Credit Rationing, Income Exaggeration, and Adverse Selection in the Mortgage Market*

Brent W. Ambrose, Pennsylvania State University†
James Conklin, University of Georgia‡
and
Jiro Yoshida, Pennsylvania State University§

October 27, 2015

Abstract

We examine the role of borrower concerns about future credit availability in mitigating the effects of adverse selection and income misrepresentation in the mortgage market. We show that the majority of additional risk associated with "low-doc" mortgages originated prior to the Great Recession was due to adverse selection on the part of borrowers who could verify income, but chose not to. We provide novel evidence that these borrowers were more likely to inflate or exaggerate their income. Our analysis suggests that recent regulations changes that have essentially eliminated the low-doc loan product would result in credit rationing against self-employed borrowers.

Key Words: *Subprime Mortgages, Default, Stated Income Loans, Adverse Selection, Reputation, Asymmetric Information, Credit Rationing, Dodd-Frank Act*
JEL Classification: G2, G01, G10, G18, D1, R2

*We thank Itzhak Ben-David, Sumit Agarwal, Manuel Adelino, Christopher Palmer, and the seminar participants at the National University of Singapore, the University of California-Berkeley, Research Institute of Capital Formation, Housing Research and Advancement Foundation of Japan, Board of Governors of the Federal Reserve System, the University of Cambridge, and the 2015 ASSA meeting in Boston for their helpful comments and suggestions. We also thank the Penn State Institute for Real Estate Studies for providing access to the New Century Mortgage database.

†Institute for Real Estate Studies and the Department of Risk Management, The Pennsylvania State University, University Park, PA 16802-3306, (814) 867-0066, bwa10@psu.edu

‡Department of Insurance, Legal Studies, and Real Estate, University of Georgia, Athens, GA, jnc152@uga.edu

§Department of Risk Management, The Pennsylvania State University, University Park, PA, juy18@psu.edu

Disclosure Statement for Brent W. Ambrose:

I thank the Penn State Institute for Real Estate Studies for providing access to the New Century Mortgage database. I receive compensation from The Pennsylvania State University and I have on-going consulting relationships with various financial institutions. However, neither I, nor any of my relatives, have received financial support from any interested party related to topics covered in this paper. Furthermore, neither I, nor any of my relatives, have any position in any relevant organization. No party had the right to review this paper prior to circulation.

Disclosure Statement for James N. Conklin:

I thank the Penn State Institute for Real Estate Studies for providing access to the New Century Mortgage database. I receive compensation from The University of Georgia. I have no consulting relationship with any financial institution. Neither I, nor any of my relatives, have received financial support from any interested party related to topics covered in this paper. Neither I, nor any of my relatives, have any position in any relevant organization. No party had the right to review this paper prior to circulation.

Disclosure Statement for Jiro Yoshida:

I receive compensation from The Pennsylvania State University. I received a research grant from the Institute for Real Estate Studies at The Pennsylvania State University. I have no consulting relationship with any financial institution. Neither I, nor any of my relatives, have received financial support from any interested party related to topics covered in this paper. Neither I, nor any of my relatives, have any position in any relevant organization. No party had the right to review this paper prior to circulation.

I. Introduction

During the Great Recession of 2007-2008, the U.S. experienced a massive increase in residential mortgage defaults and foreclosures not seen since the Great Depression. For example, the Financial Crisis Inquiry Commission reports that 9.7% of all mortgages were in default by the end of 2009 compared to approximately 1% at the start of the decade.[1] While the decline in house prices between 2007 and 2009 is obviously one of the primary causes for the significant number of mortgage defaults registered during the crisis, financial economists have only recently begun to examine the role of mortgage fraud and adverse selection in exacerbating the consequences of the 2007-2009 housing bust. Evidence is mounting that the great mortgage expansion that accompanied the rise in home prices coincided with increases in mortgage fraud related to misrepresentations of borrower income ((Jiang et al., 2014a), and (Mian and Sufi, 2015)), borrower assets ((Garmaise, 2015)), inflated appraisals ((Ben-David, 2011), (Agarwal et al., Forthcoming), (Agarwal et al., 2014), and (Griffin and Maturana, 2015)), and second liens and owner-occupancy status ((Piskorski et al., 2015)).[2] As a result, regulators and policy makers have implemented new rules to combat perceived abuses in mortgage lending.[3] Thus, the purpose of this paper is to shed light on how borrower heterogeneity with respect to employment status contributed to income misrepresentation and adverse selection, and how lender actions and concerns by borrowers about preserving future access to credit mitigated these risks. From a policy perspective, our results echo the concerns raised by Keys et al. (2009), Rajan et al. (2010), and Piskorski et al. (2010), among others, concerning the need to carefully weigh the costs and benefits of new financial regulations.

With respect to income misrepresentation, we present several novel insights. First, by comparing individual incomes within job titles, we provide new evidence that is highly suggestive that income misrepresentation was concentrated primarily among borrowers who originated low-documentation loans but could have easily originated full-documentation mortgages instead. Second, unlike previous studies that indicated that misrepresentation in mortgage originations resulted from lender actions at origination,[4] we find that income

[1]See U.S. (2011), page 215. Default is defined as "90-days or more past due or in foreclosure."

[2]In addition to mispresentation at the loan origination level, Piskorski et al. (2015) find evidence suggesting that misrepresentation was endemic in the secondary market (between originators and investors) as well. Furthermore, Agarwal and Evanoff (2013) provide evidence of systematic predatory lending practices by loan originators. These practices may have exacerbated the consequences of mortgage fraud.

[3]For example, the Consumer Financial Protections Bureau adopted the "Ability to Repay Rule" that requires lenders to provide greater documentation of borrower income, and the Federal Housing Finance Agency, in conjunction with the New York Attorney General's office, issued the Home Valuation Code of Conduct (HVCC) that was designed to reduce the incidences of inflated appraisals.

[4]Piskorski et al. (2015) is a notable exception. The authors provide evidence that borrowers misrepre-

falsification was essentially a borrower level phenomenon.[5] Thus we document that excesses in the mortgage market in the last decade resulted from both borrower and lender actions. Third, we provide new evidence about lender actions in response to potential borrower income falsification. Finally, we provide additional analysis examining the role of borrower income falsification in facilitating the expansion in mortgage credit. As a result, our analysis provides new insights into one of the possible causes of the Great Recession.

The role of borrower income misrepresentation leading up to the financial crisis is the source of considerable debate. For example, Mian and Sufi (2009) and Mian and Sufi (2015) argue that borrower income falsification was a leading culprit in facilitating the expansion of mortgage credit during the 2002 to 2006 housing boom period. Supporting this argument, Jiang et al. (2014a) show that income falsification occurred on low-documentation loans resulting in elevated defaults, particularly for loans originated through the wholesale channel. By focusing on differences in employment status, we show that the majority of adverse selection and income falsification is confined to a specific borrower group that was never intended to utilize the low-documentation product. Thus, our results show that broad policies designed to eliminate activities associated with excesses in mortgage originations during the housing boom may have unintended consequences.

Since the potential for mortgage fraud and adverse selection has always been present, lenders have long relied on underwriting guidelines to limit this risk. However, Burke et al. (2012) illustrate how lender screening to reject higher risk applicants results in greater adverse selection.[6] One such underwriting metric is the debt-to-income (DTI) ratio that limits the loan amount based on the borrower's income. This metric, in combination with the loan-to-value (LTV) ratio, serves to limit the borrower's housing consumption. As a result, borrowers seeking to maximize their housing consumption or investment have an incentive to exaggerate their income in order to reduce their DTI ratio thereby qualifying for a higher loan amount.

Recognizing the borrower's incentive to circumvent these metrics, mortgage lenders require proof of reported assets and incomes in order to verify that the borrower is capable of repaying the debt. Of course, verification of borrower income and assets comes at a cost. Not only do lenders bear costs associated with verification activities, but borrowers also bear costs of collecting and reporting incomes and assets to the lender. For some borrowers,

sented occupancy status on mortgage applications.

[5] Note that we are not arguing that the originator was unaware of the misrepresentation, only that borrowers are complicit in falsifying income.

[6] However, the presence of adverse selection at mortgage origination is not universally accepted. For example, Agarwal et al. (2012) rely on differences in loan performance between prime and subprime markets to claim that adverse selection was less severe in the subprime market.

these costs are relatively minor and involve simply submitting the prior two-years W2 tax documents from their employer along with their past two months paystubs. Unfortunately, the costs of verifying income and assets are not so trivial for many other potential borrowers. For example, self-employed individuals would need to provide full tax returns for the previous two-years. However, self-employed individuals often file for tax return extensions due to the complexity of their tax situation and as a result, the returns are not available to the lender. Furthermore, lenders will require current profit and loss statements along with bank statements for several months in order to prove sufficient cash flow to service the debt. In order to comply with underwriting debt-to-income guidelines, lenders may require additional documentation from self-employed borrowers to determine the nature of deposits and withdrawals to ascertain those expenses that are personal versus those associated with their business.[7]

Over time the mortgage industry developed different products designed to cater to borrowers with varying degrees of information verification costs. For example, the traditional mortgage referred to as a "full documentation" (or full-doc) loan is designed for borrowers who can easily and with low cost document their financial situation. However, recognizing that many self-employed borrowers would be effectively credit rationed in the traditional loan market due to the costs associated with documenting income and assets, the mortgage industry developed an alternative low-documentation (low-doc), or stated-income stated-asset loan.[8] Unfortunately, the low-doc product provides an avenue for some borrowers to inflate or exaggerate their incomes in order to qualify for larger mortgages. While borrowers are still subject to civil or criminal legal actions for providing inaccurate information, the costs associated with pursuing borrowers who fraudulently overstate income or assets often exceed the possible claims, particularly if the loan is still performing. Herein lies the tension in the low-doc product: as long as the borrower is making payments, the lender does not have an incentive to take actions against the borrower for falsely representing their income or assets.

To clarify the constraints facing borrowers, we present the mortgage rate sheet for New Century Financial Corporation (Figure 1). The rate sheet lists the interest rates charged on mortgages (as of July 10, 2006) originated by New Century based on whether the borrower was willing to verify income and assets ("Full Doc") or did not provide tax returns and bank

[7]Anecdotal discussions with mortgage brokers and other industry participants provide examples of the verification costs self-employed borrowers face. For example, lenders may require that self-employed borrowers provide written explanations for every deposit over the previous year. For a business with just two transactions per week, that would necessitate over 100 separate written explanations. Furthermore, many self-employed borrowers face serious confidentiality issues in revealing client names as the source of deposits.

[8]See Paley and Tzioumis (2011) and LaCour-Little and Yang (2013). We use the terms low-doc, no-doc, and stated-income interchangeably.

accounts to verify income and assets ("Stated Doc"). Each block in the rate sheet represents a borrower risk class ("AAA through C") that is based on the number of late payments, prior default records, or bankruptcy filings. Shaded areas without interest rates indicate that the loan product is not offered to borrowers that have credit scores in those risk categories.

To illustrate how borrower information verification costs and loan performance could interact to result in credit rationing, consider a high information cost borrower rated "A+" with a credit score of 660 who seeks an 85% loan-to-value (LTV) ratio mortgage.[9] Since this borrower finds it costly to verify income, he applies under the "Stated Doc" product type and is quoted a contract interest rate of 8.200%. The impact of reputation becomes apparent if the borrower is downgraded to the "B" category (e.g., by a 60-day late experience) before seeking to refinance into a new mortgage. Under the "B" category, New Century does not offer a stated doc loan at an 85% LTV; i.e., the borrower is effectively credit constrained unless he is willing to move to a lower LTV mortgage at a higher contract rate. In contrast, a comparable low information cost borrower that experienced a similar downgrade could easily switch to a full doc product with the same LTV. Since both borrowers are aware of this difference in borrowing constraints, reputation is relatively more valuable to the high verification cost borrower.[10]

To confirm that the insights obtained from the New Century rate sheet were common across the mortgage industry during the period prior to the Great Recession, we also collected wholesale rate sheets for several other mortgage lenders originating loans during that period. Figures A.1 - A.5 in the Internet Appendix show the wholesale rate sheets for First Franklin on July 10, 2006, Countrywide on August 16, 2006, and Argent Mortgage on July 21, 2006. Although it seems implausible in the context of current day mortgage underwriting practices, these three rate sheets and the New Century rate sheet display similar pricing patterns and reveal that full-documentation loans were available to borrowers who had declared bankruptcy or had a mortgage default within 2-years of the origination date. In contrast, the pricing matrices also clearly show that low-documentation loans were not available to borrowers with these characteristics at any price reinforcing the expectation that borrowers who could not easily verify income via a low-cost W2 would face credit rationing as a result of a prior bankruptcy or mortgage default.

Figure 2 demonstrates why understanding the role of future credit concerns in limited

[9]The "A+" category indicates that this borrower was 30-days late on a previous mortgage only once in the last twelve months.

[10]Although our example assumes borrowers accurately report their income, we recognize that the low information cost borrower may have falsified their income and thus not have sufficient "true" or verifiable income to qualify for a full doc product with the same LTV after a downgrade. However, this does not alter our intuition above because the LTV available to this borrower is still higher than the LTV available on a low-doc mortgage to a comparable high verification cost borrower.

information contracts is particularly important for self-employed borrowers. Using data from one of the largest subprime lenders in the run-up to the crisis, the figure shows the proportion of low-doc loans to self-employed and W2 borrowers by origination year. Roughly 80% of self-employed borrowers obtain low-doc loans, compared to only 30% for W2 borrowers. Clearly, low-doc loans are favored by the type of borrowers that they were originally intended for: the self-employed. Stated differently, limited information debt contracts are an important source of credit for borrowers that are likely to be credit rationed under full information (full-doc) mortgage contracts.

To better understand the link between mortgage type and borrower employment status, we theoretically and empirically demonstrate that low-doc loans experience higher *ex post* default rates than full-doc loans, and the relationship is strongest for low-doc W2 loans – the borrowers with the ability to access the full documentation origination channel. In other words, we find that the majority of the additional risk associated with low-doc loans is due to adverse selection on the part of borrowers with verifiable income. We conjecture that these borrowers likely selected into low-doc loans in order to inflate income to increase housing consumption. Thus, our analysis is connected to the theoretical insights developed in Diamond (1989) and Diamond (1991) regarding the role of borrower reputation in ameliorating adverse selection and income falsification.

Our results are related to an important recent attempt by Jiang et al. (2014a) to quantify the amount of income inflation on low-doc loans to W2 borrowers. Their results suggest that W2 borrowers with low-doc loans exaggerated income by 20% to 25%. Using a similar methodology on loans originated by a different lender, we estimate that income inflation ranged between 7% and 13% on low-doc loans to W2 borrowers. Thus, our study provides an additional point estimate for the level of income overstatement on so-called "liars' loans." Additionally, to our knowledge, we are the first to provide evidence that relative to W2 employees, self-employed borrowers refrain from overstating income when applying for mortgage loans. In fact, our regression result shows no evidence that self-employed borrowers selecting low-doc loans reported incomes that were above predictions from an income estimation model. Furthermore, we show that income inflation is directly related to *ex post* mortgage default for W2 borrowers, but the connection is less clear for the self-employed, which suggests that income falsification is most problematic on low-doc loans originated by W2 borrowers.

One of the unique features of our data is that we have information on loan applications, thus we also investigate lender actions to mitigate borrower adverse selection by documenting that the probability of lender loan application rejection was positively associated with borrowers most likely to engage in income falsification. Additionally, we provide evidence that premiums were set at a level that allowed adverse selection and untruthful reporting to persist in equilibrium. We also show that the low-doc effect on mortgage performance is reduced for borrowers with established positive credit reputation (e.g. borrowers with a high FICO score or a history of mortgage repayment). Taken together, these results suggest that reputation can mitigate adverse selection and private information in debt contracts. Finally, supporting the findings of Mian and Sufi (2015), we document that mortgages to borrowers who were the most likely to overstate income were concentrated in lower income neighborhoods.

Our findings are particularly important in light of the Consumer Financial Protections Bureau's (CFPB) "Ability to Repay Rule," which went into effect in January of 2014. This rule implements sections 1411 and 1412 of the Dodd-Frank Wall Street Reform and Consumer Protection Act (Dodd-Frank Act), requiring that lenders verify and document a potential borrower's ability to repay the loan.[11] Loans that do not meet the rule leave the lender exposed to significant litigation risks, effectively eliminating the low-doc loan market.

Unfortunately, eliminating the low-doc market likely results in regulator-imposed credit rationing against self-employed borrowers. Consistent with this idea, Green (2014, p.19) provides a telling description of the current mortgage market: "[W]hile people who draw regular salaries and receive W-2 forms from the Internal Revenue Service at the end of each year have fairly ready access to mortgage credit, self-employed people find it very difficult to obtain a mortgage. This is even true for people who can document a long history of self-employment income." Furthermore, this credit rationing against self-employed borrowers can have significant negative consequences for the economy. For example, Adelino et al. (2015b) provide direct evidence that employment in small businesses is related home price appreciation. Their analysis suggests that rising home prices allowed mortgage credit to expand via the collateral channel, which in turn created equity that could be used as working capital in small businesses. As a result, eliminating the low-doc loan market may have adverse consequences on future employment growth. However, this credit rationing against self-employed borrowers is likely unnecessary. We argue that the low-doc loan channel provides access to credit for self-employed borrowers, without a large increase in default risk, since self-employed borrowers' concerns for future credit significantly reduce the problems of adverse selection and income exaggeration endemic in low-doc loans originated by W2 borrowers. As a result, our analyses confirm the intuition embedded in models of reputation in financial contracting (e.g., Diamond (1989)).

[11]The Dodd-Frank Act is available online at https://www.sec.gov/about/laws/wallstreetreform-cpa.pdf. Information of the "Ability to Repay Rule" is available at http://www.consumerfinance.gov/regulations/ability-to-repay-and-qualified-mortgage-standards-under-the-truth-in-lending-act-regulation-z/#rule.

Our paper proceeds as follows. In section II, we discuss the interaction of borrower type based on income verification costs and mortgage product selection to develop a stylized model that motivates our empirical analysis. Section III discusses the data and summary statistics. Section IV presents the empirical results linking mortgage performance to borrower concerns over future credit. Section V provides evidence documenting the extent of borrower income misrepresentation and its impact on mortgage performance. We present robustness checks to control for income differences across job types (section V.A) and income differences within job types (section V.B). Next, section VI presents an analysis of lender responses to potential borrower income falsification. Specifically, we focus on lender screening at the time of application (section VI.A), links between observable credit reputation to mortgage performance (section VI.B), and loan pricing (section VI.C). In section VII, we highlight several important policy implications by examining the role of borrower income misrepresentation in facilitating the expansion of mortgage credit. Finally, section VIII concludes.

II. A Simple Model

To formulate testable hypotheses concerning the presence of adverse selection and borrower future access to credit, we first categorize mortgage contracts into high and low information loans based on the amount and extent of borrower information collected by the lender during the underwriting process. High information contracts represent full-doc mortgages where the loan originator collects and verifies the borrower's financial information (income and assets) as reported on the loan application. In contrast, low information contracts represent low-doc mortgages where the originator does not independently verify the borrower's claims concerning assets or income.

Next, we categorize borrowers with respect to information verification costs. For example, borrowers who are self-employed often face high information verification costs since they are unable to provide lenders with a W2 tax document from an employer. In contrast, borrowers who are employed by a third party have low information verification costs since they can easily produce an employer generated W2 statement that documents their income.

