Local Gambling Preference and Excessive Mortgage Leverage

Jiawei Hu *

March 2026

Abstract

I study the role of borrower behavioral heterogeneity in excessive mortgage leverage. Using misreported simultaneous second liens as a revealed measure of excessive mortgage leverage, I show that such borrowing was more prevalent in areas with stronger local gambling preference. The effect is stronger where borrower incentives to obtain leverage are greater, including owner-occupied and purchase loans, and is not primarily driven by lender facilitation around the FICO 620 securitization cutoff. The findings suggest that borrower behavioral traits, distinct from broad aggregate borrowing conditions, contributed to excessive mortgage leverage and financial fragility.

Keywords: Excessive mortgage leverage, Household finance, Local gambling preference

*Jiawei.Hu@utdallas.edu, Naveen Jindal School of Management, University of Texas at Dallas, 800 West Campbell Road, Richardson, Texas, 75080.

1 Introduction

Household leverage is central to both household welfare and financial stability. A large literature shows that borrowing decisions in household credit markets can have important macro-financial consequences, especially through housing markets and mortgage debt (Campbell, 2006; Mian and Sufi, 2009). Recent evidence from the New York Fed's household credit program likewise underscores an important distinction between aggregate conditions and cross-sectional outcomes: even when mortgage risk appears contained in the aggregate, debt burdens and delinquency outcomes can vary substantially across households. Yet much of the academic literature focuses on aggregate borrowing conditions rather than which households are most willing to take on especially excessive mortgage leverage in the first place. This paper studies that heterogeneity.

I examine this question in the housing boom preceding the 2008 financial crisis, a setting in which mortgage leverage expanded rapidly and its consequences became fully visible. Existing work emphasizes aggregate drivers such as credit supply expansion, securitization, and overly optimistic beliefs about future house prices (Mian and Sufi, 2009; Adelino et al., 2016; Justiniano et al., 2019). Yet these explanations do not fully account for why some households were more willing than others to stretch for leverage in the same borrowing environment. I study this cross-sectional variation by investigating how local socioeconomic characteristics shape the incidence of excessive mortgage leverage.

I capture excessive mortgage leverage using misreported simultaneous second liens. A standard first-lien mortgage already places a household in a leveraged position, and a simultaneous second lien increases leverage further by reducing the borrower's effective equity stake. When that second lien is concealed at the time of first-lien origination, the borrower can make the application appear safer than it actually is and thereby obtain mortgage leverage that would have been more difficult to secure under truthful disclosure (Griffin and Maturana, 2016). In this sense, misreported simultaneous second liens reflect not merely high leverage, but especially excessive mortgage leverage. This interpretation is consistent with evidence that concealed second liens were concentrated in fragile loans and that mortgage-backed securities containing more such loans performed substantially worse ex post (Piskorski et al., 2015; Yavas and Zhu, 2024).

Among local socioeconomic characteristics, I focus on local gambling preference, measured by the county-level Catholic-to-Protestant ratio following Kumar et al. (2011). I show that, after controlling for a rich set of local characteristics as well as detailed loan and borrower fundamentals, misreported simultaneous second liens are significantly more prevalent in areas with higher Catholic-to-Protestant ratios. In contrast, other local factors related to household risk preferences, ethics, culture, and financial literacy, such as religiosity and education, play much more limited roles in explaining this form of leverage taking. These results suggest that specific behavioral traits, rather than broad local conditions alone, help explain which households were more willing to obtain excessive mortgage leverage.

The interpretation is that local gambling preference affects borrowers' financial decisions by making them more willing to pursue opportunities with right-skewed payoffs. From the borrower's perspective, concealing a simultaneous second lien has this feature. The upside is access to homeownership and greater exposure to housing-market gains with a smaller effective equity stake. If house prices continue to rise, the gains can be substantial. The downside, by contrast, is relatively limited in the short run, consisting mainly of application costs, the effort of switching lenders, and the relatively weak legal penalties for mortgage fraud before the crisis. This combination of potentially large gains and comparatively modest losses makes excessive mortgage leverage through second-lien misrepresentation particularly attractive to gambling-prone borrowers.

I next examine whether the evidence is consistent with a borrower-side mechanism. If gambling preference affects borrowers' willingness to stretch for leverage, the effect should be stronger when the incentive to obtain the loan is greater. Consistent with this prediction, the effect of local gambling preference is more pronounced among owner-occupied mortgages and purchase loans, settings in which loan approval is especially valuable because it is tied directly to obtaining or retaining housing. The effect is also stronger in subsamples where borrower incentives to obtain leverage are plausibly higher, including those defined by income, credit scores, and prior local house-price appreciation. Together, these patterns support the view that local gambling preference shapes borrower demand for excessive mortgage leverage.

An alternative explanation is that the effect reflects lender behavior rather than borrower behavior. This concern is important because previous studies suggest that lenders could be aware of hidden second liens and may at times have facilitated mortgage misreporting. To address this possibility, I exploit the securitization discontinuity around FICO 620 for low-documentation loans following Keys et al. (2010). In the pooled sample, lenders appear to facilitate simultaneous second liens in general, but not especially misreported ones. More importantly, when I compare high- and low-gambling-preference areas using a difference-in-discontinuities (diff-in-disc) design, jumps in both loan counts and default rates for misreported seconds are smaller, not larger, in high-gambling-preference areas. Overall, these results provide little support for the view that the documented relationship is primarily driven by more permissive lenders in gambling-prone areas.

Finally, I study whether excessive mortgage leverage taken in high-gambling-preference environments is associated with worse ex post outcomes. Loans with misreported simultaneous second liens default at substantially higher rates than loans without second liens, and this default penalty is larger in high-gambling-preference areas. Thus, local gambling preference is associated not only with a greater incidence of excessive mortgage leverage, but also with leverage that proves more fragile ex post. A back-of-the-envelope calculation indicates that the incremental default penalty in high-gambling-preference areas corresponds to roughly $1.34 billion in additional first-lien exposure, suggesting that the economic consequences are meaningful.

To reinforce the findings and interpretation, I conduct a series of additional tests. The main results are robust to alternative functional-form assumptions, measures of gambling preference, and definitions of loan performance. Additional analyses also support the validity of the lender-channel tests and show that the conclusions are not driven by the particular outcome definition used in those specifications. Taken together, these checks further support the interpretation that local gambling preference is linked to borrower demand for excessive mortgage leverage.

This paper contributes to three strands of literature. First, it contributes to the literature on household leverage and the financial crisis by showing that behavioral heterogeneity helps explain which households were more willing to take on excessive mortgage leverage during the housing boom (Mian and Sufi, 2009; Adelino et al., 2016; Justiniano et al., 2019). Second, it contributes to the literature on mortgage misreporting and hidden leverage by showing that borrower-side preferences, not only lender-side incentives, help explain concealed simultaneous second liens (Piskorski et al., 2015; Griffin and Maturana, 2016; Yavas and Zhu, 2024). Third, it contributes to the literature on gambling preference in financial decision-making by extending it to the mortgage market and linking local gambling attitudes to leverage choice and subsequent loan performance (Barberis and Huang, 2008; Kumar, 2009; Kumar et al., 2011; Shu et al., 2012; Chen et al., 2014).

The rest of the paper proceeds as follows. Section 2 describes the data and variable construction. Section 3 presents the baseline evidence. Section 4 examines whether the effect of gambling preference reflects borrower behavior or lender practices. Section 5 studies loan performance across gambling environments. Section 6 reports robustness tests. Section 7 concludes.

2 Data and Summary Statistics

My sample combines three main categories of data from 2005–2007: loan-related records, religiosity data, and local demographic information. Loan-related records come from BlackBox Logic and Equifax. Religiosity data are obtained from the American Religion Data Archive (ARDA). Local demographic information at the county and ZIP code levels is drawn from the U.S. Census Bureau. I additionally collect crime data from the Uniform Crime Reporting (UCR) Program Data and voter registration data from the National Neighborhood Data Archive (NaNDA), both available through the ICPSR website, along with house price data from the Federal Housing Finance Agency (FHFA) and unemployment data from the U.S. Bureau of Labor Statistics.

2.1 Second-lien Misrepresentation Measure

Second-lien misrepresentation occurs when a first-lien loan backed by property is reported as having no associated higher lien but is actually financed with a simultaneously originated second mortgage identified in credit bureau data. A simultaneous second lien increases mortgage leverage by reducing the borrower's effective equity stake, and concealing that lien allows the borrower to obtain leverage that would have been more difficult to secure under truthful disclosure. I identify second-lien misrepresentation by comparing loan-level mortgage data from BlackBox Logic with borrower-level credit report information from Equifax. Following the procedure in Piskorski et al. (2015), I first retain first-lien loans with a merge confidence interval greater than or equal to 0.89.

Second, I select loans with a new second lien originating within one month before or after the first lien. Third, I classify a loan as having a misreported simultaneous second lien if its reported cumulative loan-to-value (CLTV) ratio is non-missing and within 1% of its reported loan-to-value (LTV) ratio. A small difference between reported CLTV and LTV indicates that no simultaneous second lien was disclosed.

2.2 Gambling Preference and Religiosity

Following Kumar et al. (2011), I measure county-level gambling preference using the ratio of Catholic residents to Protestant residents (*CPRATIO*). Direct measures of local gambling preference are generally unavailable, so I infer local gambling propensity from religious composition, which reflects differing attitudes toward gambling. In the United States, Protestant churches generally oppose gambling, whereas Catholic churches are more tolerant of moderate gambling. Prior studies show that these differences extend to financial decision-making and asset markets (Kumar, 2009; Kumar et al., 2011; Han and Kumar, 2013; Chen et al., 2014). Accordingly, counties with higher Catholic-to-Protestant ratios are expected to exhibit stronger gambling preference.

I also construct county-level religiosity (*REL*) following Kumar et al. (2011). Hilary and Hui (2009) argue that religiosity is associated with risk attitudes, with higher values of *REL* indicating greater risk aversion. Religiosity may also capture social norms and ethical behavior that can deter mortgage misrepresentation (Conklin et al., 2022). Including county religiosity therefore helps separate local gambling preference from religion-related risk aversion and ethical norms.

To construct these variables, I use the *Longitudinal Religious Congregations and Membership File, 1980–2010 (County Level)* from ARDA. I define the Protestant population as the sum of adherents in the Evangelical Protestant, Mainline Protestant, and Black Protestant categories, and the Catholic population as the number of adherents in the Catholic category. Total religiosity is measured as total religious adherents divided by total county population. Because the data are available by decade from 1980 to 2010, I linearly interpolate missing years following prior studies (Alesina and La Ferrara, 2000; Hilary and Hui, 2009; Kumar et al., 2011; Chen et al., 2014), and then calculate *CPRATIO* and *REL* for each year.

2.3 Other Socioeconomic Characteristics

In addition to gambling preference and religiosity, I include a rich set of local socioeconomic characteristics covering both demographic factors and local economic conditions. Specifically, I use county-level data from 2000 on education, marriage, urban residence, population, age, sex ratio, and minority share, as well as ZIP code-level income from the U.S. Census Bureau. I also incorporate county-level crime data from 2005 to 2007 and voter registration data from 2006 obtained from the ICPSR. These variables capture local demographic and social characteristics that may be related to risk preferences, ethical norms, culture, and financial literacy.

To capture local economic conditions, I use county-level monthly unemployment rates from the U.S. Bureau of Labor Statistics and annual ZIP code-level house price indices (HPI) to construct measures of house price appreciation.

2.4 Mortgage Microdata

In addition to the second-lien misrepresentation measure, I obtain other loan-level mortgage information from BlackBox Logic. The database provides a comprehensive set of origination characteristics, including loan pricing, borrower credit quality, loan size, leverage, amortization type, documentation status, interest rate type, prepayment penalty, loan purpose, occupancy status, and delinquency information. Detailed definitions of variables used in the analysis are reported in Appendix A.

2.5 Descriptive Statistics

I study second-lien misrepresentation in a sample that includes loans without simultaneous second liens, loans with correctly reported simultaneous second liens, and loans with misreported simultaneous second liens. I focus on the 2005–2007 period, during which second-lien misrepresentation can be identified from the matched datasets and is sufficiently prevalent for analysis. Table 1 reports summary statistics for the main variables.

The final sample contains 3,031,921 loans for which county-level *CPRATIO* is available[1]. Among these loans, 7.15 percent have misreported simultaneous second liens and 13.91 percent have correctly reported simultaneous second liens, implying that 21.06 percent of loans in the sample have simultaneous second liens. This share is comparable to that reported by Griffin and Maturana (2016), who use a different dataset and broader loan sample. Figure 1 plots the nationwide distribution of county-level second-lien misrepresentation rates, showing substantial variation both across and within states.

My primary explanatory variable is local gambling preference, measured by *CPRATIO*. Although there are 3,090 U.S. counties during 2005–2007, I retain only county-year observations for which loan-level data are available, yielding 8,716 county-year observations. The mean and

[1] I also exclude loans with loan-to-value ratios (LTV) greater than 99%, since the misrepresentation measure is constructed using LTV. Such values are likely erroneous or unlikely to involve simultaneous second liens.

median values of *CPRATIO* are 0.51 and 0.19, respectively, indicating substantial right skewness[2]. To assess whether *CPRATIO* captures local preferences in areas with meaningful religious presence, I also examine religiosity (*REL*). Among the 8,716 county-year observations, the 10th percentile of *REL* is 37 percent and the median is 59 percent, suggesting that religious composition is quantitatively important in most counties.

CPRATIO is only weakly correlated with other local socioeconomic characteristics, suggesting that it captures distinct variation rather than proxying for common local observables. Correlations among the remaining local controls are generally low to moderate and do not indicate serious multicollinearity concerns. The full correlation matrix is reported in Appendix A.

3 Local Socioeconomic Characteristics and Second-lien Misrepresentation

This section examines how local socioeconomic characteristics are associated with second-lien misrepresentation. I estimate loan-level regressions of second-lien misrepresentation on local socioeconomic characteristics using the following specification:

$$\begin{aligned} \text{Misreported second} = \quad & \alpha + \beta_1 \times \text{Local socioeconomic characteristics} \\ & + \beta_2 \times \text{Loan characteristics} + \text{Fixed effects} \end{aligned} \tag{1}$$

where *Misreported second* is an indicator equal to one if the first-lien loan has a simultaneous second lien that is misreported at origination. *Local socioeconomic characteristics* include the vector of local variables, such as *CPRATIO*, *REL*, and *Education*. *Loan characteristics* include

[2]Including all counties, the mean and median *CPRATIO* values are close to those reported in Kumar et al. (2011).

loan- and borrower-level controls at origination, such as *Interest rate*, *Balance*, and *FICO*. *Fixed effects* include originator, state, and half-year fixed effects[3]. Because most local variables are measured at the county level, I cluster heteroskedasticity-robust standard errors by county in all specifications. All continuous variables are winsorized at the 0.5 percent level and standardized so that coefficient magnitudes are comparable. Following Piskorski et al. (2015) and Yavas and Zhu (2024), I use OLS in the main tests despite the binary nature of the dependent variable, and rely on probit models and causal forest methods as robustness checks.[4]

I include a broad set of local socioeconomic characteristics to capture alternative channels through which local conditions may affect excessive mortgage leverage. Guided by the existing literature, I group these variables into four categories: social environment, economic status, local market conditions, and household structure. This grouping serves two purposes. First, it allows me to assess whether several important local factors emphasized in prior work are associated with second-lien misrepresentation. Second, it allows me to evaluate whether *CPRATIO* captures distinct behavioral variation rather than simply proxying for broader local observables.

