1 at the 1% level. There is nothing in the model that restricts the estimate of the discount factor, so it is very reassuring to see that our estimated discount factor is a plausible number. Interestingly, the monthly interest rate implied by the estimated discount factor is higher than the average monthly interest rates these borrowers pay on their mortgages in the data. In other words, these borrowers seem to discount the future more than what the market interest rate suggests. When a researcher cannot estimate the discount factor and instead has to assume an arbitrary number, a typical choice for the discount factor in the literature is a number that corresponds to some economy-wide asset return. Our empirical results suggest that such a choice in our setting would have been incorrect, leading to bias in the other structural estimates.

5.2 Alternative Specifications

As noted previously, unobserved heterogeneity in borrowers’ preferences is a potential concern. Certain kinds of unobserved heterogeneity may be borrower-specific, such as differences across individuals in their degree of aversion to defaulting. Other sources of unobserved heterogeneity may vary both across borrowers and over the duration of each loan. For example, because we only observe income at the beginning of the loan, income volatility over time (such as due to loss of employment) would manifest itself as unobserved variation over time in the borrower’s ability to make monthly payments.

The importance of incorporating unobserved heterogeneity has been well understood since the early work by Heckman and Singer (1984). Recent literature has demonstrated the possibility of extending the random effects framework of Heckman and Singer (1984) to nonlinear dynamic models. In particular, Shum and Hu (2012) show that models with unobserved heterogeneity can be identified even when the unobserved shocks are serially correlated.

However, these results do not imply that estimators for structural parameters in models with unobserved heterogeneity are always guaranteed to exist or perform well. Moreover, Khan and Nekipelov
(2011) show that in many parametric discrete-choice and game-theoretic models with unobserved heterogeneity, the Fisher information for structural parameters is zero. In other words, the shape of the likelihood function in the vicinity of the true parameter value is nearly flat, leading to irregular identification of those parameters.\textsuperscript{26} In particular for our model, these findings imply that identification is not robust to specific assumptions about the distribution of unobserved heterogeneity, unless there are observables that can credibly be excluded from the structural equations that shift the distribution of the heterogeneity. Intuitively, without such exclusion restrictions, identification is mostly driven by extreme realizations of unobserved heterogeneity (such that the decision to pay, prepay or default is almost predetermined even prior to including the covariates), meaning that the tail behavior of unobserved heterogeneity has a large effect on the estimates.\textsuperscript{27} Although unobserved heterogeneity is formally identified in many settings similar to ours, the above features have been underappreciated in the literature.

Given that this non-robustness of the discrete choice model with unobserved heterogeneity in the absence of a reliable exclusion restriction is a fundamental problem, we propose to explore some alternative specifications of our model instead of attempting to estimate the distribution of unobserved heterogeneity. Our key goal is to explore whether the structural estimates remain interpretable under these alternative specifications.

Our first of two alternative specifications is a mixture model design in which borrowers have heterogeneous types that are determined entirely by observables. The second alternative specification incorporates liquidity shocks, a plausible source of unobserved heterogeneity, by imposing a parametric distributional assumption on the probability of a liquidity constraint binding in a particular month after origination. We explain the logic behind each approach and discuss the results.

The mixture model design assumes that outcomes are determined by a mixture distribution, with heterogeneous borrower types that are determined by the observable, borrower-specific variables $c$. For instance, the observable type of a borrower could be defined by whether the borrower intends to reside in the house (i.e., owner-occupiers versus investors). One might think that owner-occupiers have a stronger emotional attachment to the house than investors, and thus have a greater aversion to defaulting, all else equal. Therefore, the parameters of the utility function could vary between owner-occupiers and investors. Clearly, this approach does not directly address unobserved heterogeneity, but captures the component of unobserved heterogeneity that is correlated with the observable types. This approach is similar to panel data models in which fixed effects are represented by an unknown function (which is the

\textsuperscript{26}That is, the estimators have non-Gaussian asymptotic distributions and do not converge at a parametric rate.

\textsuperscript{27}Moreover, Khan and Nekipelov (2012) show that possible approaches to making the estimators more robust, such as trimming the support of the distribution of unobserved heterogeneity or restricting the class of possible distributions, are infeasible because they require a priori knowledge of the structure of underlying distributions, such as their tail behavior.
same for all cross-sectional units) of included covariates.

We split the data into different borrower “types” based on owner-occupier status and loan purpose (home purchase versus refinancing). The estimates for the various observable types are reported in Table 9.

Table 9 shows how the utility functions differ across borrower types. For instance, the coefficient on the monthly payment variable is negative and statistically significant for borrowers whose loans are for home purchases, but is not significantly different from zero for borrowers with refinance loans. This suggests that the level of monthly payments is an important factor in default decisions for borrowers who are still paying down their original mortgages, but is not so for borrowers who have obtained a refinance mortgage. For most borrowers in the latter category, the payment was probably already reduced to a more manageable level when the borrowers refinanced their original mortgages.

We also see that investors tend to be more sensitive to financial incentives in their default decisions than are owner-occupiers, as evidenced by the larger magnitudes of the coefficients on house value and payment. This difference between the types may reflect a certain “ruthlessness” to investors’ default decisions, in contrast to owner-occupiers, who may have more emotional attachment to their homes and may have costs to moving out of the house in case of default. On the other hand, coefficients on some variables proxying for borrower riskiness—such as FICO scores, unemployment rate, income, and low documentation—are larger for owner-occupiers than for investors, presumably indicating that defaults driven by liquidity shocks are more common for owner-occupiers than for investors, who likely have more resources available.

In our second alternative specification, we attempt to control for the effects of liquidity shocks, an important source of unobserved heterogeneity, through parametric assumptions. A sufficiently large negative liquidity shock might force a borrower to default regardless of his strategic incentives. We can think of the borrower as facing a binding liquidity constraint at some random point in time over the course of the loan. We model the realization of this random time as being drawn from an exogenous distribution, and then adjust the choice-specific value functions to reflect the fact that default may be driven by either the borrower’s strategic incentives, as modeled in our base specification, or by this liquidity constraint.

We assume that the liquidity shocks are orthogonal to the payoff shocks. From a formal perspective, this assumption is necessary for identification. From a conceptual viewpoint, it requires the borrowers to base their default decisions only on the present discounted value of the mortgage (as opposed to expectations about future liquidity shocks) as long as an unobserved liquidity constraint is currently

32
nonbinding. Default may occur because of either “strategic” incentives—as outlined by our main model—or a negative liquidity shock, e.g., loss of job. This structure leads to a mixed hitting time model, studied in Abbring (2012).

We characterize liquidity-based default as follows. Each borrower $i$ is assumed to have a fixed default risk $\rho_i$, which characterizes the unconditional probability of the borrower defaulting due to liquidity shocks in a given period. The probability that this borrower does not default due to liquidity shocks before period $t$ of the loan is $(1-\rho_i)^{t-1}$. Assuming that the borrowers are atomic and that the conditional distribution of $\rho_i$ for a borrower of type $c_i$ is $f(\rho_i|c_i)$, the aggregate probability of illiquidity-driven default at a loan age of $t$, conditional on $c_i$ and conditional on survival to $t$, is

$$P_t^L(c_i) = \frac{\int \rho (1-\rho)^{t-1} f(\rho|c_i) d\rho}{\int (1-\rho)^{t-1} f(\rho|c_i) d\rho}.$$  

Then, if $f(\rho|c_i)$ follows a Beta distribution, $P_t^L(c_i)$ will have a binomial form in $t$, which can further be approximated by the Poisson distribution with argument $t$.

Our approach to the empirical analysis of the mixed hitting time model is based on the following observation regarding the behavior of strategic default probabilities $\sigma_{0,t}(s;c)$.  

**LEMMA 2** Under Assumption 1, if the density of the distribution of choice-specific shocks has at least exponential tails and the utility function is strictly monotone in at least one of the state variables, then there exist constants $A_1$ and $A_2 > 0$, and $B_1$ and $B_2 > 0$, such that $A_1 t^{-A_2} < \sigma_{0,t}(s;c) < B_1 t^{-B_2}$.  

This lemma establishes that the probability of strategic default declines at the geometric rate as the mortgage approaches maturity. This is due to the fact that with the finite planning horizon, the benefits of defaulting decline over time. On the other hand, we assumed that the probability of illiquidity-driven default declines at the exponential rate as the mortgage approaches maturity. This means that, starting at some period after mortgage origination, the probability of default will be dominated by the strategic default probability.

To incorporate this result into our estimation procedure, we use a parametric approach. In particular, we make an explicit assumption that the reduced-form liquidity-based default probability $P_t^L(\tau_{i,L} = t | c_i)$ is well-approximated by the Poisson distribution, where $\tau_{i,L}$ is the instant in time when borrower $i$ defaults due to a binding liquidity constraint. Based on the result from Lemma 2, the strategic default probability is monotone decreasing for fixed values of the state variables and for a particular borrower type, given Assumption 1. Therefore, if the estimated default probability $P_t^d(s;c)$, which is inclusive of both strategic default and illiquidity-driven default, is not monotone over $t$ for fixed $s$ and $c$, then the nonmonotonicity necessarily reflects the impact of illiquidity-driven default. In particular, we can recover the function

33
\[ \hat{\lambda}(s; c) = \max_{t=1, \ldots, T} P_{t}^{d}(s; c), \] which corresponds to the rounded parameter of the Poisson distribution. Thus we can restore the strategic default probabilities using the recovered parameter of the Poisson distribution:

\[ \sigma_{0,t}(s_{i,t}; c_{i}) = \frac{P_{t}^{d}(s_{i,t}, c_{i}) - \hat{\lambda}(s; c)^{t} e^{\hat{\lambda}(s; c)/t!}}{1 - \hat{\lambda}(s; c)^{t} e^{\hat{\lambda}(s; c)/t!}}. \]

Upon recovering the strategic default probabilities, we perform the second stage of our estimation as before.

Table 10 reports the estimates that result from netting out the impact of liquidity shocks through the Poisson assumption. The magnitudes of the coefficient estimates (except the discount factor) tend to be much larger compared with the baseline results in Table 8. This difference in magnitude is to be expected, because the financial incentives that make default less attractive, such as the costs of damaged credit, would be more relevant to the decision-making of borrowers who are not constrained by illiquidity than borrowers who struggle to meet monthly obligations (and thus see few options other than defaulting on the mortgage).

6 Simulations and Welfare Analysis

In this section, we consider how borrowers’ default behavior would change under various scenarios and compute the corresponding changes in borrower welfare. Our analysis uses the first-step policy function estimates to simulate borrowers’ behavior under various “counterfactual” regimes. This approach is unconventional and deserves discussion. When we conduct counterfactual analysis using the first-step policy function estimates, one might worry about the Lucas critique. However, given the individual-level, panel structure of our data, our setup is not subject to this critique so long as two requirements are met. First, for each scenario, the forward-simulated distribution of state variables must remain in the empirical support of the policy functions. Note that in our panel data, we have much contemporaneous variation in home prices, monthly payments, credit quality, and so on. Our policy function is a valid description of optimal behavior for any realization of state variables in the empirical support. Therefore, we can use the first-step policy function estimates to determine how outcomes would change so long as the scenario under consideration does not lead to simulated values of state variables lying outside the empirical support. For example, we could predict the trajectory of borrowers’ default behavior for a scenario in which all housing markets experience the same evolution of home prices as a particular MSA observed in the data. More generally, we can study any scenario that involves situations that actually
occurred for some subset of borrowers in the data. A second requirement is that the transitions of the state variables (and expectations about them) must be unchanged under the counterfactuals, because the first-step policy function estimates are implicitly conditioned on the transition functions. We judiciously choose our simulation exercises that meet these requirements.\textsuperscript{28}

Throughout our counterfactual exercises, we maintain a couple of key assumptions. First, we assume that the evolution of the macro state variables follows an exogenous process. This partial-equilibrium approach makes the problem more tractable. A general equilibrium model wherein home prices and interest rates are endogenously determined is beyond the scope of this paper. In particular, such a general approach would require modeling how default decisions would feed into the determination of home prices, which is not an easy task.

