BANK OF CANADA
BANQUE DU CANADA

Working Paper/Document de travail
2015-1

Immigrants and Mortgage Delinquency in the United States

by Zhenguo Lin, Yingchun Liu and Jia Xie

www.bank-banque-canada.ca

Bank of Canada Working Paper 2015-1

January 2015

Immigrants and Mortgage Delinquency in the United States

by

Zhenguo Lin,[1] Yingchun Liu[2] and Jia Xie[3]

[1]Department of Finance
Mihaylo College of Business and Economics
California State University, Fullerton
Fullerton, CA 92834-6848, U.S.A.

[2]Department of Finance, Insurance and Real Estate
Laval University
Quebec City, QC, Canada G1V 0A6

[3]Financial Stability Department
Bank of Canada
Ottawa, Ontario, Canada K1A 0G9
xiej@bankofcanada.ca

ISSN 1701-9397

Acknowledgements

We would like to thank Jason Allen and Natalia Kyui for helpful comments and suggestions. This paper was previously circulated under the title “Immigrants and Mortgage Delinquency.”

Abstract

We investigate the relationship between immigrant status and mortgage delinquency in the United States. We find that after controlling for observables, newly arrived immigrants are likely to have a higher delinquency rate on mortgages than natives, while immigrants who have resided in the United States for more than 20 years are no different than natives in this regard. In addition, there is no evidence that the second generation of immigrants has a higher delinquency rate than the third-or-higher generation. Our results are robust to potential sample selection bias and functional misspecifications.

JEL classification: J15, G21
Bank classification: Debt management; Financial stability

Résumé

Les auteurs étudient les liens éventuels entre le statut d'immigrant et la probabilité de défaillance sur les prêts hypothécaires aux États-Unis. Après avoir tenu compte d'un certain nombre de paramètres observables, ils constatent que les nouveaux immigrés sont susceptibles d'avoir un taux de défaillance sur leurs prêts hypothécaires plus élevé que les personnes ayant toujours vécu aux États-Unis, alors qu'ils n'ont observé aucune différence chez les immigrants qui y habitent depuis plus de 20 ans. De plus, rien n'indique que les enfants d'immigrants de seconde génération aient un taux de défaillance supérieur à celui de descendants de troisième génération ou d'une génération subséquente. Ces résultats ne sont pas sensibles à des biais de sélection ni à des erreurs de spécification fonctionnelles possibles.

Classification JEL : J15, G21
Classification de la Banque : Gestion de la dette; Stabilité financière

Non-technical summary

We investigate the relationship between immigrant status and mortgage delinquency rates in the United States by using data from the 2009 Panel Study of Income Dynamics (PSID). We find that immigrants have a higher mortgage delinquency rate than natives. However, much of the immigrant-native difference can be explained by the lower financial status of immigrants and the riskier mortgage products they hold than natives.

After controlling for household and mortgage characteristics, there remains a 2.5 percentage-point immigrant-native difference in mortgage delinquency rates. Four potential explanations for this difference are examined: (i) unobservable financial constraints; (ii) local housing market conditions; (iii) unobservable metropolitan area characteristics; and (iv) imperfect integration of immigrants into the host society. Integration is a process wherein immigrants adapt to the communities in which they settle. There are several channels through which imperfect integration may cause a higher delinquency rate among immigrant households. For example, during the financial crisis, governments and mortgage lenders provided programs to help financially stressed homeowners pay their mortgages. However, due to differences in culture and language, immigrants may not be informed of such programs. Alternatively, if immigrants have less perception of belonging, they may be less attached to their houses and communities, and may therefore have a higher delinquency rate.

We find evidence contradicting the first three explanations, and our favored explanation is the fourth, for which we find two pieces of supporting evidence. First, although newly arrived immigrants are likely to have a higher delinquency rate than natives, immigrants who have resided in the United States for more than 20 years are indistinguishable from natives. Second, there is no evidence that second-generation immigrants have a higher delinquency rate than third or subsequent generations of immigrants.

1 Introduction

According to U.S. census data, between 1970 and 2010, the proportion of United States. residents born in another country increased from 4.8 to 12.5 percent. This rise in the immigrant population can be explained by both an increase in the inflow of immigrants and a reduction in native birth rates in the United States. Immigrants and their children will account for as much as two-thirds of population growth from 1995 to 2050 (Day, 1996). Many have conjectured that this large influx of immigrants has had an enormous impact on the United States and its population in various ways (Dail, 2009; Corwin, 2012).

A number of studies describe the housing condition of immigrants. Immigrants tend to live in rental housing (JCHS, 2000) and in housing units of lower quality, especially during the period just after they arrive (Friedman, Rosenbaum and Schill, 1998). Crowding is also more common in immigrant households (Myers, Baer and Choi, 1996).

Research also exists on the home ownership gap between immigrant and native households (Kochhar, Gonzalez-Barrera and Dockterman, 2009). Home ownership has long been viewed as an important mechanism for wealth creation. Consequently, presidents have been promoting home ownership since 1934, when the Federal Housing Administration was created by Franklin D. Roosevelt to insure mortgages, in part so that low-income borrowers could qualify. Through the years, administrations touted home ownership as a way to put immigrant and low-income families on a path to social and financial stability by promoting a more involved citizenry.[1] Home ownership has also been seen as an important policy goal for immigrant households. It has been well documented that immigrants who arrived in the 1980s and 1990s are relatively less skilled and have lower wages than their native

[1]The Clinton and Bush administrations launched ambitious programs to promote home ownership, especially for low-income households. For example, in 1995, President Clinton's National Homeownership Strategy set a goal of allowing millions of families to own homes, in part by making financing more available, affordable and flexible. President George W. Bush famously said in 2002 that "We can put light where there's darkness, and hope where there's despondency in this country. And part of it is working together as a nation to encourage folks to own their own home," and in a 2004 speech he said again that "We're creating... an ownership society in this country, where more Americans than ever will be able to open up their door where they live and say, welcome to my house, welcome to my piece of property."