Table 2 reports the regression results. Columns (1)–(4) introduce local socioeconomic characteristics by group, while column (5) includes all local variables. Several patterns emerge. Among social-environment variables, *CPRATIO* is positive and highly significant both when entered with its own group and in the full specification, while *REL* is small and insignificant. *Crime rate* is positively associated with misrepresentation, whereas *Republican* is not robust. Among economic-status variables, *Income* is negatively associated with misrepresentation, while *Education* is at most

[3]Although both originators and underwriters play important roles (Griffin and Maturana, 2016), I control for originator fixed effects because originators directly interact with borrowers and may therefore affect borrower decisions, whereas underwriters do not influence borrower decision-making in the same way.

[4]OLS also has the practical advantage of accommodating multiple fixed effects. By contrast, nonlinear models such as probit or logit can yield inconsistent estimates when more than one fixed effect is included in large samples (Wooldridge, 2010).

weakly significant. Among local market conditions, *HPA* and *Total population* are both positive and highly significant, whereas *Unemployment* is insignificant and *Urban* becomes negative only in the full specification. Among household-structure variables, most coefficients are small or unstable, although *Over65* becomes positive and significant with richer controls.

The main result in Table 2 is the effect of *CPRATIO*. Across specifications, counties with higher Catholic-to-Protestant ratios exhibit higher rates of second-lien misrepresentation. Because *REL* is included separately and remains insignificant, the effect of *CPRATIO* is unlikely to reflect religion-induced risk aversion or ethical norms alone (Hilary and Hui, 2009; Conklin et al., 2022). Instead, the evidence is consistent with *CPRATIO* capturing local gambling preference, a behavioral trait that affects financial decisions when payoffs are positively skewed (Kumar, 2009; Kumar et al., 2011; Han and Kumar, 2013; Chen et al., 2014). Economically, in the full specification, a one-standard-deviation increase in *CPRATIO* raises the probability of second-lien misrepresentation by 0.14 percentage points. This effect is comparable to borrower-level fundamentals such as the interest rate and FICO score, and remains important even after controlling for other local characteristics, detailed loan and borrower fundamentals, and fixed effects.

Several other local variables also behave as expected. In particular, *HPA* is positive and highly significant, consistent with prior evidence that stronger house price appreciation encourages households to take on more leverage in order to capture potential gains (Adelino et al., 2016). *Total population* is also positive and highly significant, suggesting that more populous counties are associated with greater second-lien misrepresentation. These results indicate that standard local borrowing conditions matter for excessive mortgage leverage. At the same time, the striking result is that *CPRATIO*, which captures a borrower behavioral trait rather than a conventional market condition, remains one of the most robust local predictors in the table.

Overall, the results show that although several local characteristics are related to second-lien misrepresentation, *CPRATIO* stands out as the most robust behavioral predictor. It remains significant when entered alongside a wide range of alternative local controls, and its low pairwise correlations with most other local characteristics further suggest that it captures distinct local variation rather than relabeling standard socioeconomic conditions. These findings motivate the paper's subsequent focus on *CPRATIO* as a proxy for local gambling preference and, more broadly, as a behavioral determinant of excessive mortgage leverage.

4 Interpreting the Effect of Gambling Preference: Borrower Behavior or Lender Practice

The previous section shows that second-lien misrepresentation is more likely in counties with higher values of *CPRATIO*. Following Kumar et al. (2011), I interpret *CPRATIO* as a measure of local gambling preference, where gambling preference refers to a willingness to accept a high probability of small losses for a small chance of a large gain (Markowitz, 1952). This behavioral tendency can affect financial decisions by making households more willing to pursue opportunities with positively skewed payoffs. In my setting, it implies that borrowers in areas with stronger gambling preferences may be more willing to obtain excessive mortgage leverage through second-lien misrepresentation.

From the borrower's perspective, second-lien misrepresentation creates such an opportunity. The upside includes access to homeownership and greater exposure to housing-market gains. Homeownership carries both emotional and financial appeal: emotionally, it is closely tied to the American Dream (Bush, 2003), while financially, it offers the potential for wealth accumulation through house price appreciation and mortgage-interest tax deductions (Bostic and Lee, 2008; Goodman and Mayer, 2018). During the housing boom, this upside appeared especially attractive because households and lenders widely expected house prices to continue rising, making the perceived payoff distribution strongly right-skewed. Even if prices rose only modestly or stagnated, borrowers could still preserve some gains or at least retain homeownership.[5] The downside of misreporting was also limited. If detected before closing, the most likely outcome was loan denial, implying only modest application costs and little switching effort in an environment with abundant credit supply. If discovered after closing, borrowers could face fines, restitution, or probation (Henning, 2009). In practice, however, enforcement during this period focused primarily on fraud for profit rather than fraud for housing (FBI, 2006). This asymmetry between large perceived gains and limited expected losses makes second-lien misrepresentation particularly attractive to borrowers in areas with stronger gambling preferences.

At the same time, because *CPRATIO* is measured at the local level rather than the borrower level, the estimated relationship may also reflect lender practices rather than borrower behavior. Prior research shows that both originators and underwriters can play important roles in mortgage misreporting and may be aware of hidden second liens (Griffin and Maturana, 2016; Piskorski et al., 2015). More recently, Yavas and Zhu (2024) show that second-lien misrepresentation tends to arise primarily in the early stages of intermediation rather than at the underwriting stage. These findings raise the possibility that the effect of local gambling preference may partly operate through lender-side facilitation.

The remainder of this section therefore evaluates whether the relationship between *CPRATIO*

[5] If house prices decline, borrowers who continue making payments can still keep their homes. If they instead default because of negative equity, the loss mainly consists of the down payment and the portion of principal already repaid. Because high leverage requires only a small initial equity stake, this total loss remains limited for borrowers who take on excessive leverage.

and second-lien misrepresentation is better explained by borrower behavior or by lender practices.

4.1 Evidence from Borrower Behavior

Christensen et al. (2018) argue that the effect of gambling preference on misreporting should be more pronounced when individuals face stronger motives. Because my interpretation emphasizes the borrower's perspective, I test whether the effect of *CPRATIO* is amplified in settings where borrowers have stronger incentives to obtain the loan.

I begin with subsamples defined by loan characteristics that directly capture borrower motives: property occupancy type (primary or non-primary) and loan purpose (purchase or refinance). Primary occupancy indicates that the property is owner-occupied and therefore tied to the borrower's housing consumption, whereas non-primary occupancy refers to second-home or investment properties. Because the motive to misreport should be stronger when stable housing is at stake, I expect the effect of gambling preference to be more pronounced in the primary-occupancy subsample. Similarly, purchase loans should involve stronger motives than refinance loans because borrowers must obtain the loan to acquire the property, whereas refinance borrowers already own the home and primarily seek to improve financing terms or extract equity.

Table 3 reports the results. Columns (1) and (2) show that the coefficient on *CPRATIO* is positive and statistically significant in the primary-occupancy subsample, but smaller and insignificant in the non-primary subsample.[6] Columns (3) and (4) show that the coefficient on *CPRATIO* is significant at the 1 percent level in the purchase subsample but only marginally significant in the refinance subsample. The point estimate is also substantially larger in the purchase subsample, and the difference between the two coefficients is statistically significant at the 5 percent level. These

[6]Wald tests of equality between the *CPRATIO* coefficients for each subsample split are reported in Appendix A.

results suggest that the relationship between gambling preference and second-lien misrepresentation is stronger when borrowers have greater incentives to obtain the loan.

I next examine subsamples based on area income, borrower credit score, and past house price appreciation, which connect the analysis to broader explanations for mortgage borrowing during the pre-crisis period while also capturing variation in borrower motives. The high- and low-income subsamples are defined relative to the sample median of ZIP code-level median income, the high- and low-credit-score subsamples use a FICO cutoff of 670,[7] and the high- and low-*HPA* subsamples are defined relative to the state-year median of ZIP code-level house price appreciation.

Area income and credit score relate naturally to the credit-expansion view of the financial crisis. Although gambling preference may matter across all groups, its effect may be stronger for low-income and low-FICO borrowers because tighter financial constraints or weaker credit access increase the incentive to misreport. Consistent with this interpretation, Table 4, columns (1)–(4), shows that the coefficient on *CPRATIO* is economically large and statistically significant across groups, with larger point estimates and smaller p-values in the low-income and low-FICO subsamples, although the differences across groups are not statistically significant. Past house price appreciation relates to the belief-distortion channel, under which households in high-*HPA* areas are more likely to extrapolate recent price increases and therefore perceive greater returns to leverage (Shiller, 2007). Consistent with this idea, Table 5, columns (1) and (2), shows that the coefficient on *CPRATIO* is larger and more precisely estimated in the high-*HPA* subsample. Taken together, these results are consistent with both the credit-expansion and belief-distortion channels while continuing to support the borrower-side interpretation of gambling preference.

[7] A FICO score of 670 is the boundary between fair and good credit quality. Using the sample median of 682 yields similar results.

Overall, the subsample evidence shows that second-lien misrepresentation rises more sharply with local gambling preference when borrowers face stronger motives to obtain leverage. This pattern is consistent with a borrower-side interpretation of *CPRATIO*.

4.2 Evidence from Lender Practice

To assess the lender-side interpretation, I use the securitization discontinuity around a FICO score of 620 documented by Keys et al. (2010). Because government-sponsored enterprise underwriting guidelines make low-documentation loans with FICO scores of 620 or above much easier to securitize, lenders face weaker incentives to screen such loans carefully. This setting therefore provides a useful shock to lender incentives. Griffin and Maturana (2016) further show that the rate of second-lien misrepresentation also jumps at this cutoff, suggesting that lender incentives to securitize are related to this form of misreporting.

I first use this lender-side shock to study whether lenders facilitate second-lien misrepresentation differently from other loan types. Specifically, I examine discontinuities in both the number of loans and default rates for all loans, loans with correctly reported simultaneous second liens, and loans with misreported simultaneous second liens. [8] If lenders specifically facilitate second-lien misrepresentation, the jump in the number of loans with misreported simultaneous second liens should be larger than the jump for other loan types. In addition, if this facilitation reflects laxer screening, the jump in the default rate for loans with misreported simultaneous second liens should exceed that for loans with correctly reported simultaneous second liens. By contrast, if the jump in misreported loans is no larger than for other loans, or if the default-rate discontinuity is not

[8]Section 6 reports results using misrepresentation rates. I focus here on loan counts because they more directly reflect how the securitization shock affects each loan type.

stronger for misreported than for correctly reported second liens, then lenders do not appear to play a distinct role in facilitating misrepresentation.

I next examine whether the estimated effect of local gambling preference is consistent with lender practices. Because the variation associated with *CPRATIO* may in principle reflect either borrower behavior or lender behavior, the securitization shock provides a way to isolate the lender channel. If lenders facilitate second-lien misrepresentation more aggressively in high-gambling-preference areas, then the discontinuities in the number of misreported loans and in their default rates should be larger in those areas than in low-gambling-preference areas. If instead these discontinuities are similar or smaller in high-gambling-preference areas, then the positive relationship between *CPRATIO* and second-lien misrepresentation is unlikely to be driven primarily by lender practices.

I implement a regression discontinuity design using local linear regressions. Let

$$\mathrm{C} = \text{Credit Score} - 620 \tag{2}$$

denote the normalized running variable for low-documentation loans, and define

$$\mathrm{D} = \begin{cases} 1, & if\ \mathrm{C} \geq 0 \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

as an indicator for observations above the threshold.

To compare high- and low-gambling-preference areas, I define

$$G = \begin{cases} 1, & if \text{ CPRATIO} \geq 1.2921 \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

where 1.2921 corresponds to the 90th percentile of the county-year *CPRATIO* distribution.[9] I use this cutoff because it identifies areas with especially high local gambling preference while still leaving enough observations for meaningful comparisons.

I use a bandwidth of 10, corresponding to FICO scores from 610 to 630. This window avoids nearby discontinuities discussed in the literature and is close to the data-driven optimal bandwidth of Calonico et al. (2020). I use a common bandwidth for comparability, and the results are unchanged when I instead use the optimal bandwidth. For the number of loans and default rates in high- and low-gambling-preference areas, I estimate

$$Y = \alpha + \beta D + \gamma C + \delta D \times C + \epsilon \quad (5)$$

where Y is either the default indicator for loan i originated at time t or the number of loans at each FICO score. The baseline specification uses a uniform kernel.

To estimate the difference in discontinuities between high- and low-gambling-preference areas, I follow Dickert-Conlin and Elder (2010) and Grembi et al. (2016) and estimate

$$Y = \beta_0 + \beta_1 D + \beta_2 C + \beta_3 D \times C + \beta_4 G + \beta_5 G \times C + \beta_6 G \times D + \beta_7 G \times C \times D + \epsilon \quad (6)$$

where β_6 captures the difference in the discontinuity at the FICO cutoff between high- and low-gambling-preference areas. When I compare jump ratios in the number of loans, I first rescale the

[9] I also use alternative cutoffs, such as 1, and obtain similar results.

data in each area by the estimated value at 620^{-} so that the discontinuities can be interpreted as proportional jumps relative to the left limit. The difference between the two discontinuities then gives the difference in jump ratios across areas.

I begin by examining the discontinuities in all areas to characterize lender behavior at the securitization cutoff. Panel A of Table 6 reports the jumps in the number of all loans, loans with correctly reported simultaneous second liens, and loans with misreported simultaneous second liens. The total number of low-documentation loans roughly doubles from 620^{-} to 620^{+}, consistent with Keys et al. (2010). Loans with simultaneous second liens, whether correctly reported or misreported, also increase sharply. This pattern suggests that the ease of securitization encourages lenders to originate more loans with simultaneous second liens in general.

Panel B of Table 6 reports the corresponding jumps in default rates. Consistent with prior work, the default rate for all loans rises by 0.039 at the cutoff, indicating weaker screening for loans just above 620. A statistically significant discontinuity also appears for loans with correctly reported simultaneous second liens, with a jump of 0.045. For loans with misreported simultaneous second liens, the estimated jump is similar in magnitude but statistically insignificant. Taken together, these results suggest that lenders facilitate loans with simultaneous second liens around the cutoff, but the evidence does not indicate a distinctly stronger screening failure for misreported than for correctly reported second liens.

I next compare high- and low-gambling-preference areas. Figure 2 shows that the jump ratio in the number of loans with misreported simultaneous second liens is smaller in high-gambling-preference areas than in low-gambling-preference areas. By contrast, the jump ratios for loans with correctly reported simultaneous second liens are broadly similar across the two groups. Panel A of Table 7 also shows that the increase in the total number of loans is similar across high- and low-gambling-preference areas. These findings suggest that lenders do not facilitate misreported simultaneous second liens more aggressively in high-gambling-preference areas.[10]

Figure 3 compares the discontinuities in default rates across high- and low-gambling-preference areas. Both correctly reported and misreported simultaneous second liens exhibit smaller default-rate jumps in high-gambling-preference areas. When loan-level controls are included, Panel B of Table 7 shows that these differences are statistically insignificant, and the same pattern holds for all loans. Overall, the evidence does not support the view that lenders in high-gambling-preference areas facilitate second-lien misrepresentation through especially lax screening. Combined with the earlier results on loan counts, these findings suggest that the positive relationship between *CPRATIO* and second-lien misrepresentation is more consistent with borrower behavior than with lender practices.