Furthermore, our model addresses default decisions of borrowers conditional on their having already obtained a mortgage. As such, our model cannot be used to predict how loan originations would change under the various scenarios.

There is no doubt that these assumptions are restrictive. Thus, we want our counterfactual results to be viewed as exercises to isolate the impact of the change under investigation on borrowers’ default behavior while holding other macro variables unaffected by the change, rather than as an attempt to provide definitive answers on how various government policies would affect the housing market in practice.

6.1 Counterfactual Scenarios

Some of our counterfactual scenarios are intended to shed light on the relative importance of the major factors that contributed to higher default rates in recent years. Others are intended to assess the impact of various foreclosure mitigation policies. Because the actual policy interventions that have been implemented since the financial crisis have tended to be very multidimensional (and have applied to subsets of borrowers according to complex rules, implying that the eligible groups cannot easily be identified in our data), the general spirit of the exercise is to explore the effects of specific aspects of the programs as opposed to any particular policy implemented by the government.

The subprime mortgage crisis of 2007 was partly characterized by an unusually large fraction of subprime mortgages originated in 2006 becoming delinquent or going into foreclosure very soon after origination. For instance, the cumulative empirical probability of default by the end of 2007 is 10.81% for mortgages originated in 2006, compared with 6.48% for mortgages originated in 2004, even though the

\textsuperscript{28}One of the advantages of this approach over the conventional approach of re-solving for the new equilibrium is a reduced computational burden. We find that it takes almost a week to run each counterfactual scenario even under our computationally lighter approach.
older loans have had more time over which to default. The first set of simulations focuses on explaining this difference in performance between loans of different vintages:

1. To quantify the importance of falling home prices in explaining the observed increase in defaults, we ask what the aggregate default rate among subprime borrowers would have been under alternative evolutionary paths for home prices. Specifically, focusing on loans that were originated in 2004, we ask how they would have fared up through the censoring date (December 2009) if in 2004 homes in all markets had experienced the same precipitous decline in value as the average Las Vegas house three years later in 2007 (Scenario 1).\textsuperscript{29} Due to the assumed AR(2) structure of the home price growth rate, house prices remain dampened for subsequent years under this scenario. By comparing the predicted default rates given the actual home price path with predicted default rates under the counterfactual scenario, we can determine how home price declines affect borrowers’ default behavior. Another similar exercise simulates the default behavior of all subprime loans that were originated in 2006 under the counterfactual scenario in which all homes nationwide experience an increase in value in the year 2006 equaling that experienced by the average Las Vegas house two years earlier, in 2004 (Scenario 2). Due to the assumed AR(2) structure of the home price growth rate, house prices remain inflated for subsequent years under this scenario.

2. In our earlier work (Bajari, Chu and Park (2011)), we found that deterioration over time in the credit quality of subprime borrowers was another major factor behind the recent increase in subprime defaults. To investigate this issue, we examine how much lower aggregate default rates would have been if the borrowers who took out loans in the later years had the same overall credit quality as the borrowers from earlier years. Specifically, we shift the distribution of FICO scores for new borrowers in 2006 upward to match the mean FICO score among new borrowers in 2004. We then consider the set of loans originated in 2006 and simulate the borrowers’ default behavior through the censoring date of December 2009 (Scenario 3).

The second set of simulations evaluates the effects of stylized foreclosure mitigation policies:

1. How effective would mortgage principal writedowns be? Scenario 4 considers the effect of a 10% principal writedown on all outstanding loans; Scenario 5 considers a 20% writedown. We

\textsuperscript{29}Mechanically, we do the following. We create a sample that contains the first month’s observation for all loans originated in 2004. We then change the current and lagged values of home prices in the state vector of the observations to match the corresponding values of the average Las Vegas house in 2007. Those modified values are used in forward-simulating the housing prices for future periods, and the forward-simulated values, along with the other state variables and random draws for the idiosyncratic payoff shocks, are used in obtaining the optimal choice of each borrower in each period. We use 30 simulation draws for each loan in order to obtain the probability of eventual default or prepayment by December 2009. We then compare the predicted default/prepayment probabilities of the loans under this scenario against the probabilities predicted based on forward-simulating from the first month’s observation for each loan but without modifying the initial housing prices.
impose principal writedowns during the first month of each loan and forward-simulate their default behavior in subsequent periods in order to understand the impact of greater housing net equity and lower monthly payments on borrowers’ incentives to default. To be sure, this scenario does not entirely match the principal writedown policies discussed in reality, as we do not restrict the principal writedowns to struggling borrowers during the crisis period. However, the exercise should still provide insights on how important principal reductions are in influencing borrowers’ default incentives.

2. What effect would a cap on the loan-to-value ratio have? It is widely believed that loosened underwriting standards, such as the relaxation of downpayment requirements, paved the way for the mortgage crisis. **Scenario 6** considers what would happen if LTVs at origination were capped at 0.8 (20% downpayment) for all borrowers whose actual LTVs at origination exceeded 0.8; **Scenario 7** caps the original LTV at 0.9 (10% downpayment). Such a stricter requirement reduces the chance of borrowers going underwater even if home prices decline, thereby reducing the incentive to default.

In Table 11, we report simulation results for the counterfactual cases, which we show alongside the baseline model predictions for comparison. The results for Scenarios 1 and 2 indicate that housing price appreciation causes a meaningful reduction in default. Our model explains this effect in part by giving borrowers more net equity at each point in time, relative to the baseline case. As well, autocorrelation over time in housing prices implies that higher current appreciation leads to higher expectations for future appreciation, which reduces the incentive to default. The reverse is true when we subject borrowers to a large price decline: borrowers are much more likely to default because the price decline immediately pushes some borrowers deep into negative net equity, reducing the loss from walking away from the loan, and also because the realized price decline creates expectations of future price declines.

[Table 11 about here]

Increasing the overall level of FICO scores of the borrower pool significantly reduces the aggregate default probability. This suggests that loosened underwriting standards, which permitted consumers with low credit quality to obtain mortgages, was a significant contributor to the higher default rates among subprime mortgages in recent years. Principal writedowns have the intended effect of reducing default and the effects are substantial. Finally, the effect of the LTV caps is qualitatively similar but smaller in magnitude than that of our principal writedown scenarios. This difference in magnitude is primarily driven by the fact that our LTV caps are binding only for a small fraction of borrowers, whereas the principal writedowns we consider are applied to the entire population.
6.2 Borrower Welfare

To understand the consequences of the above scenarios for borrower welfare, we compute borrower welfare in two ways. First, we compute the ex ante value function of each borrower in each time period under both the baseline and counterfactual scenarios using (1). As a summary measure of how borrowers’ welfare changes between the two scenarios, we compare the mean of the ex ante values across borrowers and over time between the two scenarios.\footnote{Due to nonstationarity of the value function, the average of the ex ante values over time does not have a clear economic interpretation. However, we think the average is still a useful summary measure for comparison of welfare between two scenarios since the degree of nonstationarity is similar between the two scenarios.} Second, we compute the compensating variation, which is the amount of additional money the borrower must receive in order to bring his utility under the counterfactual scenario back to his initial utility level.

In Table 12, we report how mortgage borrowers’ welfare changes under the various scenarios using the first approach. With exception of Scenario 3, our model predicts that borrowers’ welfare should increase in scenarios in which the aggregate default rate goes down, and decrease when the aggregate default rate goes up. The only scenario in which the relationship is theoretically ambiguous is Scenario 3. When borrowers’ credit quality goes up, there are two opposing effects. On the one hand, borrowers are less likely to default, which leads to greater welfare. On the other hand, higher FICO scores imply greater utility loss in the case of default, because the costs of default are higher for borrowers with better credit. This could lead to lower welfare given the same rate of default. The welfare figure in Table 12 shows that the former effect dominates, with borrowers’ overall welfare increasing when their credit quality goes up.

Although Table 12 indicates the direction and relative magnitudes of the borrower welfare changes under each scenario, we cannot assign any meaningful scale to the numbers because ex ante values are unitless (Indeed, as shown, some of the welfare levels are negative). Compensating variation, on the other hand, can be expressed in dollars.

We define the “one-time compensating variation” (OTCV) as the lump-sum compensation (either a reduction or increase in payment) needed to bring the borrower’s utility under the considered counterfactual scenario to the original utility level. By assumption, the OTCV does not affect the transition path of income—in other words, a mechanism for implementing the OTCV would be to present consumers with a lump-sum while informing them that the compensation is for one time only. The OTCV is thus a variant of the standard compensating variation used in microeconomics, and can be applied to a dynamic model with randomness in borrowers’ utility functions. If $u_k(\cdot)$ is the per-period utility of a consumer
from choosing option $k$, the OTCV in the case of extreme-value distributed unobserved shocks is implied by the following equality:

$$V_{i,t}(w_{i,t}, y_{i,t}) = \log \left( \sum_{k=0}^{2} \exp \left( u_k(w_{i,t}, y_{i,t} - OTCV_{i,t}) + \beta E \left[ V_{i,t+1}(w_{i,t+1}, y_{i,t+1}) | w'_{i,t}, y'_{i,t}, k \right] \right) \right),$$

where $V_{i,t}(\cdot, \cdot)$ is the ex ante value function, $y_{i,t}$ is the income of the borrower, and $w_{i,t}$ is the vector of all other state variables. The values $(w_{i,t}, y_{i,t})$ and $(w'_{i,t}, y'_{i,t})$ correspond to the values before and after the counterfactual change, respectively. Table 13 reports the average one-time compensating variation across borrowers in the sample for each scenario.\(^{31}\)

[Table 13 about here]

As expected, we see that scenarios that would lead to a large reduction in the probability of default have large OTCV. For instance, our results show that a uniform 10% reduction in outstanding mortgage balance for the pool of borrowers in our sample would reduce the overall default probability by 22%, and that each borrower would be willing to pay up to $16,643 on average to enjoy the principal writedown instead of going back to the original circumstances. Under a uniform 20% reduction in outstanding mortgage balance, the overall default probability would go down by 38%, and each borrower would be willing to pay up to $29,511 on average to enjoy the principal writedown instead of going back to the original circumstances.\(^{32,33}\)

It is interesting to note that the average dollar amount of the 10% principal writedown, computed using the original loan balance, is $17,723, which is very close to our one-time compensating variation figure of $16,643. There is nothing in our model that constrains our welfare figure to be close to the actual dollar amount of the policy. The increase in consumer welfare under the policy simply comes from the fact that borrowers make lower payments and accordingly have lower default probability today and in subsequent periods. It is thus reassuring that our estimate of the welfare effect of lower payments and lower default probability is close to the dollar value of the policy.

\(^{31}\)In computing the OTCV, we determine the utility of money using the estimated coefficient on monthly payments instead of the estimated coefficient on income for two reasons. First, some of the stated incomes may be misrepresented in the data. Especially in the case of low-documentation loans, income falsification is a concern (see Jiang, Nelson and Vyltaclil (2011)). Second, the coefficient on income may reflect not just the effect of income in the utility of each action, but also additional, confounding effects such as the impact of income on the ability to refinance. The monthly payment variable is less likely to suffer from such problems.

\(^{32}\)In Table 11, the changes in default probability under scenarios 1 and 6 are of similar magnitudes, whereas Tables 12 and 13 indicate that the welfare change under scenario 1 is about twice as large in magnitude as under scenario 6. This seeming “inconsistency” is due to the fact that the two scenarios are not directly comparable because the sets of loans examined are different and thus the baseline default probability differs. For instance, a 1 percentage point increase in default probability will have a different impact on overall welfare depending on whether the baseline default probability is 5% or 50%. Since the baseline default probability is much higher under scenario 6, it is not surprising that the welfare change in the absolute level is smaller under scenario 6 than under scenario 1.