counterparts—there is about a 10 to 20 percent wage gap between them. Increased home ownership may not only build wealth for immigrant households, but perhaps equally important, it is a signal of assimilation and achievement of the "American Dream." As a result, the expansion of housing credit in the United States from the mid-1990s to the mid-2000s was largely cheered, and home ownership by households of immigrants and others reached record-high rates in the mid-2000s. According to the Census, the home ownership rate among immigrant households increased from 46.5 percent in 1995 to 53.3 percent in 2006. Meanwhile, the home ownership rate among native-born Americans increased from 66.1 percent in 1995 to 68.8 percent in 2006. In other words, the gap in home ownership between immigrant and native households dropped from 19.6 percent in 1995 to 15.5 percent in 2006.

As the housing and economic crises developed in 2007–09, however, immigrants were blamed by the media for the large increase in delinquencies, defaults and foreclosures in the housing market that helped to trigger the housing crisis and ultimately facilitated the bankruptcies or near-bankruptcies of multiple financial institutions (Malkin, 2008).[2] Should immigrants really be blamed for the current housing crisis? In particular, are immigrants likely to have a higher delinquency rate on mortgages than natives and, if so, why?

To shed light on these and related questions, we investigate the mortgage delinquency behavior of immigrant households by using the 2009 Panel Study of Income Dynamics (PSID) data. We find that immigrant households are likely to have a higher delinquency rate than natives, even after controlling for a wide range of household demographic, socioeconomic and mortgage characteristics. This result is consistent with several potential hypotheses. For example, although we have controlled for household socioeconomic status, immigrants may still face unobservable, tighter financial constraints than natives. Alternatively, immigrants may disproportionately live in places where the local housing markets suffered greatly from the 2007 housing crisis. It is well known that immigrants tend to cluster in gateway cities

[2]In 2008, 2.3 million homes were foreclosed, and in 2011Q2, foreclosure sales accounted for 31 percent of all residential sales in the United States (source: RealtyTrac). An initial estimate of the total dollar value of losses for the recent housing crisis was $2.4 trillion (Zandi, 2009).

that function as "ports of entry" for immigrants, such as Los Angeles, New York, Miami and Chicago. These gateway cities share certain characteristics, such as high crime rates, high population density and the sharp correction of house prices during the recent housing crisis, which could trigger greater rates of mortgage delinquency. However, we did not find any evidence consistent with these hypotheses.

A large body of literature has shown that the level of immigrant integration varies dramatically over the duration of their stay, and this correlates to a number of socioeconomic outcomes.[3] We in fact find supportive evidence that as immigrants are more integrated into U.S. society, their mortgage delinquency behavior becomes no different than that of natives. We examine this hypothesis along two dimensions.

The first dimension relates to the duration of immigrants' stays in the United States. There is ample evidence that immigrant integration improves over time the longer immigrants stay in the United States. (Osili and Xie, 2009; Coulson, 2011; Jimenez, 2011). If the immigrant-native difference in the mortgage delinquency rate is caused (in part) by imperfect short-term integration of immigrants into the host country, one would expect the difference to be larger for newly arrived immigrants and smaller for those who arrived in previous decades. Since all of the immigrant households reported in the PSID data came to the United States before 1999, the duration of immigrants' stays ranges from 10 to 40 years. Although we lack data on the most recent immigrants, i.e., those who have been in the United States for less than 10 years, the range of immigrants' stays is wide enough for us to test this hypothesis. Indeed, we find that the relatively high delinquency rate of immigrant households is mainly driven by recent immigrants who have been in the United States for less than 20 years. No evidence suggests that immigrants who migrated to the United States 20 years ago are different than natives in their rates of mortgage delinquency. We should note that, given the assimilation effect of immigration, including the most recent immigrants

[3]Over time, as immigrants gain more American experience, they increasingly tend to resemble natives in terms of housing conditions and rates of home ownership (Callis, 1997; Myers and Park, 1999; Borjas, 2002; Coulson, 2011). Hu (2000) also finds that immigrants are gradually assimilated.

would likely reinforce our main results.

The second dimension of our hypothesis relates to second-generation immigrants, defined as native-born households where at least one of the parents of the head was born outside of the United States. These are the native-born households formed by children of immigrant households when they split from their parents. Members of second-generation households are born, raised and educated in the United States, and therefore tend to be well integrated into society, just like members of other native-born households. If the immigrant-native difference in delinquency is driven by the imperfect integration of immigrants, then one would expect the difference to disappear among second-generation and other native-born households. Indeed, we find no evidence that the second generation is more likely to fall behind on mortgage payments than other native-born households.

Our results are robust to potential sample selection bias and functional misspecification. First, our results are subject to potential sample selection bias because we only observe the delinquency decisions of mortgage-indebted households, who may not represent a random selection of all home-owning households. We use the Heckman probit model to correct the potential sample selection bias, and the results are consistent with our earlier findings. Second, our results may be subject to functional misspecification if the key identifying assumption is invalid—that the effect of covariates on the outcome is linear. We use the propensity score matching (PSM) method instead of a linear regression model to ensure that the identification of parameters comes from matching households with the same covariates. Again, our findings are consistent with earlier findings.

The rest of the paper is organized as follows. In the next section, we describe our data sources and present some summary statistics. In Section 3, we present our empirical findings on the role of immigrant status in mortgage delinquency. Section 4 examines potential explanations. Section 5 conducts a robustness check. The last section concludes the paper.