Obviously, the lender understands that low information contracts are *ex ante* riskier and prices them accordingly. Furthermore, since the level of borrower income is often a critical component in determining the maximum loan amount, the lender is aware of the possibility that some borrowers may inflate their reported income using the low information contract in order to secure a higher loan amount than would otherwise be available.

In the spirit of the Diamond (1991) model, we introduce three aspects of borrower heterogeneity into the borrowers' contract selection decision: information verification costs, reputation concerns, and the loan demand relative to income.[12] Within the context of our model, "reputation" embodies the borrower's concerns about and expectations for future access to credit. Thus, a borrower who loses reputation due to defaulting on an existing debt or failing a lender audit to verify submitted financial information faces higher future credit costs or is credit rationed.

We specify the borrower's reduced-form objective function on the basis of the amount of debt originated today and at some future date.[13] Specifically, the borrower's utility is expressed by the following equation:

$$U = u\left(L_1; \mu\right) - C_1 + \rho E\left[u\left(L_2; \mu\right) - C_2\right], \tag{1}$$

where L_t and C_t ($t = \{1, 2\}$) denote the debt amount and costs associated with the loan at period t. Parameter $\rho \in [0, 1]$ represents the borrower's probability of originating a future loan; a borrower has no concerns about future credit access if $\rho = 0$ and a maximal concern if $\rho = 1$. We assume $u\left(L; \mu\right)$ is a felicity function with $\partial u / \partial L > 0$, $\partial^2 u / \partial L^2 < 0$, and $\partial u / \partial \mu > 0$.[14] Parameter μ represents the borrower's loan demand. Loan demand is large if a borrower expects larger income growth, puts a higher utility weight on housing consumption, or is more tolerant of higher amounts of leverage.

Two types of loans are available for a borrower: full-doc and low-doc loans. For a full-doc loan, a borrower must prepare an income document. A borrower's true income y is private information, but the lender can verify this income by obtaining an appropriate document. For borrowers who have W2 tax documents, the cost of producing an income verification document is low $\left(c^L\right)$. In contrast, self-employed borrowers incur a high income verification costs $\left(c^H\right)$. These documentation costs are measured in the unit of utility and we normalize c^L to be zero.

The lender uses the borrower's reported income to determine the loan amount. For a full-doc loan, the loan amount (L^F) is a linear function of the borrower's true income: $L^F = \alpha y$, where α is a constant debt-to-income ratio. For a low-doc loan, a borrower reports her stated income y^S. The stated income can deviate from the true income by an unobserved positive factor $x : y^S = xy$. The variable x represents the degree of the borrower's income

[12] Key differences between two models are that we consider: (1) the borrower's optimal choice, (2) debt instruments with complete and incomplete monitoring, (3) liquidity default, and (4) two borrowing opportunities for an individual borrower.

[13] This two-loan objective function can be derived from a standard consumption choice model, in which a borrower gains utility by intertemporally smoothing consumption or by owning a house that better matches her unique personal taste.

[14] We use $u(L; \mu) = \mu\sqrt{L}$ for analytical convenience, but another concave function such as a log utility function gives essentially the same result.

exaggeration. For a low-doc loan, the lender uses an alternative debt-to-income ratio β to determine the loan amount: $L^N(x) = \beta y^S = \beta xy$.[15]

For simplicity, we model the mortgage as similar to a discount bond; the borrower receives the loan amount and pays the entire interest cost at origination, and pays back the total loan amount at maturity. Between origination and maturity, the borrower regularly sets aside part of her income in a sinking fund to pay off the loan at maturity. The borrower will default at maturity if she cannot build a sufficient fund due to negative income shocks during the loan term. We abstract from stochastic income and collateral processes to keep the model simple. Instead, we assume that the probability of default $D \in (0, 1)$ is an increasing function of the relative debt-to-income ratio: $D'(z) > 0, z \equiv \beta x/\alpha$.[16] When $z = 1$, the default probability is the same for the low-doc loan and full-doc loans because the ratio of the sinking fund payment to the initial true income is identical. As z increases, the borrower is less likely to accumulate a sufficient repayment fund because the annuity payment is large relative to the initial true income.

The lender cannot infer the borrower's loan demand from the loan amount because a large loan amount can arise from large loan demand or large income. Without verification, the lender has no information about the borrower's true income. The lender cannot infer the borrower's loan demand from a default event because a non-exaggerating borrower may also default on a loan. However, based on the inference about the average loan demand of a borrower group, the lender determines the loan interest rate. The interest rate for a full-doc loan is normalized to zero, and the interest spread for a low-doc loan is rL^N.

A W2 or self-employed low-doc borrower may face higher future credit costs or be credit rationed after originating the first loan with probability p, due to the lender's random audit.[17] However, the W2 borrower can still arrange a standard full-doc loan in the second period (possibly from another lender). In contrast, a self-employed borrower can only arrange a smaller low-doc loan: $L_D^N = \beta y$. Furthermore, the borrower additionally pays a penalty that depends on the degree of income exaggeration in the first period: $\gamma x L_D^N$.

The utility gains from full-doc loans for W2 $\left(U_W^F\right)$ and self-employed borrowers $\left(U_S^F\right)$ are,

[15] Technically, x could be less than one if the borrower wanted to under report income. However, we view this as a relatively uninteresting and rare case since loan amounts are jointly determined by the borrower's DTI ratio and the LTV ratio. If a borrower were to under report, then the DTI would be higher for a given loan amount. All else being equal, lenders view loans with higher DTIs as having higher default risk and subject them to increased underwriting scrutiny. As a result, these loans would face either elevated probability of lender rejection or higher interest rates due to risk-based pricing reducing the incentive to under report.

[16] For example, if $D(z) = (1+\delta/z)^{-1}$, where δ is a positive constant, then $D(z)$ has the following properties: $\lim_{z\to 0} D(z) = 0, \lim_{z\to\infty} D(z) = 1, D(1) = 1/(1+\delta)$.

[17] If p represents the default probability, our result will be enhanced because a borrower who exaggerates income and subsequently defaults will be more likely to face higher future credit costs or credit rationing.

respectively,

$$U_W^F = u\left(L^F;\mu\right) + \rho u\left(L^F;\mu\right), \tag{2}$$

and

$$U_S^F = u\left(L^F;\mu\right) - c^H + \rho\left(u\left(L^F;\mu\right) - c^H\right). \tag{3}$$

The utility gains from low-doc loans for W2 $\left(U_W^N\right)$ and self-employed borrowers $\left(U_S^N\right)$ are, respectively,

$$U_W^N(x) = u\left(L^N(x);\mu\right) - rL^N(x) + \rho\left[pu\left(L^F;\mu\right) + (1-p)\left(u\left(L^N(x);\mu\right) - rL^N(x)\right)\right], \tag{4}$$

and

$$\begin{aligned} U_S^N(x) &= u\left(L^N(x);\mu\right) - rL^N(x) \\ &\quad + \rho\left[p\left(u\left(L_D^N;\mu\right) - rL_D^N - \gamma x L_D^N\right) + (1-p)\left(u\left(L^N(x);\mu\right) - rL^N(x)\right)\right]. \end{aligned} \tag{5}$$

We first analyze a borrower's utility-maximizing choice of income exaggeration for a low-doc loan, given a loan cost r. Then we analyze the borrower's choice between a low-doc and full-doc loan. The details of the solution are outlined in the Internet Appendix. We obtain the following three propositions.

Proposition 1: *The level of income exaggeration is:*

$$x_W = \frac{\mu^2}{4r^2\beta y}$$

for W2 borrowers and self-employed borrowers without concerns over future credit access, and

$$x_S = Ax_W, \text{ where } A \equiv \left[1 + \frac{\rho p \gamma}{(1+\rho(1-p))\,r}\right]^{-2} \in (0,1]$$

for self-employed borrowers with concerns about future credit rationing. Thus, $X_S < X_W$.

The degree of income exaggeration (x) is small if the loan demand (μ) is small, the interest cost (r) is large, and the borrower can arrange a large loan amount on the basis of true income (βy). The difference in income exaggeration between a self-employed borrower and other borrowers is greater if the penalty for untruthful reporting is more severe (γ is larger), the probability of detection is greater (p is larger), or the self-employed borrower has greater concerns about future access to credit (ρ is greater). Note that the amount of a low-doc loan does not depend on β because the borrower can adjust her stated income in response to the lender's debt-to-income criterion. It is straightforward to link the degree of income exaggeration to the probability of default.

Proposition 2: *The probability of default is smaller for a self-employed borrower who has greater concerns about future access to credit than for an otherwise identical W2 borrower or a self-employed borrower without concerns about future credit availability. Specifically, the default probability is:*

$$D\left(\frac{\beta x_W}{\alpha}\right) \geq D\left(\frac{\beta x_S}{\alpha}\right).$$

The equation holds with equality if $\rho p \gamma = 0$.

A borrower chooses between a full-doc loan and a low-doc loan on the basis of the relative utility benefit. The utility benefit of a low-doc loan over a full-doc loan for a W2 borrower is:

$$B_W^N(\mu) \equiv U_W^N(x_W) - U_W^F = (1 + \rho\,(1-p))\left(\frac{\mu^2}{4r} - \sqrt{\alpha y}\mu\right). \tag{6}$$

For a self-employed borrower, the utility benefit is:

$$B_S^N(\mu) \equiv U_S^N(x_S) - U_S^F = \theta_1\mu^2 + \theta_2\mu + \theta_3, \tag{7}$$

where $\theta_1 > 0, \theta_2 < 0,$ and $\theta_3 \equiv (1+\rho)c^H - \rho pr\beta y$ are specified in the Internet Appendix. Both equations are convex quadratic functions of μ. The former takes a value of zero when $\mu = \mu^* \equiv 4r\sqrt{\alpha y}$. The latter exhibits the following properties: $B_S^N(0) = \theta_3$ and $\min B_S^N(\mu) = \theta_3 - \theta_2^2/4\theta_1$. Depending on the value of θ_3, the solution to $B_S^N(\mu) = 0$ has zero, one, or two roots. Using these properties, we obtain the following proposition.

Proposition 3: *W2 borrowers, irrespective of their future credit availability concerns, choose low-doc loans if and only if loan demand μ is greater than $\mu_W^* \equiv 4r\sqrt{\alpha y}$. Self-employed borrower choice depends on the cost of income verification: For $c^H > (\rho pr\beta y + \theta_2^2/4\theta_1)/(1+\rho)$, all self-employed borrowers choose low-doc loans. For $c^H < \rho pr\beta y/(1+\rho)$, a self-employed borrower chooses a low-doc loan if and only if $\mu > \mu_S^* = -\frac{\theta_2}{2\theta_1} + \sqrt{\frac{\theta_2^2}{4\theta_1^2} + \frac{\theta_3}{\theta_1}}$. Otherwise, a self-employed borrower chooses a low-doc loan if and only if $\mu > \mu_S^*$ or $\mu < \mu_S^{**} = -\frac{\theta_2}{2\theta_1} - \sqrt{\frac{\theta_2^2}{4\theta_1^2} + \frac{\theta_3}{\theta_1}}$.*

On the basis of the comparative statics of μ_W^*, μ_S^*, and μ_S^{**}, more borrowers will choose low-doc loans if the low-doc loan is less costly (r is smaller) or a full-doc loan amount is small (αy is small). In addition, more self-employed borrowers will choose low-doc loans if the income verification cost is larger (c^H is large). As a consequence, when the income verification cost is sufficiently large, the use of low-doc loan is more prevalent in a self-employed sample than in a W2 sample.

In equilibrium, the lender will charge a positive interest rate premium for low-doc loans by recognizing that borrowers who have stronger incentives to exaggerate income will select low doc loans. Moreover, the rate premium will be greater for W2 borrowers because the average default risk of the W2 low-doc borrowers is higher than that of self-employed low-doc borrowers. Furthermore, if the lender can estimate the level of income falsification of an individual borrower, the lender may charge a larger rate premium for a high estimated value of income falsification. Although the rate premium will mitigate the adverse selection and untruthful income reporting, it will not completely eliminate the problems. By increasing a spread, the lender faces a trade-off between the benefit of mitigating the problems and the cost of decreasing the total loan volume. By charging a high spread to completely eliminate the problems, the lender will lose opportunities to extend low-doc loans to the borrowers who only moderately exaggerate income. Thus, the problems of adverse selection and untruthful reporting will persist in equilibrium.

To summarize, based on the insights derived from our theoretical model, we develop the following empirical predictions. First, low-doc loans will be preferred by borrowers with high information verification costs, e.g, self-employed (section III). Second,the ex post probability of default will be lower for self-employed low-doc borrowers than for W2 low-doc borrowers (section IV). Third, borrowers will on average exaggerate income for low-doc mortgages, and the level of income falsification will be higher in the sample of the W2 borrowers than self-employed borrowers (sectionV). Fourth, there will be a positive mortgage rate premium for low-doc loans, and the premium will be larger for W2 low-doc borrowers than for self-employed low-doc borrowers (section VI). Finally, a rate premium will be positively related to income falsification (section VI). These predictions are summarized in Table I.

III. Data and Summary Statistics

A. Data

The main dataset used in the analysis contains loans originated by New Century Financial Corporation (New Century). New Century was one of the largest subprime lenders in the run-up to the recent mortgage crisis, with a large portion of its business originated through independent mortgage brokers. Along with originations, New Century also serviced mortgage loans and held a portfolio of loans as investments. New Century collected detailed borrower and collateral information at the time of origination, as well as contractual features of the loans. Also, for the loans that New Century serviced, monthly mortgage performance data is available.

From the loan origination records, we identify the borrower's employment type (e.g. W2 versus self-employed), as well as the level of income documentation (e.g. full-doc versus stated income.)[18] We focus only on first-lien loans with complete servicing data that were originated through the mortgage broker channel between 1998 and 2005.[19] Following Conklin (Forthcoming), to limit the effect of outliers and data entry errors we exclude loans where (1) total fees are negative or greater than 15% of the loan amount; (2) the yield spread premium paid from the bank to the broker is negative or greater than 5% of the loan amount; (3) the combined loan to value at origination is negative or greater than 125%; (4) the borrower's FICO score is less than 450 or greater than 850; (6) the debt-to-income ratio is negative or greater than 60%; (7) the borrower's monthly income is negative or greater than $26,900 and (8) borrower age is less than 18 or greater than 99. The final sample includes 458,872 funded mortgage loans.

We also obtain data from several supplemental sources. First, market interest rate data come from the Federal Reserve Bank of St. Louis's Federal Reserve Economic Data and Freddie Mac's Primary Mortgage Market Survey. Second, monthly MSA level unemployment rates are obtained from the Bureau of Labor and Statistics. Time varying MSA-level house price indices come from the Federal Housing Finance Agency. ZIP code level income information is obtained from the 2000 Census and IRS individual income tax statistics. Finally, the Pahl Index for mortgage broker regulations at the state level is collected from Pahl (2007) where higher values of the Pahl index indicate stricter regulation of brokers at the state level.

B. Summary Statistics

Table II presents the summary statistics for the sample separated by employment status and loan type. We note that 21% of the borrowers are self-employed, with the remainder classified as W2. Consistent with New Century's concentration in the subprime market niche, nearly 40% of the mortgages are low-doc loans. In comparison, Paley and Tzioumis (2011) state that roughly one third of all loans originated between 2000 - 2007 were low/no

[18] Although proof of income is not required on low-doc loans, verification of employment is required. For W2 borrowers, this usually entails a verbal verification of employment from the borrower's employer. For self-employed borrowers, lenders typically require a signed letter from a CPA or copies of business licenses. The New Century dataset contains a field indicating whether the borrower is self-employed. Throughout the paper we will refer to all borrowers that are not self-employed as W2 borrowers.

[19] We focus on brokered loans since the majority of New Century's originations were through brokers.

doc loans. We also note that 5% of the loans fall at least 60 days behind on their mortgage within the first 24 months after origination. Since New Century sold the majority of its loans within six months of origination, the observed default is a lower bound on the actual default rate.[20] Furthermore, our loan sample covers loans originated from 1998 to 2005, with performance data ending in early 2007. Since this covers the early period prior to the financial crisis when house prices were generally rising, most of the loans in the sample had not yet experienced significant declines in house prices to trigger negative equity induced default.[21]

Turning to loan characteristics, the average interest rate spread is 4.72%, and an overwhelming majority are adjustable rate mortgages (ARMs).[22] The mean loan amount is $193,000 with a combined loan to value ratio (CLTV) at origination of 83%. Furthermore, 34% of the loans are originated to purchase a home, while 56% are refinance loans with the borrower extracting equity (CASH).[23] The average FICO score is 613. Taken together, the summary statistics clearly reflect the fact that New Century was primarily a subprime lender with mortgages originated to higher risk borrowers.

In terms of observable borrower characteristics, Table II shows that the average borrower is 43 years old with an income of $6,200 per month. In addition, we note that 40% of the borrowers are minorities, and a large share (44%) were originated in the West region of the United States as classified by the U.S. Census Bureau. Since New Century began its operations in California, the strong focus in the West is not surprising. Furthermore, consistent with the entire subprime market, New Century experienced significant growth

[20]Some of the loans that exit the sample due to the transfer of servicing rights likely defaulted at a later period. Unfortunately, we cannot distinguish between loans that prepaid and loans where the servicing rights were transferred. Thus, standard techniques for handling competing risks with censored data cannot be employed.

[21]We also confirmed that the reported default rate in the New Century data is roughly comparable to the default rates on subprime mortgages as reported in the BlackBox (BBX) data. For example, for subprime loans originated in 2004, BBX reports an average 24-month default rate of 7.3%, compared to the average default rate of 5% in the New Century data. In addition, to assuage any concern about the representativeness of the New Century loans to the over all subprime market, we compared our sample to the loans in Demyanyk and Van Hemert (2011), a highly cited paper on the subprime mortgage crisis. Their sample spans many subprime lenders and covers roughly half of the subprime mortgage market. Table A2 in the Internet Appendix compares the descriptive statistics from loans originated in 2004 and 2005 (the years with the most originations in our data) in Table 1 of Demyanyk and Van Hemert (2011) with the New Century loans. The samples appear to be quite similar, however, the New Century data does include a larger proportion of low-doc loans.

[22]The rate spread is the initial contract rate minus the two year constant maturity Treasury rate at the time of origination. The average note rate on the mortgages is 7.68%, and the ARMs are actually "hybrid ARMs," with an initial fixed rate period (typically two years) with the interest rate adjusting every six months thereafter.

[23]The remaining 10% of loans are for rate/term refinances. These are cases where generally the borrower is refinancing to obtain an interest rate lower than the rate on the current mortgage.

from 2000 through 2005 (Chomsisengphet and Pennington-Cross (2006)).

Table II reveals several key differences across the borrower groups. First, consistent with predictions 1 and 2 in section II, loans to self-employed borrowers are much more likely to be low-doc (79% of the self-employed subsample are low-doc loans, compared to 30% for the W2 subsample.) This is not surprising since the low-doc product was designed specifically for borrowers with difficult to verify financial situations. Also, the average loan amount in the W2 subsample is $46,000 lower than the average for the self-employed group. Consistent with the difference in average loan sizes, the self-employed report a higher average income. Finally, the average FICO score is higher in the self employed subsample.