\(^{33}\)When we make an alternative normalization setting default utility to zero, our welfare figures become smaller, but the results remain qualitatively similar.
7 Conclusion

Our paper makes a few methodological contributions. First, we propose an estimation method for dynamic discrete-choice problems with a finite horizon that is intuitive and easy to implement. Second, we prove that we can identify the discount factor in finite-horizon dynamic discrete choice models without relying on exclusion restrictions, identification at infinity, or data from the final period.

To illustrate our estimation methodology, we have estimated and simulated a dynamic structural model of mortgage default. Using our model, we have quantified the importance of home price declines and looser underwriting standards in creating the conditions that led to the recent wave of mortgage defaults. We have also used the model to investigate the impact of various aspects of interventions that have been proposed by regulators in response to the mortgage crisis.

References


Table 1: Definition of Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Owner-Occupier</td>
<td>= 1 if the loan is made with the intent to occupy the house, = 0 if for investment.</td>
</tr>
<tr>
<td>Purchase Loan</td>
<td>= 1 if purpose of the loan is house purchase, = 0 if refinance loan without cash-out.</td>
</tr>
<tr>
<td>Low Doc</td>
<td>= 1 if the loan was done with no or low documentation, = 0 otherwise.</td>
</tr>
<tr>
<td>Multiple Liens</td>
<td>= 1 if the borrower has other, junior mortgages, = 0 otherwise.</td>
</tr>
<tr>
<td>FICO</td>
<td>FICO score, a credit score developed by Fair Issac &amp; Co.</td>
</tr>
<tr>
<td></td>
<td>Scores range between 300 and 850. Higher scores indicate higher credit quality.</td>
</tr>
<tr>
<td>Income</td>
<td>Monthly income.</td>
</tr>
<tr>
<td>Contractual Interest Rate</td>
<td>Interest rate specified by the mortgage.</td>
</tr>
<tr>
<td>Market Rate</td>
<td>Current market interest rate. We impute borrower-specific market rate.</td>
</tr>
<tr>
<td>Housing Value</td>
<td>Current housing value.</td>
</tr>
<tr>
<td>Net Equity</td>
<td>Current housing value - Outstanding loan balance.</td>
</tr>
<tr>
<td>Payment</td>
<td>Monthly payment due.</td>
</tr>
<tr>
<td>Prepayment Penalty</td>
<td>= 1 if the loan has prepayment penalty, = 0 otherwise.</td>
</tr>
<tr>
<td>Unemployment Rate</td>
<td>Monthly unemployment rate at the county level.</td>
</tr>
</tbody>
</table>
Table 2: Summary Statistics for Estimation Sample

<table>
<thead>
<tr>
<th>Time-invariant loan-level variables</th>
<th>Defaulted</th>
<th>Prepaid</th>
<th>Censored</th>
<th>All loans</th>
<th>Mean</th>
<th>Std. Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Owner-Occupier Loans</td>
<td>87.5</td>
<td>75.8</td>
<td>63.5</td>
<td>73.9</td>
<td>43.9</td>
<td></td>
</tr>
<tr>
<td>% Purchase Loans</td>
<td>59.9</td>
<td>60.0</td>
<td>55.3</td>
<td>58.6</td>
<td>49.2</td>
<td></td>
</tr>
<tr>
<td>FICO Score</td>
<td>626.4</td>
<td>674.6</td>
<td>691.7</td>
<td>672.6</td>
<td>74.2</td>
<td></td>
</tr>
<tr>
<td>Monthly Payment</td>
<td>1287.3</td>
<td>1406.1</td>
<td>1236.5</td>
<td>1339.9</td>
<td>1163.6</td>
<td></td>
</tr>
<tr>
<td>% Low Documentation Loans</td>
<td>36.4</td>
<td>41.7</td>
<td>38.8</td>
<td>40.1</td>
<td>49</td>
<td></td>
</tr>
<tr>
<td>Contractual Interest Rate</td>
<td>8.528</td>
<td>7.854</td>
<td>6.722</td>
<td>7.624</td>
<td>1.75</td>
<td></td>
</tr>
<tr>
<td>Appraisal Value of House ($m)</td>
<td>0.214</td>
<td>0.272</td>
<td>0.274</td>
<td>0.264</td>
<td>0.218</td>
<td></td>
</tr>
<tr>
<td>Monthly Income</td>
<td>5239.4</td>
<td>6342.9</td>
<td>6340.8</td>
<td>6182.7</td>
<td>6497.8</td>
<td></td>
</tr>
<tr>
<td>% Loans with Multiple Liens</td>
<td>23.8</td>
<td>9.9</td>
<td>19.1</td>
<td>14.6</td>
<td>35.3</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time-dependent loan-level variables, as last observed for each loan</th>
<th>Defaulted</th>
<th>Prepaid</th>
<th>Censored</th>
<th>All loans</th>
<th>Mean</th>
<th>Std. Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loan Age in Months</td>
<td>31.44</td>
<td>24.41</td>
<td>63.38</td>
<td>36.7</td>
<td>23.7</td>
<td></td>
</tr>
<tr>
<td>% Loans with Prepayment Penalty</td>
<td>43.8</td>
<td>30</td>
<td>2.5</td>
<td>24.0</td>
<td>42.7</td>
<td></td>
</tr>
<tr>
<td>Net Equity ($m)</td>
<td>0.038</td>
<td>0.133</td>
<td>0.079</td>
<td>0.103</td>
<td>0.146</td>
<td></td>
</tr>
<tr>
<td>Current Housing Value ($m)</td>
<td>0.207</td>
<td>0.327</td>
<td>0.255</td>
<td>0.289</td>
<td>0.245</td>
<td></td>
</tr>
<tr>
<td>Monthly HPI Growth Rate (%)</td>
<td>-0.2</td>
<td>0.8</td>
<td>-0.4</td>
<td>0.3</td>
<td>1.7</td>
<td></td>
</tr>
<tr>
<td>Market Interest Rate</td>
<td>8.26</td>
<td>7.202</td>
<td>6.221</td>
<td>7.066</td>
<td>1.89</td>
<td></td>
</tr>
<tr>
<td>Unemployment Rate</td>
<td>6.666</td>
<td>5.348</td>
<td>9.668</td>
<td>6.787</td>
<td>2.714</td>
<td></td>
</tr>
<tr>
<td>Number of Loans</td>
<td>1729</td>
<td>6770</td>
<td>3456</td>
<td>11955</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3: Summary Statistics by Origination Year

<table>
<thead>
<tr>
<th></th>
<th>2000</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Loans that Default by 12/2009</td>
<td>15.7</td>
<td>11.35</td>
<td>10.6</td>
<td>6.86</td>
<td>8.47</td>
<td>22.83</td>
<td>36.01</td>
<td>31.88</td>
</tr>
<tr>
<td>% Loans that Prepay by 12/2009</td>
<td>81.83</td>
<td>85.10</td>
<td>78.5</td>
<td>58.88</td>
<td>48.23</td>
<td>30.42</td>
<td>15.5</td>
<td>11.88</td>
</tr>
<tr>
<td>Avg Duration to Default in Months</td>
<td>27.69</td>
<td>32.19</td>
<td>33.64</td>
<td>40</td>
<td>39.3</td>
<td>33.52</td>
<td>25.8</td>
<td>21.18</td>
</tr>
<tr>
<td>Avg Duration to Prepayment in Months</td>
<td>25.61</td>
<td>24.79</td>
<td>23.88</td>
<td>26.35</td>
<td>23.05</td>
<td>20.52</td>
<td>19.57</td>
<td>17.64</td>
</tr>
<tr>
<td>% Owner-Occupier Loans</td>
<td>85.74</td>
<td>94.4</td>
<td>78.04</td>
<td>50.2</td>
<td>62.73</td>
<td>84.08</td>
<td>91.02</td>
<td>87.83</td>
</tr>
<tr>
<td>% Purchase Loans</td>
<td>78.2</td>
<td>53.7</td>
<td>49.3</td>
<td>51.2</td>
<td>67.8</td>
<td>52.2</td>
<td>64.1</td>
<td>57.4</td>
</tr>
<tr>
<td>FICO Score</td>
<td>640.7</td>
<td>656.2</td>
<td>669.6</td>
<td>697.4</td>
<td>702.1</td>
<td>655.7</td>
<td>648.4</td>
<td>656</td>
</tr>
<tr>
<td>Monthly Payment</td>
<td>1262.6</td>
<td>1727</td>
<td>1320.1</td>
<td>1190</td>
<td>1328.1</td>
<td>1206.6</td>
<td>1551.8</td>
<td>1556.8</td>
</tr>
<tr>
<td>% Low Documentation Loans</td>
<td>34.3</td>
<td>38.9</td>
<td>35</td>
<td>42.9</td>
<td>41.1</td>
<td>43.9</td>
<td>39.1</td>
<td>52.2</td>
</tr>
<tr>
<td>Contractual Interest Rate</td>
<td>10.193</td>
<td>8.971</td>
<td>7.768</td>
<td>6.625</td>
<td>6.554</td>
<td>6.984</td>
<td>7.778</td>
<td>7.816</td>
</tr>
<tr>
<td>Appraisal Value of House ($m)</td>
<td>0.184</td>
<td>0.309</td>
<td>0.251</td>
<td>0.273</td>
<td>0.292</td>
<td>0.246</td>
<td>0.284</td>
<td>0.293</td>
</tr>
<tr>
<td>Monthly Income</td>
<td>5209.6</td>
<td>7099</td>
<td>6049.6</td>
<td>5682.5</td>
<td>7262.3</td>
<td>5360.6</td>
<td>6554.2</td>
<td>7667</td>
</tr>
<tr>
<td>% Loans with Multiple Liens</td>
<td>0.6</td>
<td>0.6</td>
<td>2</td>
<td>10.7</td>
<td>24.3</td>
<td>27.4</td>
<td>42</td>
<td>19.7</td>
</tr>
<tr>
<td>% Loans with Prepayment Penalty</td>
<td>33.8</td>
<td>35.5</td>
<td>28.9</td>
<td>15.7</td>
<td>13.1</td>
<td>24.8</td>
<td>25</td>
<td>42</td>
</tr>
<tr>
<td>Net Equity ($m)*</td>
<td>0.073</td>
<td>0.141</td>
<td>0.124</td>
<td>0.153</td>
<td>0.118</td>
<td>0.046</td>
<td>0.021</td>
<td>0.016</td>
</tr>
<tr>
<td>Monthly HPI Growth Rate (%)*</td>
<td>0.632</td>
<td>0.775</td>
<td>0.811</td>
<td>0.43</td>
<td>0.229</td>
<td>-0.4</td>
<td>-0.765</td>
<td>-0.699</td>
</tr>
<tr>
<td>Number of Loans</td>
<td>1431</td>
<td>1215</td>
<td>1726</td>
<td>2697</td>
<td>2007</td>
<td>1476</td>
<td>1058</td>
<td>345</td>
</tr>
</tbody>
</table>