2 Data

We use the 2009 wave of the Panel Study of Income Dynamics (PSID), which is collected by the University of Michigan Survey Center. PSID is a longitudinal household survey that started in 1968, with a sample of over 18,000 individuals living in over 5,000 families in the United States. Individuals in each household were followed annually from 1968 to 1997, and biannually after 1997.

The PSID data set is unique for the current study in several respects. First, the data set contains detailed household demographic information (i.e., age, gender, race, marital status and geographic location) and socioeconomic characteristics (i.e., education, employment status, income, assets, house value and debt). The PSID also contains mortgage information for each mortgage-indebted household, including the number of mortgages, the mortgage type (i.e., adjustable or fixed rate, recourse or nonrecourse), purpose (i.e., for refinance or purchase), term, age, rate, current balance and payment. From this information, we can calculate the current loan-to-value ratio (LTV) and the current debt-service ratio (DSR),[4] both of which have been broadly identified as the two key factors determining the probability of default and have been widely used as mortgage-underwriting variables in practice.[5] More interesting is that in the 2009 survey, mortgage-indebted households were asked for the first time whether they were currently delinquent on their mortgages—i.e., behind on mortgage payments for at least one month.

Second, each household in the data set is assigned a unique identification number by which we can distinguish immigrant households from native households. In particular, a sample of 511 immigrant families was added to the PSID in 1997 and 1999 in order to keep the data representative of the U.S. population at that time. In addition, each immigrant household was asked about the year of their arrival in the United States, which enables us to

[4]The current LTV is calculated as the ratio of the current balance of all mortgages to the current value of the house. The current DSR is calculated as the ratio of the sum of the current mortgage principal plus interest payments, property taxes and insurance to current family income.

[5]See, for example, Haughwout, Peach and Tracy (2008); Foote, Gerardi and Willen (2008); Mayer, Pence and Sherlund (2009); and Campbell and Cocco (2012), among others.

compute the duration of stay in the United States. We should note that all of the immigrant households in the PSID came to the United States before 1999. Hence, the duration of their stay ranges from 10 to 40 years.

Third, all of the households surveyed (immigrant and native-born) were asked about the birthplaces of the parents of the head, from which we can identify *second-generation households*, defined as native-born households, but with at least one of the parents born outside the United States. The rest of the native-born households are considered *third-or-higher generations*.

Definitions of the variables can be found in Appendix I. We restrict the data used in the current study as follows: the sample includes only mortgage-indebted households, i.e., those who own (rather than rent) their primary residences and who have at least one mortgage on their primary residence. After omitting observations with missing values, the final data include information on 2,383 households. Around 6.7 percent (159) of the households are immigrant households, and around 5.6 percent (125) of native-born households are second-generation households. Table 1 provides summary statistics for these groups of households. The first two columns show mean values of variables for immigrant and native-born households, respectively, with the p-value of the difference-in-means in the third column. The corresponding results for second-generation and third-or-higher generations are shown in the fourth to the sixth columns. The difference in the mortgage delinquency rates between immigrants (15.7%) and natives (4.4%) is significant. The corresponding difference between the second-generation and third-or-higher generations, however, is small and not significant. In addition, compared with native-born households, immigrants, on average, have less education, are more likely to be unemployed, and have lower income and wealth. Immigrants also have higher LTVs, DSRs and mortgage rates, and are more likely to have adjustable rates. These results suggest that, on average, immigrants have a worse financial status and riskier mortgage products than natives. The systematic differences in these observables highlight the importance of controls in the analysis we conduct.

Table 1: Summary Statistics

	Immigrant	Native	p-value	Second-generation	Third-or-higher generation	p-value
Delinquency rates	0.157	0.044	0.000	0.024	0.045	0.261
Demographic information						
age	45.786	46.229	0.665	50.064	46.000	0.000
male	0.830	0.848	0.536	0.856	0.848	0.809
married	0.723	0.745	0.544	0.800	0.742	0.147
race						
black	0.075	0.210	0.000	0.032	0.221	0.000
white	0.516	0.768	0.000	0.832	0.765	0.083
others	0.409	0.022	0.000	0.136	0.015	0.000
region						
northeast	0.138	0.145	0.812	0.144	0.145	0.968
north central	0.088	0.289	0.000	0.240	0.292	0.213
south	0.239	0.393	0.000	0.296	0.399	0.022
west	0.535	0.171	0.000	0.312	0.162	0.000
Socioeconomic status (non-financial factors)						
education						
less than high school	0.340	0.074	0.000	0.080	0.074	0.799
high school degree	0.421	0.571	0.000	0.448	0.578	0.004
college degree or higher	0.239	0.355	0.003	0.472	0.348	0.005
employment status						
unemployed	0.088	0.041	0.005	0.048	0.041	0.681
not in labor force	0.063	0.128	0.016	0.136	0.127	0.775
employed	0.849	0.831	0.564	0.816	0.832	0.636
Socioeconomic status (financial factors)						
yearly family income	90705	111555	0.119	140000	110000	0.045
income growth rate	0.478	0.319	0.115	0.286	0.321	0.761
total assets	154826	335282	0.063	505930	325120	0.106
total debts other than mortgages	11306	14709	0.201	13445	14784	0.660
monetary transfers to relatives	1588	1020	0.288	1776	975	0.141
Mortgage information						
CLTV	0.776	0.644	0.000	0.627	0.645	0.523
DSR	0.249	0.164	0.000	0.163	0.164	0.930
ARM	0.151	0.088	0.008	0.120	0.086	0.189
interest rate	6.248	6.022	0.050	5.979	6.024	0.726
mortgage term	25.340	23.505	0.005	23.512	23.504	0.991
mortgage age	4.314	4.616	0.118	4.616	4.616	0.999
refinance	0.516	0.459	0.167	0.416	0.462	0.320
two mortgages	0.120	0.151	0.286	0.192	0.148	0.183
recourse	0.465	0.768	0.000	0.720	0.770	0.195
Pseudo R-squared	36%			12%		
p-value	0.000			0.000		
Number of observations	159	2224		125	2099	