Since the summary statistics suggest that differences exist among the four borrower and loan product groups (low-doc self-employed, low-doc W2, full-doc self-employed, and full-doc W2), we report the kernel density distributions for borrower and mortgage characteristics in Figure 3. First, we see that the credit risk distribution for full-doc loans (W2 and self-employed) are wider and skewed lower than the low-doc borrower distributions. This is consistent with the lender imposing a higher underwriting screen on low-doc mortgages where borrowers have a greater opportunity to embellish their debt payment capacity. Second, the borrower income distribution for full-doc W2 loans is skewed lower than the other groups. In terms of borrower age, we see little difference in the kernel density distributions across the groups. Turning to loan characteristics, Figure 3 reveals a sizable difference in the distribution of mortgage amounts between the full-doc W2 borrowers and the other three groups. Figure 3 also reveals an interesting difference in loan pricing across the four groups. First, it appears that full-doc W2 borrowers have a higher proportion of high-fee mortgages. Second, the interest rate spread on full-doc loans (regardless of whether to a W2 borrower or self-employed borrower) are essentially the same. However, the interest rate spread distribution for the low-doc W2 borrowers is skewed higher. Thus, it appears that from a pricing perspective, the lender did anticipate that borrowers with W2s who selected low-doc loans were potentially higher risk and priced them accordingly. Yet, full-doc W2 borrowers tended to pay higher origination fees (as a percentage of their loan amount) than low-doc borrowers.

To summarize, Table II and Figure 3 indicate that several important differences exist between full-doc and low-doc loans according to borrower information verification cost type. First, the data supports our theoretical prediction that borrowers with high information verification costs (self-employed) will prefer low-doc loans. Second, borrowers with low information verification costs (W2 borrowers) that select the low-doc loan product have higher average reported incomes and loan amounts than similar borrowers who select the full documentation loans. Third, low-doc loans for the W2 borrowers experience higher levels of *ex post* default. Fourth, we do not observe a similar pattern for borrowers with high information verification costs. For self-employed borrowers, the average income and loan amount are similar regardless of the loan type. Furthermore, low-doc loans to self-employed borrowers do not have higher average default rates. Thus, the summary statistics provide preliminary evidence that is consistent with the popular narrative that low-doc loans were "liar's loans," but the role of borrower concerns for preserving access to credit may have ameliorated this tendency as low-doc loans to self-employed borrowers do not appear to have the same issues of income overstatement, loan amount distortion, or increased mortgage default risk.

IV. Borrower Type and Mortgage Performance

Since the univariate analysis confirms our first prediction that low-doc loans are preferred by self-employed borrowers, we now turn to a multivariate analysis to confirm our second theoretical prediction that W2 low-doc borrowers will be riskier than comparable self-employed borrowers. The unconditional analysis in the previous section supports this prediction. Therefore, our analysis in this section compares the *ex post* default rates *conditional* on borrower characteristics observable at loan origination as well as macro-economic factors and changes in house prices and interest rates after origination. Thus, we estimate the following loan-level regression of mortgage default:

$$\begin{aligned} Pr(DEFAULT_i) &= \Phi(\alpha + \beta_1 W2_i + \beta_2 Lowdoc_i + \beta_3 W2_i \times Lowdoc_i \\ &\quad + \delta X_i + \theta R + \vartheta W + \gamma T), \end{aligned} \tag{8}$$

where $DEFAULT_i$ is an indicator for mortgage default for loan i and Φ is the standard normal cumulative distribution function.[24] X_i represents information collected and recorded on the loan application. This information includes loan characteristics (fees charged on the loan, loan amount, combined loan-to-value ratio, whether the loan has a prepayment penalty, purchase or refinance, cash-out or rate/term refinance, and whether the payments are interest-only), property characteristics (two-unit, condominium, owner-occupied or investment property), and borrower characteristics (FICO score, borrower age, borrower income, debt-to-income ratio, whether the borrower met in person with the loan officer, and minority status). R captures market interest rates at the time of origination. The area

[24]The default variable takes a value of one if the loan becomes 60 or more days delinquent within 24 months of origination. In robustness checks reported in Table A4, we used alternate windows for delinquency (12 and 36 months) and the results were qualitatively unchanged. Unfortunately data limitations prevent us from observing loan performance in the mortgage crisis since the payment history is only available through the beginning of 2007 in the New Century database.

characteristics, W, include the monthly MSA unemployment rate, the level of broker competition, the Pahl index capturing the level of broker regulation at the state level, and the census region (West, Midwest, South, Northeast, or Pacific).[25] Since mortgage defaults are clearly related to house prices, W also includes MSA-level house price changes in the two years leading up to origination as well as MSA-level house price changes between origination and the last month the loan is observed in the performance data.[26] T is a set of variables denoting mortgage origination year to control for loan cohort effects. Throughout the analysis, unless otherwise stated, the reported standard errors are robust to heteroskedasticity and within cluster correlation of errors at the MSA level.

The parameters β_1, β_2, and β_3 are the primary coefficients of interest and capture the differential effect of borrower concerns about future credit access on the probability of default. β_1 represents the difference in outcome for borrowers with low information verification costs (i.e., when the employment type is W2.) β_2 captures the change in outcome when the loan type is low-doc. Finally, $\beta_1 + \beta_2 + \beta_3$ reflects the effect of borrowers with the least concern about future credit access as it captures borrowers with low information verification costs ($W2 = 1$) who originate a low information content mortgage ($Lowdoc = 1$).

Table III presents the estimated marginal effects from the maximum likelihood estimation of equation (8). Since Ai and Norton (2003), Williams (2012) and Buis (2010) note that reporting and interpreting a single marginal effect of an interaction term in a nonlinear model can be problematic and misleading, we follow Williams (2012) and report the marginal effects of low-doc at representative values for borrower employment type (e.g. at values of zero and one for *W2*).[27] In column [1], the marginal effects indicate that low-doc loans are associated with higher *ex post* default rates, regardless of employment type. This is consistent with the increased risk associated with low-doc loans and supports the pricing effect observed in Table II. However, the difference in magnitude between the effects for self-employed and W2 borrowers shows a more complex relationship and is consistent with borrower concerns about future credit access mitigating default risk. First, for borrowers with the highest concern (self-employed borrowers), the marginal effect of *Lowdoc* is modest (0.53%). To place this in

[25] Broker competition is computed as the quarterly Herfindahl-Hirchman Index in each MSA as in Ambrose and Conklin (2014).

[26] Although we report results using the pre-origination MSA house price changes over a two year period, results are insensitive to other window lengths (e.g. 1, 3, and 5 year house price changes).

[27] In unreported results we calculate a single estimate for the marginal effect of the interaction term using marginal effects at the sample means and the results are qualitatively unchanged. Williams provides a detailed discussion of the differences between average marginal effects and marginal effects at the mean http://www3.nd.edu/~rwilliam/stats/Margins01.pdf. As an additional robustness check, we employed a linear probability model of default. Consistent with the findings reported in Table III, we find that the relationship between low-doc and default is driven by W2 borrowers. Table A3 in the Internet Appendix reports the estimated marginal effects for this specification.

perspective, dividing the marginal effect by the mean default rate (0.0053/.0512) indicates that self-employed borrowers originating low-doc loans have a 10.4% higher probability of default than the reference group (self-employed borrowers originating full-doc loan.) In contrast, for borrowers with the least concern about future credit access (W2 borrowers) moving from a full-doc to a low-doc mortgage is associated with a 25.9% increase about the mean in mortgage default, *ceteris paribus*.[28] In other words, low-doc loans to self-employed borrowers pose modest additional default risk, consistent with the theory that borrowers with high information verification costs value the ability to obtain credit. However, low-doc loans to W2 borrowers have substantially higher default rates, in line with the hypothesis that they have less concern about being credit rationed in the future since they can easily switch to full-doc mortgages in the future where reputation is less important.

Although we include time-varying controls at the MSA level to account for local economic conditions (e.g. pre- and post-origination house price changes and unemployment), the possibility remains that unobserved time-constant geographic effects are driving the observed effect. Thus, as a robustness check, we include MSA fixed effects to address this concern (column [2]).[29] The results are virtually identical and confirm that low-doc mortgages have a higher likelihood of default, but the marginal effect is much larger for W2 borrowers.

Although our regression framework controls for all observable information available at loan origination, there remains the possibility of an omitted variables bias. Thus, as a robustness check, we present two additional specifications in columns [3] and [4]. First, in column [3] we use a propensity score matching approach. We match low-doc W2 observations with full-doc W2 observations using a nearest neighbor propensity score based on observable loan, borrower, and geographic characteristics. We also use the same matching procedure for self-employed low-doc observations. After creating our matched sample, we repeat the estimation of equation (8) and note that the results remain qualitatively unchanged. Finally, for a subsample of the borrowers, we are able to observe the total of the borrower's liquid assets (e.g. checking, savings, stocks, etc.). Thus, we repeat our main default regression controlling for borrower liquid assets (column [4]) and again, our primary results remain unchanged.

To summarize, Table III provides several key insights that are consistent with our second theoretical prediction. First, full-doc loans to self-employed borrowers are, *ex post*, marginally riskier than full-doc loans to W2 borrowers. This makes sense as income for self-employed borrowers is likely more volatile. Second, low-doc loans, in general, are riskier

[28]This is calculated by taking the ratio of the marginal effect to the average W2 borrower default rate (0.0124/0.0478.)

[29]Since several MSAs had no defaults, the number of observations included in the regression in column [2] is lower than in column [1].

than full-doc loans. Third, and most importantly, a distinction exists between low-doc loans originated to self-employed borrowers and low-doc mortgages originated by W2 borrowers. Consistent with our theoretical prediction that preserving access to future capital is valuable, the magnitude of the change in default risk is considerably larger for W2 borrowers originating low-doc loans.

V. Income Exaggeration and Mortgage Performance

In the previous section, we established that the majority of the elevated risk associated with low-doc mortgages resulted from the set of borrowers that were clearly capable of verifying income at a relatively low cost by providing a W2 statement. Having established that the problems documented with the low-doc product arose from a particular set of borrowers, we now test our third theoretical prediction by exploring the interaction of adverse selection and expectations of future access to credit with respect to borrower income falsification as a possible causal link for this increased risk.

We measure income exaggeration following the method outlined in Jiang et al. (2014a) and estimate a semi-log model of borrower income as a function of borrower characteristics (credit rating, race, sex, and age), area characteristics (income per capita measured at the borrower's ZIP code and house price growth over the previous two-years in the borrower's MSA), loan amount, an indicator for whether the property is an investment property, origination year dummies, and state dummies.[30] Section A.2 in the Internet Appendix reports the results for the borrower income model estimation.[31]

Table IV presents descriptive statistics for INC_EXAG across employment and documentation type. For low information verification cost borrowers (W2 borrowers) originating low-doc loans, the average estimated income overstatement is approximately 8%. In com-

[30] We recognize two potential issues that may result from including loan amount as a control variable. First, by including the loan amount as an explanatory variable in the income regression, we introduce a conservative bias into our estimation of income falsification. This bias may arise since borrower income is one of the metrics used in mortgage underwriting to determine the loan amount. Thus, in estimating income falsification, our method will tend to have higher predicted incomes for low-doc loans (and thus under estimate income falsification) if these borrowers used inflated incomes to qualify for higher loan amounts. A second, and closely related concern is that loan amount is endogenous. Since we are primarily interested in predictive accuracy, we do not view this as a major concern. Results in later sections that rely on our income estimates are not materially affected if we exclude loan amount from the income regressions.

[31] We use the coefficients reported in Table A1 in the Internet Appendix to compute estimates of income for the full-doc (in-sample) and low-doc (out-of-sample) loan borrowers. To calculate an estimate of income exaggeration (INC_EXAG), we subtract the estimated income from the reported income. Since estimated and reported income are both in logs, INC_EXAG represents the percentage difference between the borrower's reported income and estimated income. INC_EXAG is winsorized at the at the 1% level, but the main results are unchanged without winsorization.

parison, the average income overstatement associated with full-doc self-employed mortgages is 1%. For both W2 and self-employed borrowers, INC_EXAG is significantly different from zero.[32]

To formally identify the extent of income falsification, we estimate the following regression:

$$INC\_EXAG_i = \alpha + \beta_1 W2_i + \beta_2 Lowdoc_i + \beta_3 W2_i \times Lowdoc_i + \delta X_i + \theta R + \vartheta W + \gamma T + \varepsilon_i. \quad (9)$$

where INC_EXAG_i is our measure of income exaggeration, and X_i, R, W, and T are defined in equation (8).[33] Equation (9) tests whether borrowers selecting low-doc loans are correlated with our measure of income exaggeration and whether this effect depends on the value of reputation. Table V reports the coefficients of the OLS estimation of equation (9).[34] First, we note that the parameter estimate for *W2* is small and not statistically significant suggesting no material difference in income exaggeration between W2 and self-employed borrowers, on average. Next, we note that *Lowdoc* is positively related to our measure of income exaggeration, but again is not statistically significant. Third, the positive and statistically significant coefficient for the interaction term (*Low-Doc x W2*) indicates that income exaggeration increases in low-doc loans when the borrower is likely to have less concern for future credit access (i.e., W2 borrowers).[35] These results are consistent with the hypothesis that borrowers with the lowest *ex ante* concern over future credit availability

[32] By construction, INC_EXAG is not different from zero for the full-doc loans.

[33] Equation (9) is analogous to explaining the residuals from equation 8, so we do not include the control variables from equation 8 in equation 9. The reason we say analogous is because INC_EXAG includes both in-sample (full-doc) and out-of-sample (low-doc) estimates. Results are qualitatively unchanged if we include controls from equation 8 in equation 9.

[34] Since some of the observations in our full sample are missing variables used in the income prediction model, regressions using our income exaggeration measure will have a somewhat smaller sample size. For example, the sex of the primary borrower, which is used as an explanatory variable in our income prediction model, is missing for 5,144 observations.

[35] As a robustness check, we estimate the following probit model of income exaggeration:

$$Pr(INC\_EXTREME_i) = \Phi(\alpha + \beta_1 W2_i + \beta_2 Lowdoc_i + \beta_3 W2_i \times Lowdoc_i + \delta X_i + \theta R + \vartheta W + \gamma T),$$

where $INC\_EXTREME_i$ is a dummy variable equal to one if INC_EXAG_i is in the top decile for the borrower's employment type, and X_i, R, W, and T are defined in equation (8). Table A5 in the Internet Appendix reports the estimated marginal effects for this regression. The results confirm that for self-employed borrowers, selection of a low-doc loan is not significantly related to the probability of extreme income exaggeration. However, W2 borrowers originating low-doc loans are significantly more likely to have extreme income overstatement. We also confirm that the results remain unchanged if we use the top quartile of *INC_EXAG* as our cutoff for *INC_EXTREME*. Finally, Table A6 in the Internet Appendix reports the estimated coefficients assuming a linear probability model of $INC\_EXTREME$. Again, the results confirm that income exaggeration increases in the low-doc loans when originated by W2 borrowers.

(W2 borrowers originating low-doc loans) are likely to inflate income. Focusing on the low-doc loan type and comparing income exaggeration across self-employed and W2 borrowers, the interaction term shows that W2 borrowers have a significantly higher level of income exaggeration than self-employed borrowers. To put our income exaggeration measure into perspective, using a sample of loans from a different lender, Jiang et al. (2014a) estimate income overstatement of 20% to 25% on low-doc loans to W2 borrowers. Although the magnitudes differ somewhat across our studies, both estimates suggest that low-doc loans to W2 borrowers are in fact "liars' loans."

Finally, to estimate the impact of income exaggeration on *ex post* mortgage default, we estimate the following regression:

$$\begin{aligned} Pr(DEFAULT_i) = \Phi(&\alpha + \beta_1 W2_i + \beta_2 Lowdoc_i + \beta_3 W2_i \times Lowdoc_i \\ &+ \lambda_1 INC\_EXAG_i + \lambda_2 W2_i \times INC\_EXAG_i + \lambda_3 Lowdoc_i \times INC\_EXAG_i + \\ &\lambda_4 W2_i \times Lowdoc_i \times INC\_EXAG_i + \delta X_i + \theta R + \vartheta W + \gamma T), \end{aligned} \quad (10)$$

where $DEFAULT_i$ measures whether the loan is 60-days delinquent over the 24-months following origination, X_i, R, W, and T are defined in equation (8). To provide more comprehensive insight into the observed effect, we compute the average marginal effects of *Lowdoc* at different levels of income exaggeration across employment types and present the results graphically in Panel A of Figure 4.[36] The horizontal axis in Panel A of Figure 4 runs from the 5th to the 95th percentile of *INC_EXAG*. Displaying the marginal effects across a range of income exaggeration levels reveals several interesting insights. We see that higher levels of income exaggeration among W2 borrowers have a larger impact on the probability of default. In contrast, the slope of the marginal effect for self-employed borrowers is negative but not significantly different from zero. Thus, to summarize, we find that income falsification is positively related to default for low-doc loans with low information verification costs (W2 borrowers). However, the same relation does not hold for low-doc self-employed borrowers. Therefore, the results support the hypothesis that borrowers with relatively low *ex ante* concern for future credit access that self select into low information mortgages are most likely to inflate income during loan origination and this risk manifests itself in higher *ex post* default

[36] Since *INC_EXAG* is a generated regressor, we use a bootstrapping procedure to estimate the marginal effects and standard errors. The first step of the bootstrapping procedure is to take a random sample of size N (with replacement) from the overall sample size of N. Next, we follow the methodology outlined above to create estimates of *INC_EXAG*. Third, we estimate Equation 10 on the sample and record the marginal effect and heteroskedasticity robust standard error estimates. We then repeat the sampling and estimating procedure 400 times and use the average marginal effects and standard errors across the 400 replications. Table A7 in the Internet Appendix presents the results.

rates.[37]

A. *Robustness Check: Job-Specific Overstatement*

To better understand the magnitude of income exaggeration, we create a second measure of income overstatement. For a subset of the applications in the New Century database, the lender recorded the borrower's line of business or job title (e.g. "TEACHER," "PRESIDENT"). Using these classifications, we can compute the average income for low-doc and full-doc borrowers *within* each job title classification. Comparing average incomes across low-doc and full-doc loans within the same job title and employment type (W2, self-employed) provides another measure of whether low-doc borrowers systematically inflate income, and whether this varies according to employment type.[38]

Table VI presents the average incomes across documentation types for the 25 most frequently used job titles by W2 borrowers.[39] In the first column, we see that there are 1,855 low-doc loans to W2 borrowers whose job title is "MANAGER," with an average income of $6,720 per month. In column [2], there are 1,794 full-doc W2 borrowers with a job title of "MANAGER" that have an average income of $5,563 per month. Column [3] presents the mean difference test across the documentation types. In column [4], the mean difference is divided by the average income for the full-doc group to create *Job-Specific Overstatement (%)*.

[37]Even though income is not verified on low-doc loans, verification of employment and employment type (W2, self-employed) is required, as described in section III.A. Thus, employment type misrepresentation is not a major concern for our study. A related, but separate concern is that low-doc borrowers may receive both W2 and self-employment income (dual-employment), even though we do not observe this in our data. This imperfect measurement of W2 status, which would be more pronounced in the low-doc sample, may potentially introduce two sources of bias. First, there could be an attenuation bias in the estimated coefficients on W2, especially the interaction term with *Low-doc*. However, this is not a major concern since we find statistically significant effects of these variables despite this potential bias. Second, our income exaggeration measure for W2 low-doc borrowers may be biased, but the effect is likely minor. For example, if a dual-employment borrower is less productive in the non-W2 occupation, then the reported total income would be evaluated as a negative exaggeration because we predict income on the basis of full-time W2 income. However, since W2 and self-employment productivity should be highly correlated within an individual borrower, this bias is likely to be small.