*: as last observed for each loan

Table 4: Duration to Default or Prepayment

<table>
<thead>
<tr>
<th>Duration to Default (in months)</th>
<th>1-6</th>
<th>7-12</th>
<th>13-18</th>
<th>19-24</th>
<th>25-30</th>
<th>31-36</th>
<th>37-42</th>
<th>43-48</th>
<th>&gt; 48 months</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.89%</td>
<td>14.11%</td>
<td>12.96%</td>
<td>11.91%</td>
<td>10.87%</td>
<td>12.32%</td>
<td>8.79%</td>
<td>6.82%</td>
<td>19.32%</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Duration to Prepayment (in months)</th>
<th>1-6</th>
<th>7-12</th>
<th>13-18</th>
<th>19-24</th>
<th>25-30</th>
<th>31-36</th>
<th>37-42</th>
<th>43-48</th>
<th>&gt; 48 months</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.91%</td>
<td>18.29%</td>
<td>23.15%</td>
<td>17.09%</td>
<td>12.02%</td>
<td>8.60%</td>
<td>6.31%</td>
<td>3.41%</td>
<td>9.23%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Default Probability</td>
<td>Prepay Probability</td>
<td>Pay Probability</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------</td>
<td>---------------------</td>
<td>--------------------</td>
<td>-----------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Prediction Data</td>
<td>Prediction Data</td>
<td>Prediction Data</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All</td>
<td>0.0046 0.0047</td>
<td>0.0178 0.0179</td>
<td>0.9775 0.9774</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FICO G1</td>
<td>0.0107 0.0109</td>
<td>0.0209 0.0209</td>
<td>0.9684 0.9682</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FICO G2</td>
<td>0.0085 0.0083</td>
<td>0.0168 0.0171</td>
<td>0.9747 0.9746</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FICO G3</td>
<td>0.0059 0.0060</td>
<td>0.0187 0.0188</td>
<td>0.9754 0.9753</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FICO G4</td>
<td>0.0025 0.0026</td>
<td>0.0194 0.0190</td>
<td>0.9780 0.9784</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FICO G5</td>
<td>0.0011 0.0010</td>
<td>0.0167 0.0166</td>
<td>0.9822 0.9824</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FICO G6</td>
<td>0.0006 0.0010</td>
<td>0.0123 0.0138</td>
<td>0.9872 0.9852</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Payment G1</td>
<td>0.0032 0.0030</td>
<td>0.0130 0.0136</td>
<td>0.9838 0.9834</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Payment G2</td>
<td>0.0041 0.0044</td>
<td>0.0132 0.0127</td>
<td>0.9826 0.9829</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Payment G3</td>
<td>0.0049 0.0047</td>
<td>0.0152 0.0154</td>
<td>0.9798 0.9799</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Payment G4</td>
<td>0.0051 0.0054</td>
<td>0.0194 0.0190</td>
<td>0.9755 0.9756</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Payment G5</td>
<td>0.0048 0.0048</td>
<td>0.0230 0.0235</td>
<td>0.9722 0.9717</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Payment G6</td>
<td>0.0050 0.0047</td>
<td>0.0247 0.0252</td>
<td>0.9704 0.9700</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low Doc = 0</td>
<td>0.0048 0.0048</td>
<td>0.0168 0.0169</td>
<td>0.9784 0.9782</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low Doc = 1</td>
<td>0.0045 0.0044</td>
<td>0.0196 0.0195</td>
<td>0.9759 0.9760</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multi Liens = 0</td>
<td>0.0041 0.0042</td>
<td>0.0188 0.0188</td>
<td>0.9771 0.9770</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multi Liens = 1</td>
<td>0.0079 0.0078</td>
<td>0.0118 0.0124</td>
<td>0.9804 0.9798</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Net Equity G1</td>
<td>0.0147 0.0147</td>
<td>0.0045 0.0041</td>
<td>0.9808 0.9812</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Net Equity G2</td>
<td>0.0067 0.0066</td>
<td>0.0117 0.0120</td>
<td>0.9816 0.9814</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Net Equity G3</td>
<td>0.0047 0.0048</td>
<td>0.0181 0.0179</td>
<td>0.9772 0.9774</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Net Equity G4</td>
<td>0.0027 0.0030</td>
<td>0.0234 0.0234</td>
<td>0.9738 0.9736</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Net Equity G5</td>
<td>0.0015 0.0014</td>
<td>0.0233 0.0244</td>
<td>0.9753 0.9742</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Net Equity G6</td>
<td>0.0007 0.0006</td>
<td>0.0181 0.0174</td>
<td>0.9812 0.9820</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Likelihood Ratio Index: 0.1525

This table examines probability of default/prepay/pay in each period. G1: bottom 10% in the specified variable; G2: 10%-25%; G3: 25%-50%; G4: 50%-75%; G5: 75%-90%; G6: top 10%.
Table 6: Out-of-Sample Fit of First-Step Estimates

<table>
<thead>
<tr>
<th></th>
<th>Default Probability</th>
<th>Prepay Probability</th>
<th>Pay Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Prediction</td>
<td>Data</td>
<td>Prediction</td>
</tr>
<tr>
<td>All</td>
<td>0.0046</td>
<td>0.0048</td>
<td>0.0179</td>
</tr>
<tr>
<td>FICO G1</td>
<td>0.0106</td>
<td>0.0111</td>
<td>0.0210</td>
</tr>
<tr>
<td>FICO G2</td>
<td>0.0080</td>
<td>0.0091</td>
<td>0.0170</td>
</tr>
<tr>
<td>FICO G3</td>
<td>0.0060</td>
<td>0.0058</td>
<td>0.0186</td>
</tr>
<tr>
<td>FICO G4</td>
<td>0.0025</td>
<td>0.0027</td>
<td>0.0197</td>
</tr>
<tr>
<td>FICO G5</td>
<td>0.0011</td>
<td>0.0010</td>
<td>0.0164</td>
</tr>
<tr>
<td>FICO G6</td>
<td>0.0007</td>
<td>0.0010</td>
<td>0.0124</td>
</tr>
<tr>
<td>Payment G1</td>
<td>0.0035</td>
<td>0.0028</td>
<td>0.0138</td>
</tr>
<tr>
<td>Payment G2</td>
<td>0.0040</td>
<td>0.0047</td>
<td>0.0137</td>
</tr>
<tr>
<td>Payment G3</td>
<td>0.0047</td>
<td>0.0049</td>
<td>0.0153</td>
</tr>
<tr>
<td>Payment G4</td>
<td>0.0051</td>
<td>0.0054</td>
<td>0.0190</td>
</tr>
<tr>
<td>Payment G5</td>
<td>0.0048</td>
<td>0.0048</td>
<td>0.0225</td>
</tr>
<tr>
<td>Payment G6</td>
<td>0.0049</td>
<td>0.0053</td>
<td>0.0251</td>
</tr>
<tr>
<td>Low Doc = 0</td>
<td>0.0047</td>
<td>0.0049</td>
<td>0.0167</td>
</tr>
<tr>
<td>Low Doc = 1</td>
<td>0.0044</td>
<td>0.0046</td>
<td>0.0198</td>
</tr>
<tr>
<td>Multi Liens = 0</td>
<td>0.0041</td>
<td>0.0043</td>
<td>0.0189</td>
</tr>
<tr>
<td>Multi Liens = 1</td>
<td>0.0077</td>
<td>0.0081</td>
<td>0.0114</td>
</tr>
<tr>
<td>Net Equity G1</td>
<td>0.0144</td>
<td>0.0144</td>
<td>0.0043</td>
</tr>
<tr>
<td>Net Equity G2</td>
<td>0.0068</td>
<td>0.0069</td>
<td>0.0119</td>
</tr>
<tr>
<td>Net Equity G3</td>
<td>0.0049</td>
<td>0.0046</td>
<td>0.0183</td>
</tr>
<tr>
<td>Net Equity G4</td>
<td>0.0025</td>
<td>0.0035</td>
<td>0.0233</td>
</tr>
<tr>
<td>Net Equity G5</td>
<td>0.0012</td>
<td>0.0017</td>
<td>0.0230</td>
</tr>
<tr>
<td>Net Equity G6</td>
<td>0.0008</td>
<td>0.0005</td>
<td>0.0188</td>
</tr>
</tbody>
</table>

This table examines probability of default/prepay/pay in each period. Also see the notes in Table 5.
Table 7: Simulated Probability of Eventual Default or Prepay by End of 2009

<table>
<thead>
<tr>
<th></th>
<th>Prob. Default</th>
<th>Duration to Default</th>
<th>Prob. Prepay</th>
<th>Duration to Prepay</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>All</strong></td>
<td>Prediction</td>
<td>0.1482</td>
<td>34.1782</td>
<td>0.5536</td>
</tr>
<tr>
<td></td>
<td>Data</td>
<td>0.1446</td>
<td>31.834</td>
<td>0.5663</td>
</tr>
<tr>
<td><strong>FICO G1</strong></td>
<td>Prediction</td>
<td>0.2968</td>
<td>28.5291</td>
<td>0.5423</td>
</tr>
<tr>
<td></td>
<td>Data</td>
<td>0.2828</td>
<td>27.7744</td>
<td>0.5584</td>
</tr>
<tr>
<td><strong>FICO G2</strong></td>
<td>Prediction</td>
<td>0.2399</td>
<td>30.8991</td>
<td>0.4972</td>
</tr>
<tr>
<td></td>
<td>Data</td>
<td>0.2438</td>
<td>30.2051</td>
<td>0.5026</td>
</tr>
<tr>
<td><strong>FICO G3</strong></td>
<td>Prediction</td>
<td>0.1778</td>
<td>33.514</td>
<td>0.5391</td>
</tr>
<tr>
<td></td>
<td>Data</td>
<td>0.1743</td>
<td>33.3803</td>
<td>0.5603</td>
</tr>
<tr>
<td><strong>FICO G4</strong></td>
<td>Prediction</td>
<td>0.085</td>
<td>37.2138</td>
<td>0.6103</td>
</tr>
<tr>
<td></td>
<td>Data</td>
<td>0.0803</td>
<td>36.7915</td>
<td>0.6178</td>
</tr>
<tr>
<td><strong>FICO G5</strong></td>
<td>Prediction</td>
<td>0.0463</td>
<td>40.5877</td>
<td>0.6022</td>
</tr>
<tr>
<td></td>
<td>Data</td>
<td>0.0353</td>
<td>39.2546</td>
<td>0.588</td>
</tr>
<tr>
<td><strong>FICO G6</strong></td>
<td>Prediction</td>
<td>0.0284</td>
<td>41.8558</td>
<td>0.4859</td>
</tr>
<tr>
<td></td>
<td>Data</td>
<td>0.0389</td>
<td>30.381</td>
<td>0.5357</td>
</tr>
<tr>
<td><strong>Net Equity G1</strong></td>
<td>Prediction</td>
<td>0.2869</td>
<td>32.3095</td>
<td>0.4235</td>
</tr>
<tr>
<td></td>
<td>Data</td>
<td>0.2861</td>
<td>28.5546</td>
<td>0.4315</td>
</tr>
<tr>
<td><strong>Net Equity G2</strong></td>
<td>Prediction</td>
<td>0.2148</td>
<td>33.7037</td>
<td>0.5125</td>
</tr>
<tr>
<td></td>
<td>Data</td>
<td>0.2057</td>
<td>31.9563</td>
<td>0.5451</td>
</tr>
<tr>
<td><strong>Net Equity G3</strong></td>
<td>Prediction</td>
<td>0.1511</td>
<td>34.3932</td>
<td>0.5502</td>
</tr>
<tr>
<td></td>
<td>Data</td>
<td>0.1466</td>
<td>32.2371</td>
<td>0.5618</td>
</tr>
<tr>
<td><strong>Net Equity G4</strong></td>
<td>Prediction</td>
<td>0.1034</td>
<td>34.2837</td>
<td>0.599</td>
</tr>
<tr>
<td></td>
<td>Data</td>
<td>0.1082</td>
<td>32.6803</td>
<td>0.614</td>
</tr>
<tr>
<td><strong>Net Equity G5</strong></td>
<td>Prediction</td>
<td>0.0635</td>
<td>35.3587</td>
<td>0.6209</td>
</tr>
<tr>
<td></td>
<td>Data</td>
<td>0.0562</td>
<td>34.1</td>
<td>0.6169</td>
</tr>
<tr>
<td><strong>Net Equity G6</strong></td>
<td>Prediction</td>
<td>0.0275</td>
<td>40.3674</td>
<td>0.5767</td>
</tr>
<tr>
<td></td>
<td>Data</td>
<td>0.0164</td>
<td>39.1111</td>
<td>0.5327</td>
</tr>
</tbody>
</table>

This table examines probability of eventual default/prepay by the end of 2009. Duration is measured in months. Also see the notes in Table 5.
Table 8: Structural Estimates of Per-Period Utility