5 Outcomes of Misrepresentation Across Gambling Preference Levels

The previous section shows that local gambling preference is positively associated with second-lien misrepresentation. I next examine whether it also shapes the consequences of misrepresentation. Prior studies show that mortgage misrepresentation is associated with worse loan performance (Piskorski et al., 2015; Griffin and Maturana, 2016), but it remains unclear whether misreported loans perform even worse in high-gambling-preference environments. I therefore estimate loan-level regressions of delinquency on indicators for misreported and correctly reported simultaneous

[10] If borrowers account for more of the misreporting in high-gambling-preference areas, then the marginal effect of the lender-side securitization shock on the number of misreported loans will naturally be smaller there.

second liens and compare the estimates across high- and low-gambling-preference areas defined in Section 4.2 using the following specification:

$$\begin{aligned}\text{Default} = \ & \alpha + \beta_1 \times \text{Misreported second} + \beta_2 \times \text{Reported second} \\ & + \beta_3 \times \text{Controls} + \text{Fixed effects}\end{aligned} \tag{7}$$

where *Default* is an indicator equal to one if the loan becomes 90 days or more delinquent under the MBA method. *Misreported second* and *Reported second* indicate first-lien loans with simultaneously originated second liens that are, respectively, misreported or correctly reported at origination. *Controls* include the full set of local characteristics and loan-level variables from the baseline analysis, except that I replace the first-lien loan-to-value ratio (*LTV*) with the reported combined loan-to-value ratio (*CLTV*), which better captures total leverage for loans with reported simultaneous second liens.[11] *Fixed effects* include originator, state, and half-year fixed effects. All specifications include the full set of controls, fixed effects, and county-clustered standard errors.

Table 8 reports the results. Column (1) presents estimates for high-gambling-preference areas, and column (2) presents estimates for low-gambling-preference areas. In both subsamples, loans with misreported simultaneous second liens default at significantly higher rates than loans without second liens. More importantly, the estimated default penalty is larger in high-gambling-preference areas, suggesting that the adverse consequences of concealed leverage are amplified in stronger gambling environments. Loans with correctly reported simultaneous second liens also default more often than loans without second liens, even after controlling for *CLTV*, and this effect is likewise larger in high-gambling-preference areas. Although my focus is on misreported simultaneous

[11] I do not include *CLTV* in the baseline misreporting regressions because misrepresentation is identified mechanically from the reported *CLTV*.

second liens as the more aggressive form of leverage, the results more broadly indicate that high local gambling preference is associated with greater fragility in high-leverage borrowing.

The economic magnitude is also meaningful. In low-gambling-preference areas, misreported simultaneous second liens are associated with an 11.3 percentage-point increase in the probability of default, compared with a 15.7 percentage-point increase in high-gambling-preference areas. The difference of 4.4 percentage points implies an additional default penalty associated with the gambling environment. Multiplying this differential by the 93,925 misreported loans originated in high-gambling-preference areas yields roughly 4,133 additional defaults. At an average first-lien balance of about $325,000, this corresponds to approximately $1.34 billion in incremental first-lien exposure. This back-of-the-envelope calculation indicates that the amplification of misreported-loan risk in gambling-prone areas is economically important.

6 Robustness

6.1 Nonlinear Model

The baseline results are estimated using a linear probability model. To ensure that the findings are not driven by this modeling choice, I also estimate a probit model. One concern is that restricting the sample to loans with simultaneous second liens changes the relevant population. To address this issue, I estimate the probit model in the full sample without including a loan-level indicator for simultaneous second liens. Instead, I control for the prevalence of simultaneous second liens at the county level. Table 9 reports the corresponding marginal effects. *CPRATIO* remains robustly significant, while the significance of other demographic variables is less stable across specifications.

6.2 Causal Forest and Causality

To strengthen the causal interpretation of the relation between gambling preference and mortgage misrepresentation, and to study how the effect of misrepresentation on default varies with gambling preference, I employ the causal forest method of Wager and Athey (2018). This approach combines augmented inverse propensity weighting with random forests under an honesty condition and is well suited for observational settings.[12]

I first study the effect of gambling preference on second-lien misrepresentation, treating misrepresentation as the outcome and *CPRATIO* as the treatment. I use the same controls as in Table 2 as matching covariates, require observations to be non-missing, and winsorize and standardize all continuous variables. Because fixed effects are not directly applicable, I include a half-year time variable and cluster by state.[13] Following Athey and Wager (2019), I grow 2,000 trees and use the default 50 percent honesty split. Table 10 shows that the estimated effect of *CPRATIO* on second-lien misrepresentation remains significant and similar in magnitude to the OLS estimates.

I next study heterogeneous treatment effects of mortgage misrepresentation on default across different levels of gambling preference, treating default as the outcome and mortgage misrepresentation as the treatment. I use the same matching variables as above and additionally include *CPRATIO*. Using the same forest settings, I compute average treatment effects of mortgage misrepresentation on default for each quartile of the sample sorted by *CPRATIO*. Figure 4 shows that the treatment effects are generally larger at higher levels of *CPRATIO*, consistent with the OLS results.

[12] See Wager and Athey (2018) for the statistical framework, Athey and Wager (2019) for an empirical application, and Gulen et al. (2021) for a finance application.

[13] The `grf` package in R requires numeric matching variables. I therefore do not encode originators numerically, which could impose an artificial distance structure, and I do not cluster by originator because the large number of originators would create excessive computational burden.

6.3 Alternative Gambling Preference Measures

Although *CPRATIO* is the main measure of gambling preference, I also consider alternative proxies to address the concern that it may capture broader religious differences rather than gambling-specific attitudes. First, I use per capita employment in the gambling industry at the core-based statistical area (CBSA) level as a supply-based measure of local gambling exposure. I regress second-lien misrepresentation on this measure together with CBSA-level controls, loan-level controls, and fixed effects. See Appendix B for data details.

Second, following Kumar et al. (2011), I adjust *CPRATIO* to incorporate the shares of Latter-Day Saints and Jewish populations. Because Jewish populations tend to resemble Catholics in their acceptance of gambling, whereas Latter-Day Saints are more similar to Protestants, this adjusted measure continues to capture gambling-related religious variation while reducing the concern that *CPRATIO* reflects other Catholic–Protestant differences. As in the baseline analysis, I regress second-lien misrepresentation on this adjusted measure together with the full set of local controls.

Table 11 reports the results. Column (1) uses per capita gambling-industry employment, and column (2) uses the adjusted *CPRATIO*. Both measures yield significantly positive coefficients, supporting the interpretation that second-lien misrepresentation is more prevalent in areas with stronger gambling preference.

6.4 RDD and Diff-in-disc Concerns

The RDD analysis is intended to show that lenders do not specifically facilitate second-lien misrepresentation relative to loans with simultaneous second liens more generally. For this interpretation to be valid, borrower and collateral fundamentals should be smooth at the FICO cutoff. Table 12 shows that local characteristics such as *CPRATIO* and *REL*, as well as collateral fundamentals such as *LTV* and *Balance*, are smooth around the cutoff.[14]

Because the main lender-side evidence relies on comparing high- and low-gambling-preference areas, I also assess the validity of the diff-in-disc design. I examine the three core assumptions discussed by Grembi et al. (2016) and evaluate sensitivity to controls, bandwidth choice, and kernel specification. The findings are robust across these exercises. See Appendix C for details. Overall, these tests support the validity of the diff-in-disc approach and continue to provide no evidence that the effect of gambling preference operates through lender practices.

6.5 Misrepresentation Rate

In the lender-practice analysis, I focus on discontinuities in the number of loans because this measure makes it easier to distinguish changes across loan types. A potential concern, however, is that loan counts are aggregated and therefore do not incorporate the same control variables used in the baseline loan-level regressions. To address this issue, I also examine discontinuities in the rates of correctly reported and misreported simultaneous second liens using specifications that include the same control variables as in the baseline analysis.[15]

Table 13 reports the raw-rate estimates, estimates with controls, and estimates with controls plus fixed effects. Correctly reported simultaneous second liens behave similarly across high- and low-gambling-preference areas, with slightly larger jumps in high-gambling-preference areas. By

[14] This validity check differs from the use of the 620 cutoff in Keys et al. (2010) and Yavas and Zhu (2024). My purpose is not to use the cutoff to establish the causal effect of securitization on lender screening per se, but to benchmark misreporting against the facilitation of simultaneous-second origination more generally. In this setting, continuity of borrower and collateral fundamentals is the key requirement. Discontinuities in contract terms such as interest rates, ARM incidence, prepayment penalties, negative amortization, cash-out refinancing, or occupancy type are part of the treatment channel rather than threats to the design.

[15] I do not calculate rates at the county level. Instead, I use the loan-level misrepresentation indicator directly so that loan-level controls can be included.

contrast, misreported simultaneous second liens exhibit smaller jumps in high-gambling-preference areas, although the difference becomes less precise once all controls and fixed effects are included. These results continue to provide no support for the view that lenders facilitate more second-lien misrepresentation in high-gambling-preference areas.

6.6 Alternative Measures of Default

Borrowers may default for different reasons and over different horizons. The baseline measure defines default as 90 days or more delinquent under the MBA method within the first three years after origination. To examine whether the results are sensitive to outcome definition, I consider three alternatives: 90 days or more delinquent within the first two years, 60 days or more delinquent within the first three years, and bankruptcy/foreclosure/REO within the first three years.

Table 14 reports the results. The effect of second-lien misrepresentation remains significant across all default definitions, and in every specification the coefficient is larger in high-gambling-preference areas. In addition, the effect is larger when the definition of default is less restrictive. Overall, the findings from Section 5 are robust to alternative measures of loan performance.

7 Conclusion

This paper studies the role of borrower behavioral heterogeneity in excessive mortgage leverage. Using misreported simultaneous second liens as a revealed measure of excessive mortgage leverage, I show that such borrowing was more prevalent in areas with stronger local gambling preference. The relationship is stronger where borrower incentives to obtain leverage are greater, including owner-occupied and purchase loans, and is not readily explained by lender facilitation around the

FICO 620 securitization cutoff. Excessive mortgage leverage in high-gambling-preference areas also exhibits worse ex post performance.

These findings suggest that the buildup of excessive mortgage leverage cannot be understood solely through broad aggregate forces such as credit supply expansion or optimistic beliefs about house prices. Borrower behavioral traits also help explain which households were more willing to stretch for leverage in the same borrowing environment. In this sense, local gambling preference contributed not only to excessive mortgage leverage, but also to the financial fragility associated with that leverage.

More broadly, the paper highlights the importance of incorporating behavioral heterogeneity into the study of household finance and mortgage markets. Even when aggregate borrowing conditions are similar, differences in local behavioral traits can shape how aggressively households respond to those conditions and how vulnerable their borrowing becomes ex post.

References

Adelino, M., Schoar, A., Severino, F., 2016. Loan originations and defaults in the mortgage crisis: The role of the middle class. Review of Financial Studies 29, 1635–1670.

Alesina, A., La Ferrara, E., 2000. Participation in heterogeneous communities. Quarterly Journal of Economics 115, 847–904.

Athey, S., Wager, S., 2019. Estimating treatment effects with causal forests: An application. Observational Studies 5, 37–51.

Barberis, N., Huang, M., 2008. Stocks as lotteries: The implications of probability weighting for security prices. American Economic Review 98, 2066–2100.

Bostic, R.W., Lee, K.O., 2008. Mortgages, risk, and homeownership among low- and moderate-income families. American Economic Review 98, 310–314.

Bush, G.W., 2003. Remarks on Signing the American Dream Downpayment Act. URL: `https://www.presidency.ucsb.edu/documents/remarks-signing-the-american-dream-downpayment-act`.

Calonico, S., Cattaneo, M.D., Farrell, M.H., 2020. Optimal bandwidth choice for robust bias corrected inference in regression discontinuity designs. Econometrics Journal 23, 192–210.

Calonico, S., Cattaneo, M.D., Titiunik, R., 2014. Robust nonparametric confidence intervals for regression-discontinuity designs. Econometrica 82, 2295–2326.

Campbell, J.Y., 2006. Household finance. Journal of Finance 61, 1553–1604.

Chen, Y., Podolski, E.J., Rhee, S.G., Veeraraghavan, M., 2014. Local gambling preferences and corporate innovative success. Journal of Financial and Quantitative Analysis 49, 77–106.

Christensen, D.M., Jones, K.L., Kenchington, D.G., 2018. Gambling attitudes and financial misreporting. Contemporary Accounting Research 35, 1229–1261.

Conklin, J., Diop, M., Qiu, M., 2022. Religion and mortgage misrepresentation. Journal of Business Ethics 179, 273–295.

Dickert-Conlin, S., Elder, T., 2010. Suburban legend: School cutoff dates and the timing of births. Economics of Education Review 29, 826–841.

Eckert, F., Fort, T.C., Schott, P.K., Yang, N.J., 2021. Imputing missing values in the us census bureau's county business patterns. Working paper.

FBI, 2006. Mortgage Fraud Report 2006. Technical Report. Federal Bureau of Investigation.

Goodman, L.S., Mayer, C., 2018. Homeownership and the american dream. Journal of Economic Perspectives 32, 31–58.

Grembi, V., Nannicini, T., Troiano, U., 2016. Do fiscal rules matter? American Economic Journal: Applied Economics 8, 1–30.

Griffin, J.M., Maturana, G., 2016. Who facilitated misreporting in securitized loans? Review of Financial Studies 29, 384–419.

Gulen, H., Jens, C.E., Page, T.B., 2021. The heterogeneous effects of default on investment: An application of causal forest in corporate finance. Working paper.

Han, B., Kumar, A., 2013. Speculative retail trading and asset prices. Journal of Financial and Quantitative Analysis 48, 377–404.

Henning, A.C., 2009. Mortgage Fraud: Federal Criminal Provisions. Technical Report. Congressional Research Service.

Hilary, G., Hui, K.W., 2009. Does religion matter in corporate decision making in america? Journal of Financial Economics 93, 455–473.

Justiniano, A., Primiceri, G.E., Tambalotti, A., 2019. Credit supply and the housing boom. Journal of Political Economy 127, 1317–1350.

Keys, B.J., Mukherjee, T., Seru, A., Vig, V., 2010. Did securitization lead to lax screening? evidence from subprime loans. Quarterly Journal of Economics 125, 307–362.

Kumar, A., 2009. Who gambles in the stock market? Journal of Finance 64, 1889–1933.

Kumar, A., Page, J.K., Spalt, O.G., 2011. Religious beliefs, gambling attitudes, and financial market outcomes. Journal of Financial Economics 102, 671–708.

Markowitz, H., 1952. The utility of wealth. Journal of Political Economy 60, 151–158.

Mian, A., Sufi, A., 2009. The consequences of mortgage credit expansion: Evidence from the u.s. mortgage default crisis. Quarterly Journal of Economics 124, 1449–1496.

Piskorski, T., Seru, A., Witkin, J., 2015. Asset quality misrepresentation by financial intermediaries: Evidence from the rmbs market. Journal of Finance 70, 2635–2678.

Shiller, R.J., 2007. Understanding Recent Trends in House Prices and Home Ownership. NBER Working Paper 13553. National Bureau of Economic Research.

Shu, T., Sulaeman, J., Yeung, P.E., 2012. Local religious beliefs and mutual fund risk-taking behaviors. Management Science 58, 1779–1796.

Wager, S., Athey, S., 2018. Estimation and inference of heterogeneous treatment effects using random forests. Journal of the American Statistical Association 113, 1228–1242.

Wooldridge, J.M., 2010. Econometric Analysis of Cross Section and Panel Data. 2 ed., MIT Press, Cambridge, MA.

Yavas, A., Zhu, S., 2024. Misreporting of second liens in portfolio mortgages and privately securitized mortgages. Real Estate Economics 52, 73–109.