3 Empirical Findings

We use the standard probit model to study the impact of immigrant status on the probability of delinquency. The results, however, are robust to various estimation methods, as we will show in Section 5. We assume that there is an unobservable latent variable z_i, with

$$z_i = \alpha\, Immigrant_i + \beta\, X_i + \epsilon_i, \tag{1}$$

where i denotes a household, and $Immigrant_i$ is an indicator variable for immigrant households. X_i is a vector of control variables, including household demographic and socioeconomic characteristics and mortgage information. Let y_i be the indicator variable of household i's delinquency status, such that household i is delinquent on mortgages if z_i is above zero, i.e.,

$$y_i = \begin{cases} 1, & \text{if } z_i > 0; \\ 0, & \text{if } z_i \leq 0 \end{cases}.$$

We estimate a series of different specifications by gradually increasing the number of controlled variables in X_i to see their effects on the probability of delinquency. The estimated coefficients and marginal effects are reported in Table 2.

We begin with the simplest specification by controlling for $Immigrant_i$ only, and report the results (i.e., coefficient, standard error, significance level and marginal effect) in column 1 of Table 2. The marginal effect indicates that, without controlling for any observables, the mortgage delinquency rate of immigrant households is 11.3 percentage points higher than that of native-born households. The gap is statistically significant at the 1% level. As a first step toward measuring the effect of immigrants on mortgage delinquency, in Specification 2 we control for household demographic characteristics, including age, gender, race, marital status and region. The immigrant-native difference in mortgage delinquency rates decreased to 9.8 percentage points, but remains significant at the 1% level. The pseudo R-squared increased from 2.8% to 8.2%, implying that it is important to control for household demographic characteristics in predicting households' delinquency behavior.

In Specification 3, we further explore the delinquency gap between immigrants and natives by controlling for certain non-financial factors in household socioeconomic status, such as education and employment status. As shown in column 3 of Table 2, after controlling for these non-financial factors in household socioeconomic status, the immigrant-native difference in mortgage delinquency rates narrows further, from 9.8 to 6.9 percentage points. The pseudo R-squared increased from 8.2% to 10.3%.

Table 2: The Impact of Immigrant Status on Mortgage Delinquency
(Estimated by probit models)

	(1)		(2)		(3)		(4)		(5)	
	Marginal effect	Coefficient	Marginal effect	Coefficient	Marginal effect	Coefficient	Marginal effect	Coefficient	Marginal effect	Coefficient
Immigrant	0.113	0.699***	0.098	0.689***	0.069	0.554***	0.040	0.445***	0.026	0.377**
		(0.096)		(0.080)		(0.091)		(0.112)		(0.159)
	The Impacts of Other Covariates									
Demographic characteristics										
age (x 1000)			9.594	111.042***	8.541	104.465***	8.491	136.402***	5.907	118.281***
				(37.869)		(37.741)		(40.570)		(44.607)
age squared (x1000)			-0.102	-1.179***	-0.090	-1.104***	-0.082	-1.319***	-0.057	-1.148**
				(0.402)		(0.400)		(0.421)		(0.454)
male			-0.005	-0.059	-0.009	-0.104	-0.002	-0.024	0.002	0.033
				(0.182)		(0.187)		(0.190)		(0.204)
married			-0.020	-0.212	-0.015	-0.173	-0.007	-0.112	-0.008	-0.152
				(0.135)		(0.139)		(0.161)		(0.156)
race (white omitted)										
black			0.060	0.524***	0.053	0.497***	0.029	0.373***	0.022	0.350***
				(0.115)		(0.120)		(0.119)		(0.130)
other			0.020	0.195*	0.024	0.244**	0.006	0.092	-0.000	-0.002
				(0.114)		(0.116)		(0.136)		(0.139)
region (northeast omitted)										
north central			0.015	0.167	0.010	0.119	-0.004	-0.065	-0.007	-0.156
				(0.164)		(0.177)		(0.186)		(0.166)
south			-0.004	-0.053	-0.006	-0.078	-0.014	-0.241	-0.014	-0.308*
				(0.144)		(0.159)		(0.177)		(0.161)
west			0.008	0.093	0.006	0.075	-0.005	-0.082	-0.017	-0.444**
				(0.155)		(0.168)		(0.185)		(0.226)
Socioeconomic status (non-financial factors)										
education (less than high school omitted)										
high school degree					-0.019	-0.226	-0.009	-0.146	-0.010	-0.188
						(0.144)		(0.165)		(0.174)
college degree or higher					-0.035	-0.472***	-0.015	-0.263	-0.014	-0.311
						(0.158)		(0.178)		(0.196)
unemployed					0.070	0.557***	0.058	0.569***	0.050	0.587***
						(0.190)		(0.171)		(0.168)
Socioeconomic status (financial factors)										
log(yearly family income)							-0.008	-0.133	-0.004	-0.076
								(0.090)		(0.118)
income growth rate (x 100)							-0.130	-2.095	-0.133	-2.665
								(4.991)		(2.734)