<table>
<thead>
<tr>
<th></th>
<th>Default</th>
<th>Prepay</th>
<th>Pay</th>
</tr>
</thead>
<tbody>
<tr>
<td>FICO</td>
<td>-0.615 (0.0819) ***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSA dummies</td>
<td>0.0216 (0.0335)</td>
<td>0.0629 (0.0182) ***</td>
<td></td>
</tr>
<tr>
<td>Housing Value</td>
<td></td>
<td>0.0205 (0.0049) ***</td>
<td>-0.0205 (0.0049) ***</td>
</tr>
<tr>
<td>Monthly Payment</td>
<td>-0.0205 (0.0049) ***</td>
<td>-0.0205 (0.0049) ***</td>
<td></td>
</tr>
<tr>
<td>Prepayment Penalty</td>
<td>0.1343 (0.0097) ***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income</td>
<td>-0.0034 (0.0010) ***</td>
<td>0.0022 (0.0002) ***</td>
<td></td>
</tr>
<tr>
<td>Unemployment Rate</td>
<td>-0.2304 (0.0051) ***</td>
<td>-0.0110 (0.0030) ***</td>
<td></td>
</tr>
<tr>
<td>Low Doc</td>
<td>0.0185 (0.0101) *</td>
<td>-0.0210 (0.0022) ***</td>
<td></td>
</tr>
<tr>
<td>Multiple Liens</td>
<td>0.0592 (0.0157) ***</td>
<td>-0.0060 (0.0042)</td>
<td></td>
</tr>
<tr>
<td>$\beta$ (coeff on $\hat{E}[V_{t+1}(s_{t,t+1})</td>
<td>s_{t,t}, a_{t,t}]$)</td>
<td>0.9815 (0.0033) ***</td>
<td>0.9815 (0.0033) ***</td>
</tr>
<tr>
<td>No. of Obs</td>
<td>400493</td>
<td>400493</td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.8675</td>
<td>0.9935</td>
<td></td>
</tr>
</tbody>
</table>

SUR with constraints that coefficients on monthly payment and $\hat{E}[V_{t+1}(s_{t,t+1})|s_{t,t}, a_{t,t}]$ are the same between the payment and prepayment equations. Bootstrapped standard errors are reported to account for estimation errors in first-step policy functions and ex ante value functions.
### Table 9: Estimates of Per-Period Utility for Various Subsamples

<table>
<thead>
<tr>
<th></th>
<th>Loans for Home Purchase</th>
<th></th>
<th>Loans for Refinancing</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Default</td>
<td>Prepay</td>
<td>Pay</td>
<td>Default</td>
</tr>
<tr>
<td><strong>FICO</strong></td>
<td>-0.94(0.11)***</td>
<td></td>
<td>-1.04(0.13)***</td>
<td></td>
</tr>
<tr>
<td><strong>MSA</strong></td>
<td>Included</td>
<td></td>
<td></td>
<td>Included</td>
</tr>
<tr>
<td>Housing Value</td>
<td>0.10(0.05)***</td>
<td>0.09(0.02)***</td>
<td>0.11(0.08)</td>
<td>0.07(0.06)</td>
</tr>
<tr>
<td>Payment</td>
<td>-0.04(0.005)***</td>
<td>-0.04(0.005)***</td>
<td>0.02(0.02)</td>
<td>0.02(0.015)</td>
</tr>
<tr>
<td>P_Penalty</td>
<td>0.005(0.02)</td>
<td></td>
<td>0.32(0.02)***</td>
<td></td>
</tr>
<tr>
<td>Income</td>
<td>-0.004(0.001)***</td>
<td>0.003(0.0003)***</td>
<td>-0.01(0.002)***</td>
<td>0.001(0.001)</td>
</tr>
<tr>
<td>Unem. Rate</td>
<td>-0.23(0.01)***</td>
<td>-0.003(0.001)***</td>
<td>-0.26(0.01)***</td>
<td>-0.02(0.008)***</td>
</tr>
<tr>
<td>Low Doc</td>
<td>0.05(0.02)***</td>
<td>-0.02(0.002)***</td>
<td>-0.11(0.02)***</td>
<td>-0.01(0.005)**</td>
</tr>
<tr>
<td>Multi-Liens</td>
<td>0.15(0.02)***</td>
<td>-0.001(0.003)</td>
<td>-0.18(0.04)***</td>
<td>-0.003(0.01)</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.99(0.003)***</td>
<td>0.99(0.003)***</td>
<td>0.97(0.008)***</td>
<td>0.97(0.008)***</td>
</tr>
<tr>
<td>No. of Obs</td>
<td>221612</td>
<td>221612</td>
<td></td>
<td>178881</td>
</tr>
<tr>
<td>R²</td>
<td>0.839</td>
<td>0.994</td>
<td></td>
<td>0.942</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Loans for Home Purchase</th>
<th></th>
<th>Loans for Refinancing</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Default</td>
<td>Prepay</td>
<td>Pay</td>
<td>Default</td>
</tr>
<tr>
<td><strong>FICO</strong></td>
<td>-0.84(0.15)***</td>
<td></td>
<td>0.05(0.5)</td>
<td></td>
</tr>
<tr>
<td><strong>MSA</strong></td>
<td>Included</td>
<td></td>
<td></td>
<td>Included</td>
</tr>
<tr>
<td>Housing Value</td>
<td>0.33(0.04)***</td>
<td>0.13(0.02)***</td>
<td>-0.28(0.26)</td>
<td>0.96(0.24)***</td>
</tr>
<tr>
<td>Payment</td>
<td>-0.05(0.005)***</td>
<td>-0.05(0.005)***</td>
<td>-0.19(0.05)***</td>
<td>-0.19(0.05)***</td>
</tr>
<tr>
<td>P_Penalty</td>
<td>0.006(0.01)</td>
<td></td>
<td>0.21(0.04)***</td>
<td></td>
</tr>
<tr>
<td>Income</td>
<td>-0.01(0.002)***</td>
<td>0.004(0.001)***</td>
<td>-0.01(0.002)***</td>
<td>0.000(0.0004)</td>
</tr>
<tr>
<td>Unem. Rate</td>
<td>-0.25(0.01)***</td>
<td>-0.009(0.004)**</td>
<td>-0.19(0.01)***</td>
<td>0.001(0.001)</td>
</tr>
<tr>
<td>Low Doc</td>
<td>-0.02(0.01)</td>
<td>-0.03(0.003)***</td>
<td>0.004(0.02)</td>
<td>0.006(0.006)</td>
</tr>
<tr>
<td>Multi-Liens</td>
<td>-0.08(0.02)***</td>
<td>-0.0002(0.01)</td>
<td>1.13(0.08)***</td>
<td>0.21(0.02)***</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.98(0.004)***</td>
<td>0.98(0.004)***</td>
<td>0.98(0.005)***</td>
<td>0.98(0.005)***</td>
</tr>
<tr>
<td>No. of Obs</td>
<td>270093</td>
<td>270093</td>
<td></td>
<td>130356</td>
</tr>
<tr>
<td>R²</td>
<td>0.856</td>
<td>0.987</td>
<td></td>
<td>0.963</td>
</tr>
</tbody>
</table>

SUR with constraints that coefficients on monthly payment and $E[V_{i+1}(s_{i,t+1}|s_{i,t}, a_{i,t})$ are the same
between the payment and prepayment equations. Bootstrapped standard errors are reported to account for estimation errors in first-step policy functions and ex ante value functions.

Table 10: Estimates after Controlling for Liquidity Shocks through Parametric Assumption

<table>
<thead>
<tr>
<th></th>
<th>Default</th>
<th>Prepay</th>
<th>Pay</th>
</tr>
</thead>
<tbody>
<tr>
<td>FICO</td>
<td>-1.2462 (0.1315) ***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSA dummies</td>
<td></td>
<td>Included</td>
<td></td>
</tr>
<tr>
<td>Housing Value</td>
<td>0.5242 (0.3923)</td>
<td>0.4394 (0.2196) **</td>
<td></td>
</tr>
<tr>
<td>Monthly Payment</td>
<td>-0.1351 (0.0530) **</td>
<td>-0.1351 (0.0530) **</td>
<td></td>
</tr>
<tr>
<td>Prepayment Penalty</td>
<td>0.1424 (0.0459) ***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income</td>
<td>0.004 (0.0024)</td>
<td>0.0066 (0.0010) ***</td>
<td></td>
</tr>
<tr>
<td>Unemployment Rate</td>
<td>-0.2368 (0.0097) ***</td>
<td>-0.0405 (0.009) ***</td>
<td></td>
</tr>
<tr>
<td>Low Doc</td>
<td>-0.1053 (0.0256) ***</td>
<td>-0.0475 (0.0091) ***</td>
<td></td>
</tr>
<tr>
<td>Multiple Liens</td>
<td>0.008 (0.0399)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[ \beta (\text{coeff on } \hat{E}[V_{t+1}(s_{i,t+1})|s_{i,t}, a_{i,t}]) \]

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta )</td>
<td>0.9336 (0.0168) ***</td>
<td>0.9336 (0.0168) ***</td>
</tr>
<tr>
<td>No. of Obs</td>
<td>198122</td>
<td>198122</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>0.7456</td>
<td>0.909</td>
</tr>
</tbody>
</table>

SUR with constraints that coefficients on monthly payment and \( \hat{E}[V_{t+1}(s_{i,t+1})|s_{i,t}, a_{i,t}] \) are the same between the payment and prepayment equations. Bootstrapped standard errors are reported to account for estimation errors in first-step policy functions and ex ante value functions.

Table 11: Counterfactual Analyses

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Counterfactual</th>
<th>Baseline</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Default</td>
<td>Prepay</td>
</tr>
<tr>
<td>Scenario 1: Home Price Decline</td>
<td>0.1129</td>
<td>0.4717</td>
</tr>
<tr>
<td>Scenario 2: Home Price Increase</td>
<td>0.2204</td>
<td>0.2635</td>
</tr>
<tr>
<td>Scenario 3: Higher Credit Quality</td>
<td>0.2329</td>
<td>0.2232</td>
</tr>
<tr>
<td>Scenario 4: 10% Principal Writedown</td>
<td>0.1167</td>
<td>0.5783</td>
</tr>
<tr>
<td>Scenario 5: 20% Principal Writedown</td>
<td>0.0914</td>
<td>0.5938</td>
</tr>
<tr>
<td>Scenario 6: LTV Cap at 0.8</td>
<td>0.1316</td>
<td>0.5692</td>
</tr>
<tr>
<td>Scenario 7: LTV Cap at 0.9</td>
<td>0.1458</td>
<td>0.556</td>
</tr>
</tbody>
</table>
This table examines probability of eventual default/prepay by the end of 2009. The first (second) column reports predicted probability of eventual default or prepay by December 2009 under the specified counterfactual (baseline) scenario. The baseline default and prepayment probabilities differ across scenarios since different scenarios examine different subgroups. For instance, Scenarios 4-7 examine all loans while Scenario 1 examines loans originated in 2004. For Scenario 1, the cumulative housing price growth rate over the sample period is 11.8% under the baseline case while it is only 5.6% under the counterfactual case. For Scenario 2, the cumulative housing price growth rate over the sample period is 4.5% under the baseline case while it is 20.9% under the counterfactual scenario.