[38]We restrict our analysis in this section to observations where there is no co-borrower or the co-borrower's income is listed as zero. On low-doc loans with multiple borrowers and multiple job types, detecting income overstatement becomes much more difficult as exaggeration could occur within either (or both) jobs.

[39]Borrower business type is not a standardized field in the New Century data. For example, Table VI shows borrower business types of "NURSE," "REGISTERED NURSE," and "RN." Although these are clearly similar (or the same) positions, we did not attempt to standardize the field for several reasons. First, there are over 39,000 unique borrower business types in the data, so manually reviewing and standardizing these is cost prohibitive. Second, any attempt to standardize the field, including fuzzy matching techniques, requires significant judgment calls on the part of the authors. Instead of letting our own biases enter into the standardization algorithm, we chose to use the field in its raw form. This is a conservative treatment as it reduces the number of observations in each category and thereby reduces the overall statistical power of our subsequent tests. As a result, our analysis is biased toward not finding income exaggeration within job titles.

As the name suggests, *Job-Specific Overstatement (%)* can be interpreted as the percentage increase in reported income for a job type when no income documentation is provided.[40]

For every job title in Table VI, the average low-doc/W2 income is significantly higher than the average full-doc/W2 income. Furthermore, the differences are significant in economic terms as well. For example, the average low-doc W2 school teacher's income is $1,458 greater per month ($17,496 annually) than the average full-doc W2 teacher's income. If we take the average full-doc income as an unbiased estimate of the average teacher's "true" income,[41] this suggests that low-doc teachers inflated their income by 24%. Within these 25 most frequently used W2 job titles, the average *Job-Specific Overstatement (%)* is 20%.

Next, we turn our attention to the 25 most frequently used job titles by self-employed borrowers. In Table VII, the same pattern of overstatement does not emerge for self-employed borrowers. First, for many of the job titles, no significant difference exists across the low-doc and full-doc groups. In addition, whereas in Table VI all of the mean differences are positive, for self-employed borrowers there are both positive and negative differences, and the average overstatement is -3%. Consistent with our previous findings, this suggests that income exaggeration is systematic for low-doc W2 borrowers, but not for the self-employed.

To ensure that our results are not driven by including only the 25 most frequently reported borrower business types, we broadened our sample to include any job titles that meet at least one of the following two requirements: 1) there are ten low-doc W2 *and* ten full-doc W2 observations with the job title or 2) there are ten low-doc self-employed *and* ten full-doc self-employed observations with the job title. 313 job titles meet the first criteria and 55 job titles meet the second. Requiring ten loans of each documentation types limits the ability of outliers to drive our results.[42] For each of these job titles, we calculated *Job-Specific Overstatement (%)* as above. The distribution of *Job-Specific Overstatement (%)* by employment type (W2, self-employed) is presented graphically in Figure 5. The distribution of job title overstatement for the 313 W2 job titles is clearly shifted to the right of the distribution for the 55 self-employed job titles, providing further evidence that income inflation is a problem on low-doc W2 loans.

To formalize the visual results in Figure 5, Table VIII presents the mean of each distribution. The average overstatement for W2 jobs is 28%, versus -0.42% for self-employed job titles. In the second row of Table VIII, we report the proportion of job titles with overstate-

[40] 0.18% of the borrowers in this subsample have a job title of "OWNER" but are not coded as self-employed in the New Century data. Since we cannot determine whether the job title or the self-employment flag was coded incorrectly, we exclude those 329 observations from the analysis.

[41] We believe this is a reasonable assumption since full-doc borrowers provided proof of income in the underwriting process.

[42] Although our choice of ten loans per documentation type within a job title is somewhat arbitrary, our results are robust to other limits (7, 12) and the use of the median rather than the mean income.

ment above zero. If borrowers report true income on low-doc loans, then we would expect this number to be 50%. For W2 job titles the number is 90%, which using a two-tailed t-test, the null hypothesis is rejected at the 1% level of confidence. Turning to self-employed job titles, where 39% have overstatement above zero, we fail to reject the null of 50%.

Next we calculate a mean difference test of average incomes across documentation types (as in column [3] of Tables VI and VII) for each borrower business type. For each job title, we test the null hypothesis of $H_0 : \theta \leq 0$ against $H_a : \theta > 0$. The third row of Table VIII reports the fraction of the mean differences for which the null hypothesis of $H_0 : \theta \leq 0$ is rejected at least at the 10% level of confidence. For the 313 W2 job titles (of which 90% had higher average income on low-doc loans), the null hypothesis is rejected 73% of the time. In comparison, the null hypothesis is rejected only on 16% of the job titles for self-employed borrowers. The results in Table VIII provide strong evidence of income inflation on low-doc loans within the W2 employment type, however, again we see no evidence of income exaggeration by the self-employed.[43]

Next, we investigate which jobs tend to have the largest income inflation. Table IX reports the top 25 job titles by employment type in terms of *Job-Specific Overstatement (%)*. For the top ranking W2 job title (PERSONAL BANKER), the average annual income for low-doc borrowers ($84,672) is more than double the average income for full-doc borrowers with the same job title ($38,412). The average low-doc W2 letter carrier reports annual income of $103,128, as compared with his full-doc W2 counterpart of $60,252. The table also shows that the largest job title overstatement for self-employed borrowers (CLEANING) has overstatement below the 25th highest job title (WELDER) for W2 (58% versus 47%), again providing evidence that overstatement is particularly problematic in the low-doc W2 job titles.

To summarize, in this section we created a second measure of income overstatement (*Job-Specific Overstatement (%)*) based on borrower business type. This variable is unique to the New Century data set, and allows us to test for differences in average income *within* a specific job title. Our results show that income overstatement is systematic on low-doc loans within W2 job titles, however, we find no evidence of the same phenomenon in self-employed job titles, consistent with our earlier results on income exaggeration. To our knowledge, our study is the first to exploit differences in income across documentation types within job titles. Similar to the estimates of 20-25% in Jiang et al. (2014b), our results suggest that on

[43] One concern is that our analysis of job title incomes does not control for differences across location. Thus, to alleviate concerns about geographic differences in incomes, Table A8 in the Internet Appendix repeats the analysis from Table VIII for borrowers located in California and Florida. Although the sample sizes are much smaller, the results are consistent with those in Table VIII leading us to conclude that geographic differences in incomes are not biasing the analysis.

average, low-doc W2 borrowers inflate income by 28%. For low-doc self-employed borrowers the average inflation is 0%.

B. Robustness Check: Income within Jobs

In section V.A we show that within the same job title, the average income for low-doc W2 borrowers is significantly higher than for full-doc borrowers, but the same relationship does not hold for self-employed borrowers. However, there are several potential concerns with that analysis. First, we limit our sample to jobs that have at least 10 full-doc and 10 low-doc observations within one of the employment types (W2, self-employed). Second, the averages reported may simply pick up systematic differences in salaries across geographic locations. For example, if most low-doc loans to W2 teachers occur in areas with relatively high teacher salaries, while the majority of full-doc loans to W2 teachers occur in regions where teachers' salaries are low, then we would incorrectly attribute differences to income falsification when the causal mechanism is actually benign. Third, we did not control for individual borrower characteristics that may be correlated with income. Thus, to address these concerns, we estimate the following loan level regression on the subsample of loans where borrower business type is not mising:

$$ln(INCOME_{ikj}) = \alpha + \beta_1 W2_i + \beta_2 Lowdoc_i + \beta_3 W2_i \times Lowdoc_i + \delta X_i + \vartheta W_k + \gamma T + \\ + \left(\sum_{j=1}^{J} \delta_j JOB\_TITLE_j + \sum_{j=1}^{J} W2_i \times \delta_j JOB\_TITLE_j \right) + \varepsilon_{ikj}, \quad (11)$$

where JOB_TITLE_j is the borrower's business type as listed in the NCEN database.[44] The other variables are as defined above. The first term in parentheses allows us to compare within job-specific income differences between low-doc and full-doc loans, while the second term controls for the possibility that W2 and self-employed borrowers in the same position might earn different incomes.

Column [1] of Table X serves as a baseline regression of equation (11) where we include no additional control variables. The estimated coefficients on *Low-doc* and W2 × *Low-Doc* are consistent with the average overstatement in Table VIII.[45] In column [2], we introduce job

[44] Due to the large number of fixed effects, we only include observations for which the job title has three or more observations. These observations can come from any of the employment type/documentation type combinations (e.g. low-doc/W2, full-doc/W2, low-doc/self-employed, full-doc/self-employed). The subsample includes 2,934 unique job titles. 448 job titles are held by both W2 and self-employed borrowers, 468 are held only by self-employed borrowers, and 2,018 job titles are only held by W2 borrowers.

[45] The average monthly incomes for full-doc self-employed and W2 observations in this subsample are \$8,363 and \$5,122, respectively. Note that these averages are higher than in the full sample used in earlier

title fixed effects and the interaction of job title fixed effects with the W2 indicator.[46] The coefficients on the employment type and income documentation variables represent income differences within a specific job title. The income difference becomes somewhat smaller in magnitude, indicating correlations between income documentation and job title. In column [3], we additionally include MSA fixed effects and origination year fixed effects to control for geographic income variation and nation-wide changes in economic conditions, respectively. The results are qualitatively similar to those in column [2].

In column [4], we further control for borrower and area characteristics. More specifically, we include the natural logarithm of FICO score, an indicator for female, the natural logarithm of age, an indicator for minority status, the natural logarithm of the ZIP code per capita income reported annually, an indicator for investment property, and the MSA level house price growth over the previous two years.[47] Although the signs and significance of the coefficients are similar to those in column [3], the magnitude of income difference is significantly smaller. For example, the coefficient on W2 is not statistically significant, indicating that income is now comparable between W2 and self-employed full-doc borrowers. The coefficient on W2 × *Low-doc* is 0.149, which is smaller than the coefficient in Column [3], but statistically significant at the 1% level. The low-doc W2 borrowers appear to over-report income by approximately 13.24% relative to the full-doc W2 mean income.[48] In contrast, the income reported by self-employed borrowers for low-doc loans is slightly lower (by 1.70%) than the full-doc self-employed income. Thus, consistent with all of our previous findings, income falsification is only problematic on low-doc loans to W2 borrowers after controlling for job title and other relevant factors.

To summarize, several important facts emerge from the results in Table X. First, job titles are important in explaining income. Although this may not be surprising, to our knowledge this is the first study to control for the borrower's job type when examining income on low-doc loans. Second, even after controlling for job titles, area characteristics, and borrower characteristics, the results are consistent with our previous findings: income overstatement appears to be problematic only on low-doc loans to W2 borrowers. However, when we control for borrower and area characteristics, as well as the borrower's job title, the average amount of income exaggeration by low-doc W2 borrowers is reduced to 13%, a smaller number than the 20 - 25% reported in Jiang et al. (2014b), as well as our results in Section V.A.

---

analysis reported in Table II. As noted above, the subsample in this section includes only observations where borrower business type is not missing and there is no co-borrower income.

[46]The omitted job title category "TEACHER."

[47]The IRS income data is not available for 1999, 2000, and 2003. Loans originated in those years are matched to IRS data from the most recent prior year for which data is available. Results remain qualitatively unchanged if we exclude loans originated in 1999, 2000, and 2003.

[48]We report $\beta_2 + \beta_3 + 1/2\times$ standard errors of $(\beta_2 + \beta_3)$.

VI. Lender Attempts at Controlling Falsification

Predictions 4 and 5 from our theoretical model imply that lenders should react to potential borrower income falsification by charging higher interest rate premiums on low-doc loans and to borrowers with low information verification costs that seek out low-doc loans. Thus, in this section we test these predictions using a unique feature of the New Century data that allows us to examine the loan applications as well as loans that were actually originated. By using loan applications, we make a novel contribution to the literature in that we are able to examine the impact of potential borrower income falsification on the underwriting decision.

A. Loan Application Rejection

Lenders make decisions on loan applications along two important margins: pricing and application acceptance. Because most mortgage databases contain information only on funded loans, previous studies on low-doc loans have focused on the former. Since the NCEN data includes data on funded and non-funded mortgage applications, we are able to help fill this gap in the literature. We ask several questions regarding the lender's accept/reject decision. First, since agents (borrowers or brokers) likely inflate income with the goal of increasing the probability of application acceptance, are low-doc loans less likely to be declined by the lender? Second, does the lender reject low-doc loans differently across employment types? If the risk of default on low-doc loans varies with employment type, the lender may base its rejection decision on this information. Finally, is income exaggeration accounted for in the lender's rejection decision? To examine these questions, we expand our sample to include 698,019 funded and non-funded applications. The percentage of loans that are funded, approved but not funded, and rejected are 67%, 19%, and 14%, respectively.[49]

To investigate whether the lender's rejection decision varies with documentation type, we first estimate a probit regression similar to equation (8) with the dependent variable taking a value of one if the loan application is rejected (see Table XI). Whereas in the default regressions we included post-origination variables to control for changing market conditions (e.g. house price changes), the explanatory variables in this regression only include information available to the lender at the time of the accept/reject decision. As in Section IV, we follow Williams (2012) and report the marginal effects of low-doc at representative values (MERs) for borrower employment type (e.g. at values of zero and one for *W2*). Table XI presents the results from this regression. For self-employed borrowers, low-doc is associated

[49]Due to missing variables, the sample size for the regression in this section is 697,020 observations. Also, our sample of funded loans is larger in this section than the sample used in Section IV, since in the default regressions, observations are dropped that are missing post-origination information, but no such requirement is made in this section.

with a 1% reduction in the probability of being rejected, or a 4.5% reduction relative to the mean for full-doc self-employed borrowers. However, the relationship reverses for W2 borrowers. The probability of application rejection is 1% higher on low-doc loans to W2 borrowers, or a 6.6% increase relative to the mean rejection rate for full-doc W2 borrowers. Clearly the documentation type affects the application rejection decision. Moreover, the results suggest that the lender recognizes that the propensity for income falsification is larger on W2 low-doc loans.

To test whether the lender incorporates income falsification into the rejection decision, we estimate a linear probability regression similar to equation (10) where we include our measure of income exaggeration. We now use the rejection indicator as the dependent variable and include all of the independent variables from equation (10) that are observable to the lender at the time of the rejection decision. Table XII reports the marginal coefficient estimates for the OLS estimation.[50] The results indicate that W2 borrowers are 2.4 percent less likely to be rejected than self-employed borrowers. Similarly, borrowers originating low-doc loans are 1.06 percent less likely to be rejected than borrowers seeking full-doc loans. However, interaction of W2 and low-doc confirms that borrowers with low costs of verifying information faced significantly higher lender scrutiny as the probability of rejection is 1.77 percent higher than for self-employed low-doc borrowers. Finally, the positive and significant coefficient for the triple interaction between low-doc, W2, and INC_EXAG indicates the lender recognized income falsification on low-doc loans to W2 borrowers and adjusted the probability of rejection accordingly.

Taken together, the results in Tables XI and XII provide several new insights on low-doc loans. First, low-doc loans are treated differently from full-doc loans with regards to loan approval. Second, the relationship varies according to employment type. Low-doc is associated with a lower likelihood of rejection for self-employed borrowers, but for W2 borrowers low-doc loans are more likely to be declined. Third, the lender appears to incorporate income exaggeration into the rejection decision for low-doc W2 borrowers, but not for self-employed borrowers. This is consistent with our previous results that income falsification appears to be problematic only on low-doc loans to W2 borrowers.

[50]Since INC_EXAG is a generated regressor, we use a bootstrapping procedure similar to the procedure outlined for Equation 10. However, the bootstrapping procedure causes problems when using a probit model on all loan application (versus funded loans in earlier sections), so we report bootstrap estimates from an OLS model in Table XII. Column (1) simply reports the coefficient estimates when we do not include the generated regressor in the model.

B. *Low-doc Loans and Credit Reputation*

Next, we examine the interaction of low-doc, employment type, and credit history. We measure credit history that is observable at origination using the borrower's credit (FICO) score, a standard risk metric used in mortgage underwriting in the United States. Over time an individual develops a reputation with creditors through credit usage and debt repayment patterns. The FICO score quantifies this reputation, with higher scores reflecting more credit-worthy borrowers, ceteris paribus. Since credit scores are widely used for lending, insurance, and employment decisions, a strong credit reputation, as indicated by a high FICO score, is a valuable asset for a borrower.

In this section we test whether observed credit reputation mitigates the default risk of borrowers that otherwise have signaled a low concern over access to future credit (W2 borrowers selecting low-doc loans). Our regression now takes the form

$$\begin{aligned}Pr(DEFAULT_i) = \Phi(&\alpha + \beta_1 W2_i + \beta_2 Lowdoc_i + \beta_3 W2_i \times Lowdoc_i \\ &+ \lambda_1 FICO_i + \lambda_2 W2_i \times FICO_i + \lambda_3 Lowdoc_i \times FICO_i + \\ &\lambda_4 W2_i \times Lowdoc_i \times FICO_i + \delta X_i + \theta R + \vartheta W + \gamma T), \quad (12)\end{aligned}$$

where $FICO_i$ is the borrower's credit score at origination. All other variables are as defined in equation (8). The three-way interaction of $W2$ with *Lowdoc* and $FICO$ allows us to test whether an established credit reputation ameliorates the additional default risk of low-doc loans.

Panel B of Figure 4 graphs the average marginal effects of low-doc, by employment type, across FICO scores.[51] For low-cost verification borrowers (W2), the downward sloping line provides some evidence that credit reputation counteracts the income exaggeration problem inherent in low-doc loans. That is, borrowers with higher FICO scores have lower default probabilities. However, the same result does not hold for self-employed borrowers. Interestingly, we note that the average marginal effect of *Lowdoc* increases over the lower range of FICO scores for self-employed borrowers. Given the wide confidence intervals, we are careful not to interpret the results in this section too strongly. However, Panel B of Figure 4 does suggest that the increased risk associated with low-doc loans is most severe for borrowers that are least likely to be concerned about future credit rationing: W2 borrowers with low FICO scores.

[51] Table A10 in the Internet Appendix reports the marginal effects of low-doc at different levels of FICO score by employment type in tabular form.

C. Reputation, Income Falsification, and Mortgage Pricing

The previous sections demonstrate that low-doc loans to borrowers with a low value for reputation are riskier due to income inflation. In this section, we examine whether the lender priced this risk. To test this hypothesis we estimate the following OLS model of pricing:

$$\begin{aligned} RATE\_SPREAD_{ij} &= \alpha + \beta_1 W2_i + \beta_2 Lowdoc_i + \beta_3\{W2_i \times Lowdoc_i\} \\ &\quad + \lambda_1 INC\_EXAG_i + \lambda_2\{W2_i \times INC\_EXAG_i\} \\ &\quad + \lambda_3\{Lowdoc_i \times INC\_EXAG_i\} \\ &\quad + \lambda_4\{W2_i \times Lowdoc_i \times INC\_EXAG_i\} \\ &\quad + \delta X_i + \theta R + \vartheta W_j + \gamma T + \varepsilon_{ij}, \end{aligned} \tag{13}$$

where $RATE\_SPREAD$ is the note rate on the mortgage minus the two year T-bill rate in the month of origination, with the control variables as defined in Section III. Column [1] of Table XIII reports the coefficient estimates from the pricing regression using the entire sample. Relative to full-doc self-employed borrowers, interest rate spreads on loans to W2 borrowers are 9.8 basis points lower. The second and third rows of column [1] suggest that the lender recognized differences in low-doc loan quality according to borrower reputation. Low-doc loans to borrowers with a low value for reputation (W2) carried an additional risk premium of 15 basis points relative to low-doc loans to self-employed borrowers. Interestingly, although most of the additional risk on low-doc loans is attributable to W2 borrowers, the majority of the low-doc premium (53 basis points) applies to all borrower types.