Continued on next page

Table 2: – *Continued from previous page*

	(1)		(2)		(3)		(4)		(5)	
	Marginal effect	Coefficient	Marginal effect	Coefficient	Marginal effect	Coefficient	Marginal effect	Coefficient	Marginal effect	Coefficient
log (total assets)							-0.013	-0.213***	-0.006	-0.116***
								(0.024)		(0.034)
log (total debts other than mortgages)							-0.001	-0.011	-0.000	-0.007
								(0.011)		(0.012)
log (monetary transfers to relatives)							0.001	0.020	0.001	0.017
								(0.015)		(0.017)
Mortgage characteristics										
loan-to-value ratio									0.038	0.754***
										(0.196)
debt-service ratio									0.024	0.487**
										(0.196)
variable rate mortgage									0.039	0.506***
										(0.126)
interest rate									0.007	0.146***
										(0.041)
mortgage term (30 years omitted)										
less than 15 years									-0.003	-0.053
										(0.227)
15 years									-0.001	-0.023
										(0.128)
mortgage age									-0.000	-0.002
										(0.015)
refinance									0.003	0.063
										(0.122)
two mortgages									0.010	0.174
										(0.166)
recourse									-0.012	-0.213
										(0.147)
Pseudo R-squared		2.8%		8.2%		10.3%		19.0%		26.5%
Number of observations		2383		2383		2383		2383		2383

Note: This table reports marginal effects and coefficients from various Probit regressions, along with robust standard errors (clustered at the state level) in parentheses. The dependent variable is an indicator variable of mortgage delinquency. * significant at 10% level, ** significant at 5% level, *** significant at 1% level. The marginal effect of a dummy variable measures the impact of a discrete change of the dummy variable from 0 to 1.

The remaining immigrant-native difference in delinquency rates may be due to the different financial constraints faced by immigrant and native-born households. Indeed, as we saw in Table 1, immigrant households, on average, have much lower levels of assets and income than native-born households, suggesting that, other things being equal, immigrant households may face tighter financial constraints. We thus control for variables of household financial status, including family income, income growth, assets and debt, and report the results in Specification 4. We also control for monetary transfers to relatives, because we notice that immigrants, on average, give \$568 more to their relatives than natives, even though their assets and income are much less (see Table 1). Transfers to relatives may have a crowding-out effect on mortgage repayment and therefore raise household delinquency rates. After controlling for household financial status, the immigrant-native difference in delinquency rates further decreases to 4.0 percentage points but is still significant at the 1% level. The pseudo R-squared almost doubles, from 10.3% to 19.0%, suggesting the importance of financial status as a determinant of delinquency probability.

The remaining possibility is that immigrant households have riskier mortgage products, which affects their ability and incentives to repay their mortgages. For example, from Table 1 we see that immigrants, on average, have higher current LTVs and DSRs. It is widely recognized in the literature that a high LTV decreases a household's incentive to repay mortgages, while a high DSR decreases the ability to repay (Haughwout, Peach and Tracy, 2008; Foote, Gerardi and Willen, 2008; Mayer, Pence and Sherlund, 2009; and Campbell and Cocco, 2012). In Specification 5, we control for mortgage characteristics, such as the current LTV and DSR, rate, term, age, and the existence of a second mortgage. In addition, we control for the mortgage purpose (i.e., for refinance or purchase) and type (i.e., adjustable or fixed rate, recourse or nonrecourse). Mayer, Pence and Sherlund (2009) find that mortgages originated for purchase have higher delinquency rates than those for refinance. Campbell and Cocco (2012) show that adjustable-rate mortgages are riskier than fixed-rate mortgages, holding everything else constant. The difference between recourse and nonrecourse mortgages refers to loss prevention and lender's rights. For a recourse mortgage, the lender can repossess the borrower's other assets or have his/her wages garnished. For a nonrecourse mortgage, however, the lender must absorb the loss. Therefore, recourse is likely to decrease the probability of delinquency or default. In fact, empirical evidence suggests that borrowers with homes appraised at \$500,000 to \$750,000 are twice as likely to default on nonrecourse mortgages (Ghent and Kudlyak, 2011).

The results, as displayed in column 5 of Table 2, suggest that differences in mortgages can account for 35% of the immigrant-native difference in delinquency rates (from 4.0 to 2.6 percentage points). Still, a 2.6 percentage-point difference remains, and it is statistically significant at the 5% level. The pseudo R-squared increases from 19.0% to 26.5%, suggesting the importance of mortgage characteristics in predicting delinquency.

The estimated coefficients and marginal effects of other variables in Specification 5 are also as expected. For example, holding other things equal, the mortgage delinquency rate of unemployed households is 5.1 percentage points higher than that of other households. Households with lower wealth, higher LTVs, higher DSRs, higher interest rates, and adjustable mortgage rates have higher delinquency probabilities.

4 Potential Explanations

A number of theories can, in principle, produce the basic pattern of results that we observe in the data. In this section, we attempt to distinguish between these potential hypotheses. There are four possible reasons why immigrants are likely to have a higher mortgage delinquency rate than natives: (i) unobservable financial constraints; (ii) local housing market conditions; (iii) unobservable metropolitan area characteristics; and (iv) imperfect integration of immigrants into the host society. We consider these four explanations in turn.

4.1 Unobservable Financial Constraints

Even though we have controlled for observable household socioeconomic status, there may exist unobservable financial constraints that affect households' willingness or ability to repay their mortgages. To the extent that these unobservable financial constraints differ across immigrant and native households, our previous estimates may suffer from omitted variable biases. To address this issue, we proxy households' unobservable financial conditions by the monetary assistance they receive from others. We expect that households who receive monetary assistance from others tend to be in a poorer financial state than those who do not receive help. We therefore re-estimate Specification 5, adding the indicator variable "Receiving financial assistance from others." The results are reported in column 1 of Table 3.

As we expected, the mortgage delinquency rate of households who receive monetary assistance from others is 2.7 percentage points higher than the rate of those who receive no help, holding everything else equal. This result confirms the existence of unobservable financial constraints and implies intuitively that monetary assistance from others is a valid proxy for unobservable financial constraints. There is no evidence that the observed immigrant-native difference in delinquency rates can be explained by unobservable financial constraints.