Table 12: Borrowers’ Welfare, Ex Ante Values

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Counterfactual</th>
<th>Baseline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scenario 1: Home Price Decline</td>
<td>0.5584</td>
<td>0.7167</td>
</tr>
<tr>
<td>Scenario 2: Home Price Increase</td>
<td>0.1716</td>
<td>-0.1353</td>
</tr>
<tr>
<td>Scenario 3: Higher Credit Quality</td>
<td>-0.0201</td>
<td>-0.1353</td>
</tr>
<tr>
<td>Scenario 4: 10% Principal Writedown</td>
<td>-0.0744</td>
<td>-0.3569</td>
</tr>
<tr>
<td>Scenario 5: 20% Principal Writedown</td>
<td>0.2291</td>
<td>-0.3569</td>
</tr>
<tr>
<td>Scenario 6: LTV Cap at 0.8</td>
<td>-0.273</td>
<td>-0.3569</td>
</tr>
<tr>
<td>Scenario 7: LTV Cap at 0.9</td>
<td>-0.3456</td>
<td>-0.3569</td>
</tr>
</tbody>
</table>

This table compares borrowers’ welfare under the specified counterfactual scenarios against welfare under the baseline. Our measure of welfare is the average ex ante values of borrowers. The baseline welfare figures differ across scenarios since different scenarios examine different subgroups. For instance, Scenarios 4-7 examine all loans while Scenario 1 examines loans originated in 2004.

Table 13: Borrowers’ Welfare, One-Time Compensating Variation

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Average OTCV per borrower</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scenario 1: Home Price Decline</td>
<td>-$8976.7</td>
</tr>
<tr>
<td>Scenario 2: Home Price Increase</td>
<td>$16812.2</td>
</tr>
<tr>
<td>Scenario 3: Higher Credit Quality</td>
<td>$6414.8</td>
</tr>
<tr>
<td>Scenario 4: 10% Principal Writedown</td>
<td>$16643.6</td>
</tr>
<tr>
<td>Scenario 5: 20% Principal Writedown</td>
<td>$29510.7</td>
</tr>
<tr>
<td>Scenario 6: LTV Cap at 0.8</td>
<td>$4932.2</td>
</tr>
<tr>
<td>Scenario 7: LTV Cap at 0.9</td>
<td>$560.2</td>
</tr>
</tbody>
</table>
Appendix
Proof of Theorem 2

Using the joint cdf of the idiosyncratic payoff shocks, we introduce the following functions:

\[
\sigma_0(z_1, z_2; c) = \int 1\{\varepsilon_0 \geq z_1 + \varepsilon_1, \varepsilon_0 \geq z_2 + \varepsilon_2\} F_{\varepsilon}(d\varepsilon | c),
\]
\[
\sigma_1(z_1, z_2; c) = \int 1\{z_1 + \varepsilon_1 \geq \varepsilon_0, z_1 + \varepsilon_1 \geq z_2 + \varepsilon_2\} F_{\varepsilon}(d\varepsilon | c),
\]
\[
\sigma_2(z_1, z_2; c) = \int 1\{z_2 + \varepsilon_2 \geq \varepsilon_0, z_2 + \varepsilon_2 \geq z_1 + \varepsilon_1\} F_{\varepsilon}(d\varepsilon | c).
\]

We note that the introduced functions are known (given that we can normalize the distribution of the idiosyncratic shocks) and are monotone and differentiable in their arguments. Our result will be based on the following technical lemma.

Lemma 1
Under our assumptions, the system of equations

\[
\sigma_0(z_1, z_2; c) = \tilde{\sigma}_0(c),
\]
\[
\sigma_1(z_1, z_2; c) = \tilde{\sigma}_1(c)
\]

has a unique solution for each \(c\) if and only if \(\tilde{\sigma}_0(c) + \tilde{\sigma}_1(c) < 1\).

Proof:
Consider partial derivatives

\[
\frac{\partial \sigma_0(z_1, z_2)}{\partial z_1} = -\int_{-\infty}^{+\infty} \frac{\partial^2 F_{\varepsilon}}{\partial \varepsilon_0 \partial \varepsilon_1} (\varepsilon_0 - z_1, \varepsilon_0 - z_2) d\varepsilon_0,
\]
\[
\frac{\partial \sigma_0(z_1, z_2)}{\partial z_2} = -\int_{-\infty}^{+\infty} \frac{\partial^2 F_{\varepsilon}}{\partial \varepsilon_0 \partial \varepsilon_2} (\varepsilon_0 - z_1, \varepsilon_0 - z_2) d\varepsilon_0.
\]

Similarly, we can find that

\[
\frac{\partial \sigma_1(z_1, z_2)}{\partial z_1} = \int_{-\infty}^{+\infty} \frac{\partial^2 F_{\varepsilon}}{\partial \varepsilon_0 \partial \varepsilon_1} (z_1 + \varepsilon_1, z_1 - z_2) d\varepsilon_1
\]
\[
+ \int_{-\infty}^{+\infty} \frac{\partial^2 F_{\varepsilon}}{\partial \varepsilon_1 \partial \varepsilon_2} (z_1 + \varepsilon_1, \varepsilon_0 - z_2) d\varepsilon_1
\]
\[
= \int_{-\infty}^{+\infty} \frac{\partial^2 F_{\varepsilon}}{\partial \varepsilon_0 \partial \varepsilon_1} + \int_{-\infty}^{+\infty} \frac{\partial^2 F_{\varepsilon}}{\partial \varepsilon_1 \partial \varepsilon_2} (\varepsilon_0 - z_1, \varepsilon_0 - z_2) d\varepsilon_0,
\]

55
We assumed that the joint distribution of errors has a continuous density with a full support on \( \mathbb{R}^3 \). Provided that \( \frac{\partial \sigma_0(z_1, z_2)}{\partial z_1} \frac{\partial \sigma_0(z_1, z_2)}{\partial z_2} > 0 \) the mapping \( z_1 \rightarrow z_2 \) implicitly defined by equation \( \sigma_0(z_1, z_2; c) = \bar{\sigma}_0(c) \) is invertible for any \( c \). Moreover, if we denote this mapping \( z_2 = m_0(z_1, \bar{\sigma}_0(c); c) \), then using the result regarding the derivative of the inverse function, we can conclude that

\[
\frac{\partial m_0(z_1, \bar{\sigma}_0(c); c)}{\partial z_1} \leq 0.
\]

Similarly, we can define a map \( z_2 = m_1(z_1, \bar{\sigma}_1(c); c) \), then using the result regarding the derivative of the inverse function, we can conclude that

\[
\frac{\partial m_1(z_1, \bar{\sigma}_1(c); c)}{\partial z_1} \geq 0.
\]

We can explore the asymptotic behavior of both maps. Consider \( m_0 \) first. Suppose that \( z_1 \rightarrow -\infty \). Then \( \lim_{z_1 \rightarrow -\infty} m_0(z_1, \bar{\sigma}_0(c); c) = z_2^* \), where \( z_2^* \) solves \( \int 1 \{ \varepsilon_0 \geq z_2^* + \varepsilon_2 \} F_\varepsilon(d\varepsilon | c) = \bar{\sigma}_0(c) \). Also let \( z_1^* \) solve \( \int 1 \{ \varepsilon_0 \geq z_1^* + \varepsilon_1 \} F_\varepsilon(d\varepsilon | c) = \sigma_0(c) \). Then \( \lim_{z_1 \rightarrow -\infty} m_0(z_1, \bar{\sigma}_0(c); c) = -\infty \).

Next consider \( m_1 \). Suppose that \( z_2^{**} \) is the solution of \( \int 1 \{ z_1 \geq z_2^{**} + \varepsilon_2 \} F_\varepsilon(d\varepsilon | c) = \bar{\sigma}_1(c) \). Then as \( z_1 \rightarrow +\infty \), the map approaches asymptotically to the line: \( m_1(z_1, \bar{\sigma}_1(c); c) \rightarrow z_1 + z_2^{**} \). Suppose that \( z_1^{**} \) is the solution of \( \int 1 \{ z_1^{**} + \varepsilon_1 \geq \varepsilon_0 \} F_\varepsilon(d\varepsilon | c) = \bar{\sigma}_1(c) \). Then \( \lim_{z_1 \rightarrow z_1^{**}} m_1(z_1, \bar{\sigma}_1(c); c) = -\infty \). Thus \( m_0 \) is a continuous strictly decreasing mapping from \( (-\infty, z_1^{**}] \) into \( (-\infty, z_2^*] \) and \( m_1 \) is a continuous strictly increasing mapping from \( [z_1^{**}, +\infty) \) into the real line.

Provided that both curves are continuous and monotone, they intersect if and only if their projections on \( z_1 \) and \( z_2 \) axes overlap. The projections on the \( z_2 \) axis are guaranteed to overlap ((\( -\infty, z_2^*] \subset \mathbb{R} \)). The projections on the \( z_1 \) axis will overlap if and only if \( z_1^{**} < z_1^* \). Given that function \( \sigma(z) = \int 1 \{ \varepsilon_0 - \varepsilon_1 \leq z \} F_\varepsilon(d\varepsilon | c) \) is strictly monotone in \( z \), then \( z_1^{**} < z_1^* \) if and only if \( \bar{\sigma}_0(c) + \bar{\sigma}_1(c) < 1 \).

This proves the statement of Lemma 1.

Next we can show that the model is nonparametrically identified. We introduce function

\[
\nu(z_1, z_2; c) = \int \left( z_1 1 \{ z_1 + \varepsilon_1 \geq \varepsilon_0, z_1 + \varepsilon_1 \geq z_2 + \varepsilon_2 \}
+ z_2 1 \{ z_2 + \varepsilon_2 \geq \varepsilon_0, z_2 + \varepsilon_2 \geq z_1 + \varepsilon_1 \} \right) F_\varepsilon(d\varepsilon | c) + u(0; c),
\]

56
where the payoff from the default option \( u(0; c) \) is a fixed known function according to the assumption of the theorem.

The observed probability distribution characterizes the conditional choice probabilities \( \{\sigma_{k,t}(\cdot;), k \in A\} \). Given the structure of the optimal solution, there is a direct link between the choice-specific value functions in period \( t \) and the choice probability which is expressed through the distribution of the idiosyncratic payoff shocks. In particular, for each \( s \in S \) and \( c \in C \) and each \( t \leq T \), we can write the system of identifying equations

\[
\begin{align*}
\sigma_0(V_t(1, s; c) - u(0, s; c), V_t(2, s; c) - u(0, s; c); c) &= \sigma_{0,t}(s; c), \\
\sigma_1(V_t(1, s; c) - u(0, s; c), V_t(2, s; c) - u(0, s; c); c) &= \sigma_{1,t}(s; c).
\end{align*}
\]

Given the result of Lemma 1, we can solve for the choice-specific value functions \( V_t(1, s; c) \) and \( V_t(2, s; c) \) over \( S \) and \( C \).

The conditional distribution \( s_{t+1} | s_t, a_t, c \) is observable. As a result, for each \( k \in \{1, 2\} \) and each \( s \) and \( c \) we can consider the system of equations in periods \( t \) and \( t' < T \):

\[
\begin{align*}
V_t(1, s; c) &= u(1, s; c) + \beta E \left[ \nu (V_{t+1}(1, s'; c) - u(0, s'; c), V_{t+1}(2, s'; c) - u(0, s'; c); c) \mid a_t = 1, s, c \right], \\
V_t(2, s; c) &= u(2, s; c) + \beta E \left[ \nu (V_{t+1}(1, s'; c) - u(0, s'; c), V_{t+1}(2, s'; c) - u(0, s'; c); c) \mid a_t = 2, s, c \right], \\
V_{t'}(1, s; c) &= u(1, s; c) + \beta E \left[ \nu (V_{t'+1}(1, s'; c) - u(0, s'; c), V_{t'+1}(2, s'; c) - u(0, s'; c); c) \mid a_{t'} = 1, s, c \right].
\end{align*}
\]

This is a system of three linear equations with three unknowns \( u(1, s; c), u(2, s; c) \) and \( \beta \). We note that to set up this system of equations, we need to have observations for at least three periods (i.e., when \( t' = t + 1 \)).