The second column of Table XIII includes INC_EXAG.[52] The coefficients on *Lowdoc* and *W2* $\times$ *Lowdoc* are nearly identical to those in column [1]. The coefficient on INC_EXAG suggests that full-doc self-employed borrowers with high income levels (relative to our model estimates) pay a rate premium. The same result holds for full-doc W2 borrowers with high levels of income. Since INC_EXAG for a full-doc borrower does not contain income falsification, this rate premium corresponds to a higher risk in a mortgage originated to a high-income individual, possibly due to a higher risk in income or collateral value. This rate premium on INC_EXAG is not significantly different for low-doc loans to self-employed or W2 borrowers.

The results in column [2] clearly show that the lender prices additional risk associated with low-doc loans to W2 borrowers, but we find no evidence that the pricing is related to income exaggeration at the loan level. However, it is important to recognize that loan

[52] Since INC_EXAG is a generated regressor, we use a bootstrapping procedure similar to the procedure outlined for Equation 10.

pricing is the result of two processes: 1) the lender's risk based pricing and 2) negotiations between the borrower and the originator. Although we cannot fully disentangle each of these effects, Figure 1 provides some insight. Clearly there is a risk-based premium moving from the "FULL DOC" to the "STATED DOC" side of the pricing sheet. In addition, the "Adjustments To Rate" section shows an additional rate premium of 30 basis points if the loan is for a "Stated Wage Earner." This indicates that the lender increased the low-doc risk premium for borrowers likely to have a low value of reputation (W2), consistent with our empirical results. Not surprisingly, the rate sheet does not contain any pricing adjustments for "income inflation," "income exaggeration," "unbelievable income," or any other variant of those phrases, since the lender would not have wanted to publicize that it had officially accepted falsified applications and that it had charged a higher rate on the basis of its imperfect assumption of income falsification. Our empirical results, combined with the New Century Rate Sheet, suggest that the lender did price reputational risk explicitly, but we find no evidence that income exaggeration was priced at the loan level.[53] As we predict, the rate differentials did not completely eliminate the problems; adverse selection and income falsification did remain in equilibrium in the mortgage market.

VII. Policy Implications: Income Falsification, Borrower Location, and Subsequent House Price Declines

As we noted in the introduction, the role of borrower income misrepresentation in facilitating the expansion of mortgage credit is controversial. The extent that borrowers (or lenders/brokers operating on behalf of borrowers) systematically inflated incomes in order to obtain larger loans is consistent with the theory that the 2002-2006 housing boom resulted from an expansion in mortgage credit due to a decline in underwriting standards. In support of this theory, Mian and Sufi (2015) examine ZIP code level differences in income growth reported on mortgage applications and the growth in IRS reported income. Their analysis confirms that areas that experienced significant growth in subprime mortgage origination activity also saw higher levels of income overstatement. Furthermore, using micro-level mortgage data compiled by Piskorski et al. (2015), Mian and Sufi (2015) document that incidents of mortgage fraud were significantly more likely in areas that were identified as having higher

[53]We are careful not to generalize from the rate sheet to our entire sample period, since rate sheets were region specific and changed frequently. However, we note that the First Franklin rate sheet (Figure A.1) contained a similar premium for "NIV Wage Earner" indicating that New Century was not alone in pricing low-doc loans to W2 borrowers.

levels of borrower income misrepresentation.

We contribute to understanding the linkage between income overstatement and mortgage fraud by conducting an analysis of income falsification by borrower employment status at the ZIP code level using a research design similar to that employed by Mian and Sufi (2015). Specifically, we regress the percentage of each ZIP code's loans that are low-doc on the natural logarithm of ZIP code median income from the 2000 Census.[54] Table XIV reports the estimated coefficients where columns [1] and [2] are the W2 borrower sample and columns [3] and [4] are the self-employed sample. Focusing first on the W2 borrowers, when looking across MSAs (column [1] without MSA fixed effects) we see that higher income areas are correlated with higher proportions of loans to low-doc W2 borrowers. However, looking within MSAs (column [2] with MSA fixed effects), the sign on the estimated coefficient becomes negative suggesting that loans to low-doc W2 borrowers are concentrated in lower income ZIP codes. Together, the results in columns [1] and [2] suggest that low-doc loans to W2 borrowers are more prevalent in wealthier (higher income) MSAs, but the origination activity is occurring in the lower income areas of those MSAs. In contrast, for the self-employed borrowers (columns [3] and [4]), the negative relation between low-doc loans and lower income ZIP codes holds regardless of whether we are looking across or within MSAs. Thus, our results support the findings of Mian and Sufi (2015) that mortgages to borrowers most likely to overstate income (W2 borrowers originating low-doc loans) are concentrated in lower income neighborhoods.

Based on the evidence linking buyer income overstatement to specific areas, Mian and Sufi (2015) argue that this expansion in the supply of mortgage credit, including low-doc loans, put upward pressure on house prices. However, this interpretation is controversial as Adelino et al. (2015a) point out that the income distribution of mortgage purchase applicants may be different from the ZIP code income distribution reported in the IRS data. Rather than reflecting income overstatement on low-doc mortgage applications, Mian and Sufi's (2015) measure may simply reflect that home buyers have higher average incomes than the average income of all individuals within a ZIP code. In other words, the link between low-doc loans and house prices remains an empirical question.

We add to this debate by examining the relationship between low-doc market share by employment type and subsequent house price growth. We ask whether greater exposure to low-doc loans, especially low-doc loans to W2 borrowers, is negatively related to house price growth rates after the housing boom. Specifically, for each MSA i from 2004:Q1 to

[54] We estimate the regression for the W2 borrowers and self-employed borrowers separately. For the W2 borrower group, we select ZIP codes that have at least 9 total loans to W2 borrowers (the median number of W2 borrowers across all ZIP codes). For the self-employed sample, we select ZIP codes that have at least 4 total loans to self-employed borrowers (the median number of self-employed borrowers across all ZIP codes.)

2005:Q4, we measure the share of low-doc loan originations ($L_i \equiv \frac{Lowdoc_i}{All_i}$ where $Lowdoc_i$ and All_i denote the number of originated low-doc and all mortgages, respectively) and the proportion of low-doc W2 borrowers in the low-doc loan originations ($W_i \equiv \frac{W2\&Lowdoc_i}{Lowdoc_i}$ where $W2\&Lowdoc_i$ denotes the number of W2 low-doc loan originations). We compute the subsequent house price change starting from 2006:Q1, which corresponds to the time when a small number of MSAs started to exhibit price declines. We use three different periods of cumulative house price changes: 2006-2007, 2006-2008, and 2006-2009 and estimate the following MSA-level equation:

$$\begin{aligned}\Delta HPI_i = \alpha + \beta_1 (L_i) + \beta_2 (W_i) + \beta_3 \left(E_i^{-1}\right) + \beta_4 \left(L_i \times E_i^{-1}\right) + \beta_5 \left(W_i \times E_i^{-1}\right) \\ + \beta_6 (L_i \times W_i) + \beta_7 \left(L_i \times W_i \times E_i^{-1}\right) + \gamma M_i + \delta \left(M_i \times E_i^{-1}\right) + \varepsilon, \quad (14)\end{aligned}$$

where ΔHPI_i denotes cumulative house price change since 2006 in MSA i measured by the Federal Housing Finance Agency (FHFA) MSA level house price index, E_i^{-1} denotes the inverse elasticity of housing supply estimated by Saiz (2010), and M_i represents the variables that control for changes in housing demand in MSA i; i.e., house price growth between 2000 and 2005 and changes in population, per capita income, and unemployment rates since 2006. We include the interaction terms between the inverse of supply elasticity and other exogenous variables because the inverse of elasticities work as the conditioning variables in the reduced-form equilibrium price equation. We only control for demand factors because the main cause of the housing bust was likely due to housing demand shocks. We require each MSA to have at least 23 loans to be included in the sample, with 95% of the MSAs meeting this requirement. [55]

Table XV presents the marginal effect of the share of low-doc loans (L) and the proportion of low-doc W2 borrowers (W), which are evaluated at the mean values of the interacted variables:

$$\left.\frac{\partial \Delta HPI_i}{\partial L_i}\right|_{\overline{W_i}, \overline{E_i^{-1}}} = \beta_1 + \beta_4 \overline{E_i^{-1}} + \beta_6 \overline{W_i} + \beta_7 \left(\overline{W_i} \times \overline{E_i^{-1}}\right), \quad (15)$$

$$\left.\frac{\partial \Delta HPI_i}{\partial W_i}\right|_{\overline{L_i}, \overline{E_i^{-1}}} = \beta_2 + \beta_5 \overline{E_i^{-1}} + \beta_6 \overline{L_i} + \beta_7 \left(\overline{L_i} \times \overline{E_i^{-1}}\right), \quad (16)$$

where the upper bar indicates the sample mean. We also evaluate how these effects vary with supply inelasticity by computing partial derivatives of equations (15) and (16) with respect to E_i^{-1}. Column [1] shows that both the share of low-doc loans and the proportion

[55] Results are qualitatively similar when we use other cutoff values for the minimum number of loans to be included in the sample.

of low-doc W2 borrowers are negatively associated with house price growth from 2006 to 2007. The estimated coefficients are statistically significant at the 1% level. The results indicate that a 10 percentage point increase in the share of low-doc loans in 2004 and 2005 is associated with a 1.87% lower house price growth rate in 2006 and 2007. A 10 percentage point increase in the proportion of low-doc W2 borrowers in 2004 and 2005 is associated with a 1.72% lower growth rate. Moreover, these effects change with supply inelasticity. The change in the effect of low-doc share by supply inelasticity is -0.601 and statistically significant at the 1% level. Thus, for an MSA that exhibits a one standard deviation larger value of the inverse elasticity of supply, a 10 percentage point increase in low-doc share is associated with a 3.75% lower house price growth rate. The change in the effect of low-doc W2 borrowers' proportion by supply inelasticity (-0.355) is marginally significant at the 11% level. This coefficients suggests that a 10 percentage point increase in low-doc W2 borrowers' proportion is associated with a 2.83% lower house growth rate in an MSA with a one standard deviation larger value of inelasticity.[56] Columns [2] and[3] indicate that the relation between the low-doc share and the subsequent house price changes becomes smaller and weaker as the recession grows in severity in 2008 and 2009. However, the relation between the low-doc W2 borrowers' proportion and house price changes remains strong and statistically significant until 2009. Although we are careful not to claim a strong causal interpretation, this result suggests that an exposure to low-doc loans, especially to W2 low-doc loans, at the peak of housing boom is closely related with the beginning of housing bust.

VIII. Conclusion

Financial economists have only recently begun to examine the role of mortgage fraud and adverse selection as contributing factors to the Great Recession of 2007-2009. For example, growing evidence suggests that misrepresentations of borrower income, borrower assets, inflated appraisals, and second liens and owner-occupancy status increased significantly during the period prior to the financial crisis. In contributing to this literature, we document how borrower heterogeneity with respect to employment status contributed to income misrepresentation and adverse selection.

Using a national dataset of subprime mortgages originated by a major financial institution during the house price boom period, we provide several novel insights into the role of borrower income misrepresentation. First, we document that income misrepresentation was concentrated among borrowers who originated low-documentation loans but could have orig-

[56] The estimated values of 3.75% and 2.83% are calculated as $10 \times (-0.187 - 0.601 \times 0.312)$ and $10 \times (-0.172 - 0.355 \times 0.312)$, respectively, where 0.312 is the standard deviation of inverse elasticity.

inated full-documentation loans instead. Second, we show that the majority of additional risk associated with low-doc mortgages was due to adverse selection and income falsification on the part of borrowers with verifiable income. We also provide evidence that these borrowers were more likely to inflate or exaggerate their income on the mortgage application. As a result, we provide new evidence showing that income misrepresentation resulted from borrower actions, which is consistent with excesses in the mortgage market arising from both borrowers and lenders. Third, we document lender actions designed to counter potential borrower income falsification. Finally, we discuss how borrower income falsification may have facilitated the expansion in mortgage credit and thus, we provide new insights into one of the possible causes of the Great Recession.

Taken together, our empirical analysis suggests a more nuanced market where borrower concerns about future credit access can mitigate the effects of adverse selection in limited information documentation mortgage contracts. From a policy perspective, our results indicate that a blanket regulation mandating "qualified" mortgages (i.e. loans that require full documentation) may be overly restrictive and lead to credit rationing for a subset of the population that faces high information verification costs. In fact, we point out that such regulations may have serious unintended consequences for the economy. Rather, our analysis suggests that regulators seeking to limit the potential of a future foreclosure crisis should rely on a more nuanced or targeted regulatory approach that limits the use of low information documentation loans by borrowers who have *ex ante* low information verification costs.

References

Adelino, Manuel, Antoinette Schoar, and Felipe Severino, 2015a, Changes in Buyer Composition and the Expansion of Credit During the Boom, *NBER Working Paper 20848* .

Adelino, Manuel, Antoinette Schoar, and Felipe Severino, 2015b, House prices, collateral, and self-employment, *Journal of Financial Economics* 117, 288 – 306.

Agarwal, Sumit, Brent W. Ambrose, and Vincent W. Yao, 2014, The Limits of Regulation: Appraisal Bias in the Mortgage Market, *Working paper,* Available at SSRN: http://ssrn.com/abstract=2487393 or http://dx.doi.org/10.2139/ssrn.2487393.

Agarwal, Sumit, Itzhak Ben-David, and Vincent Yao, Forthcoming, Collateral valuation and borrower financial constraints: Evidence from the residential real estate market, *Management Science* 0, null.

Agarwal, Sumit, Yan Chang, and Abdullah Yavas, 2012, Adverse Selection in Mortgage Securitization, *Journal of Financial Economics* 105, 640–660.

Agarwal, Sumit, and Douglas D. Evanoff, 2013, Loan product steering in mortgage market, *Working paper,* Available at SSRN: http://ssrn.com/abstract=2204400 or http://dx.doi.org/10.2139/ssrn.2204400.

Ai, Chunrong, and Edward C. Norton, 2003, Interaction Terms in Logit and Probit Models, *Economics Letters* 80, 123 – 129.

Ambrose, Brent W., and James N. Conklin, 2014, Mortgage Brokers, Origination Fees, Price Transparency and Competition, *Real Estate Economics* 42, 363–421.

Ben-David, Itzhak, 2011, Financial Constraints and Inflated Home Prices During the Real Estate Boom, *American Economic Journal: Applied Economics* 3, 55–87.

Buis, Maarten L, 2010, Stata tip 87: Interpretation of Interactions in Non-linear Models, *The Stata Journal* 10, 305–308.

Burke, Jeremy M., Curtis R. Taylor, and Liad Wagman, 2012, Information Acquisition in Competitive Markets: An Application to the US Mortgage Market, *American Economic Journal - Microeconomics* 4, 65–106.

Chomsisengphet, Souphala, and Anthony Pennington-Cross, 2006, The Evolution of the Subprime Mortgage Market, *Federal Reserve Bank of St. Louis Review* 88, 31–56.

Conklin, James N., Forthcoming, Financial literacy, broker-borrower interaction, and mortgage default, *Real Estate Economics* .

Demyanyk, Yuliya, and Otto Van Hemert, 2011, Understanding the Subprime Mortgage Crisis, *Review of Financial Studies* 24, 1848–1880.

Diamond, Douglas W., 1989, Reputation Acquisition in Debt Markets, *Journal of Political Economy* 97, 828–862.

Diamond, Douglas W., 1991, Monitoring and Reputation - The Choice Between Bank Loans and Directly Placed Debt, *Journal of Political Economy* 99, 689–721.

Garmaise, Mark J., 2015, Borrower Misreporting and Loan Performance, *Journal of Finance* Forthcoming.

Green, Richard K., 2014, *Introduction to Mortgages and Mortgage Backed Securities* (Academic Press, Waltham, MA).

Griffin, John M., and Gonzalo Maturana, 2015, Who Facilitated Misreporting in Securitized Loans?, *Journal of Finance* Forthcoming.

Jiang, Wei, Ashlyn Aiko Nelson, and Edward Vytlacil, 2014a, Liar's loan? effects of origination channel and information falsification on mortgage delinquency, *Review of Economics and Statistics* 96, 1–18.

Jiang, Wei, Ashlyn Aiko Nelson, and Edward Vytlacil, 2014b, Liar's Loan? Effects of Origination Channel and Information Falsification on Mortgage Delinquency, *Review of Economics and Statistics* 96, 1–18.

Keys, Benjamin J., Tanmoy Mukherjee, Amit Seru, and Vikrant Vig, 2009, Financial Regulation and Securitization: Evidence from Subprime Loans, *Journal of Monetary Economics* 56, 700 – 720, Carnegie-Rochester Conference Series on Public Policy: Distress in Credit Markets: Theory, Empirics, and Policy, November 14-15, 2008.

LaCour-Little, Michael, and Jing Yang, 2013, Taking the Lie Out of Liar Loans: The Effect of Reduced Documentation on the Performance and Pricing of Alt-A and Subprime Mortgages, *Journal of Real Estate Research* 35, 507–553.

Mian, Atif, and Amir Sufi, 2009, The Consequences of Mortgage Credit Expansion: Evidence from the U.S. Mortgage Default Crisis, *The Quarterly Journal of Economics* 124, 1449–1496.

Mian, Atif, and Amir Sufi, 2015, Fraudulent Income Overstatement on Mortgage Applications During the Credit Expansion of 2002 to 2005, *NBER Working Paper 20947* .

Pahl, Cynthia, 2007, A Compilation of State Mortgage Broker Laws and Regulations 1996-2006, *Federal Reserve Bank of Minneapolis, Community Affairs Report #2007-2* .

Paley, Irina, and Konstantinos Tzioumis, 2011, Rethinking stated-income loans: Separating the wheat from the chaff, *Available at SSRN 1773824* .

Piskorski, Tomasz, Amit Seru, and Vikrant Vig, 2010, Securitization and distressed loan renegotiation: Evidence from the subprime mortgage crisis, *Journal of Financial Economics* 97, 369 – 397, The 2007-8 financial crisis: Lessons from corporate finance.

Piskorski, Tomasz, Amit Seru, and James Witkin, 2015, Asset Quality Misrepresentation by Financial Intermediaries: Evidence from the RMBS Market, *Journal of Finance* Forthcoming.

Rajan, Uday, Amit Seru, and Vikrant Vig, 2010, Statistical default models and incentives, *American Economic Review (Papers and Proceedings),* 100, 506–210.

Saiz, Albert, 2010, The geographic determinants of housing supply, *The Quarterly Journal of Economics* 125, 1253–1296.

U.S., Financial Crisis Inquiry Commission, 2011, *The Financial Crisis Inquiry Report: Final Rport of the National Commission on the Causes of the Financial and Economic Crisis in the United States (revised corrected copy)* (Government Printing Office).

Williams, Richard, 2012, Using the Margins Command to Estimate and Interpret Adjusted Predictions and Marginal Effects, *Stata Journal* 12, 308.