Table 1
Descriptive Statistics

	N	Mean	SD	P10	P25	P50	P75	P90
Simultaneous second liens (loan level)								
Misreported second (%)	3,031,921	7.15	25.76					
Reported second (%)	3,031,921	13.91	34.61					
Simul. second (%)	3,031,921	21.06	40.77					
Loan characteristics (loan level)								
Interest rate (%)	3,026,958	6.41	2.12	2.75	5.88	6.50	7.55	8.75
FICO	2,993,630	681.78	72.60	578.00	632.00	688.00	739.00	776.00
Balance (ln)	3,031,921	12.40	0.73	11.44	11.87	12.40	12.98	13.31
LTV (%)	3,031,921	74.82	13.34	56.65	70.00	80.00	80.00	90.00
CLTV (%)	3,031,921	78.99	15.71	57.86	72.00	80.00	90.00	100.00
ARM (%)	3,031,921	58.87	49.21					
Option ARM (%)	3,031,921	0.24	4.84					
Negative amortization (%)	3,031,921	13.58	34.26					
Low or no doc. (%)	3,031,921	65.42	47.56					
Prepayment penalty (%)	3,031,921	40.99	49.18					
Cash-out (%)	3,031,921	40.66	49.12					
No-cash-out (%)	3,031,921	14.73	35.44					
Investment (%)	3,031,921	10.27	30.36					
Second-home (%)	3,031,921	3.55	18.49					
Default (%)	3,014,278	23.51	42.40					
Local characteristics (county-year level)								
CPRATIO	8,716	0.51	0.88	0.01	0.04	0.19	0.54	1.29
REL (%)	8,716	59.59	17.64	36.62	45.90	58.51	72.42	89.21
Crime rate (per 10,000 residents)	8,228	279.34	150.46	113.14	161.38	244.84	358.32	496.85
Local characteristics (county-month level)								
Unemployment (%)	71,156	5.14	1.73	3.20	3.90	4.90	6.00	7.50
Local characteristics (county level)								
Education (%)	3,035	16.64	7.51	9.70	11.20	14.40	19.30	26.60
Married (%)	3,035	60.39	5.10	53.90	57.90	61.30	63.90	66.00
Urban (%)	3,035	40.62	30.54	0.00	12.93	40.32	64.65	84.67
Total population (ln)	3,035	10.49	1.11	9.52	9.52	10.15	11.05	12.09
Over65 (%)	3,035	14.71	4.05	9.90	12.10	14.40	17.00	20.00
Male-female ratio	3,035	0.98	0.06	0.92	0.94	0.97	1.00	1.05
Minority (%)	3,035	15.32	15.80	2.02	3.34	8.71	22.96	37.90
Republican	3,015	0.23	0.42					
Local characteristics (ZIP code-year level)								
HPA	76,535	18.45	14.28	4.23	7.55	13.93	26.56	39.99
Local characteristics (ZIP code level)								
Income (ln)	26,517	10.57	0.34	10.17	10.34	10.53	10.77	11.02

The table presents descriptive statistics for the variables used in the analysis, covering the sample period from 2005 to 2007. For all continuous variables, I report the number of observations, mean, standard deviation, 10th percentile, 25th percentile, 50th percentile, 75th percentile, and 90th percentile. For all dummy variables, I report the number of observations, mean, and standard deviation. The variables are winsorized at the 0.5 percent level. (ln) indicates that the value of the variable is the natural logarithm of the original value. (%) indicates that the value of the variable is expressed as a percentage.

Table 2
Local Socioeconomic Characteristics and Second-lien Misrepresentation

	(1)	(2)	(3)	(4)	(5)
CPRATIO	0.0017***				0.0014***
	(0.000)				(0.000)
REL	0.0001				-0.0001
	(0.000)				(0.000)
Crime rate	0.0008***				0.0008**
	(0.000)				(0.000)
Republican	0.0003				0.0009
	(0.001)				(0.001)
Education		0.0004			0.0005*
		(0.000)			(0.000)
Income		-0.0008***			-0.0004**
		(0.000)			(0.000)
Unemployment			-0.0002		-0.0008
			(0.000)		(0.001)
HPA			0.0017***		0.0017***
			(0.001)		(0.001)
Urban			-0.0003		-0.0007**
			(0.000)		(0.000)
Total population			0.0014***		0.0010***
			(0.000)		(0.000)
Married				-0.0001	0.0002
				(0.000)	(0.000)
Over65				0.0002	0.0006**
				(0.000)	(0.000)
Male-female ratio				-0.0003	-0.0000
				(0.000)	(0.000)
Minority				0.0010**	0.0006
				(0.000)	(0.000)
Simul. Second	0.3125***	0.3130***	0.3126***	0.3126***	0.3130***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Interest rate	-0.0072***	-0.0074***	-0.0073***	-0.0073***	-0.0072***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
FICO	-0.0063***	-0.0063***	-0.0062***	-0.0063***	-0.0062***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Balance	0.0049***	0.0051***	0.0048***	0.0050***	0.0048***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
LTV	0.0056***	0.0055***	0.0056***	0.0056***	0.0056***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ARM	0.0087***	0.0088***	0.0086***	0.0086***	0.0088***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Negative amortization	-0.0037***	-0.0044***	-0.0038***	-0.0040***	-0.0041***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Option ARM	-0.0227***	-0.0227***	-0.0227***	-0.0227***	-0.0227***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Prepayment penalty	-0.0079***	-0.0080***	-0.0079***	-0.0078***	-0.0082***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Low or no doc.	-0.0078***	-0.0079***	-0.0080***	-0.0078***	-0.0080***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Table 2 (continued)

	(1)	(2)	(3)	(4)	(5)
Cash-out	0.0184***	0.0181***	0.0184***	0.0184***	0.0182***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
No-cash-out	0.0156***	0.0156***	0.0158***	0.0157***	0.0158***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Investment	0.0259***	0.0260***	0.0259***	0.0259***	0.0258***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Second-home	0.0157***	0.0158***	0.0157***	0.0157***	0.0158***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Originator FE	Y	Y	Y	Y	Y
State FE	Y	Y	Y	Y	Y
Half-year FE	Y	Y	Y	Y	Y
Observations	2,453,157	2,389,104	2,481,611	2,486,167	2,353,016
Adj. R^2	0.317	0.318	0.317	0.317	0.318

The table shows OLS estimates from regressions where the dependent variable takes a value of one if the loan is recognized as second-lien misreported and zero otherwise. All continuous variables are winsorized at the 0.5 percent level and then standardized. Standard errors clustered at the county level are reported in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Table 3
Gambling Preference and Second-lien Misrepresentation - Borrower Motives

	(1) Primary	(2) Non-primary	(3) Purchase	(4) Refinance
CPRATIO	0.0014***	0.0008	0.0022***	0.0006*
	(0.000)	(0.001)	(0.001)	(0.000)
REL	-0.0001	-0.0003	-0.0005	0.0001
	(0.000)	(0.001)	(0.001)	(0.000)
Crime rate	0.0010**	0.0009	0.0016**	0.0003
	(0.000)	(0.001)	(0.001)	(0.000)
Republican	0.0009	0.0005	0.0012	0.0005
	(0.001)	(0.001)	(0.001)	(0.000)
Education	0.0007**	-0.0011*	0.0008	0.0003
	(0.000)	(0.001)	(0.001)	(0.000)
Income	-0.0011***	0.0016***	-0.0003	-0.0003**
	(0.000)	(0.000)	(0.000)	(0.000)
Unemployment	-0.0008	-0.0011*	-0.0010	-0.0001
	(0.001)	(0.001)	(0.001)	(0.000)
HPA	0.0018***	-0.0000	0.0021**	0.0019***
	(0.001)	(0.001)	(0.001)	(0.000)
Urban	-0.0008**	0.0010	-0.0008	-0.0006**
	(0.000)	(0.001)	(0.001)	(0.000)
Total population	0.0011***	-0.0010	0.0020***	0.0007**
	(0.000)	(0.001)	(0.001)	(0.000)
Married	0.0004	-0.0004	0.0008	-0.0002
	(0.000)	(0.001)	(0.001)	(0.000)
Over65	0.0006**	0.0004	0.0001	0.0005**
	(0.000)	(0.000)	(0.000)	(0.000)
Male-female ratio	-0.0001	0.0007	-0.0001	0.0003
	(0.000)	(0.001)	(0.000)	(0.000)
Minority	0.0007	-0.0006	0.0009	0.0003
	(0.000)	(0.001)	(0.001)	(0.000)
Loan chars controls	Y	Y	Y	Y
Originator FE	Y	Y	Y	Y
State FE	Y	Y	Y	Y
Half-year FE	Y	Y	Y	Y
Observations	2,052,270	300,095	1,007,280	1,344,733
Adj. R^2	0.312	0.418	0.315	0.382

The table shows OLS estimates from regressions where the dependent variable takes a value of one if the loan is recognized as second-lien misreported and zero otherwise. I look at subsamples by borrower motives: primary or non-primary (column (1) and (2)) and purchase or refinance (column (3) and (4)). All continuous variables are winsorized at the 0.5 percent level and then standardized. Standard errors clustered at the county level are reported in parentheses. $^*p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Table 4
Gambling Preference and Second-Lien Misrepresentation – Credit Expansion

	(1) High Income	(2) Low Income	(3) High FICO	(4) Low FICO
CPRATIO	0.0013**	0.0017***	0.0012**	0.0017***
	(0.001)	(0.001)	(0.001)	(0.000)
REL	0.0001	-0.0010**	-0.0008*	0.0003
	(0.000)	(0.000)	(0.000)	(0.000)
Crime rate	0.0008	0.0012**	0.0012**	0.0005
	(0.000)	(0.001)	(0.001)	(0.000)
Republican	0.0013*	0.0002	0.0007	0.0011
	(0.001)	(0.001)	(0.001)	(0.001)
Education	0.0007**	-0.0002	0.0003	0.0009**
	(0.000)	(0.001)	(0.000)	(0.000)
Income	0.0001	-0.0017**	-0.0001	-0.0008**
	(0.000)	(0.001)	(0.000)	(0.000)
Unemployment	-0.0006	-0.0012**	-0.0013**	-0.0002
	(0.001)	(0.001)	(0.001)	(0.001)
HPA	0.0025***	-0.0004	0.0021***	0.0022***
	(0.001)	(0.001)	(0.001)	(0.001)
Urban	-0.0004	-0.0012**	-0.0003	-0.0011***
	(0.000)	(0.000)	(0.000)	(0.000)
Total population	0.0007	0.0019***	0.0016***	0.0004
	(0.000)	(0.001)	(0.001)	(0.001)
Married	0.0000	0.0007	0.0007	-0.0007*
	(0.000)	(0.001)	(0.001)	(0.000)
Over65	0.0007**	0.0002	0.0003	0.0005
	(0.000)	(0.000)	(0.000)	(0.000)
Male-female ratio	0.0003	-0.0006*	0.0005	-0.0003
	(0.000)	(0.000)	(0.000)	(0.000)
Minority	0.0009*	-0.0000	0.0007	0.0003
	(0.001)	(0.001)	(0.001)	(0.000)
Loan chars controls	Y	Y	Y	Y
Originator FE	Y	Y	Y	Y
State FE	Y	Y	Y	Y
Half-year FE	Y	Y	Y	Y
Observations	1,809,909	542,366	1,308,243	1,044,300
Adj. R^2	0.315	0.329	0.298	0.358

The table shows OLS estimates from regressions where the dependent variable takes a value of one if the loan is recognized as second-lien misreported and zero otherwise. I look at subsamples related to credit expansion argument: high or low income ((1) and (2)) and high or low credit score ((3) and (4)). The high or low income subsamples are defined by comparing local household median income with the sample median at the ZIP code level. The high or low credit score subsamples are divided by a FICO score of 670. All continuous variables are winsorized at the 0.5 percent level and then standardized. Standard errors clustered at the county level are reported in parentheses. $*p < 0.10$, $**p < 0.05$, and $***p < 0.01$.

Table 5
Gambling Preference and Second-Lien Misrepresentation – Housing Expectation

	(1) High HPA	(2) Low HPA
CPRATIO	0.0018***	0.0008
	(0.000)	(0.001)
REL	-0.0010***	0.0006
	(0.000)	(0.000)
Crime rate	0.0002	0.0013**
	(0.000)	(0.001)
Republican	0.0027***	-0.0016*
	(0.001)	(0.001)
Education	0.0008*	0.0007*
	(0.000)	(0.000)
Income	-0.0007**	-0.0000
	(0.000)	(0.000)
Unemployment	-0.0004	-0.0004
	(0.001)	(0.001)
HPA	0.0005	0.0055***
	(0.001)	(0.001)
Urban	-0.0011**	-0.0004
	(0.000)	(0.000)
Total population	0.0017***	0.0008
	(0.000)	(0.001)
Married	0.0002	0.0004
	(0.000)	(0.001)
Over65	0.0003	0.0007*
	(0.000)	(0.000)
Male-female ratio	-0.0004	0.0006
	(0.000)	(0.000)
Minority	0.0008	0.0004
	(0.001)	(0.001)
Loan chars controls	Y	Y
Originator FE	Y	Y
State FE	Y	Y
Half-year FE	Y	Y
Observations	1,352,931	999,126
Adj. R^2	0.321	0.315

The table shows OLS estimates from regressions where the dependent variable takes a value of one if the loan is recognized as second-lien misreported and zero otherwise. I look at subsamples related to housing expectation: high or low HPA areas ((1) and (2)). The high and low HPA subsamples are defined based on the state-year median HPA, calculated from all available ZIP code-level data. All continuous variables are winsorized at the 0.5 percent level and then standardized. Standard errors clustered at the county level are reported in parentheses. $*p < 0.10$, $**p < 0.05$, and $***p < 0.01$.

Table 6
Discontinuities for Low-documentation Loans Around the Credit Threshold in All Areas

	(1) All loan	(2) Reported second	(3) Misreported second
Panel A. Jump Ratio in Number of Loans			
Est. 620+/620-	2.266	4.939	3.838
Panel B. Default Rate			
FICO≥620 (β)	0.038	0.045	0.046
t-stat	(3.89)	(2.18)	(-0.84)

The table reports the discontinuities for low-documentation loans around a FICO score of 620 in all areas. Panel A presents the estimates from regressions where the dependent variable is the number of loans at each FICO score. Using local linear regressions of the RDD approach, I estimate the number of loans at FICO 620^- and FICO 620^+ and compute the ratio of the estimated number of loans at FICO 620^+ to 620^-. Panel B presents the estimates from regressions where the dependent variable takes a value of one if the loan becomes 90 days or more delinquent using the MBA method in the first three years after origination and zero otherwise. Using local linear regressions of the RDD approach with control variables, I estimate the difference in default rates between FICO 620^- and FICO 620^+. I perform the estimation for all loans, loans with correctly reported simultaneous seconds, and loans with misreported simultaneous seconds, and the results are reported in columns (1) to (3), respectively. Bias-corrected t-values (standard errors clustered at the county level) following Calonico et al. (2014) are reported in parentheses.