Then simple differencing solves for the discount factor:

\[
\beta = \left( E \left[ \nu (V_{t+1}(1, s'; c) - u(0, s'; c), V_{t+1}(2, s'; c) - u(0, s'; c); c) \mid a_t = 1, s, c \right] \\
- E \left[ \nu (V_{t'+1}(1, s'; c) - u(0, s'; c), V_{t'+1}(2, s'; c) - u(0, s'; c); c) \mid a_{t'} = 1, s, c \right] \right)^{-1} \\
\times [V_t(1, s; c) - V_{t'}(1, s; c)].
\]

The denominator in this expression is not equal to zero because of the assumption of the theorem that \( \sigma_{k,t}(s; c) \neq \sigma_{k,t'}(s; c) \) for at least two consecutive periods \( t \) and \( t' \). We also can recover the per-period
utility function as

\[
\begin{align*}
 u(1, s; c) &= \left( E \left[ \nu (V_{t+1}(1, s'; c) - u(0, s'; c), V_{t+1}(2, s'; c) - u(0, s'; c); c) \mid a_t = 1, s, c \right] \\
 &\quad - E \left[ \nu (V_{t+1}(1, s'; c) - u(0, s'; c), V_{t+1}(2, s'; c) - u(0, s'; c); c) \mid a_t = 1, s, c \right] \right)^{-1} \\
 &\quad \times \left( V_t(1, s; c) E \left[ \nu (V_{t+1}(1, s'; c) - u(0, s'; c), V_{t+1}(2, s'; c) - u(0, s'; c); c) \mid a_t = 1, s, c \right] \\
 &\quad - V_t(1, s; c) E \left[ \nu (V_{t+1}(1, s'; c) - u(0, s'; c), V_{t+1}(2, s'; c) - u(0, s'; c); c) \mid a_t = 1, s, c \right] \right).
\end{align*}
\]

Similarly, we can explicitly recover the utility \( u(2, s; c) \) for any \( s \) and \( c \) in their support. We note that if the discount factor is fixed, then we can identify the utility function from just two periods.

If the per-period payoff function in the last period coincides with the per-period payoff function in the previous periods \( (u(\cdot, \cdot; \cdot) = u_T(\cdot, \cdot; \cdot)) \), then in the last period the choice specific values coincide with the utilities, meaning that \( V_T(1, s; c) = u(1, s; c) \) and \( V_T(2, s; c) = u(2, s; c) \). Both these utilities are recovered from just observing the last period choice probabilities. Then we can re-construct the ex ante value function of the last period using these estimates. We can take the first equation of the system considered before

\[
V_{T-1}(1, s; c) = u(1, s; c) + \beta E \left[ \nu (u(1, s'; c) - u(0, s'; c), u(2, s'; c) - u(0, s'; c); c) \mid a_{T-1} = 1, s, c \right],
\]

where the only remaining unknown is the discount factor. We can recover it as

\[
\beta = \frac{V_{T-1}(1, s; c) - u(1, s; c)}{E \left[ \nu (u(1, s'; c) - u(0, s'; c), u(2, s'; c) - u(0, s'; c); c) \mid a_{T-1} = 1, s, c \right]}.
\]

Therefore, in the special case where \( u = u_T \), the discount factor as well as the utility functions are identified from just two periods, if one of the observed periods is the period of mortgage maturity \( T \).

**Proof of Theorem 3**

Consider normalization \( u(0, s; c) \). Then, for instance

\[
\sigma_{0,t}(s; c) = \int \mathbf{1} \{ u(0, s; c) + \varepsilon_{0,t} \geq V_t(1, s; c) + \varepsilon_{1,t}, u(0, s; c) + \varepsilon_{0,t} \geq V_t(2, s; c) + \varepsilon_{2,t} \} F_c(d\varepsilon) \mid c).
\]

Using the notation in the proof of the previous theorem, we can then express

\[
\sigma_{0,t}(s; c) = \sigma_0 (V_t(1, s; c) - u(0, s; c), V_t(2, s; c) - u(0, s; c); c).
\]
We can also provide similar expressions for other choices. Then we solve the system of equations

\[
\sigma_{0,t}(s; c) = \sigma_0 \left( V_t(1, s; c) - u(0, s; c), V_t(2, s; c) - u(0, s; c); c \right), \\
\sigma_{1,t}(s; c) = \sigma_1 \left( V_t(1, s; c) - u(0, s; c), V_t(2, s; c) - u(0, s; c); c \right)
\]

to recover the value functions. We note that the recovered differences \( V_t(k, s; c) - u(0, s; c) \) are invariant to the choice of \( u(0, s; c) \) as they are directly recovered from the data.

The ex ante value function can be recovered as

\[
V_t(s; c) = \int \left( (V_t(1, s; c) - u(0, s; c)) \mathbf{1}\{V_t(1, s; c) - u(0, s; c) + \varepsilon_{1,t} \geq \varepsilon_{0,t}, V_t(1, s; c) + \varepsilon_{1,t} \geq V_t(2, s; c) + \varepsilon_{2,t} \} \\
+ (V_t(2, s; c) - u(0, s; c)) \mathbf{1}\{V_t(2, s; c) - u(0, s; c) + \varepsilon_{2,t} \geq \varepsilon_{0,t}, V_t(2, s; c) + \varepsilon_{2,t} \geq V_t(1, s; c) + \varepsilon_{1,t} \} \right) F_{\varepsilon}(d\varepsilon | c) + u(0, s; c) = V_t(s; c) + u(0, s; c),
\]

where function \( V_t(\cdot; \cdot) \) is invariant to the choice of the default payoff (it is determined by the observed choice probabilities). The system of equations identifying the per-period payoffs takes the form

\[
V_t(1, s; c) = u(1, s; c) + \beta E \left[ V_{t+1}(s' ; c) + u(0, s'; c) \right| a_t = 1, s, c], \\
V_t(2, s; c) = u(2, s; c) + \beta E \left[ V_{t+1}(s' ; c) + u(0, s'; c) \right| a_t = 2, s, c], \\
V_{t'}(1, s; c) = u(1, s; c) + \beta E \left[ V_{t'+1}(s' ; c) + u(0, s'; c) \right| a_{t'} = 1, s, c].
\]

Provided that \( u(0, s; c) \) is normalized to a fixed function, taking the difference between the first and the third equation allows us to express the discount factor as

\[
\beta = \frac{V_t(1, s; c) - u(0, s; c) - (V_{t'}(1, s; c) - u(0, s; c))}{E \left[ V_{t+1}(s' ; c) \right| a_t = 1, s, c] - E \left[ V_{t'+1}(s' ; c) \right| a_{t'} = 1, s, c].
\]

Provided that the numerator and the denominator are invariant with respect to the choice of the default utility (as they can be directly recovered from the observed choice probabilities), the discount factor will not depend on that choice either. Having obtained the discount factor, we can focus on the first two
equations of the considered system and re-cast them to the form

\[ V_t(1, s; c) - u(0, s; c) - \beta E \left[ V_{t+1}(s'; c) \mid a_t = 1, s, c \right] = u(1, s; c) - u(0, s; c) + \beta E \left[ u(0, s'; c) \mid a_t = 1, s, c \right], \]
\[ V_t(2, s; c) - u(0, s; c) - \beta E \left[ V_{t+1}(s'; c) \mid a_t = 2, s, c \right] = u(2, s; c) - u(0, s; c) + \beta E \left[ u(0, s'; c) \mid a_t = 2, s, c \right]. \]

The left-hand side of this system is immune to the choice of the default utility. As result, the right-hand side should be invariant too. In other words, for any two choices \( u(0, s; c) \) and \( u'(0, s; c) \),

\[ u(k, s; c) - u(0, s; c) + \beta E \left[ u(0, s'; c) \mid a_t = k, s, c \right] = u'(k, s; c) - u'(0, s; c) + \beta E \left[ u'(0, s'; c) \mid a_t = k, s, c \right]. \]

Thus, the gap between the utilities from different options depends on the choice of the default utility

\[ u(k, s; c) - u(0, s; c) - (u'(k, s; c) - u'(0, s; c)) = \beta E \left[ u'(0, s'; c) - u(0, s'; c) \mid a_t = k, s, c \right] \]

for \( k = 1, 2 \).

If the last period choice probability is observed and \( u = u_T \), then we can complement the system of Bellman equations with the expressions for the choice probabilities of the last period. Those choice probabilities identify the differences \( u(1, s; c) - u(0, s; c) \) and \( u(2, s; c) - u(0, s; c) \). Thus, we can identify the pair of conditional expectations \( E \left[ u(0, s'; c) \mid a_t = 1, s, c \right] \) and \( E \left[ u(0, s'; c) \mid a_t = 2, s, c \right] \) from the system of linear equations that we constructed before:

\[ \beta E \left[ u(0, s'; c) \mid a_t = 1, s, c \right] = V_t(1, s; c) - u(0, s; c) - \beta E \left[ V_{t+1}(s'; c) \mid a_t = 1, s, c \right] - u(1, s; c) + u(0, s; c), \]
\[ \beta E \left[ u(0, s'; c) \mid a_t = 2, s, c \right] = V_t(2, s; c) - u(0, s; c) - \beta E \left[ V_{t+1}(s'; c) \mid a_t = 2, s, c \right] - u(2, s; c) + u(0, s; c). \]

This system of equations represents a familiar nonlinear instrumental variable problem which is known to be an ill-posed inverse problem. It has a unique solution and, thus, the utility from default is identified, under the completeness condition that is discussed in Newey and Powell (2003), and Chen, Chernozhukov, Lee and Newey (2011). Therefore, the utilities from all choices \( u(0, s; c) \), \( u(1, s; c) \), and \( u(2, s; c) \) are identified along with the discount factor \( \beta \) in this case.

**Proof of Theorem 4**

In this proof by \( n \) we denote the sample size corresponding to the borrowers observed with \( t \) periods from mortgage origination with the heterogeneity characteristic equal to \( c \). We introduce the notation for
the trinomial logit function \( \ell(z_1, z_2) = \frac{\exp(z_1)}{1+\exp(z_1)+\exp(z_2)} \). By \( \tilde{\sigma}^L_{k,t}(s; c) \) we denote the choice probability

\[
\tilde{\sigma}^L_{k,t}(s; c) = \ell(\tilde{r}^L(t, k, c) q^L(s), \tilde{r}^L(t, j, c) q^L(s)), \; j \neq k
\]

where \( \tilde{r}^L(t, k, c) \) are the coefficients of the projection of the probability ratio \( \log \frac{\sigma_{k,t}(s; c)}{\sigma_{0,t}(s; c)} \) on \( L \) first orthogonal polynomials. We also denote \( \tilde{\sigma}^L_{1,t}(s; c) = 1 - \tilde{\sigma}^L_{0,t}(s; c) - \tilde{\sigma}^L_{2,t}(s; c) \). We note that \( \frac{\partial}{\partial z_1}, \frac{\partial}{\partial z_2} \leq \frac{1}{2} \).

Thus, \( \frac{1}{2} \) is a uniform Lipschitz constant and

\[
\sup_{s \in S} |\tilde{\sigma}^L_{k,t}(s; c) - \sigma_{k,t}(s; c)| = \sup_{s \in S} \left| \ell \left( \log \frac{\tilde{\sigma}^L_{k,t}(s; c)}{\sigma_{k,t}(s; c)} \right) - \ell \left( \log \frac{\tilde{\sigma}^L_{1,t}(s; c)}{\sigma_{1,t}(s; c)} + \log \frac{\tilde{\sigma}^L_{2,t}(s; c)}{\sigma_{2,t}(s; c)} \right) \right| \leq \frac{1}{2} \sup_{s \in S} \left| \log \frac{\tilde{\sigma}^L_{1,t}(s; c)}{\sigma_{1,t}(s; c)} - \log \frac{\tilde{\sigma}^L_{2,t}(s; c)}{\sigma_{2,t}(s; c)} \right|^2 = O(L^{-a}).
\]

This guarantees the quality of approximation of the choice probability using a logit transformation of the series expansion.