Table I: Summary of Model Predictions for Mortgage Type Preference, Income Exaggeration, Default Rate, and Interest Rate Premium

	Information Verification Cost	
Mortgage Type	Low (W2)	High (Self-Employed)
High Information (*Full-Doc*)	Preferred	–
	No Income Exaggeration	No Income Exaggeration
	Low Default Rate	Low Default Rate
	No Rate Premium	No Rate Premium
Low Information (*Low-Doc*)	–	Preferred
	Large Income Exaggeration	Moderate Income Exaggeration
	High Default Rate	Moderate Default Rate
	Large Rate Premium	Moderate Rate Premium

Table II: Summary Statistics for Loans by Employment and Documentation Type

	[1] W2 Full-Doc		[2] W2 Low-Doc		[3] Self-employed Full-Doc		[4] Self-employed Low-Doc	
	Mean	**Std. Dev.**	**Mean**	**Std. Dev**	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev**
Default	0.0478	0.2133	0.0525	0.2230	0.0512	0.2203	0.0505	0.2189
Loan Characteristics								
The difference between the rate on the mortgage and the two year Treasury rate (*RATE_SPREAD*)	4.6351	1.5288	4.9516	1.4065	4.5239	1.4989	4.7337	1.4566
Combined loan to value ratio at origination (*CLTV*)	82.9510	14.2817	83.5768	14.0507	83.6182	13.4387	83.6585	14.9095
Loan amount at origination (*LOAN_AMOUNT*)	174,019	107,689	207,296	115,962	224,821	138,936	231,555	134,122
Fees as a percentage of loan amount (*FEES*)	0.0410	0.0197	0.0375	0.0180	0.0358	0.0186	0.0359	0.0182
An indicator set to one if the loan was an adjustable rate mortgage (*ARM*)	0.7234	0.4473	0.8311	0.3746	0.7316	0.4431	0.8285	0.3769
An indicator set to one if the loan was a prepayment penalty on the loan (*PREPAY*)	0.7900	0.4073	0.7528	0.4314	0.8125	0.3903	0.7484	0.4340
An indicator variable set to one if the loan was for a home purchase (*PURCH*)	0.2802	0.4491	0.4186	0.4933	0.3203	0.4666	0.4327	0.4954
An indicator variable set to one if the loan was a cash-out refinance (*CASH*)	0.5974	0.4904	0.5063	0.5000	0.5669	0.4955	0.4986	0.5000
The FICO score of the primary borrower at origination (*FICO*)	600.5364	59.7866	628.0193	61.7492	613.4986	59.2660	632.4919	63.2627
An indicator set to one if the loan had interest only payments (*IO*)	0.1212	0.3263	0.1470	0.3541	0.1431	0.3501	0.2031	0.4023
Months since origination (*MONTHS*)	6.7830	7.0646	6.7887	6.7864	7.1651	7.4309	6.9174	6.8689
Property Characteristics								
An indicator set to one if the property was an investment property (*INVESTMENT*)	0.0575	0.2329	0.0849	0.2787	0.1263	0.3322	0.1227	0.3281
An indicator set to one if the property was a two-unit property (*TWO_UNIT*)	0.0587	0.2350	0.1021	0.3028	0.0705	0.2560	0.0823	0.2748
An indicator set to one if the property was a condo (*CONDO*)	0.0617	0.2406	0.0723	0.2589	0.0606	0.2387	0.0694	0.2541
Borrower Characteristics								
The age of the primary borrower (*AGE*)	43.1441	11.6602	40.6433	10.8312	43.9410	11.1064	42.9128	11.0918
An indicator set to one if the borrower was a minority (*MINORITY*)	0.4126	0.4923	0.4187	0.4934	0.3248	0.4683	0.3405	0.4739
The total monthly income of the borrowers (*INCOME*)	5,396	3,014	6,398	3,199	8,465	5,245	8,251	4,403
Face-to-face interview between broker and borrower (*FACE*)	0.3947	0.4888	0.4545	0.4979	0.4214	0.4938	0.4448	0.4970
Interest Rate Environment								
The average monthly prime 30-year fixed rate at the time of origination (*RATE_30*)	6.0803	0.5642	6.0194	0.4903	6.1870	0.6614	6.0332	0.5119
Area Characteristics								
Monthly unemployment rate at the MSA level (*UNEMP*)	5.3590	1.5017	5.3759	1.4690	5.3100	1.5263	5.2813	1.5256
MSA level Herfindahl-Hirschman index for broker competition (*HHI*)	0.0892	0.1281	0.0682	0.1084	0.0975	0.1422	0.0753	0.1203
Pahl-Index for state level broker regulations (*REG*)	7.6612	3.5965	8.0793	3.6356	8.1008	3.4888	8.2463	3.5411
MSA house price growth over previous two years (*HPI_2YR*)	0.2404	0.1619	0.2789	0.1665	0.2522	0.1676	0.2886	0.1774
MSA house price growth since origination (*HPI_GROWTH*)	0.0685	0.0957	0.0753	0.0945	0.0723	0.0982	0.0749	0.0938
MIDWEST	0.2088	0.4064	0.1588	0.3655	0.1406	0.3477	0.1276	0.3336
SOUTH	0.2543	0.4355	0.2198	0.4141	0.2280	0.4195	0.2330	0.4227
NORTHEAST	0.1219	0.3272	0.1674	0.3733	0.0941	0.2920	0.1197	0.3246
PACIFIC	0.0079	0.0887	0.0073	0.0851	0.0122	0.1096	0.0099	0.0990
Origination Year								
1999	0.0216	0.1455	0.0137	0.1163	0.0363	0.1870	0.0216	0.1452
2000	0.0273	0.1630	0.0173	0.1304	0.0468	0.2112	0.0193	0.1374
2001	0.0464	0.2104	0.0366	0.1878	0.0663	0.2489	0.0344	0.1823
2002	0.0972	0.2962	0.0836	0.2768	0.1004	0.3005	0.0702	0.2555
2003	0.2238	0.4168	0.1924	0.3942	0.1973	0.3980	0.1923	0.3941
2004	0.2803	0.4491	0.3537	0.4781	0.2353	0.4242	0.2695	0.4437
2005	0.2921	0.4547	0.2923	0.4548	0.2956	0.4563	0.3815	0.4858
N	256,747		107,621		19,823		74,681	

Note: This table presents summary statistics for the funded loans from the New Century database.

Table III: Relationship Between Low-doc, Employment Type, and Mortgage Performance

	[1]	[2]	[3]	[4]
	Baseline Model		**Propensity Score Matching**	**Controlling For Borrower's Liquid Assets**
Dependent Variable: Default	**M.E.**	**M.E.**	**M.E.**	**M.E.**
Marginal Effects of Low-Doc				
Self-Employed (*Lowdoc*)	0.0053***	0.0054***	0.0043**	0.0062
	(0.0019)	(0.0019)	(0.0020)	(0.0060)
W2 (*Lowdoc*)	0.0124***	0.0119***	0.0111***	0.0090***
	(0.0013)	(0.0012)	(0.0012)	(0.0022)
Loan Characteristics	Yes	Yes	Yes	Yes
Property Characteristics	Yes	Yes	Yes	Yes
Borrower Characteristics	Yes	Yes	Yes	Yes
Interest rate environment	Yes	Yes	Yes	Yes
Area Characteristics	Yes	Yes	Yes	Yes
Origination Year Fixed Effects	Yes	Yes	Yes	Yes
MSA Fixed Effects	No	Yes	Yes	Yes
N	455,546	455,058	269,662	50,231
Mean Default Rate (Full-doc Self-Employed)	0.0512	0.0512	0.0505	0.0426
Mean Default Rate (Full-doc W2)	0.0478	0.0478	0.0412	0.0422
Log Likelihood	-77,364	-76,716	-45,197	-7,853

Note: This table presents the marginal effects of low-doc loan status on default by employment type. Columns [1] and [2] are the baseline model without and with MSA fixed effects, respectively. Column [3] presents the marginal effects of low-doc loan status on default by employment type after matching low-doc loans with full-doc loans based on the propensity score. The propensity score is derived using loan, property, borrower, and area characteristics as well as MSA and origination year. Column [4] presents marginal effects of low-doc loan status on default by employment type for the subsample of loans where the borrower's liquid assets field is not missing in the New Century data. The marginal effects are derived from a probit model of mortgage performance on income documentation, employment type, an interaction term between income documentation and employment type, loan characteristics, borrower characteristics, property characteristics, and area characteristics for the funded loans from the New Century database. Heteroskedasticity-robust standard errors adjusted for clustering at the MSA level are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Table IV: Descriptive Statistics for Income Exaggeration Measure by Employment and Documentation Type

	[1] W2 Income Exaggeration				[2] Self-employed Income Exaggeration				[3] Mean Difference	
	N	**Mean**	**Std.Err.**	**t-stat**	**N**	**Mean**	**Std. Err.**	**t-stat**	**[1] - [2]**	**t-stat**
Doc Type										
Full-Doc	253,114	0.0004	(0.0007)	0.59	19,509	0.0009	(0.0031)	0.27	-0.0005	-0.17
Low-Doc	106,351	0.0799	(0.0010)	83.51	73,830	0.0106	(0.0013)	7.78	0.0693	43.01

Note: This table presents descriptive statistics for the income exaggeration measure. Column 1 presents the mean and standard deviation of income exaggeration for W2 borrowers split according to income documentation. The null hypothesis for the test statistic is that income exaggeration is equal to zero. Column 2 presents the mean and standard deviation of income exaggeration for self-employed borrowers split according to income documentation. The null hypothesis for the test statistic is that income exaggeration is equal to zero. Column 3 presents the mean difference across columns 1 and 2. The null hypothesis for the test statistic is that the mean difference is zero. The t-stat in column 3 assumes unequal variances across samples.

Table V: Explaining Income Exaggeration

	[1]	
	Income Exaggeration	
Dependent Variable: Income Exaggeration	**Coeff.**	**Std. Err.**
Employment Type		
W2	0.0002	(0.0012)
Income Documentation		
Low-Doc	0.0046	(0.0108)
Low-Doc × W2	0.0662***	(0.0094)
Loan Characteristics	Yes	
Property Characteristics	Yes	
Borrower Characteristics	Yes	
Interest rate environment	Yes	
Area Characteristics	Yes	
Origination Year Fixed Effects	Yes	
MSA Fixed Effects	Yes	
Constant	-0.2293***	(0.0195)
N	449,916	
Adj. R^2	0.03	

Note: This table presents coefficient estimates from an OLS regression of INC_EXAG on employment type, type of income documentation, loan characteristics, borrower characteristics, property characteristics, and area characteristics for the funded loans for the funded loans from the New Century Database. Heteroskedasticity-robust standard errors adjusted for clustering at the MSA level are reported in parentheses.. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Table VI: Average Income by Documentation Status for the Most Frequently Used Borrower Business Types (W2)

Borrower Business Type		[1] Low-Doc Income			[2] Full-Doc Income		[3] Difference		[4] Job-Specific Overstatment (%)
	N	Mean	Std. Dev.	N	Mean	Std. Dev.	[1] - [2]	t-stat	
MANAGER	1,855	$6,720	($2,831)	1,794	$5,563	($2,946)	$1,157	12.10	17%
SUPERVISOR	706	$6,378	($2,280)	945	$4,879	($2,200)	$1,499	13.49	24%
TEACHER	389	$6,162	($2,552)	4704	$4,704	($1,655)	$1,458	15.89	24%
DRIVER	431	$5,582	($2,165)	896	$4,399	($1,919)	$1,183	10.08	21%
SALES	413	$6,491	($3,448)	624	$5,590	($3,196)	$901	4.31	14%
TRUCK DRIVER	317	$5,996	$2,141	575	$4,582	($2,085)	$1,414	9.60	24%
OFFICE MANAGER	459	$5,830	($2,270)	410	$4,704	($2,707)	$1,126	6.67	19%
CONSTRUCTION	344	$5,743	($2,584)	355	$4,653	($2,404)	$1,090	5.78	19%
GENERAL MANAGER	279	$8,478	($3,793)	386	$7,788	($3,871)	$690	2.29	8%
SALES MANAGER	259	$8,326	($3,825)	390	$7,787	($4,191)	$539	1.66	6%
NURSE	194	$6,685	($2,724)	444	$5,236	($2,475)	$1,449	6.59	22%
FOREMAN	252	$6,777	($2,509)	378	$5,226	($2,612)	$1,551	7.42	23%
MECHANIC	184	$5,219	($1,918)	400	$4,375	($1,814)	$844	5.13	16%
REGISTERED NURSE	95	$8,993	($3,387)	403	$6,812	($2,722)	$2,181	6.69	24%
MACHINE OPERATOR	220	$5,200	($1,651)	269	$3,685	($1,513)	$1,515	10.57	29%
EDUCATION	143	$5,540	($2,483)	314	$4,579	($1,915)	$961	4.52	17%
LABORER	143	$4,261	($1,704)	296	$3,334	($1,430)	$927	5.97	22%
MEDICAL	120	$5,894	($3,505)	297	$5,224	$3,527	$670	1.76	11%
ELECTRICIAN	133	$6,303	($2,347)	283	$4,949	($2,173)	$1,354	5.78	21%
CUSTOMER SERVICE	142	$4,345	($1,788)	268	$3,377	($1,293)	$968	6.29	22%
RN	75	$7,872	($2,830)	311	$5,908	($2,287)	$1,964	6.36	25%
MAINTENANCE	158	$4,693	($1,637)	221	$3,425	($1,501)	$1,268	7.81	27%
CARPENTER	169	$5,837	($2,359)	205	$4,329	($1,921)	$1,508	6.81	26%
POLICE OFFICER	62	$7,581	($2,946)	310	$6,157	($2,210)	$1,424	4.36	19%
ENGINEER	83	$9,426	($3,735)	289	$6,511	($2,667)	$2,915	7.97	31%

Note: This table presents summary statistics by income documentation status for the top 25 W2 borrower business types in the subsample of loans where business type was not missing. Columns [1] and [2] report the average income for each business type for low-doc and full-doc loans, respectively. Column [3] includes the mean difference and a t-statistic from a mean difference test assuming unequal variance across the two groups. Column [4] lists *Job-Specific Overstatement* which is the difference in Column[3] divided by the mean income for full-doc in column [2].

Table VII: Average Income by Documentation Status for the Most Frequently Used Borrower Business Types (Self-Employed)

Borrower Business Type	N	[1] Low-Doc Income Mean	Std. Dev.	N	[2] Full-Doc Income Mean	Std. Dev.	[3] Difference [1] - [2]	t-stat	[4] Job-Specific Overstatement (%)
OWNER	9,447	$8,540	($4,400)	1,988	$9,431	($5,389)	-$891	-7.87	-9%
PRESIDENT	739	$10,989	($5,355)	181	$11,133	($6,438)	-$144	-0.31	-1%
REALTOR	546	$10,149	$5,087	141	$11,408	($6,090)	-$1,259	-2.51	-11%
CONSTRUCTION	360	$7,829	($3,955)	127	$8,253	($5,278)	-$424	-0.95	-5%
OWNER/OPERATOR	395	$7,965	($4,104)	92	$8,137	($4,401)	-$172	-0.36	-2%
TRUCK DRIVER	236	$6,870	($2,844)	69	$6,547	($3,308)	$323	0.80	5%
REAL ESTATE AGENT	223	$10,809	($5,431)	57	$10,504	($5,255)	$305	0.38	3%
MANAGER	224	$7,412	($3,508)	45	$7,863	($4,227)	-$451	-0.76	-6%
SALES	220	$7,387	($3,584)	42	$9,170	($6,812)	-$1,783	-2.49	-19%
REAL ESTATE	167	$9,600	($5,094)	67	$8,770	($5,326)	$830	1.11	9%
CONTRACTOR	182	$7,126	$3,109	44	$7,997	($4,334)	-$871	-1.53	-11%
CONSULTANT	156	$8,944	($4,408)	33	$8,059	($3,597)	$885	1.08	11%
LANDSCAPING	153	$6,514	($2,593)	29	$7,362	($4,794)	-$848	-1.38	-12%
OWNER/MANAGER	123	$9,640	($4,144)	27	$10,587	($5,748)	-$947	-1.00	-9%
OWNER OPERATOR	107	$8,272	($3,976)	36	$8,573	($6,232)	-$301	-0.34	-4%
DRIVER	125	$6,369	($2,784)	17	$4,547	($2,143)	$1,822	2.59	40%
LOAN OFFICER	106	$10,038	($4,155)	33	$12,521	($5,944)	-$2,483	-2.69	-20%
ATTORNEY	92	$12,635	($5,176)	41	$11,039	($5,411)	$1,596	1.62	14%
HAIR STYLIST	106	$5,843	($2,560)	22	$4,544	($1,973)	$1,299	2.24	29%
HANDYMAN	119	$4,696	($1,794)	7	$7,916	($3,387)	-$3,220	-4.35	-41%
VICE PRESIDENT	104	$10,620	($4,713)	14	$9,356	($4,395)	$1,264	0.95	14%
CARPENTER	97	$6,259	($2,605)	14	$7,551	($5,077)	-$1,292	-1.50	-17%
SALES MANAGER	97	$9,077	($3,562)	12	$10,273	($6,138)	-$1,196	-1.00	-12%
PAINTER	84	$5,909	($2,611)	11	$7,300	($1,961)	-$1,391	-1.70	-19%
OWNER/PRESIDENT	83	$11,363	($5,386)	12	$13,078	($7,818)	-$1,715	-0.97	-13%

Note: This table presents summary statistics by income documentation status for the top 25 self-employed borrower business types in the subsample of loans where business type was not missing. Columns [1] and [2] report the average income for each business type for low-doc and full-doc loans, respectively. Column [3] includes the mean difference and a t-statistic from a mean difference test assuming unequal variance across the two groups. Column [4] lists *Job-Specific Overstatement* which is the difference in Column[3] divided by the mean income for full-doc in column [2].

Table VIII: Job-Specific Overstatement Summary Statistics by Employment Type

Borrower Business Types that have at least 10 full- and 10 low-doc observations	[1] W2	[2] Self-employed
Average Job-Specific Overstatement	28.10%	-0.42%
% of Borrower Business Types with Job-Specific Overstatement > 0	90.10%	41.82%
% of Borrower Business Types with Job-Specific Overstatement Significantly > 0	72.52%	16.36%
N	313	55

Note: This table presents summary statistics by employment type for *Job-Specific Overstatement*. Column [1] includes borrower business types that had at least 10 full-doc/W2 and 10 low-doc/W2 observations. Column [2] includes borrower business types that had at least 10 full-doc/self-employed and 10 low-doc/self-employed observations. *% of Borrower Business Types with Job-Specific Overstatement* > 0 is significantly different from 50% at the 1% level of confidence for W2 borrower business types, but is not significantly different from 50% for self-employed business types. We perform a one tail mean difference test with the null hypothesis that the low-doc average income is less than or equal to the full-doc income for each of the borrower business types. *% of Borrower Business Types with Job-Specific Overstatement Significantly* > 0 reports the percentage of borrower business types for which we were able to reject the null hypothesis.