Table 7
Discontinuities for Low-documentation Loans Around the Credit Threshold in High and Low Gambling Preference Areas

	(1) High	(2) Low	(3) High-Low
Panel A. Jump Ratio in Number of Loans (Est. 620+/620-)			
All loan	2.240	2.289	-0.049
			(-0.52)
Reported second	5.498	4.581	0.917
			(1.23)
Misreported second	2.763	4.826	-2.063
			(-2.83)
Panel B. Default Rate			
All loan	0.039	0.037	0.001
	(3.45)	(2.29)	(0.11)
Reported second	0.012	0.064	-0.052
	(0.22)	(2.64)	(-1.36)
Misreported second	0.017	0.077	-0.061
	(0.12)	(-1.09)	(-1.31)

The table reports the discontinuities for low-documentation loans around a FICO score of 620 separately in high and low gambling preference areas. Panel A presents the estimates from regressions where the dependent variable is the number of loans at each FICO score. Using local linear regressions of the RDD approach, we estimate the number of loans at FICO 620^{-} and FICO 620^{+} and compute the ratio of the estimated number of loans at FICO 620^{+} to 620^{-}. Panel B presents the estimates from regressions where the dependent variable is the default dummy variable. Using local linear regressions of the RDD approach with control variables, I estimate the difference in default rates between FICO 620^{-} and FICO 620^{+}. Using the diff-in-disc approach, we estimate the difference between high and low gambling preference areas. I perform the estimation for all loans, loans with correctly reported simultaneous seconds, and loans with misreported simultaneous seconds. The results for loans in high, low, and the difference between high and low gambling preference areas are reported in columns (1) to (3), respectively. For discontinuities, bias-corrected t-values (standard errors clustered at the county level) following Calonico et al. (2014) are reported in parentheses. For the difference between high and low gambling preference areas, t-statistics based on heteroskedasticity-consistent standard errors are reported in parentheses.

Table 8
Outcomes of Misrepresentation Across Gambling Preference Levels

	(1) High CPRATIO	(2) Low CPRATIO
Misreported second	0.157***	0.113***
	(0.004)	(0.006)
Reported second	0.111***	0.049***
	(0.006)	(0.003)
REL	-0.006	-0.001
	(0.004)	(0.002)
Crime rate	0.018***	0.001
	(0.005)	(0.003)
Republic	-0.014**	-0.007
	(0.007)	(0.006)
Education	0.004	0.001
	(0.004)	(0.003)
Income	-0.011***	-0.014***
	(0.003)	(0.002)
Unemployment	-0.004	0.009***
	(0.004)	(0.002)
HPA	0.042***	0.065***
	(0.006)	(0.007)
Urban	0.010**	-0.008***
	(0.004)	(0.002)
Total population	-0.014**	0.027***
	(0.006)	(0.004)
Married	0.021***	0.016***
	(0.003)	(0.003)
Over65	0.010*	0.000
	(0.005)	(0.003)
Male-female ratio	0.000	-0.002
	(0.005)	(0.002)
Minority	0.007	0.006
	(0.005)	(0.004)
Interest rate	0.024***	0.035***
	(0.002)	(0.002)
FICO	-0.091***	-0.101***
	(0.002)	(0.002)
Balance	0.022***	0.026***
	(0.003)	(0.001)
CLTV	0.077***	0.063***
	(0.003)	(0.004)
ARM	0.045***	0.034***
	(0.003)	(0.002)
Negative amortization	0.016***	0.042***
	(0.004)	(0.006)
Option ARM	0.049***	0.036***
	(0.011)	(0.009)
Prepayment penalty	0.064***	0.049***
	(0.003)	(0.002)
Low or no doc.	0.055***	0.052***
	(0.002)	(0.002)

Table 8 (continued)

	(1) High CPRATIO	(2) Low CPRATIO
Cash-out	0.004	-0.013***
	(0.003)	(0.002)
No-cash-out	0.040***	0.006
	(0.005)	(0.004)
Investment	0.031***	0.072***
	(0.005)	(0.004)
Second-home	0.012**	0.019***
	(0.005)	(0.005)
Originator FE	Y	Y
State FE	Y	Y
Half-year FE	Y	Y
Observations	1,149,680	1,201,389
Adj. R^2	0.259	0.211

The table shows OLS estimates from regressions where the dependent variable takes a value of one if the loan becomes 90 days or more delinquent using the MBA method in the first three years after origination and zero otherwise. I look at subsamples by gambling preference: high or low CPRATIO (column (1) and (2)). All continuous variables are winsorized at the 0.5 percent level and then standardized. All specifications include state, half-year, and originator fixed effects. Standard errors clustered at the county level are reported in parentheses. *$p < 0.10$, **$p < 0.05$, and ***$p < 0.01$.

Table 9
Local Socioeconomic Characteristics and Mortgage Misrepresentation - Probit Model

	(1)	(2)	(3)	(4)	(5)
CPRATIO	0.0017**				0.0014**
	(0.001)				(0.001)
REL	-0.0004				-0.0005
	(0.000)				(0.000)
Crime rate	0.0001				0.0018***
	(0.000)				(0.001)
Republican	0.0007				-0.0007
	(0.001)				(0.001)
Education		0.0003			0.0001
		(0.001)			(0.001)
Income		-0.0003			-0.0003
		(0.000)			(0.000)
Unemployment			-0.0020***		-0.0026***
			(0.000)		(0.001)
HPA			0.0041***		0.0040***
			(0.001)		(0.001)
Urban			0.0000		-0.0004
			(0.000)		(0.000)
Total population			0.0003		0.0006
			(0.000)		(0.000)
Married				0.0012*	0.0022***
				(0.001)	(0.001)
Over65				-0.0008**	0.0000
				(0.000)	(0.000)
Male-female ratio				-0.0005	0.0000
				(0.000)	(0.000)
Minority				0.0006	0.0005
				(0.001)	(0.001)
Prevalence of simul.	0.1751***	0.1780***	0.1761***	0.1723***	0.1695***
second	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)
Loan chars controls	Y	Y	Y	Y	Y
Originator FE	Y	Y	Y	Y	Y
State FE	Y	Y	Y	Y	Y
Half-year FE	Y	Y	Y	Y	Y
Observations	2,342,532	2,280,584	2,368,874	2,372,952	2,247,430

The table shows OLS estimates from regressions where the dependent variable takes a value of one if the loan is recognized as second-lien misreported and zero otherwise. All continuous variables, except for the crime rates, are winsorized at the 0.5 percent level and then standardized. Standard errors clustered at the county level are reported in parentheses. $*p < 0.10$, $**p < 0.05$, and $***p < 0.01$.

Table 10
Gambling Preference and Mortgage Misrepresentation - Causal Forest

	(1) Misreported second
CPRATIO	0.0016**
	(0.0007)

The table shows causal forest estimates from regressions where the dependent variable takes a value of one if the loan is recognized as second-lien misreported and zero otherwise. All control variables in Table 2 are used for growing trees and forests. Before growing the trees, all continuous variables are winsorized at the 0.5 percent level and then standardized. A half-year variable is created, starting from the first half of 2005 as 1, and used as a variable in growing trees and forests. Observations are clustered by state, and each unit is given the same weight (so that larger clusters receive more weight). 2000 trees are grown in the causal forest. The fraction used for determining splits (honesty) is 50 percent. Standard errors are reported in parentheses. *p < 0.10, **p < 0.05, and ***p < 0.01.

Table 11
Alternative Gambling Preference Measures

	(1)	(2)
Gambling Industry Employment	0.0031***	
	(0.001)	
CPRATIO (adjusted)		0.0011**
		(0.000)
REL	0.0025*	0.0001
	(0.001)	(0.000)
Crime rate	0.0039**	0.0008*
	(0.002)	(0.000)
Republican	-0.0016	0.0008
	(0.002)	(0.001)
Education	-0.0004	0.0004
	(0.002)	(0.000)
Income	0.0005	-0.0004*
	(0.002)	(0.000)
Unemployment	-0.0011	-0.0008
	(0.001)	(0.001)
HPA	0.0016	0.0017***
	(0.002)	(0.001)
Urban	0.0005	-0.0004
	(0.002)	(0.000)
Total population	0.0009	0.0009**
	(0.002)	(0.000)
Married	0.0003	0.0002
	(0.001)	(0.000)
Over65	-0.0000	0.0005*
	(0.001)	(0.000)
Male-female ratio	0.0020	-0.0000
	(0.001)	(0.000)
Minority	-0.0026*	0.0006
	(0.001)	(0.000)
Loan chars controls	Y	Y
Originator FE	Y	Y
State FE	Y	Y
Half-year FE	Y	Y
Observations	1,032,077	2,308,542
Adj. R^2	0.318	0.318

The table shows OLS estimates from regressions where the dependent variable takes a value of one if the loan is recognized as second-lien misreported and zero otherwise. The local controls in column (1) are aggregated to CBSA level. The local controls in column (2) follow the baseline specification and remain at the same geographical level as in the main regressions. All continuous variables are winsorized at the 0.5 percent level and then standardized. Standard errors are clustered at the CBSA level for column (1) and at the county level for column (2). Standard errors are reported in parentheses. $*p < 0.10$, $**p < 0.05$, and $***p < 0.01$.

Table 12
Local Characteristics and Loan Collateral Fundamentals around Discontinuity in Low-Documentation Loans

	FICO≥620 (β)	t-stat
CPRATIO	-0.041	(-0.035)
REL	0.007	(0.741)
Crime rate	-5.528	(-0.416)
Republican	0.011	(0.049)
Education (%)	0.619	(1.561)
Income	0.024	(1.460)
Unemployment (%)	0.024	(-0.210)
HPA	-0.002	(-0.102)
Urban (%)	-0.278	(0.330)
Total population (ln)	-0.045	(0.038)
Married (%)	0.191	(0.135)
Over65 (%)	-0.386	(-0.883)
Male-female ratio	0.001	(0.147)
Minority (%)	-0.032	(0.123)
Balance	0.045	(0.868)
LTV (%)	0.637	(-0.329)

The table reports estimates from regressions of local characteristics and loan collateral fundamentals around a FICO score of 620 for low-documentation loans. Local linear regressions using the RDD approach are employed to estimate the differences in these predetermined variables between FICO 620^{-} and FICO 620^{+}. The bias-corrected t-values (standard errors clustered at the county level) following Calonico et al. (2014) are reported in parentheses.

Table 13
Discontinuities in Reported and Misreported Second-Lien Rates around the Credit Threshold in Low-Documentation Loans

		(1) All	(2) High	(3) Low	(4) High-Low
Panel A. Misreported second					
Raw	FICO≥620 (β)	0.042	0.014	0.065	-0.051
	t-stat	(6.49)	(0.69)	(9.13)	(-5.19)
Control	FICO≥620 (β)	0.032	0.009	0.052	-0.043
	t-stat	(5.52)	(0.26)	(7.86)	(-4.49)
Control+FEs	FICO≥620 (β)	0.010	0.006	0.015	-0.008
	t-stat	(2.97)	(0.97)	(3.33)	(-1.14)
Panel B. Reported second					
Raw	FICO≥620 (β)	0.090	0.094	0.087	0.006
	t-stat	(11.22)	(7.38)	(8.58)	(0.57)
Control	FICO≥620 (β)	0.078	0.087	0.070	0.017
	t-stat	(11.13)	(7.32)	(8.84)	(1.85)
Control+FEs	FICO≥620 (β)	0.062	0.066	0.058	0.008
	t-stat	(8.35)	(5.84)	(6.23)	(0.89)

The table reports discontinuities in reported and misreported second-lien rates around the FICO score threshold of 620 for low-documentation loans. Panel A presents the estimates for misreported second-lien rates, where the dependent variable is the dummy for misreported seconds, and Panel B presents the estimates for correctly reported second-lien rates, where the dependent variable is the dummy for correctly reported seconds. Local linear regressions from the RDD approach are used to estimate the difference in rates between FICO 620^{-} and FICO 620^{+}. Using the diff-in-disc approach, I further estimate the difference between high and low gambling preference areas. Estimations are conducted without controls, with controls, and with both controls and the fixed effects from the baseline regressions. For discontinuities, bias-corrected t-values (standard errors clustered at the county level) following Calonico et al. (2014) are reported in parentheses. For the difference between high and low gambling preference areas, t-statistics based on heteroskedasticity-consistent standard errors are reported in parentheses.

Table 14
Alternative Default Measures for Loan Performance Tests

	(1)	(2)	(3)	(4)	(5)	(6)
	90+ delinq in 2 year		60+ delinq in 3 year		B/F/REO in 3 year	
	High	Low	High	Low	High	Low
Misreported second	0.084***	0.063***	0.160***	0.116***	0.137***	0.103***
	(0.004)	(0.004)	(0.004)	(0.006)	(0.004)	(0.006)
Reported second	0.067***	0.023***	0.110***	0.051***	0.092***	0.040***
	(0.004)	(0.002)	(0.006)	(0.003)	(0.005)	(0.003)
REL	-0.004	-0.001	-0.006*	0.001	-0.003	-0.001
	(0.002)	(0.001)	(0.004)	(0.002)	(0.003)	(0.002)
Crime rate	0.010***	0.003	0.017***	0.000	0.018***	0.003
	(0.004)	(0.002)	(0.005)	(0.003)	(0.004)	(0.002)
Republic	-0.003	-0.003	-0.016**	-0.008	-0.012**	-0.004
	(0.005)	(0.004)	(0.007)	(0.006)	(0.006)	(0.005)
Education	0.002	0.002	0.002	0.000	0.003	0.002
	(0.003)	(0.002)	(0.004)	(0.003)	(0.003)	(0.003)
Income	-0.007***	-0.010***	-0.012***	-0.014***	-0.008***	-0.011***
	(0.002)	(0.001)	(0.003)	(0.002)	(0.002)	(0.002)
Unemployment	0.000	0.010***	-0.005	0.008***	-0.003	0.009***
	(0.003)	(0.002)	(0.004)	(0.002)	(0.004)	(0.002)
HPA	0.022***	0.038***	0.041***	0.065***	0.034***	0.050***
	(0.005)	(0.005)	(0.006)	(0.007)	(0.005)	(0.005)
Urban	0.007**	-0.005***	0.011**	-0.009***	0.006	-0.005**
	(0.003)	(0.002)	(0.004)	(0.002)	(0.004)	(0.002)
Total population	-0.007*	0.016***	-0.014**	0.027***	-0.013***	0.018***
	(0.003)	(0.003)	(0.006)	(0.004)	(0.005)	(0.004)
Married	0.011***	0.011***	0.021***	0.017***	0.017***	0.014***
	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
Over65	0.008**	0.001	0.010*	-0.000	0.008*	0.003
	(0.004)	(0.002)	(0.005)	(0.003)	(0.005)	(0.003)
Male-female ratio	-0.001	-0.000	0.000	-0.002	-0.001	-0.002
	(0.003)	(0.002)	(0.005)	(0.002)	(0.004)	(0.002)
Minority	0.003	0.003	0.006	0.007	0.003	0.002
	(0.003)	(0.003)	(0.005)	(0.005)	(0.005)	(0.004)
Loan chars controls	Y	Y	Y	Y	Y	Y
Originator FE	Y	Y	Y	Y	Y	Y
State FE	Y	Y	Y	Y	Y	Y
Half-year FE	Y	Y	Y	Y	Y	Y
Observations	1,149,680	1,201,389	1,149,680	1,201,389	1,149,680	1,201,389
Adj. R^2	0.175	0.150	0.267	0.223	0.180	0.144

The table shows OLS estimates from regressions in which the dependent variable uses different measures of default. I look at subsamples by gambling preference: high or low CPRATIO. Column (1) and (2) uses 90 days or more delinquency using MBA method in the first two years after origination. Column (3) and (4) uses 60 days or more delinquency using MBA method in the first two years after origination. Column (5) and (6) uses bankruptcy/foreclosure/REO in the first three years after origination. All continuous variables are winsorized at 0.5 percent level and then standardized. Standard errors clustered at the county level are reported in the parentheses. $*p < 0.10$, $**p < 0.05$, and $***p < 0.01$.