Table 3: The Impact of Immigrant Status with Additional Controls
(Estimated by probit models)

	(1)		(2)		(3)	
	Marginal effect	Coefficient	Marginal effect	Coefficient	Marginal effect	Coefficient
Immigrant	0.029	0.408**	0.025	0.361**	0.024	0.356**
		(0.167)		(0.160)		(0.182)
Receiving financial assistance from others	0.027	0.393***				
		(0.113)				
House price growth rate since the last mortgage			-0.0002	-0.004		
				(0.003)		
Living in a gateway city					0.004	0.074
						(0.120)
Other variables controled						
demographic characteristics		Yes		Yes		Yes
socioeconomic status (non-financial factors)		Yes		Yes		Yes
socioeconomic status (financial factors)		Yes		Yes		Yes
mortgage characteristics		Yes		Yes		Yes
Pseudo R-squared		27.1%		26.6%		26.5%
Number of observations		2383		2383		2383

Note: This table reports marginal effects and coefficients from various probit regressions, along with robust standard errors (clustered at the state level) in parentheses. The dependent variable is an indicator variable of mortgage delinquency.
* significant at 10% level, ** significant at 5% level, *** significant at 1% level. The marginal effect of a dummy variable measures the impact of a discrete change of the dummy variable from 0 to 1.

4.2 Local Housing Market Conditions

The real estate literature has provided evidence that local housing market conditions affect households' incentives to repay mortgages. Everything else being equal, households living in a rising market are less likely to be delinquent than those in a declining market (Campbell and Cocco, 2012; Lin, Rosenblatt and Yao, 2009; Deng, Quigley and Van Order, 2000; and Vandell, 1978). If immigrants and natives tend to live in different states, where the housing market suffered to different extents from the 2007 housing crisis, then they will have dissimilar incentives to repay mortgages. We therefore re-estimate Specification 5, controlling for the growth rate of the state house price index since the last mortgage.[6] The results are reported in column 2 of Table 3. As we expected, the direction of the effect suggests a negative correlation between house price growth and delinquency rates. The size of the effect, however, is both economically negligible and statistically insignificant, suggesting that, holding everything else constant, a home price decline may not necessarily trigger borrowers' default, which is in fact consistent with the recent finding by Foote, Gerardi, and Willen (2008) and Campbell and Cocco (2012). In particular, both studies find that it is not simply negative equity that leads to default, but a combination of negative equity and reduced monthly cash flows (congruent with a reduction in monthly disposable income). Foote, Gerardi, and Willen (2008) state: "From a borrower's perspective, the decision to default hinges on how onerous the monthly mortgage payment is, relative to the possibility that the house's value will eventually exceed the balance on the mortgage." The bursting of

[6]The state house price indexes used in this paper are purchase-only indexes from the Federal Housing Finance Agency.

the housing bubble provided the necessary negative equity condition in many cases, but not all homeowners had reduced monthly cash flows.

4.3 Unobservable Metropolitan Area Characteristics

Another possibility is that immigrants, especially those who are newly arrived, tend to cluster in gateway cities that function as "ports of entry," such as New York City, Los Angeles, Miami and Chicago. Indeed, in our PSID data, while 42% of immigrants resided in these gateway cities, less than 13% of native-born households lived there. These gateway cities share certain characteristics, such as high crime rates, high population density, sharp house price corrections during the housing crisis, etc., which may cause immigrant households to fall behind on mortgage payments more frequently than do native-born households. Our definition of gateway cities includes New York City, Los Angeles, San Francisco, Phoenix, Miami, Chicago and Philadelphia. The results, after controlling for these gateway cities, are reported in column 3 of Table 3. Again, we do not find evidence that living in gateway cities is correlated with higher delinquency rates, holding everything else equal. Nor does living in gateway cities explain the immigrant-native difference in mortgage delinquency rates: the difference remains at around 2.5 percentage points and is significant at the 5% level.

4.4 Integration of Immigrants into the Host Society

Another plausible explanation as to why immigrants have higher mortgage delinquency rates is that they are not well integrated into the host society, due to differences in culture and language, and their lack of experience and perceptions of belonging. If immigrant households, especially those who are newly arrived, cannot form a strong connection with the host society, then they may have different delinquency propensities.

In this subsection, we examine this issue along two dimensions. The first dimension relates to the duration of the immigrant household's stay in the United States. The second dimension relates to second-generation immigrants, i.e., families formed by children from immigrant households. We report the results in Table 4.

Table 4: The Duration Impact and the Long-Term Impact of Immigration on Mortgage Delinquency
(Estimated by probit models)

	(1)		(2)	
	Marginal effect	Coefficient	Marginal effect	Coefficient
Duration of immigrants' stay in the United States				
10-20 years	0.064	0.676***		
		(0.117)		
21-30 years	0.016	0.248		
		(0.229)		
31-40 years	0.014	0.228		
		(0.341)		
Second generation			-0.004	-0.094
				(0.324)
Other variables controlled				
demographic characteristics		Yes		Yes
socioeconomic status (non-financial factors)		Yes		Yes
socioeconomic status (financial factors)		Yes		Yes
mortgage characteristics		Yes		Yes
Pseudo R-squared		26.7%		23.9%
Number of observations		2383		2224

Note: This table reports marginal effects and coefficients from various probit regressions, along with robust standard errors (clustered at the state level) in parentheses. The dependent variable is an indicator variable of mortgage delinquency. * significant at 10% level, ** significant at 5% level, *** significant at 1% level. The marginal effect of a dummy variable measures the impact of a discrete change of the dummy variable from 0 to 1.

4.4.1 The Duration Impact of Immigration

There is abundant evidence that immigrants assimilate and are more and more likely to resemble natives over time as they accumulate United States experience (Osili and Xie, 2009; and Coulson, 2011). If the immigrant-native difference in mortgage delinquency rates is caused (in part) by imperfect integration of immigrants into the host country, then one would expect the difference to decrease with the duration of the immigrant's stay in the United States: the difference should be larger for newly arrived immigrants than for those who arrived decades ago.