Table IX: Borrower business types with the largest Job-Specific Overstatement (%)

	[1] W2			[2] Self-Employed	
Ranking	Business Type	Job-Specific Overstatement (%)	Ranking	Business Type	Job-Specific Overstatement (%)
1	PERSONAL BANKER	120%	1	CLEANING	47%
2	PHLEBOTOMIST	110%	2	DRIVER	40%
3	FORKLIFT OPERATOR	97%	3	HAIR SALON	36%
4	KITCHEN MANAGER	85%	4	DAY CARE	30%
5	ASST MGR	82%	5	HAIR STYLIST	29%
6	DENTAL ASSISTANT	77%	6	INSTALLER	28%
7	LETTER CARRIER	71%	7	DENTIST	24%
8	NURSING	70%	8	LAW	23%
9	BUS OPERATOR	69%	9	CHILD CARE	18%
10	MEDICAL BILLER	69%	10	ATTORNEY	14%
11	PROGRAM MANAGER	66%	11	CONSULTING	14%
12	SENIOR ACCOUNTANT	64%	12	VICE PRESIDENT	14%
13	DEALER	62%	13	CHILD CARE PROVIDER	12%
14	FINISHER	62%	14	CONSULTANT	11%
15	MEDICAL ASSISTANT	62%	15	REAL ESTATE	9%
16	CAN	61%	16	CEO	8%
17	CSR	61%	17	SUB CONTRACTOR	7%
18	CARRIER	61%	18	OPERATOR	6%
19	OFFICER	60%	19	TRUCK DRIVER	5%
20	PLUMBING	60%	20	REAL ESTATE SALES	5%
21	FORMAN	59%	21	REAL ESTATE AGENT	3%
22	COURIER	59%	22	INSUARNCE	1%
23	CONSTRUCTION SUPERVISOR	58%	23	TRUCKING	0%
24	WELDER	58%	24	OWNER.	-1%
25	BUYER	58%	25	PRESIDENT	-1%

Note: This table includes the 25 borrower business types with the largest values for *Job-Specific Overstatement* by employment type. Column [1] includes borrower business types that had at least 10 full-doc/W2 and 10 low-doc/W2 observations. Column [2] includes borrower business types that had at least 10 full-doc/self-employed and 10 low-doc/self-employed observations.

Table X: Explaining Income *within* Job Title

	[1]		[2]		[3]		[4]	
Dependent Variable:	**Log Income**		**Log Income**		**Log Income**		**Log Income**	
Log Borrower Monthly Income	**Coeff.**	**Std. Err.**	**Coeff.**	**Std. Err.**	**Coeff.**	**Std. Err.**	**Coeff.**	**Std. Err.**
Employment Type								
W2	-0.5349***	(0.0080)	-0.2796**	(0.1193)	-0.2311**	(0.1103)	-0.0978	(0.0878)
Income Documentation								
Low-Doc	-0.0428***	(0.0082)	-0.0162**	(0.0075)	-0.0343***	(0.0070)	-0.0170***	(0.0056)
W2 x Low-Doc	0.2815***	(0.0093)	0.2721***	(0.0086)	0.2275***	(0.0079)	0.1494***	(0.0063)
Job Title Fixed Effects	No		Yes		Yes		Yes	
W2 x Job Title Fixed Effects	No		Yes		Yes		Yes	
MSA Fixed Effects	No		No		Yes		Yes	
Origination Year Fixed Effects	No		No		Yes		Yes	
Borrower Characteristics	No		No		No		Yes	
Area Characteristics	No		No		No		Yes	
Constant	8.9432***	(0.0075)	8.6742***	(0.1187)	8.6440***	(0.1136)	-0.5838***	(0.1247)
N	86,717		86,717		86,717		86,717	
Adj. R^2	0.15		0.34		0.44		0.65	

Note: This table presents coefficient estimates from an OLS regression of log borrower income controlling for the borrower's job title for the subsample of New Century loans where borrower business type was not missing. Standard errors are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Table XI: Relationship Between Low-doc, Employment Type, and Application Rejection

	[1]	
	Rejected	
Dependent Variable: Application Rejected	**M.E.**	**Std. Err.**
Marginal Effects of Stated Income Documentation		
Self-Employed (Low-Doc)	-0.0083***	(0.0028)
W2 (Low-Doc)	0.0093***	(0.0018)
Loan Characteristics	Yes	
Property Characteristics	Yes	
Borrower Characteristics	Yes	
Interest rate environment	Yes	
Area Characteristics	Yes	
Origination Year Fixed Effects	Yes	
MSA Fixed Effects	Yes	
N	697,009	
Mean Rejection Rate (Full-doc Self-Employed)	0.1840	
Mean Rejection Rate (Full-doc W2)	0.1407	
Log Likelihood	-254,462	

Note: This table presents marginal effects of no-doc on application rejection by employment type. The marginal effects are derived from a probit model of application rejection on income documentation, employment type, an interaction term between income documentation and employment type, loan characteristics, borrower characteristics, property characteristics, and area characteristics at the time of the accept/reject decision for the loan applications from the New Century database. Heteroskedasticity-robust standard errors adjusted for clustering at the MSA level are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Table XII: Relationship Between Low-doc, Employment Type, and Application Rejection (OLS)

	[1]		[2]	
	Rejected		**Rejected**	
Dependent Variable: Rejected	**Coeff.**	**Std. Err.**	**Coeff.**	**Std. Err.**
W2	-0.0209***	(0.0027)	-0.0243***	(0.0024)
Low-Doc	-0.0096***	(0.0030)	-0.0106***	(0.0022)
Low-Doc x W2	0.0176***	(0.0027)	0.0177***	(0.0024)
INC_EXAG			0.0028	(0.0046)
INC_EXAG x W2			-0.0089*	(0.0049)
Low-Doc x INC_EXAG			0.0009	(0.0054)
Low-Doc x W2 x INC_EXAG			0.0162***	(0.0061)
Loan Characteristics	Yes		Yes	
Property Characteristics	Yes		Yes	
Borrower Characteristics	Yes		Yes	
Interest rate environment	Yes		Yes	
Area Characteristics	Yes		Yes	
Origination Year	Yes		Yes	
MSA Fixed Effects	Yes		Yes	
N	697,020		687,199	
Adj. R^2	0.11		0.10	

Note: This table presents marginal coefficient estimates from a linear probability model of application rejection on income documentation, employment type, an interaction term between income documentation and employment type, loan characteristics, borrower characteristics, property characteristics, and area characteristics at the time of the accept/reject decision for the loan applications from the New Century database. Heteroskedasticity-robust standard errors adjusted for clustering at the MSA level are reported in parentheses in Column [1]. Bootstrapped standard errors are reported in parentheses in Column [2]. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Table XIII: Relationship Between Low-doc, Employment Type, and Mortgage Pricing

	[1]		[2]	
	Rate Spread		**Rate Spread**	
Dependent Variable: Rate Spread	**Coeff.**	**Std. Err.**	**Coeff.**	**Std. Err.**
W2	-0.0981***	(0.0083)	-0.0717***	(0.0077)
Low-Doc	0.5305***	(0.0173)	0.5308***	(0.0073)
Low-Doc x W2	0.1508***	(0.0113)	0.1509***	(0.0079)
INC_EXAG			0.0852***	(0.0172)
INC_EXAG x W2			0.0614***	(0.0149)
Low-Doc x INC_EXAG			0.0235	(0.0167)
Low-Doc x W2 x INC_EXAG			0.0291	(0.0201)
Loan Characteristics	Yes		Yes	
Property Characteristics	Yes		Yes	
Borrower Characteristics	Yes		Yes	
Interest rate environment	Yes		Yes	
Area Characteristics	Yes		Yes	
Origination Year	Yes		Yes	
MSA Fixed Effects	Yes		Yes	
N	455,687		449,917	
Adj. R^2	0.69		0.69	

Note: This table presents coefficient estimates from an OLS regression of mortgage pricing on type of income documentation, loan characteristics, borrower characteristics, property characteristics, and area characteristics for the funded loans from the New Century database. The average *RATE_SPREAD* for the reference group (Full-Doc/Self-employed) is 4.5672 holding all other control variables at average values. Heteroskedasticity-robust standard errors adjusted for clustering at the MSA level are reported in parentheses in Column [1]. Bootstrapped standard errors are reported in parentheses in Column [2]. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Table XIV: Low-Doc Penetration and Zip Code Median Income

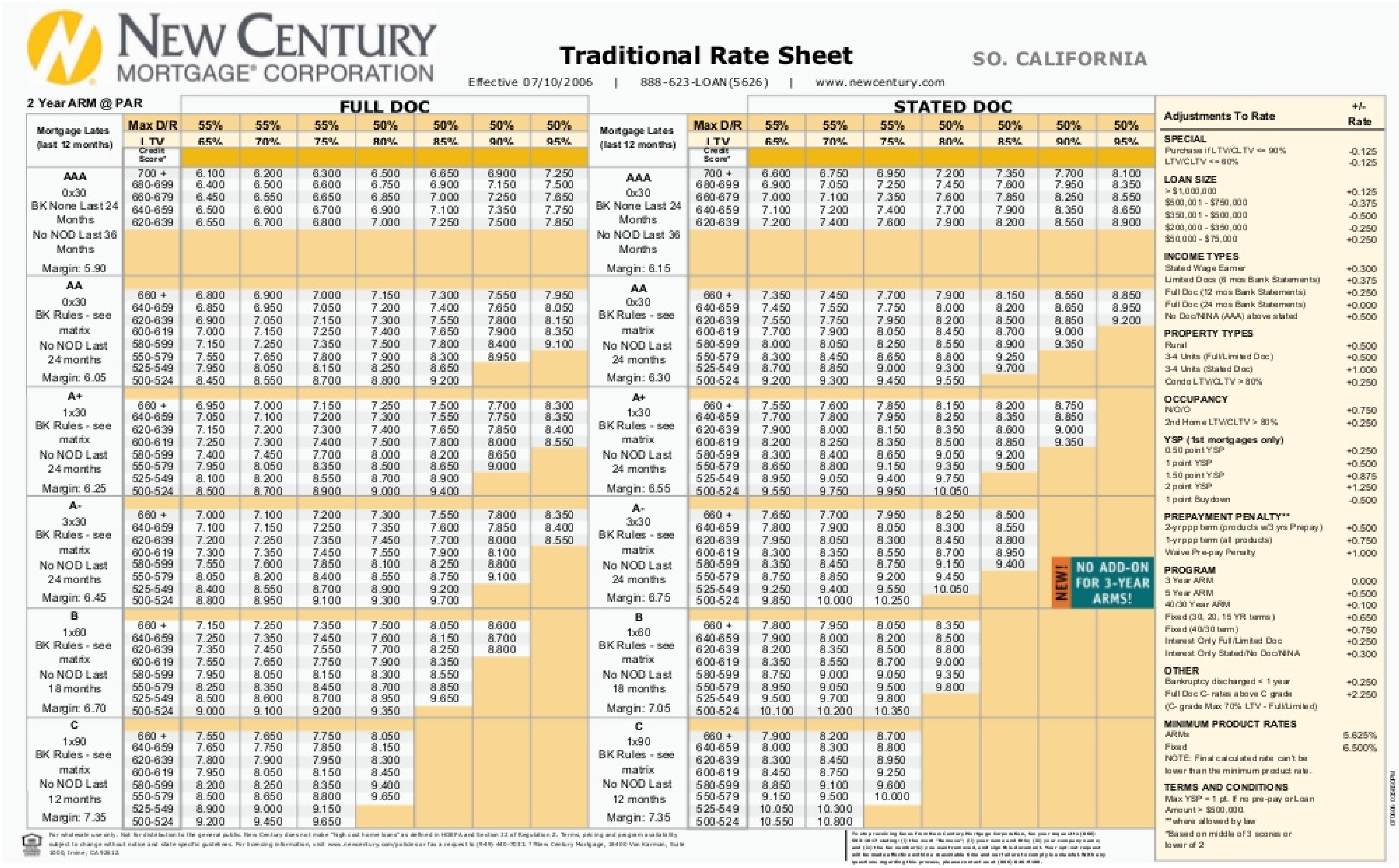
	[1]		[2]		[3]		[4]	
	% Low-Doc W2		**% Low-Doc W2**		**% Low-Doc Self**		**% Low-Doc Self**	
Dependent Variable: Low-Doc Share in Zip	**Coeff.**	**Std. Err.**	**Coeff.**	**Std. Err.**	**Coeff.**	**Std. Err.**	**Coeff.**	**Std. Err.**
Zip Median Income	0.0328***	(0.0046)	-0.0202***	(0.0042)	-0.0233***	(0.0060)	-0.0505***	(0.0065)
MSA Fixed Effects	No		Yes		No		Yes	
N	7,063		7,033		5,464		5,464	
Adjusted R^2	0.01		0.36		0.00		0.13	

Note: This table presents coefficient estimates from an OLS regression of the percentage of a zip code's loans that are low-doc on the natural logarithm of zip code median income from the 2000 Census. The dependent variable in Columns [1] and [2] is the number of loans in a zip code that were low-doc to W2 borrowers divided by the total number of W2 loans originated in that zip code. The dependent variable in Columns [3] and [4] is the number of loans in a zip code that were low-doc to self-employed borrowers divided by the total number of loans originated to self-employed borrowers in that zip code. Zip codes included in Columns [1] and [2] must have at least nine total loans to W2 borrowers (the median for all the zip codes), whereas Columns [3] and [4] include zip codes that have at least four originations to self-employed borrowers (the median for all the zip codes). The total number of zip codes in our sample is 14,090. Standard errors are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Table XV: MSA House Price Changes and Low-Doc Share of Originations

	[1]	[2]	[3]
	ΔHPI	ΔHPI	ΔHPI
Dependent Variable	(2006-2007)	(2006-2008)	(2006-2009)
Marginal Effect of the Share of Low-Doc Loans (Equation (15))	-0.187***	-0.106*	-0.023
	(0.053)	(0.062)	(0.072)
Change in the Effect by Supply Ineslasticity	-0.601***	-0.345**	0.074
	(0.144)	(0.159)	(0.206)
Marginal Effect of Proportion of W2 (Equation (16))	-0.172***	-0.184***	-0.138**
	(0.056)	(0.068)	(0.070)
Change in the Effect by Supply Ineslasticity	-0.355	-0.527**	-0.376
	(0.221)	(0.234)	(0.243)

Note: This table reports the marginal effects on the subsequent house price changes of the share of low-doc loans and the proportion of low-doc loans to that are to W2 borrowers as defined by Equations (15) and (16). The marginal effect is evaluated at the mean values of the interacted variables. The regression result is presented in Table A9 in the Appendix. White's heteroskedasticity robust standard errors are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.



NEW CENTURY MORTGAGE® CORPORATION

Traditional Rate Sheet SO. CALIFORNIA

Effective 07/10/2006 | 888-623-LOAN(5626) | www.newcentury.com

2 Year ARM @ PAR

FULL DOC

Mortgage Lates (last 12 months)	Max D/R / LTV / Credit Score*	55% / 65%	55% / 70%	55% / 75%	50% / 80%	50% / 85%	50% / 90%	50% / 95%
AAA	700 +	6.100	6.200	6.300	6.500	6.650	6.900	7.250
0x30	680-699	6.400	6.500	6.600	6.750	6.900	7.150	7.500
BK None Last 24 Months	660-679	6.450	6.550	6.650	6.850	7.000	7.250	7.650
No NOD Last 36 Months	640-659	6.500	6.600	6.700	6.900	7.100	7.350	7.750
Margin: 5.90	620-639	6.550	6.700	6.800	7.000	7.250	7.500	7.850
AA	660 +	6.800	6.900	7.000	7.150	7.300	7.550	7.950
0x30	640-659	6.850	6.950	7.050	7.200	7.400	7.650	8.050
BK Rules - see matrix	620-639	6.900	7.050	7.150	7.300	7.550	7.800	8.150
No NOD Last 24 months	600-619	7.000	7.150	7.250	7.400	7.650	7.900	8.350
Margin: 6.05	580-599	7.150	7.250	7.350	7.500	7.800	8.400	9.100
	550-579	7.550	7.650	7.800	7.900	8.300	8.950	
	525-549	7.950	8.050	8.150	8.250	8.650		
	500-524	8.450	8.550	8.700	8.800	9.200		
A+	660 +	6.950	7.000	7.150	7.250	7.500	7.700	8.300
1x30	640-659	7.050	7.100	7.200	7.300	7.550	7.750	8.350
BK Rules - see matrix	620-639	7.150	7.200	7.300	7.400	7.650	7.850	8.400
No NOD Last 24 months	600-619	7.250	7.300	7.400	7.500	7.800	8.000	8.550
Margin: 6.25	580-599	7.400	7.450	7.700	8.000	8.200	8.650	
	550-579	7.950	8.050	8.350	8.500	8.650	9.000	
	525-549	8.100	8.200	8.550	8.700	8.900		
	500-524	8.500	8.700	8.900	9.000	9.400		
A-	660 +	7.000	7.100	7.200	7.300	7.550	7.800	8.350
3x30	640-659	7.100	7.150	7.250	7.350	7.600	7.850	8.400
BK Rules - see matrix	620-639	7.200	7.250	7.350	7.450	7.700	8.000	8.550
No NOD Last 24 months	600-619	7.300	7.350	7.450	7.550	7.900	8.100	
Margin: 6.45	580-599	7.550	7.600	7.850	8.100	8.250	8.800	
	550-579	8.050	8.200	8.400	8.550	8.750	9.100	
	525-549	8.400	8.550	8.700	8.900	9.200		
	500-524	8.800	8.950	9.100	9.300	9.700		
B	660 +	7.150	7.250	7.350	7.500	8.050	8.600	
1x60	640-659	7.250	7.350	7.450	7.600	8.150	8.700	
BK Rules - see matrix	620-639	7.350	7.450	7.550	7.700	8.250	8.800	
No NOD Last 18 months	600-619	7.550	7.650	7.750	7.900	8.350		
Margin: 6.70	580-599	7.950	8.050	8.150	8.300	8.550		
	550-579	8.250	8.350	8.450	8.700	8.850		
	525-549	8.500	8.600	8.700	8.950	9.650		
	500-524	9.000	9.100	9.200	9.350			
C	660 +	7.550	7.650	7.750	8.050			
1x90	640-659	7.650	7.750	7.850	8.150			
BK Rules - see matrix	620-639	7.800	7.900	7.950	8.300			
No NOD Last 12 months	600-619	7.950	8.050	8.150	8.450			
Margin: 7.35	580-599	8.200	8.250	8.350	9.400			
	550-579	8.500	8.650	8.800	9.650			
	525-549	8.900	9.000	9.150				
	500-524	9.200	9.450	9.650				

STATED DOC

Mortgage Lates (last 12 months)	Max D/R / LTV / Credit Score*	55% / 65%	55% / 70%	55% / 75%	50% / 80%	50% / 85%	50% / 90%	50% / 95%
AAA	700 +	6.600	6.750	6.950	7.200	7.350	7.700	8.100
0x30	680-699	6.900	7.050	7.250	7.450	7.600	7.950	8.350
BK None Last 24 Months	660-679	7.000	7.100	7.350	7.600	7.850	8.250	8.550
No NOD Last 36 Months	640-659	7.100	7.200	7.400	7.700	7.900	8.350	8.650
Margin: 6.15	620-639	7.200	7.400	7.600	7.900	8.200	8.550	8.900
AA	660 +	7.350	7.450	7.700	7.900	8.150	8.550	8.850
0x30	640-659	7.450	7.550	7.750	8.000	8.200	8.650	8.950
BK Rules - see matrix	620-639	7.550	7.750	7.950	8.200	8.500	8.850	9.200
No NOD Last 24 months	600-619	7.700	7.900	8.050	8.450	8.700	9.000	
Margin: 6.30	580-599	8.000	8.050	8.250	8.550	8.900	9.350	
	550-579	8.300	8.450	8.650	8.800	9.250		
	525-549	8.700	8.850	9.000	9.300	9.700		
	500-524	9.200	9.300	9.450	9.550			
A+	660 +	7.550	7.600	7.850	8.150	8.200	8.750	
1x30	640-659	7.700	7.800	7.950	8.250	8.350	8.850	
BK Rules - see matrix	620-639	7.900	8.000	8.150	8.350	8.600	9.000	
No NOD Last 24 months	600-619	8.200	8.250	8.350	8.500	8.850	9.350	
Margin: 6.55	580-599	8.300	8.400	8.650	9.050	9.200		
	550-579	8.650	8.800	9.150	9.350	9.500		
	525-549	8.950	9.050	9.400	9.750			
	500-524	9.550	9.750	9.950	10.050			
A-	660 +	7.650	7.700	7.950	8.250	8.500		
3x30	640-659	7.800	7.900	8.050	8.300	8.550		
BK Rules - see matrix	620-639	7.950	8.050	8.300	8.450	8.800		
No NOD Last 24 months	600-619	8.300	8.350	8.550	8.700	8.950		
Margin: 6.75	580-599	8.350	8.450	8.750	9.150	9.400		
	550-579	8.750	8.850	9.200	9.450			
	525-549	9.250	9.400	9.550	10.050			
	500-524	9.850	10.000	10.250				
B	660 +	7.800	7.950	8.050	8.350			
1x60	640-659	7.900	8.000	8.200	8.500			
BK Rules - see matrix	620-639	8.200	8.350	8.500	8.800			
No NOD Last 18 months	600-619	8.350	8.550	8.700	9.000			
Margin: 7.05	580-599	8.750	9.000	9.050	9.350			
	550-579	8.950	9.050	9.500	9.800			
	525-549	9.500	9.700	9.800				
	500-524	10.100	10.200	10.350				
C	660 +	7.900	8.200	8.700				
1x90	640-659	8.000	8.300	8.800				
BK Rules - see matrix	620-639	8.300	8.450	8.950				
No NOD Last 12 months	600-619	8.450	8.750	9.250				
Margin: 7.35	580-599	8.850	9.100	9.600				
	550-579	9.150	9.500	10.000				
	525-549	10.050	10.300					
	500-524	10.550	10.800					

NEW! NO ADD-ON FOR 3-YEAR ARMS!