Now we omit index \( t \) in the variables (whenever the period of time under consideration is known) and construct function

\[
\rho(a_i, s_i; r^L_1, r^L_2) = (1\{a_i = 1\} - 1\{a_i = 0\}) \ell (r^L_1 q^L(s_i), r^L_2 q^L(s_i)) + (1\{a_i = 2\} - 1\{a_i = 0\}) \ell (r^L_2 q^L(s_i), r^L_1 q^L(s_i)).
\]

Then we can express the sample quasi-likelihood as

\[
\tilde{Q}(r^L_1, r^L_2) = E_n \left[ \rho(a_i, s_i; r^L_1, r^L_2) \right] + E_n [1\{a_i = 0\}],
\]

where we adopted the notation from the empirical process theory where \( E_n[\cdot] = \frac{1}{n} \sum_{i=1}^{n} \cdot \). Also introduce the population likelihood with the series expansion

\[
Q(r^L_1, r^L_2) = E \left[ \rho(a_i, s_i; r^L_1, r^L_2) \right] + E [1\{a_i = 0\}].
\]

Consider function

\[
f(a_i, s_i; r^L_1, r^L_2, \tilde{r}^L_1, \tilde{r}^L_2) = \rho(a_i, s_i; r^L_1, r^L_2) - \rho(a_i, s_i; \tilde{r}^L_1, \tilde{r}^L_2) - E[\rho(a_i, s_i; r^L_1, r^L_2)] + E[\rho(a_i, s_i; \tilde{r}^L_1, \tilde{r}^L_2)].
\]
Provided that we established that function \( \ell(\cdot, \cdot) \) is Lipschitz, we can evaluate

\[
\text{Var}\left(f(a_i, s_i; r^L_1, r^L_2, \tilde{r}^L_1, \tilde{r}^L_2)\right) = O\left(L \sup_{p=1, 2, 1 \leq L} \| r_{p,l} - \tilde{r}_{p,l} \|\right) = O(L).
\]

Next we impose a technical assumption that allows us to establish consistency of estimator (3).

**ASSUMPTION 4**

Consider the class of functions indexed by \( n \)

\[
\mathcal{F}_n = \left\{ f(\cdot, \cdot; r^L_1, r^L_2, \tilde{r}^L_1, \tilde{r}^L_2) - E[f(\cdot, \cdot; r^L_1, r^L_2, \tilde{r}^L_1, \tilde{r}^L_2)], r_{l,p} \in \Theta, l \leq L_n, p = 1, 2 \right\},
\]

where \( \Theta \) is the compact subset of \( \mathbb{R} \) and \( \tilde{r}_1^{L_n} \) and \( \tilde{r}_2^{L_n} \) are the coefficients of projections of population probability ratios on \( L_n \) series terms. Then for each \( L_n \to \infty \) such that \( n/(L_n \log n) \to \infty \) the \( L_1 \) covering number for class \( \mathcal{F}_n, N, \) has the following bound

\[
\log N(\delta, \mathcal{F}_n, L_1) \leq A\frac{r_0}{\log 1/n},
\]

where \( 0 < r_0 \leq \frac{3}{4} \) and \( r_0 \downarrow 0 \) is assumed to correspond to the factor \( \log n \).

This is the condition restricting the complexity of the functions created by logit transformations of series expansions. By construction any \( f \in \mathcal{F}_n \) is bounded \(|f| < 1 < \infty\). We established that \( \text{Var}(f) = O(L_n) \) for \( f \in \mathcal{F}_n \). The symmetrization inequality (30) in Pollard (1984) holds if \( \varepsilon_n/(16n \mu^2_n) \leq \frac{1}{2} \). This will occur if \( \frac{\varepsilon_n}{n^{3/2}} \to \infty \). Provided that the symmetrization inequality holds, we can follow the steps of Theorem 37 in Pollard (1984) to establish the tail bound on the deviations of the sample average of \( f \) via a combination of the Hoeffding inequality and the covering number for the class \( \mathcal{F}_n \). As a result, we obtain that

\[
P\left( \sup_{f \in \mathcal{F}_n} \frac{1}{n} \left\| E_n[f(\cdot)] \right\| > 8 \mu_n \right) \\
\leq 2 \exp \left( A\frac{r_0}{\mu_n} \log \frac{1}{\mu_n} \right) \exp \left( -\frac{1}{128} \frac{n \mu^2_n}{L_n} \right) + P\left( \sup_{f \in \mathcal{F}_n} \frac{1}{n} \left\| E_n[f(\cdot)] \right\|^2 > 64L_n \right).
\]

The second term can be evaluated with the aid of Lemma 33 in Pollard (1984):

\[
P\left( \sup_{f \in \mathcal{F}_n} \frac{1}{n} \left\| E_n[f(\cdot)] \right\|^2 > 64L_n \right) \leq 4 \exp \left( A\frac{2r_0}{L_n} \log \frac{1}{L_n} \right) \exp(-nL_n).
\]
As a result, we find that

\[
P \left( \sup_{f \in \mathcal{F}_n} \frac{1}{n} \left\| \mathbb{E}_n[f(\cdot)] \right\| > 8 \mu_n \right)
\leq 2 \exp \left( An \log \frac{1}{\mu_n} \right) \exp \left( -\frac{1}{128} \frac{n \mu_n^2}{L_n} \right) + 4 \exp \left( An \log \frac{1}{L_n} - nL_n \right).
\]

We start the analysis with the first term. Consider the case with \( r_0 > 0 \). Then the log of the first term takes the form

\[
An \log \frac{1}{\mu_n} - \frac{1}{128} \frac{n \mu_n^2}{L_n}.
\]

Then one needs that \( \frac{n \mu_n^2}{L_n \log{n}} \to \infty \) if \( r_0 > 0 \) and \( \frac{n \mu_n^2}{L_n \log{n}} \to \infty \) if \( r_0 \downarrow 0 \). Hence the first term is of \( o(1) \). This condition also guarantees that the second term vanishes. We note also that the CLT applies to the term \( n \log \frac{1}{L_n} \to 0 \), if \( r_0 > 0 \) and \( n \log \frac{1}{L_n} \to 1 \), if \( r_0 \leq 0 \). Hence the first term is of \( o(1) \). This condition also guarantees that the second term vanishes. We note also that the CLT applies to the term \( \mathbb{E}_n[\{a_i = 0\}] = E[\{a_i = 0\}] + O_p(\frac{1}{\sqrt{n}}) \). Now for some slowly diverging sequence \( \delta_n \to \infty \) such that \( \mu_n = \delta_n \sqrt{\frac{L_n n_{r_0} \log{n}}{n}} \to 0 \), we establish that

\[
\sup_{(r_1^n, r_2^n) \in \Theta_n \times \Theta_n} \left\| \tilde{Q}(r_1^n, r_2^n) - Q(r_1^n, r_2^n) + \tilde{Q}(\tilde{r}_1^n, \tilde{r}_2^n) - Q(\tilde{r}_1^n, \tilde{r}_2^n) \right\| = O_p \left( \mu_n + \frac{1}{\sqrt{n}} \right) = o_p(1).
\]

Thus, the sample quasi-likelihood converges uniformly to the population quasi-likelihood and the estimated choice probabilities are uniformly consistent over \( S \). To establish the rate for the estimated choice probabilities, we consider a neighborhood of the population projections defined by \( \sup_{p=1,2, t \leq L_n} \left\| r_{p,t} - \tilde{r}_{p,t} \right\| \leq \varepsilon \). Using Lemma 2.3.1 from van der Vaart and Wellner (1996), we can find that

\[
E \left[ \sup_{f \in \mathcal{F}_n} \sqrt{n} \mathbb{E}_n[f(\cdot)] \right] \leq C n^{r_0/2} \sqrt{L_n} \varepsilon \log \frac{1}{\sqrt{n} \varepsilon},
\]

for some constant \( C \). Using Theorem 3.4.1 from van der Vaart and Wellner (1996) and the derived inequality, we can express the convergence rate for the estimated parameters of the approximated choice probabilities as \( \rho_n^2 n^{r_0/2} \sqrt{L_n} \frac{1}{\rho_n} \log \frac{\rho_n}{\sqrt{L_n}} \leq \sqrt{n} \). Then

\[
\sup_{s \in S} \left\| \tilde{\sigma}_{k,t}(s; c) - \hat{\sigma}_{k,t}(s; c) \right\| = O_p \left( \frac{L_n}{\rho_n} \right).
\]

To attain the rate \( o_p(n^{-1/4}) \) we need to assure that \( \frac{L_n}{\rho_n} = o(n^{-1/4}) \). To assure the \( n^{-1/4} \) we choose \( \delta_n \to 0 \) and set \( L_n = \delta_n n^{-1/4} \rho_n \). Then the rate constraint can be re-written as

\[
\rho_n^2 n^{-3/8 + r_0/2} \frac{\log \frac{n^{1/4} \sqrt{\rho_n}}{\sqrt{\delta_n}}}{\frac{n^{1/4} \sqrt{\rho_n}}{\sqrt{\delta_n}}} \leq \sqrt{n}.
\]

63
Provided that \( \lim_{x \to \infty} \log x / x = 0 \), we conclude that \( \rho_n = O(n^{7/8 - r_0/2}) \), meaning that \( L_n = o(n^{5/8 - r_0/2}) \).

We note that the slowest rate for the choice of \( L_n \) has to satisfy

\[
\frac{n^{1-r_0} \mu_n^2}{L_n \log n} \to \infty,
\]

for \( \mu_n \to 0 \). Thus, the estimator with rate \( o(n^{-1/4}) \) is plausible if \( r_0 < 3/4 \). Using the triangle inequality and our previous result, we find that

\[
\sup_{s \in S} \| \sigma_k,t(s; c) - \sigma_k,t(s; c) \| = O_p \left( \frac{L_n}{\rho_n} + L_n^{-\alpha} \right) = o_p \left( n^{-1/4} \right),
\]

if \( \alpha \geq 1 \).

**Expression for the Asymptotic Variance of the Structural Estimator**

We introduce

\[
J_k \left( \sigma_0,t, \sigma_1,t, \sigma_0,t+1, \sigma_1,t+1, s; c \right) = \left( \frac{\partial F_k}{\partial \sigma_0,t}, \frac{\partial F_k}{\partial \sigma_1,t}, \frac{\partial F}{\partial \sigma_0,t+1}, \frac{\partial F}{\partial \sigma_1,t+1} \right)
\]

and \( J(s; c) = (J_1(s; c), J_2(s; c))' \),

\[
M(s; c) = E \left[ \begin{vmatrix} \frac{\partial u(s,t; \theta(1,c))}{\partial \theta(1,c)} & 0 & - \frac{\partial u(s,t; \theta(0,c))}{\partial \theta(0,c)} + \beta \frac{\partial u(s+1,t; \theta(0,c))}{\partial \theta(0,c)} F(s+1,t; c) + u(s+1,t; \theta(0,c)) \\ 0 & \frac{\partial u(s,t; \theta(2,c))}{\partial \theta(2,c)} & - \frac{\partial u(s,t; \theta(0,c))}{\partial \theta(0,c)} + \beta \frac{\partial u(s+1,t; \theta(0,c))}{\partial \theta(0,c)} F(s+1,t; c) + u(s+1,t; \theta(0,c)) \end{vmatrix} \right| s_t = s, c,
\]

as well as

\[
\Omega(s; c) = \text{Var} \left( (\sigma_0,t(s_t; c), \sigma_1,t(s_t; c), \sigma_0,t+1(s_{t+1}; c), \sigma_1,t+1(s_{t+1}; c))' \right| s_t = s, c.
\]

Then, the variance of the second-stage estimates is determined by the sampling noise and the error from the first stage estimates:

\[
V = E \left[ M(s_t; c)^{-1} \left( E \left[ Z_t | \text{Var}(\epsilon_t | s_t) Z_t \right| s_t \right] + J(s_t; c) \Omega(s_t; c) J(s_t; c)' \right) M(s_t; c)^{-1} \right| c.
\]

As an alternative to using the asymptotic formula, we can use the subsampling approach to estimate the variance.