In our data, the duration of stay of immigrant households ranges from 10 to 40 years. Therefore, to test the above hypothesis, we classify immigrant households into the following three duration categories: 10 to 20 years, 21 to 30 years and 31 to 40 years. In part, the issue is to assure enough immigrant households in each category: around 25% (47) of immigrant households have been in the United States for 10 to 20 years, 46% (87) for 21 to 30 years, and 29% (54) for 31 to 40 years. We estimate a Probit model with the following latent variable specification:

$$z_i = \alpha_1\, duration1020_i + \alpha_2\, duration2030_i + \alpha_3\, duration3040_i + \beta\, X_i + \epsilon_i, \qquad (2)$$

where X_i includes all the covariates in Specification 5, and $duration1020_i$, $duration2030_i$ and $duration3040_i$ are three indicator variables of the immigrant's duration of stay. All three indicator variables equal zero for native-born households. We report the coefficients and marginal effects of $duration1020_i$, $duration2030_i$ and $duration3040_i$ in column 1 of Table 4.

The results show that the estimated impact of immigrant status on delinquency is mainly driven by those newly arrived immigrants who have been in the United States for less than 20 years. The estimated delinquency gap between newly arrived immigrants and natives is 6.4 percentage points, compared with the 2.6 percentage points reported in Table 2. The gap is statistically significant at the 1% level. For those immigrants who have resided in the United States for more than 20 years, delinquency gaps are not only much smaller in magnitude (1.6% and 1.4%) but also statistically insignificant. Therefore, no evidence indicates that immigrants who have been in the United States for more than 20 years are different from natives in the incidence of mortgage delinquency.

The above delinquency pattern of immigrants appears to be consistent with other potential theories, one of which is the cohort effect that immigrants who migrated to the United States during different periods are from different regions of the world. For example, it is likely that the early immigrants are mainly from Europe, and the recent immigrants are mainly from Asia, Mexico and South America. Therefore, it is possible that the observed delinquency pattern among immigrant households is driven by where the immigrants originate from (i.e., the cohort effect) instead of by how long they have been in the United States (i.e., the duration effect). To test this hypothesis, we re-estimate Specification (5) by controlling for the continent from which the immigrant households originally came. However, there is no clear pattern in the results to support the cohort effect hypothesis.

Another theory that seems to be consistent with the observed delinquency pattern of immigrant households is the decentralization effect—that immigrants tend to cluster in big metropolitan areas when they first arrive but disperse from the port-of-entry for other, smaller cities as they stay longer in the United States. However, as suggested by the result in Section 4.2, there is no evidence of the decentralization effect.

For other control variables, the results, although not reported, are consistent with earlier findings. High current LTVs and DSRs, adjustable mortgage rates, high mortgage rates, nonrecourse mortgages, less education, unemployment and black ethnicity are all factors that positively contribute to increased delinquency.

4.5 The Long-Term Impact of Immigration

The second dimension along which we analyze the integration effect is related to the second-generation households, i.e., the native-born households formed by children from immigrant households. Second-generation households share similar characteristics with not only their parents (i.e., immigrant households) but also with other native-born households (i.e., third-or-higher-generation households). On one hand, they tend to speak the same languages and have similar religious beliefs as their parents. On the other hand, like other native-born households, second-generation households are often born, raised and educated in the United States. Therefore, like other native-born households, they tend to be well integrated into society. If the immigrant-native difference in delinquency behavior is driven by the integration effect, then one would expect the difference to disappear among second-generation and third-or-higher-generation households.

To identify the delinquency difference between second-generation and third-or-higher-generation households, we estimate the following probit model, using the subsample of native-born households:

$$z_i = \theta_1 \, secgen_i + \theta_2 \, X_i + \epsilon_i, \tag{3}$$

where $secgen_i$ is a second-generation indicator variable that equals 1 for second-generation households, and 0 for third-or-higher-generation households. Immigrant households are excluded in this step of the analysis.

The estimated coefficient and the marginal effect of $secgen_i$ are reported in column 2 of Table 4. The results indicate that the delinquency gap between the second-generation and third-or-higher-generation is negligible (-0.4%) and statistically insignificant. Therefore, consistent with the theory of the integration effect, there is no evidence that the second-generation is different from the third-or-higher generations in mortgage delinquency.

5 Robustness Check

In this section, we use a Heckman-probit model to correct for potential sample selection bias and the propensity score matching (PSM) method to correct for potential function misspecifications. Our main results are robust to both models.

5.1 Heckman Probit Model

We observe delinquency behavior only for mortgage-indebted households. If the subsample of mortgage-indebted households is not a random sample of the entire population of home-owning households, then our previous estimators are likely to suffer from sample selection bias. Indeed, in our data, while 85% of immigrant home-owning households have mortgages, the number for native home-owning households is only 72%, and the difference is significant at the 1% level.

To correct this potential bias, we implement the Heckman probit model, also known as a bivariate probit with sample selection (Cameron and Trivedi, 2005). The Heckman probit model uses the full sample of home-owning households, including both mortgage-indebted and mortgage-free homeowners.