Adjustments To Rate	+/- Rate
SPECIAL	
Purchase if LTV/CLTV <= 90%	-0.125
LTV/CLTV <= 60%	-0.125
LOAN SIZE	
> $1,000,000	+0.125
$500,001 - $750,000	-0.375
$350,001 - $500,000	-0.500
$200,000 - $350,000	-0.250
$50,000 - $75,000	+0.250
INCOME TYPES	
Stated Wage Earner	+0.300
Limited Docs (6 mos Bank Statements)	+0.375
Full Doc (12 mos Bank Statements)	+0.250
Full Doc (24 mos Bank Statements)	+0.000
No Doc/NINA (AAA) above stated	+0.500
PROPERTY TYPES	
Rural	+0.500
3-4 Units (Full/Limited Doc)	+0.500
3-4 Units (Stated Doc)	+1.000
Condo LTV/CLTV > 80%	+0.250
OCCUPANCY	
N/O/O	+0.750
2nd Home LTV/CLTV > 80%	+0.250
YSP (1st mortgages only)	
0.50 point YSP	+0.250
1 point YSP	+0.500
1.50 point YSP	+0.875
2 point YSP	+1.250
1 point Buydown	-0.500
PREPAYMENT PENALTY**	
2-yr ppp term (products w/3 yrs Prepay)	+0.500
1-yr ppp term (all products)	+0.750
Waive Pre-pay Penalty	+1.000
PROGRAM	
3 Year ARM	0.000
5 Year ARM	+0.500
40/30 Year ARM	+0.100
Fixed (30, 20, 15 YR terms)	+0.650
Fixed (40/30 term)	+0.750
Interest Only Full/Limited Doc	+0.250
Interest Only Stated/No Doc/NINA	+0.300
OTHER	
Bankruptcy discharged < 1 year	+0.250
Full Doc C- rates above C grade	+2.250
(C- grade Max 70% LTV - Full/Limited)	
MINIMUM PRODUCT RATES	
ARMs	5.625%
Fixed	6.500%
NOTE: Final calculated rate can't be lower than the minimum product rate.	
TERMS AND CONDITIONS	
Max YSP = 1 pt. If no pre-pay or Loan Amount > $500,000.	
**where allowed by law	
*Based on middle of 3 scores or lower of 2	

Figure 1: **New Century Rate Sheet.** This figure includes a pricing matrix from New Century Mortgage Corporation for the Southern California Region as of 07/10/2006. The original image is available at http://www.slideshare.net/devonauerswald/new-century-subprime-matrix-80-5 .

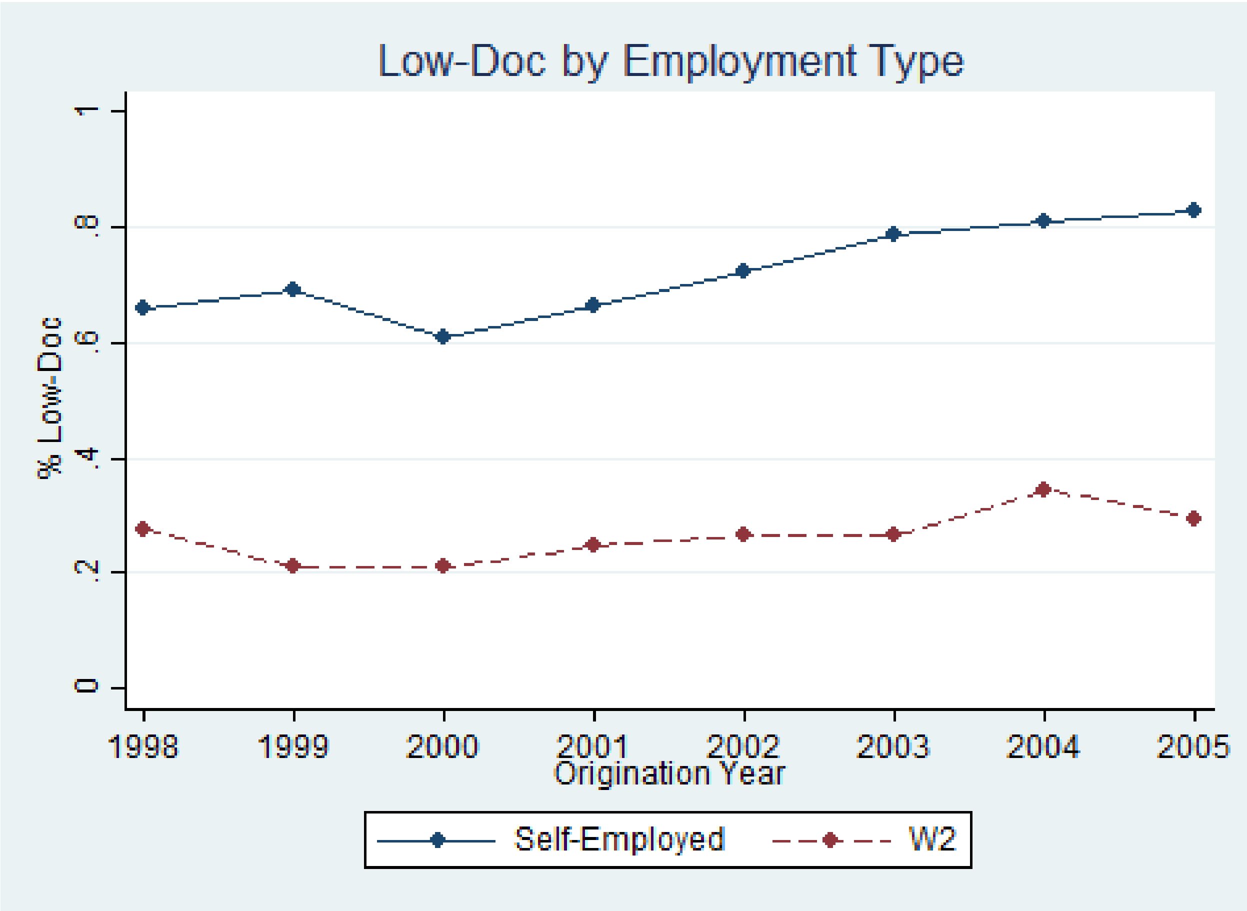


Figure 2: **Share of Originations that are Low-Doc by Employment Type.** This figure shows the proportion of originated loans that are low-doc by employment type in each origination year. The sample includes funded loans from the New Century database as described in Section III.A.

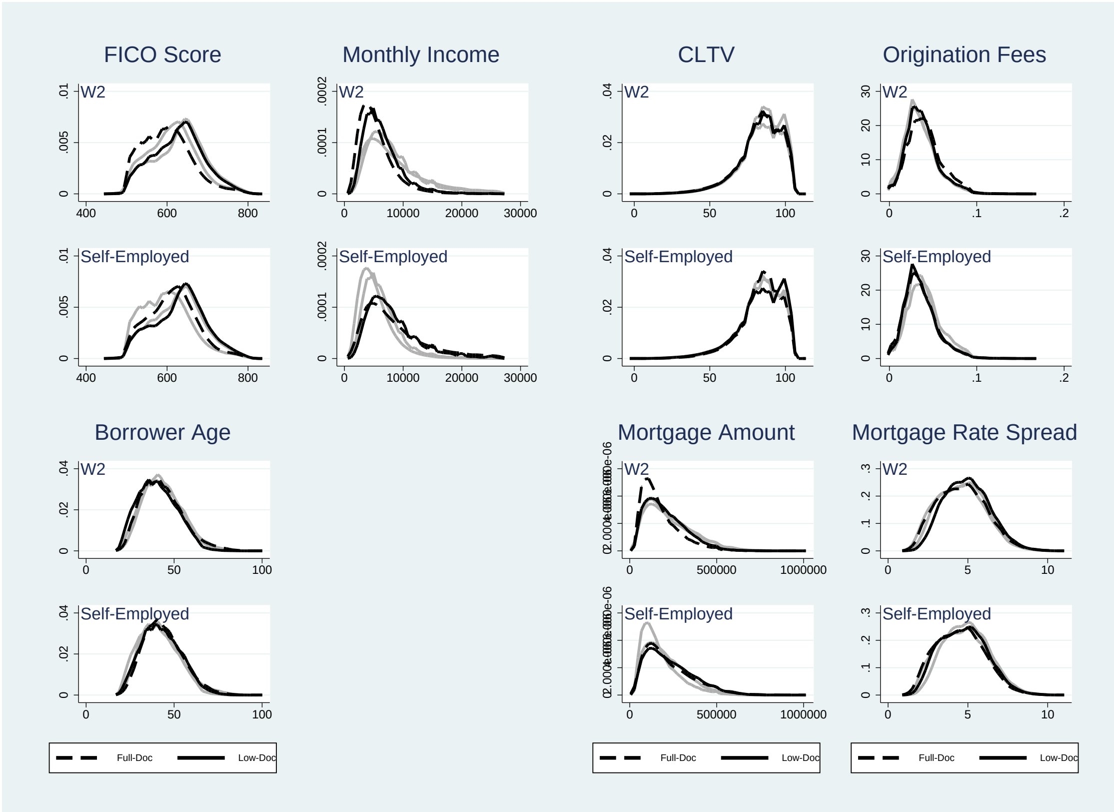


Figure 3: **Distribution of Borrower and Loan Characteristics.** This figure shows the kernel densities of borrower FICO score, reported income, borrower age, combined loan to value ratio (CLTV), origination fees, mortgage amount, and the rate spread (contract rate minus the two year constant maturity Treasury rate) at origination in the sample of funded loans from New Century. Each panel includes the densities for all combinations of employment type and income documentation (W2/full-doc; W2/low-doc; self-employed/full-doc; self-employed/low-doc). The top and bottom panels for each borrower characteristic are identical, however, the top highlights the densities for W2 borrowers while the bottom highlights the densities for self-employed borrowers.

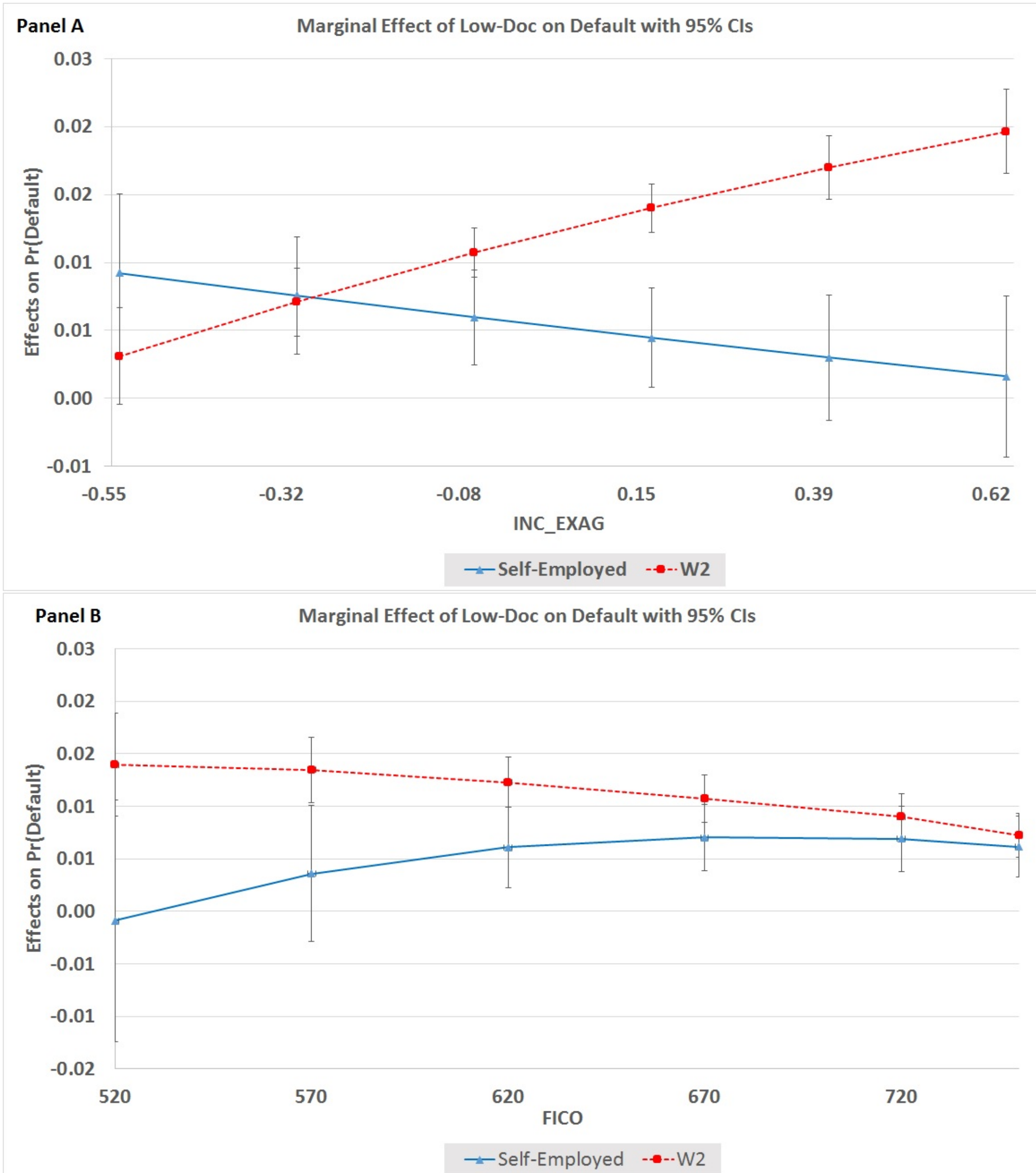


Figure 4: **Marginal Effect of Low-Doc on Default by Income Exaggeration and FICO Score.** Panel A shows the average marginal effects of low-doc at different levels of estimated income exaggeration by employment type. -0.57 and 1.02 are the 5th and 95th percentiles of income exaggeration, respectively. The marginal effects are derived the probit model of mortgage default described in equation (10) for the funded loans from the New Century database. Panel B shows the average marginal effects of low-doc across different FICO scores by employment type. The marginal effects are derived from the probit model of mortgage default described in equation (12) for the funded loans from the New Century database.

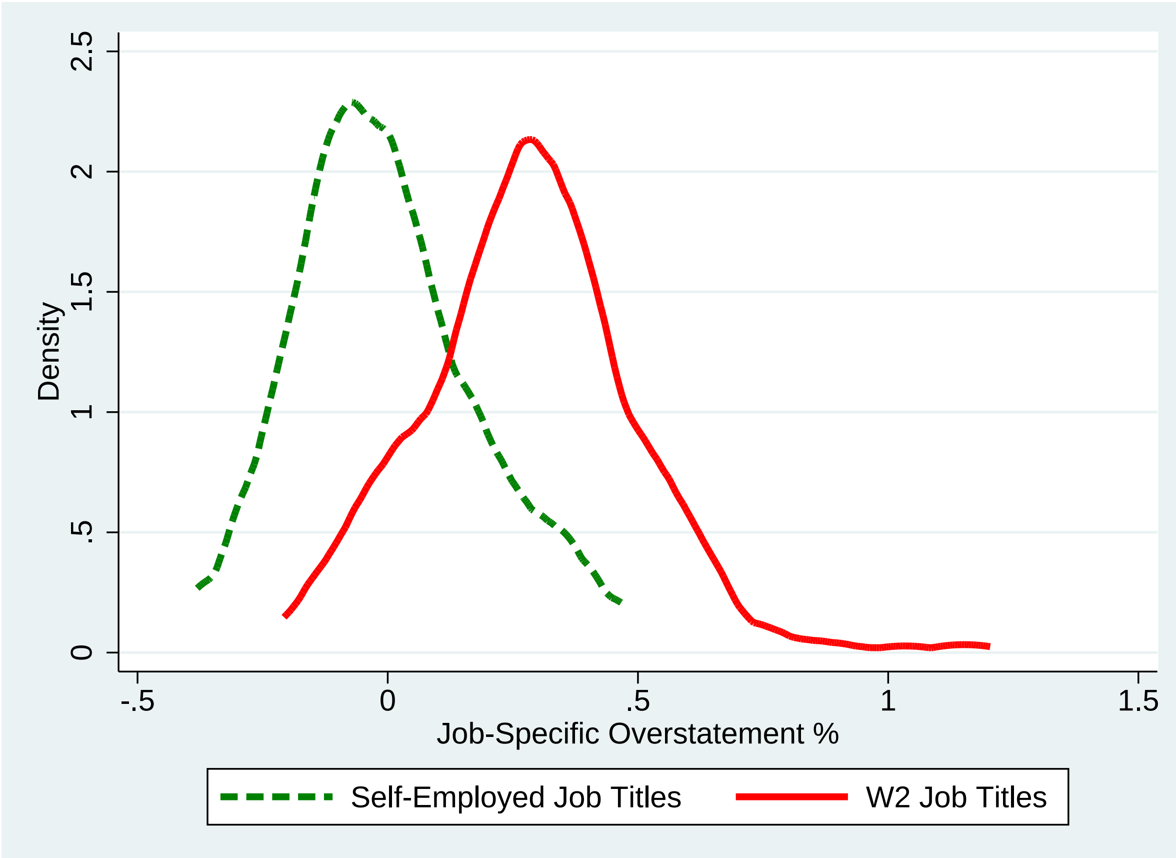


Figure 5: **Distribution of Job-Specific Overstatement.** This table presents the distribution on *Job-Specific Overstatement (%)* for the 313 W2 job titles that had at least 10 low- and 10 full-doc observations as well as the 55 self-employed job titles that had at least 10 full- and 10 low-doc observations.