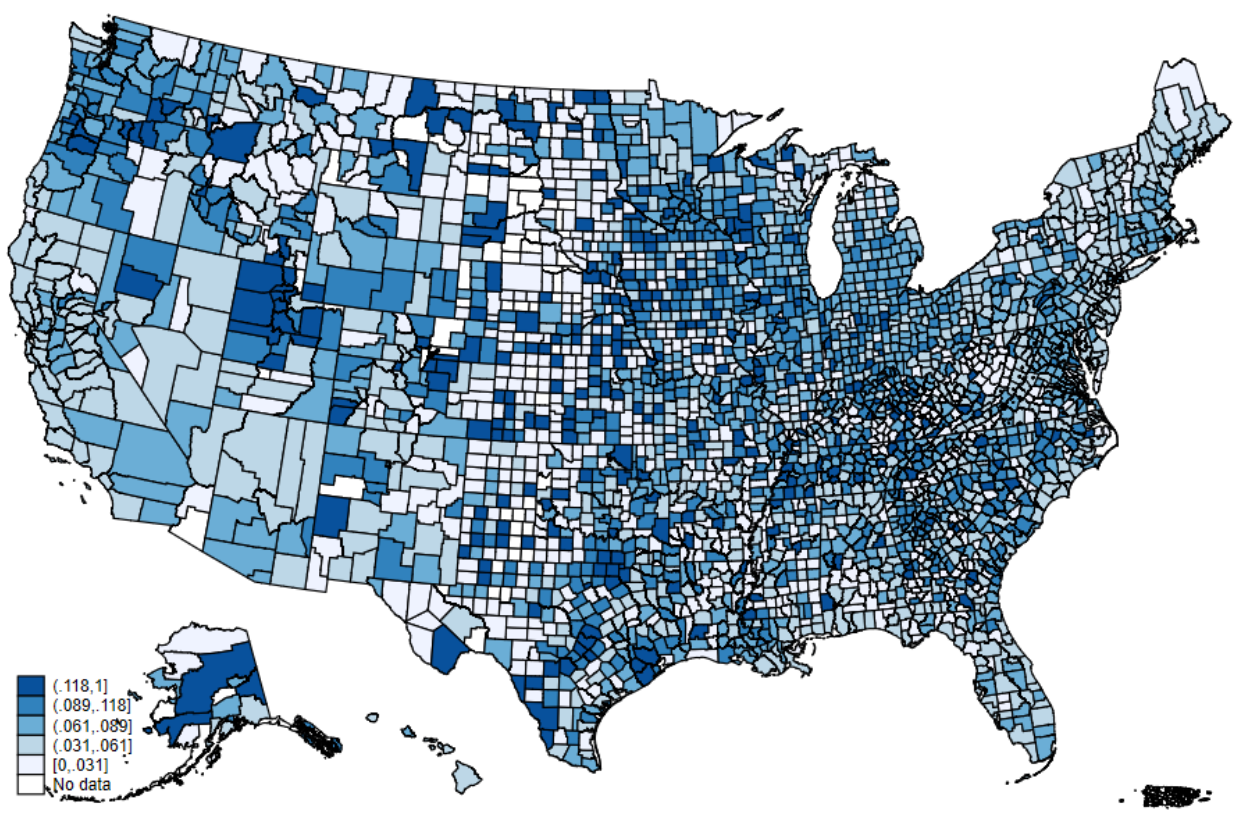

Figure 1
Second-lien Misrepresentations Distribution
The figure plots the county level proportion of second-lien misrepresentation in all loans.

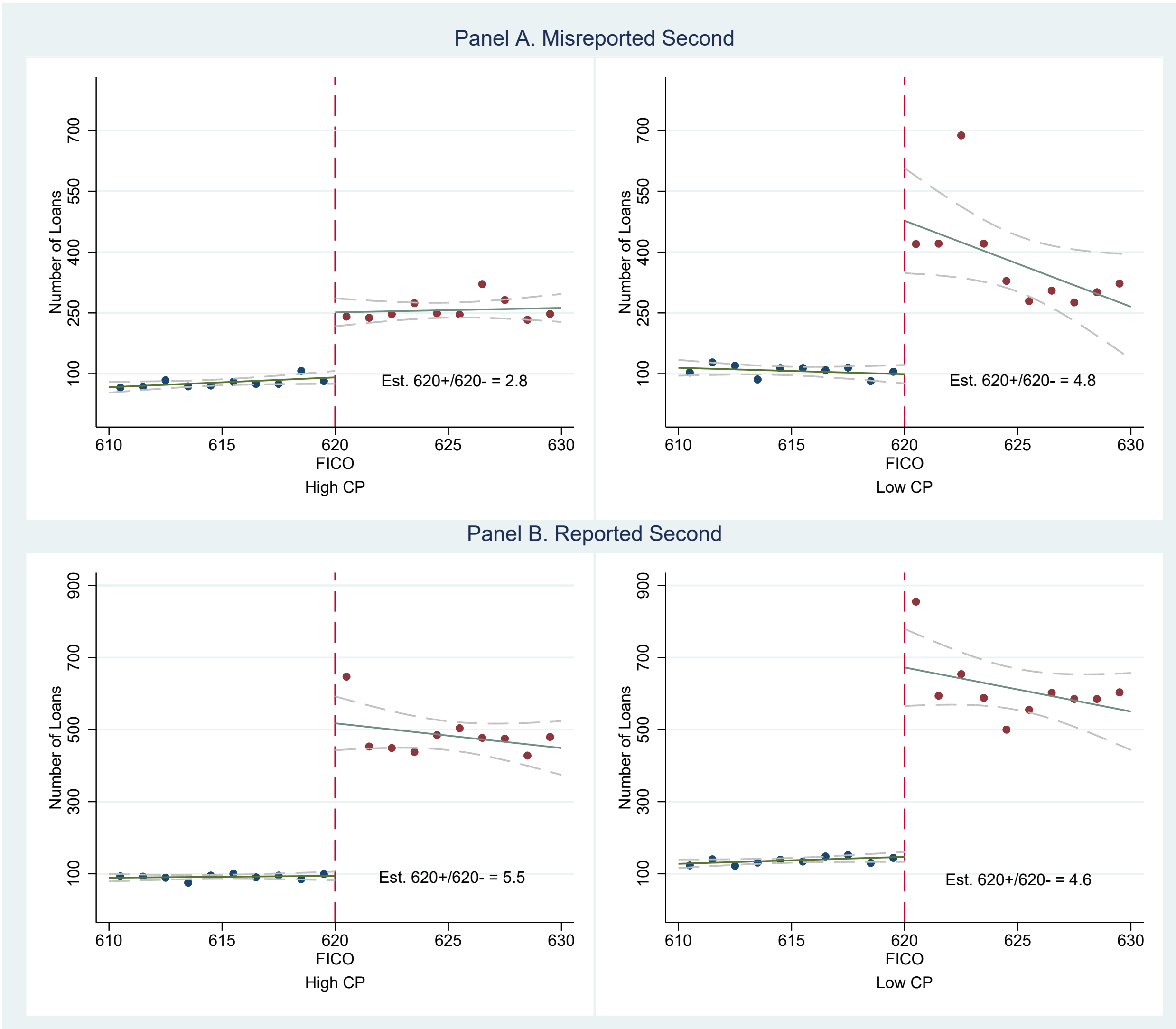

Figure 2

Discontinuities in the Number of Low-Documentation Loans Around Credit Score Thresholds by Gambling Preference

The figure plots discontinuities in the number of low-documentation loans around the FICO score threshold of 620 in high and low gambling preference areas. Panel A reports the loans with misreported seconds, and Panel B reports loans with correctly reported seconds. The circles represent the average number of loans for each credit score between 610 and 629. The solid lines show the linear fit on either side of the credit score threshold, with dashed lines denoting the 95% confidence interval. The left subfigures report results for loans in high gambling preference areas, while the right subfigures present results for loans in low gambling preference areas. As shown, the jump ratio for loans with misreported seconds is much smaller in high gambling preference areas than in low gambling preference areas, while the jump ratio for loans with correctly reported seconds is similar across the two sets of areas.

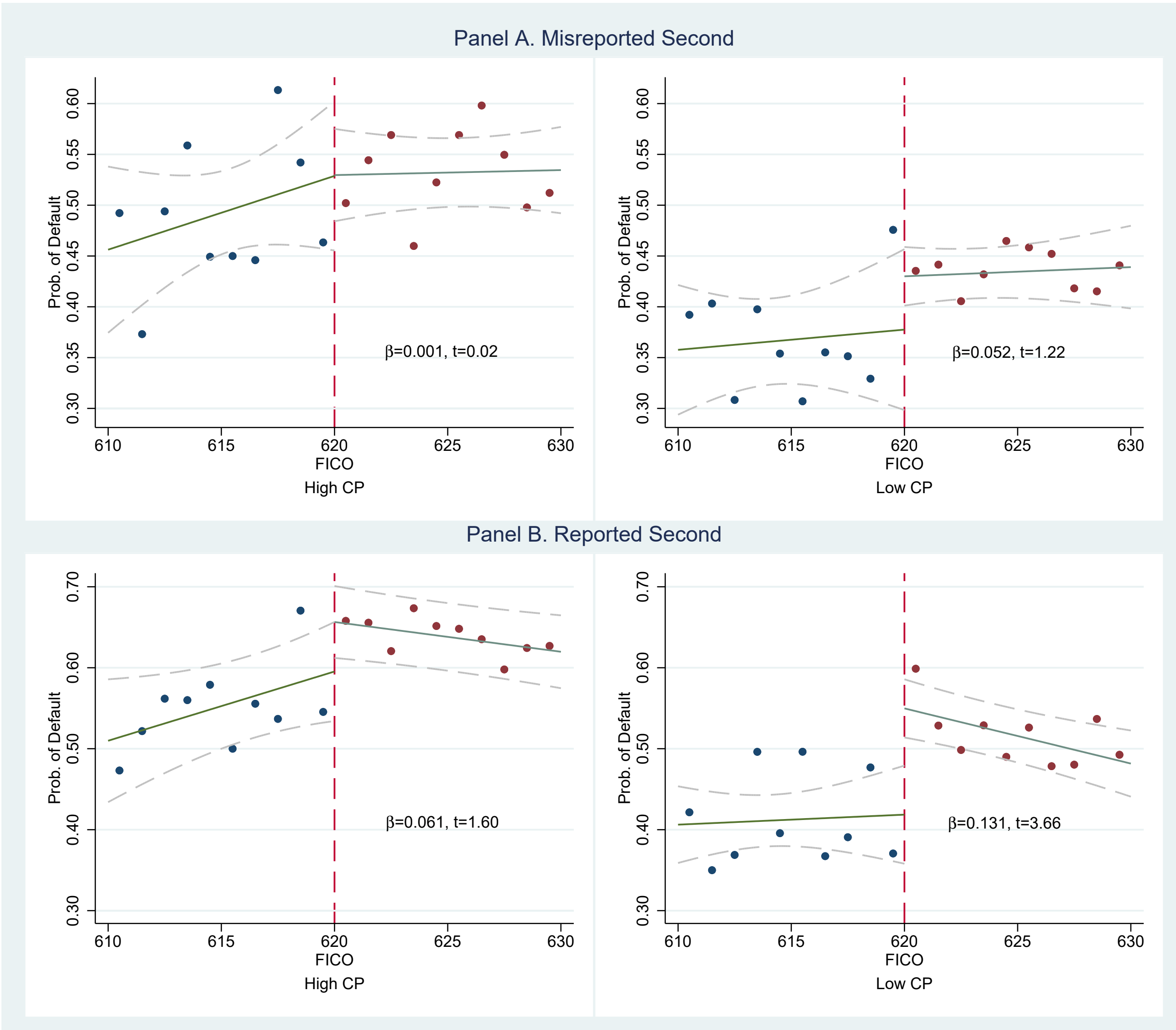

Figure 3

Discontinuities in Default Rates for Low-Documentation Loans Around Credit Score Thresholds by Gambling Preference

The figure plots discontinuities in default rates for low-documentation loans around the FICO score threshold of 620 in high and low gambling preference areas. Panel A reports the loans with misreported seconds, and Panel B reports loans with correctly reported seconds. The circles represent the average number of loans for each credit score between 610 and 629. The solid lines show the linear fit on either side of the credit score threshold, with dashed lines denoting the 95% confidence interval. The left subfigures report results for loans in high gambling preference areas, while the right subfigures present results for loans in low gambling preference areas. As shown, for loans with misreported seconds, the jump in default rates is insignificant in both groups, although it is larger in magnitude in low gambling preference areas. For loans with correctly reported seconds, the jump in default rates is also greater in low gambling preference areas.

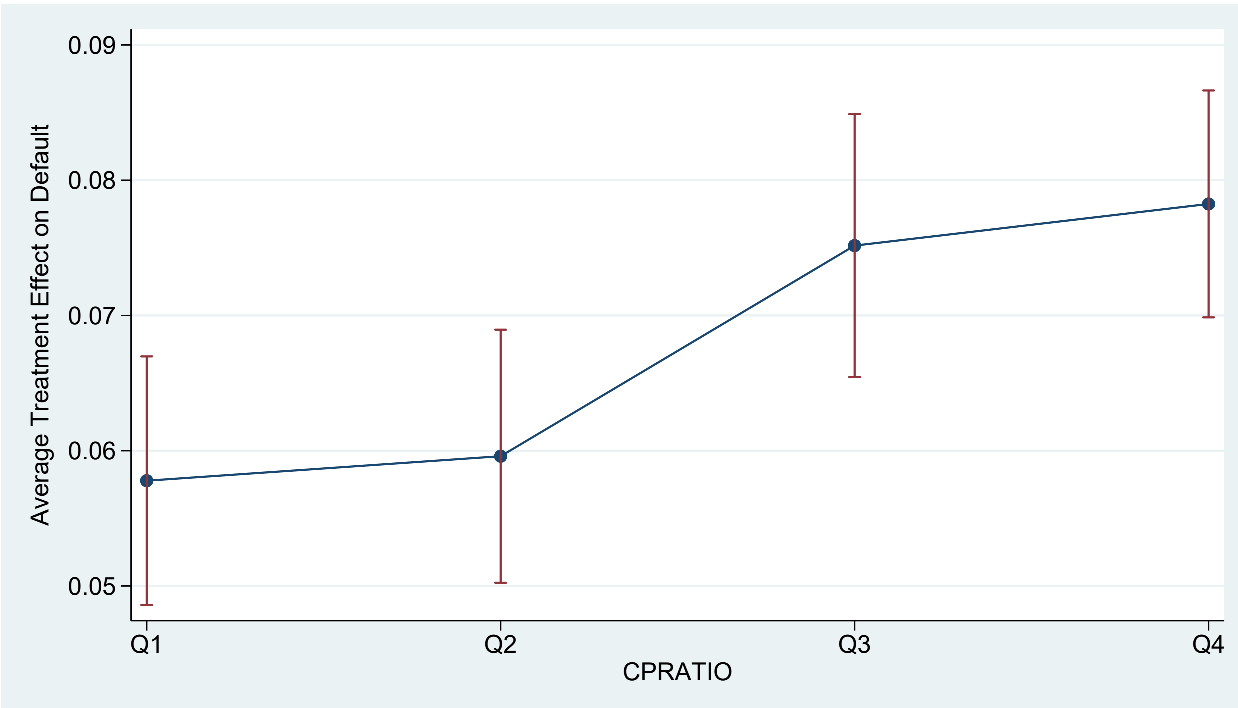

Figure 4
Heterogeneous Treatment Effects of Second-lien Misrepresentation on Default Conditioned on Different Levels of Gambling Preference
The figure plots the heterogeneous treatment effects of second-lien misrepresentation on default, conditioned on different levels of gambling preference, using the causal forest approach. The dependent variable is an indicator that takes a value of one if the loan becomes 90 days or more delinquent using the MBA method in the first three years after origination. The average treatment effect of mortgage misrepresentation is estimated at different levels of local gambling preference (four quarters divided by CPRATIO). All control variables in column (3) of Table 8 are used for growing trees and forests. Before growing the trees, all continuous variables are winsorized at the 0.5 percent level and then standardized. A half-year variable is created, starting from the first half of 2005 as 1, and used as a variable in growing trees and forests. Observations are clustered by state, and each unit is given the same weight (so that larger clusters receive more weight). 2000 trees are grown in the causal forest. The fraction used for determining splits (honesty) is 50 percent.

Appendix

A Variable Definitions, Correlations, and Additional Supporting Tests

This appendix presents supplementary material for the main text. It first reports detailed definitions of the variables used in the analysis. It then provides additional descriptive evidence, including pairwise correlations among local socioeconomic characteristics, to show that *CPRATIO* captures distinct local variation rather than proxying for common local observables. Finally, it reports supporting tests for the subsample comparisons and lender-channel designs discussed in the paper.

Table A1
Variable Definitions

Variable name	Definition
Mortgage Reporting	
Misreported second	Indicator that equals one if the borrower has simultaneous second liens and misreported the second liens on the loan application
Reported second	Indicator that equals one if the borrower has simultaneous second liens and correctly reported the second-liens on the loan application
Simul. second	Indicator that equals one if the borrower have simultaneous second liens regardless the reporting status on the loan application
Local Characteristics	
CPRATIO	Ratio of the county's Catholic residents to Protestant residents
REL	Proportion of the county population that are religious adherents
Education	Proportion of the county population over age 25 that has completed a bachelor's degree or higher
Married	Proportion of the county population over age 15 that is married
Income	Natural logarithm of the ZIP code-level median household income
Urban	Proportion of the county population that lives in urban area
Total population	Natural logarithm of the county total population
Over65	Proportion of the county population over age 65
Male-female ratio	Ratio of the county's male residents to female residents
Minority	Proportion of the county non-white residents
Crime rate	Crime rates of part I offenses excluding arson expressed as the number of crimes per 10,000 residents per year
Republican	Indicator that equals one if the Republican partisanship index is equal or greater than 50%
HPA	ZIP code house price appreciation in the two years prior to loan origination year. County level data used if ZIP code index is not available
Unemployment	Unemployment rate of the county
Loan characteristics	
Interest rate	Loan interest rate at origination
FICO	Natural logarithm of borrower's FICO credit score at loan origination
Balance	Natural logarithm of the initial loan balance
LTV	Loan-to-value ratio at loan origination
CLTV	Reported Combined loan-to-value at loan origination
ARM	Indicator that equals one if the loan is an adjustable rate mortgage
Option ARM	Indicator that equals one if the loan has an ARM convertibility clause
Negative amortization	Indicator that equals one if the loan allows negative amortization
Low or no doc.	Indicator that equals one if the loan is originated with no or limited documentation
Prepayment penalty	Indicator that equals one if the loan would be assessed a penalty on any early voluntary prepayment
Cash-out	Indicator that equals one if the loan purpose is cash out refinancing
No-cash-out	Indicator that equals one if the loan purpose is no cash out refinancing
Investment	Indicator that equals one if the loan occupancy status is investment
Second-home	Indicator that equals one if the loan occupancy status is second-home
Default	Indicator that equals one if the loan becomes 90 days or more delinquent using MBA method in the first three years after origination

The table reports the variable definitions used in the empirical analysis part.

Table A2
Local Socioeconomic Characteristics Correlation

		1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	CPRATIO	1.00													
2	REL	-0.01	1.00												
3	Crime rate	0.01	-0.13	1.00											
4	Republican	-0.28	0.07	-0.19	1.00										
5	Education	0.26	-0.09	0.14	-0.18	1.00									
6	Income	0.27	-0.11	-0.01	-0.05	0.64	1.00								
7	Unemployment	-0.06	-0.20	0.13	-0.17	-0.40	-0.34	1.00							
8	HPA	0.26	-0.29	0.03	-0.09	0.22	0.18	-0.21	1.00						
9	Urban	0.33	-0.01	0.49	-0.18	0.51	0.41	-0.17	0.15	1.00					
10	Total population	0.33	-0.18	0.47	-0.25	0.52	0.51	-0.08	0.23	0.74	1.00				
11	Married	-0.20	0.12	-0.44	0.43	-0.29	0.13	-0.14	-0.08	-0.40	-0.37	1.00			
12	Over65	-0.15	0.30	-0.26	0.07	-0.33	-0.42	0.03	0.00	-0.33	-0.37	0.23	1.00		
13	Male-female ratio	0.05	-0.21	-0.19	0.09	-0.01	0.11	-0.08	0.12	-0.17	-0.19	0.17	-0.21	1.00	
14	Minority	0.08	-0.09	0.43	-0.29	0.01	-0.15	0.22	0.13	0.20	0.23	-0.52	-0.29	-0.12	1.00

The table shows the correlation between local characteristics variables at the county-year level. Income and HPA at the county level are obtained from the same data sources as the ZIP code-level data, which are the U.S. Census Bureau and the Federal Housing Finance Agency, respectively. Unemployment is the average of the monthly unemployment rate within each year.

Table A3
Subsample Estimates and Tests of Equality — Second-Lien Misrepresentation

Split	(1) Group 1 (β_1)	(2) Group 2 (β_2)	(3) $\beta_1 - \beta_2$	(4) p-value ($\beta_1 - \beta_2$)
Primary vs Non-primary	0.0014***	0.0008	0.0006	0.458
	(0.000)	(0.001)	(0.001)	
Purchase vs Refinance	0.0022***	0.0006*	0.0016**	0.018
	(0.001)	(0.000)	(0.001)	
High vs Low Income	0.0013**	0.0017***	-0.0004	0.543
	(0.001)	(0.001)	(0.001)	
High vs Low FICO	0.0012**	0.0017***	-0.0005	0.497
	(0.001)	(0.000)	(0.001)	
High vs Low HPA	0.0018***	0.0008	0.0010	0.240
	(0.000)	(0.001)	(0.001)	

The table reports the estimated coefficients on *CPRATIO* from regressions estimated separately within each subsample in Table 4 and 5. Columns (1) and (2) report the coefficients for the two subsample groups. Column (3) reports the difference between these coefficients, and Column (4) reports the p-value for a Wald test of equality of the two coefficients. Standard errors clustered at the county level are reported in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Table A4
Subsample Estimates and Tests of Equality — Mortgage Default

Split	(1) Group 1 (β_1)	(2) Group 2 (β_2)	(3) $\beta_1 - \beta_2$	(4) p-value ($\beta_1 - \beta_2$)
High vs Low CPRATIO	0.157***	0.113***	0.044***	0.000
	(0.004)	(0.006)	(0.007)	

The table reports the estimated coefficients on *Misreported second* from regressions estimated separately within each subsample in Table 8. Columns (1) and (2) report the coefficients for the two subsample groups. Column (3) reports the difference between these coefficients, and Column (4) reports the p-value for a Wald test of equality of the two coefficients. Standard errors clustered at the county level are reported in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

B Alternative Gambling Preference Measures

This appendix provides additional details on the use of gambling-industry activity as an alternative measure of local gambling preference.

B.1 Data and Measure Construction

Because *CPRATIO* may capture broader religious influences, such as attitudes toward forgiveness and dishonesty, I also use more direct measures of local gambling preference. Specifically, I construct CBSA-level measures based on per capita employment in the gambling industry.

I begin with employment data from the County Business Patterns (CBP) dataset provided by the U.S. Census Bureau. I extract CBSA-level observations for 2005–2007 and retain NAICS code 7132, which includes 71321 for Casinos (except Casino Hotels) and 71329 for Other Gambling Industries. I then normalize employment by total CBSA population. Population for 2005–2007 is linearly interpolated using CBSA-level data from the 2000 and 2010 Census.

I also construct supplementary proxies. Because CBP classifies establishments by primary activity, casino hotels (NAICS code 72110) are excluded from the gambling-industry category. Since these businesses also provide gambling services, I separately collect casino-hotel employment and normalize it by total population.

Because CBP covers only employer-based businesses, some small gambling venues without formal employees may be omitted. To capture this margin, I use the Nonemployer Statistics (NES) datasets from the U.S. Census Bureau and retain observations with NAICS code 7132. I normalize the number of nonemployer gambling establishments by CBSA population.

A key limitation of the CBP employment data is that many cells are suppressed for confidentiality (Eckert et al., 2021). To improve CBSA-level coverage, I impute suppressed employment values by adapting the approach in Eckert et al. (2021) to the CBSA setting. When employment is reported and nonzero, I keep the reported value. Otherwise, I estimate employment using establishment counts within employment-size intervals and the corresponding interval medians, conservatively assigning 5,000 to the open-ended interval, and then retain the larger of that estimate and the lower bound implied by the employment-level flag.

The NES data have a similar limitation: many cells are suppressed for confidentiality or data-quality reasons. In such cases, the number of establishments is recorded as zero even though the establishment flag indicates that the value is not truly zero. I therefore replace zeros with 2 for observations flagged `D` and with the state-level average for the corresponding industry category for observations flagged `S`. [1]

For geographic control variables, the U.S. Census Bureau does not report CBSA-level data for 2000, so I aggregate county-level variables to the CBSA level. For ratio variables, I first aggregate numerators and denominators across counties and then compute the ratio from the aggregated values. I apply the same procedure to annual average county-level unemployment data from the U.S. Bureau of Labor Statistics and use a one-year lag in the regressions. For HPA, I use CBSA-level house price indices directly from the FHFA.

Table B1 reports summary statistics for the alternative gambling preference measures.

B.2 Validation Tests

To validate the alternative measures and examine how they relate to *CPRATIO*, I conduct two sets of tests.

[1] The smallest observable value is 3 and replacing suppressed values with 1 does not materially affect the results.

First, I perform a univariate sorting test. CBSAs are sorted into quintiles based on *CPRATIO*, and I calculate equal-weighted quintile averages for the alternative gambling preference measure and the two additional control variables. Table B2 reports the results. Gambling-industry employment generally increases with *CPRATIO*, and the difference between the highest and lowest quintiles is statistically significant. This pattern suggests that gambling-industry employment captures a construct broadly consistent with *CPRATIO*.[2]

The two additional control variables display different patterns. Employment in casino hotels does not increase systematically with *CPRATIO*. This likely reflects the fact that casino hotels are influenced not only by local gambling demand but also by tourism-oriented gambling activity. In addition, legal exceptions such as tribal casinos may permit casino hotel operations in areas with relatively low local gambling preference. By contrast, the number of nonemployer gambling establishments tends to rise with *CPRATIO*, suggesting that these businesses complement minor local gambling demand and provide an additional proxy for community-level gambling activity.

Second, I estimate a series of OLS regressions to examine whether *CPRATIO* explains variation in the alternative gambling preference measure and the additional control variables. In each regression, the dependent variable is regressed on *CPRATIO* and CBSA-level local controls, with state and year fixed effects. All continuous variables are winsorized at the 0.5 percent level and standardized before estimation. Table B3 reports the results. Per capita gambling-industry employment is positively and significantly associated with *CPRATIO*, supporting the use of *CPRATIO* as a proxy for local gambling preference. By contrast, per capita casino-hotel employment is negatively associated with *CPRATIO*, which supports treating casino-hotel variables as controls

[2]The unconditional correlation between *CPRATIO* and gambling-industry employment is not large. One reason is that although Catholic doctrine does not view gambling as inherently unjust, Catholic bishops' conferences have often opposed the expansion of the gambling industry because of its social costs. As a result, areas with high *CPRATIO* do not necessarily have high gambling-industry employment.

for tourism-oriented gambling demand rather than as proxies for local gambling preference. Finally, nonemployer gambling establishments are also positively and significantly associated with *CPRATIO*, further supporting its relevance as an indicator of local gambling preference.

B.3 Additional Tests

In the main text, I report results using per capita employment in the gambling industry as the alternative measure of local gambling preference. Here I add controls for casino-hotel employment, which may reflect tourist-driven gambling demand, and for nonemployer gambling establishments, which may capture complementary local gambling demand omitted from the primary measure because of data limitations.

Table B4 reports the regression results. The coefficient on gambling-industry employment remains positive and significant, consistent with a strong association between local gambling preference and second-lien misrepresentation. By contrast, the casino-hotel variables, which reflect a mix of local and tourist demand, are generally insignificant. The coefficients on nonemployer gambling establishments are also insignificant, consistent with the idea that these businesses provide only limited gambling access. Overall, the results continue to support the interpretation that second-lien misrepresentation is more prevalent in areas with stronger gambling preference.

B.4 Additional Illustration

This subsection addresses several concerns about the alternative measure of local gambling preference.

First, although gambling-industry employment and establishment data may more directly reflect local gambling preference, they are still supply-side measures rather than direct measures of household demand. They are also affected by local legality and economic conditions, which introduce additional noise. Coverage is another important limitation. The alternative measures cover only 460 CBSAs, or about 1,187 counties, and roughly 1 million loans. By contrast, *CPRATIO* covers nearly 3,000 counties and almost the full sample of 3 million loans. This broader coverage is particularly important for analyses of lender behavior and borrower heterogeneity.

Second, I construct the alternative measures at the CBSA level rather than the county level because residents in most metropolitan areas can commute easily across counties within the same local labor market. As a result, low county-level gambling activity does not necessarily imply low gambling preference among local residents. The CBSA level therefore provides a more stable proxy for local gambling preference, while cross-CBSA gambling related to tourism is partly captured by the casino-hotel measures.

Third, I restrict the sample to CBSAs in which all three categories are present: gambling-industry establishments, casino hotels, and nonemployer gambling businesses. This restriction reduces false zeros that may arise from legal prohibitions, data suppression, or missing supply. In untabulated regressions, when I instead replace missing values with zeros and include all CBSAs, the coefficient on the gambling preference measure remains positive but becomes only marginally significant. Moreover, when I split the sample by per capita establishment density, the coefficient is strongly significant in the high-establishment subsample but insignificant in the low-establishment subsample. This pattern suggests that second-lien misrepresentation rises with stronger local gambling preference, while very low values of the alternative measure may partly reflect measurement noise rather than genuinely weak gambling tendencies.

Table B1
Descriptive Statistics for Alternative Gambling Preference Measures

	N	Mean	SD	P10	P25	P50	P75	P90
Gambling Industries (NAICS 7132, CBSA-year level)								
Employment (per 10,000 residents)	1,201	19.60	42.74	0.22	0.69	2.53	16.41	59.50
Casino Hotel (NAICS 72110, CBSA-year level)								
Employment (per 10,000 residents)	258	118.49	264.87	0.07	0.50	16.90	77.41	389.73
Nonemployer Gambling (NAICS 7132, CBSA-year level)								
Establishment (per 10,000 residents)	1,239	0.38	0.50	0.00	0.14	0.27	0.45	0.76

The table presents descriptive statistics for the alternative local gambling preference measure and its related controls, covering the sample period from 2005 to 2007. I report the number of observations, mean, standard deviation, 10th percentile, 25th percentile, 50th percentile, 75th percentile, and 90th percentile.

Table B2
Univariate Sorting

Quintile	Gambling Industries Employment	Casino Hotel Employment	Nonemployer Gambling Establishment
Low	16.2	114.20	0.28
Q2	11.77	140.31	0.35
Q3	18.67	49.83	0.43
Q4	26.47	95.67	0.40
High	24.92	219.91	0.46
High-Low	8.66	105.72	0.18
t-Statistic	2.02	1.72	4.01

The table presents univariate sorting results for three gambling preference related measures: per capita employment in the gambling industry (Gambling Industries Employment), per capita employment in the casino hotel industry (Casino Hotel Employment), and per capita number of nonemployer gambling establishments (Nonemployer Gambling Establishment). CBSAs are sorted into quintiles based on CPRATIO, and for each quintile, the equal-weighted average of the five gambling preference related measures is computed. The table reports White robust t-statistics for the difference in mean values between the High and Low CPRATIO quintiles.

Table B3
Alternative Measures and CPRATIO

	(1) Gambling Industry Employment	(2) Casino Hotel Employment	(3) Nonemployer Gambling Establishment
CPRATIO	0.103**	-0.161*	0.126***
	(0.043)	(0.084)	(0.038)
REL	-0.161***	-0.086	-0.159**
	(0.059)	(0.151)	(0.071)
Crime rate	-0.113***	-0.134	-0.028
	(0.040)	(0.131)	(0.038)
Republican	-0.005	-0.088	-0.212***
	(0.062)	(0.144)	(0.075)
Education	0.004	-0.819***	-0.094
	(0.055)	(0.241)	(0.070)
Income	-0.227***	1.056***	0.130
	(0.067)	(0.294)	(0.082)
Unemployment	0.018	-0.121	0.001
	(0.060)	(0.088)	(0.067)
HPA	0.017	0.124	0.085
	(0.044)	(0.092)	(0.052)
Urban	0.084	-0.452***	-0.030
	(0.067)	(0.119)	(0.052)
Total population	-0.072	-0.011	-0.017
	(0.048)	(0.126)	(0.061)
Married	0.067	-0.540***	0.004
	(0.045)	(0.167)	(0.059)
Over65	0.084**	0.045	0.034
	(0.040)	(0.114)	(0.051)
Male-female ratio	0.062	-0.000	-0.039
	(0.044)	(0.085)	(0.036)
Minority	0.040	0.192*	-0.042
	(0.045)	(0.116)	(0.048)
State FE	Y	Y	Y
Half-year FE	Y	Y	Y
Observations	1,015	227	1,042
Adj. R^2	0.330	0.541	0.281

The table shows OLS estimates from regressions where the dependent variables are the alternative gambling preference measure and related control variables. CPRATIO and other geographic controls are aggregated to CBSA level. All continuous variables are winsorized at the 0.5 percent level and then standardized. White robust standard errors are reported in parentheses. $*p < 0.10$, $**p < 0.05$, and $***p < 0.01$.

Table B4
Alternative Gambling Preference Measures and Second-lien Misrepresentation

	(1)
Gambling Industry Employment	0.0030***
	(0.001)
Casino Hotel Employment	-0.0013
	(0.002)
Nonemployer Gambling Establishment	0.0010
	(0.002)
REL	0.0022
	(0.002)
Crime rate	0.0040**
	(0.002)
Republican	-0.0012
	(0.002)
Education	-0.0006
	(0.002)
Income	0.0006
	(0.002)
Unemployment	-0.0011
	(0.001)
HPA	0.0016
	(0.002)
Urban	0.0006
	(0.002)
Total population	0.0005
	(0.002)
Married	0.0001
	(0.001)
Over65	-0.0002
	(0.001)
Male-female ratio	0.0020
	(0.002)
Minority	-0.0025*
	(0.002)
Loan chars controls	Y
Originator FE	Y
State FE	Y
Half-year FE	Y
Observations	1,032,077
Adj. R^2	0.318

The table shows OLS estimates from regressions where the dependent variable takes a value of one if the loan is recognized as second-lien misreported and zero otherwise. The local controls are aggregated to CBSA level. All continuous variables are winsorized at the 0.5 percent level and then standardized. Standard errors clustered at the county level are reported in the parentheses. $*p < 0.10$, $**p < 0.05$, and $***p < 0.01$.

C Diff-in-disc Concerns

This appendix assesses the validity of the diff-in-disc approach used to compare loan volumes and default rates between high- and low-gambling-preference areas.

C.1 Assumptions for Applying Diff-in-disc

According to Grembi et al. (2016), the validity of the diff-in-disc approach rests on three assumptions. First, potential outcomes must be continuous at the cutoff. In my setting, this requires that FICO scores are not manipulated around 620. Although I do not observe the full U.S. distribution of FICO scores, Keys et al. (2010) show, using anonymous credit-bureau data, that the distribution of FICO scores is smooth around this threshold. They also show, using the reversal of anti-predatory lending laws in Georgia and New Jersey, that borrowers were either unaware of differential screening around 620 or unable to manipulate their FICO scores quickly.[3]

Second, in the absence of treatment, the effect of the confounding policy around the cutoff must be constant across the "diff" dimension. To assess this assumption, I examine full-documentation loans around FICO 620. These loans are exposed to similar confounding factors but do not receive the securitization treatment. Table C1 reports the results for the number of loans (Panel A) and default rates (Panel B). For full-documentation loans, the jump in the number of loans is small, and the difference between high- and low-gambling-preference areas is limited. Likewise, there is no meaningful jump in default rates for any loan type, and the differences across gambling-preference areas are minor. These findings suggest that without the ease of securitization, crossing the 620 threshold does not generate differential changes in loan volumes or screening outcomes across high-

[3] According to Fair Isaac, strategic manipulation of FICO scores is difficult.

and low-gambling-preference areas.

Third, the treatment effect at the cutoff must not depend on the confounding policy. In this setting, this implies that predetermined outcomes and covariates should exhibit similar jumps, if any, across high- and low-gambling-preference areas. To examine this assumption, I estimate equation 6 using predetermined outcomes and covariates as dependent variables, while including the same controls, bandwidth, and kernel as in the main tests. Table C2 reports the results. Column (1), which focuses on loan characteristics, shows that most variables have similar discontinuities across high- and low-gambling-preference areas. The shares of negative-amortization loans and second-home loans are statistically significant, but these differences are unlikely to affect the conclusions. First, although the absolute discontinuities differ, the jump ratios are very similar: negative amortization rises from 6.74/2.97 percent to 15.82/6.61 percent in high/low-gambling-preference areas, corresponding to 2.35 and 2.23 times, respectively. Second, these loan types are quantitatively rare in the local range; second homes account for only 1.2 percent of all loans and 0.6 percent of misreported loans. Column (2), which focuses on local characteristics, likewise shows similar discontinuities for most variables. Although the shares of highly educated residents, older residents, and the male-female ratio are statistically significant, their magnitudes are economically trivial. In the local range, the average share of highly educated residents is 25.01 percent, the average share of older residents is 12.2 percent, and the average male-female ratio is 28.89 percent, while the differences between high- and low-gambling-preference areas are only -0.331 percent, 0.322 percent, and 0.003 percent, respectively.

C.2 Sensitivity to Regression Choices

Beyond the three identifying assumptions, I also examine whether the results are sensitive to modeling choices, including the inclusion of control variables, bandwidth selection, and kernel choice. Specifically, I estimate models with and without covariates, use bandwidths of 8, 10, and 12, and apply either a uniform or triangular kernel. Table C3 reports the estimated jump ratios at FICO 620^{+} relative to 620^{-} for the number of loans.[4] The results are similar to the baseline findings: the ease of securitization increases the number of all loan types, while the jump for loans with misreported simultaneous second liens remains smaller in high-gambling-preference areas than in low-gambling-preference areas. This pattern continues to suggest that lenders do not play a larger role in increasing second-lien misrepresentation in high-gambling-preference areas.

Table C4 reports the corresponding results for the default rate of loans with misreported simultaneous second liens. In all specifications, the estimated jump is small and statistically insignificant. Moreover, although the differences are not statistically significant, the jump is smaller in high-gambling-preference areas than in low-gambling-preference areas in most cases. Taken together, these sensitivity checks continue to provide no evidence that lenders in high-gambling-preference areas are more likely to facilitate second-lien misrepresentation.

[4]Because the number of loans is aggregated at the level of gambling-preference areas while the control variables are measured at the county level, covariates cannot be included in these specifications.

Table C1
Discontinuities of Full-documentation Loans Around the Credit Threshold

	(1) All	(2) High	(3) Low	(4) High-Low
Panel A. Number of loans (Est. 620+/620-)				
All loan	1.186	1.244	1.154	0.090
				(2.27)
Reported second	1.161	1.183	1.152	0.031
				(0.58)
Misreported second	1.251	1.325	1.216	0.109
				(1.03)
Panel B. Default rate				
All loan	0.007	0.001	0.009	-0.007
	(0.36)	(0.24)	(0.08)	(-0.58)
Reported second	-0.007	-0.026	0.002	-0.027
	(-0.39)	(-1.03)	(0.06)	(-0.83)
Misreported second	0.015	-0.016	0.029	-0.045
	(-0.49)	(-0.89)	(0.03)	(-1.12)

The table presents the results of the estimations in Table 6 and Table 7 using full-documentation loans. Panel A reports the estimates for the number of loans at each FICO score. Panel B reports the estimates for loans that become 90 days or more delinquent using the MBA method in the first three years after origination. Using local linear regressions of the RDD approach, we estimate the number of loans and the default rate (with control variables) at FICO 620^- and FICO 620^+ and compute the ratio of the estimated number of loans at FICO 620^+ to 620^-. Using the diff-in-disc approach, I estimate the difference between high and low gambling preference areas. I perform the estimation for all loans, loans with correctly reported simultaneous seconds, and loans with misreported simultaneous seconds. The results for loans in all areas, loans in high gambling preference areas, loans in low gambling preference areas, and the difference between high and low gambling preference areas are reported in columns (1) to (4), respectively. For discontinuities, bias-corrected t-values (standard errors clustered at the county level) following Calonico et al. (2014) are reported in parentheses. For the difference in the ratio of the number of loans/default rate between high and low gambling preference areas, t-statistics based on heteroskedasticity-consistent standard errors/standard errors clustered at the county level are reported in parentheses.

Table C2
Estimates of the discontinuities in observable covariates

	(1) loan characteristics		(2) local characteristics
Interest rate (%)	-0.006	REL	-0.005
	(0.041)		(0.003)
Balance	0.015	Crime rate	-3.650
	(0.013)		(4.730)
LTV (%)	0.170	Republican	0.000
	(0.367)		(0.014)
ARM	-0.010	Education (%)	-0.331*
	(0.013)		(0.199)
Negative amortization	0.025***	Income	0.005
	(0.007)		(0.007)
Option ARM	0.001	Unemployment (%)	0.046
	(0.002)		(0.067)
Prepayment penalty	0.005	HPA	-0.005
	(0.015)		(0.004)
Cash-out	0.003	Urban (%)	0.529
	(0.013)		(0.336)
No-cash-out	0.002	Total population (ln)	0.030
	(0.007)		(0.027)
Investment	0.007	Married (%)	-0.076
	(0.007)		(0.095)
Second-home	0.007**	Over65 (%)	0.322***
	(0.003)		(0.106)
		Male-female ratio	0.003**
			(0.001)
		Minority (%)	-0.415
			(0.309)

The table reports the results of diff-in-disc regression for pre-determined outcomes and covariates using Equation 6 with other controls. The dependent variables are the pre-determined outcomes and covariates. Column (1) shows the tests for loan characteristics, and column (2) shows the tests for borrower characteristics. As in the main test, the bandwidth is set to 10 and the kernel is uniform. Standard errors clustered at the county level are reported in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Table C3
Sensitivity of Results for Number of Loans

	Uniform				Triangular			
	High	Low	All	H-L	High	Low	All	H-L
Bandwidth = 8								
All loan	2.230	2.242	2.237	-0.013	2.250	2.204	2.225	0.046
				(-0.13)				(0.44)
Reported second	5.469	4.693	4.999	0.776	5.761	5.235	5.446	0.526
				(0.96)				(0.56)
Misreported second	2.709	5.095	3.957	-2.385	2.496	5.128	3.826	-2.632
				(-2.85)				(-3.21)
Bandwidth = 10								
All loan	2.240	2.289	2.266	-0.049	2.240	2.234	2.237	0.006
				(-0.52)				(0.06)
Reported second	5.498	4.581	4.939	0.917	5.616	4.971	5.228	0.645
				(1.23)				(0.74)
Misreported second	2.763	4.826	3.838	-2.063	2.609	5.133	3.900	-2.524
				(-2.83)				(-3.20)
Bandwidth = 12								
All loan	2.199	2.245	2.224	-0.047	2.237	2.262	2.250	-0.025
				(-0.50)				(-0.27)
Reported second	5.209	4.602	4.846	0.607	5.520	4.845	5.114	0.676
				(0.91)				(0.84)
Misreported second	2.656	4.388	3.590	-1.732	2.668	4.971	3.858	-2.302
				(-2.60)				(-3.20)
Bandwidth = 14								
All loan	2.197	2.199	2.198	-0.002	2.222	2.257	2.241	-0.036
				(-0.02)				(-0.39)
Reported second	5.278	4.666	4.915	0.612	5.424	4.788	5.043	0.636
				(0.95)				(0.84)
Misreported second	2.515	3.940	3.275	-1.426	2.645	4.755	3.747	-2.110
				(-1.99)				(-3.18)

The table reports the sensitivity of RDD and diff-in-disc regression for the number of loans to the choice of bandwidth and estimation kernel. For the difference between high and low gambling preference areas, t-statistics based on heteroskedasticity-consistent standard errors are reported in parentheses.

Table C4
Sensitivity of Results for Default Rate of Loans with Misreported Second

	Uniform				Triangular			
	High	Low	All	H-L	High	Low	All	H-L
Bandwidth = 8								
No control	0.005	0.019	-0.004	-0.014	-0.007	-0.016	-0.031	0.009
	(-0.24)	(-0.83)	(-1.18)	(-0.23)	(0.81)	(-1.43)	(-0.89)	(0.13)
With control	0.029	0.042	0.031	-0.013	-0.002	-0.017	-0.008	0.015
	(-0.48)	(-1.34)	(-1.31)	(-0.26)	(0.42)	(-1.90)	(-1.06)	(0.24)
Bandwidth = 10								
No control	0.001	0.052	0.014	-0.052	-0.004	0.011	-0.014	-0.015
	(-0.07)	(-0.86)	(-1.15)	(-0.98)	(0.32)	(-1.15)	(-1.07)	(-0.25)
With control	0.017	0.077	0.046	-0.061	0.012	0.021	0.015	-0.008
	(0.12)	(-1.09)	(-0.84)	(-1.31)	(0.01)	(-1.56)	(-1.14)	(-0.16)
Bandwidth = 12								
No control	-0.022	0.065	0.014	-0.087	-0.003	0.033	0.000	-0.036
	(0.30)	(-0.30)	(-0.50)	(-1.71)	(0.09)	(-1.00)	(-1.14)	(-0.65)
With control	0.003	0.081	0.045	-0.078	0.014	0.046	0.029	-0.033
	(0.29)	(-0.08)	(0.04)	(-1.79)	(0.07)	(-1.26)	(-0.91)	(-0.68)
Bandwidth = 14								
No control	0.007	0.071	0.033	-0.064	-0.007	0.046	0.007	-0.053
	(-0.47)	(0.11)	(-0.72)	(-1.39)	(0.06)	(-0.59)	(-0.88)	(-1.03)
With control	0.023	0.082	0.054	-0.059	0.012	0.058	0.036	-0.046
	(-0.26)	(0.65)	(0.19)	(-1.40)	(0.09)	(-0.57)	(-0.43)	(-1.04)

The table reports the sensitivity of RDD and diff-in-disc regression for the default rate of loans with misreported seconds to the inclusion of control variables, the choice of bandwidth, and the choice of estimation kernel. For discontinuities, bias-corrected t-values (standard errors clustered at the county level) following Calonico et al. (2014) are reported in parentheses. For the difference between high and low gambling preference areas, t-statistics based on heteroskedasticity-consistent standard errors are reported in parentheses.