The Heckman probit model is estimated using the maximum likelihood method. Assume that y_1 is the indicator variable of having mortgages, and X_1 is the set of covariates that affect y_1. Also assume that y_2 is the indicator variable of mortgage delinquency conditional on having a mortgage, and X_2 is the set of covariates that effect y_2.[7] There are three types of observations in our sample, with the following probabilities:

$$\begin{aligned} y_1 = 0 &\quad Prob(y_1 = 0) = \Phi(-\beta_1 X_1), \\ y_1 = 1, y_2 = 1 &\quad Prob(y_1 = 1, y_2 = 1) = \Phi_2(\beta_1 X_1, \beta_2 X_2, \rho), \\ y_1 = 1, y_2 = 0 &\quad Prob(y_1 = 1, y_2 = 0) = \Phi(\beta_1 X_1) - \Phi_2(\beta_1 X_1, \beta_2 X_2, \rho), \end{aligned}$$

where Φ is the standard normal cumulative distribution function and Φ_2 is the two-dimensional standard normal cumulative distribution function, with ρ denoting the correlation coefficient

[7] As with the standard Heckman model (Heckman, 1979), we need to include at least one variable in X_1 that does not appear in X_2. This is the so-called "exclusive restriction." To satisfy the exclusive restriction, we include state fixed effects and the house value in X_1 that do not appear in X_2. All of the other covariates are the same in X_1 and X_2.

between the two standard normal distributions. The maximum-likelihood method finds values of β_1, β_2, and ρ to maximize the following joint-likelihood function:

$$\ln(L(\beta_1, \beta_2, \rho)) = \sum_{i=1}^{n} \{y_{1,i} y_{2,i} \ln(\Phi_2(\beta_1 X_1, \beta_2 X_2, \rho)) + \\ y_{1,i}(1 - y_{2,i}) \ln[\Phi(\beta_1 X_1) - \Phi_2(\beta_1 X_1, \beta_2 X_2, \rho)] + (1 - y_{1,i}) \ln(\Phi(-\beta_1 X_1))\}.$$

We report the coefficient (β_2), the marginal effect of X_2, and the correlation coefficient ρ in Table 5. The results reported in column 1 of Table 5 correspond to the specification in column 5 of Table 2. Although the Heckman probit estimator of the delinquency gap is smaller than the corresponding probit regression result (1.2% vs. 2.6%), it is still statistically significant at the 5% level. Note, however, that the estimated ρ is only significant at the 10% level, indicating only weak evidence of sample selection bias.

Table 5: The Impact of Immigrant Status
(Estimated by Heckman probit)

	(1)		(2)		(3)		(4)	
	Marginal effect	Coefficient	Marginal effect	Coefficient	Marginal effect	Coefficient	Marginal effect	Coefficient
Immigrant	0.012	0.356**	0.014	0.387**	0.012	0.339*	0.011	0.335*
		(0.177)		(0.178)		(0.179)		(0.181)
Receiving financial assistance from others			0.013	0.384**				
				(0.158)				
House price growth rate since the last mortgage					-0.0009	-0.004		
						(0.003)		
Living in a gateway city							0.002	0.076
								(0.143)
ρ		-0.368*		-0.359		-0.395*		-0.390*
		(0.201)		(0.207)		(0.235)		(0.234)
Other variables Controlled								
demographic characteristics				Yes		Yes		Yes
socioeconomic status (non-financial factors)				Yes		Yes		Yes
socioeconomic status (financial factors)				Yes		Yes		Yes
mortgage characteristics				Yes		Yes		Yes
Log likelihood		-1463		-1460.086		-1462.424		-1462.878
Number of observations		3287		3287		3287		3287

Note: This table reports marginal effects and coefficients from various Heckman probit regressions, along with robust standard errors in parentheses. The dependent variable is an indicator variable of mortgage delinquency. * significant at 10% level, ** significant at 5% level, *** significant at 1% level. The marginal effect of a dummy variable measures the impact of a discrete change of the dummy variable from 0 to 1.

Columns 2 to 4 of Table 5 correspond to the specifications in columns 1 to 3 of Table 3, respectively. Specifically, we use receipt of monetary assistance to proxy unobservable financial constraints, use house price changes since mortgage origination or since refinance to proxy state housing market conditions, and use the gateway city indicator to proxy for unobservable characteristics of cities where immigrants tend to cluster. Consistent with the previous probit results, the Heckman probit results provide no evidence of either unobservable financial constraints or a local housing market effect.

We next use the Heckman probit model to test the integration effect, by controlling for the duration of stay of the immigrant households and second-generation status. The results are reported in Table 6, which corresponds to the probit regression results reported in Table 4. Similar to the probit regression results, the delinquency gap is not evenly distributed among immigrants: the gap between the newly arrived immigrants and natives is 3.3% and is significant at the 5% level, while for the immigrants who have been in the United States for over 20 years, the gaps are negligible and insignificant. In addition, the delinquency gap between second generations and third-or-higher generations is small (-0.3%) and statistically insignificant. In summary, our main results are robust to possible sample selection bias.

Table 6: The Duration Impact and the Long-Term Impact of Immigration
(Estimated by Heckman probit model)

	(1)		(2)	
	Marginal effect	Coefficient	Marginal effect	Coefficient
Duration of the immigrant's stay in the United States				
10-20 years	0.033	0.658**		
		(0.280)		
21-30 years	0.007	0.234		
		(0.225)		
31-40 years	0.006	0.193		
		(0.290)		
Second generation			-0.002	-0.098
				(0.322)
ρ		-0.396*		-0.347
		(0.236)		(0.214)
Other variables controlled				
demographic characteristics		Yes		Yes
socioeconomic status (non-financial factors)		Yes		Yes
socioeconomic status (financial factors)		Yes		Yes
mortgage characteristics		Yes		Yes
Log likelihood		-1462.031		-1362.102
Number of observations		3287		3099

Note: This table reports marginal effects and coefficients from various Heckman probit regressions, along with robust standard errors in parentheses. The dependent variable is an indicator variable of mortgage delinquency. * significant at 10% level, ** significant at 5% level, *** significant at 1% level. The marginal effect of a dummy variable measures the impact of a discrete change of the dummy variable from 0 to 1.

5.2 Propensity Score Matching Estimation

Both the probit and the Heckman probit models assume linear impacts of covariates on the latent variable. If this assumption is invalid, our previous estimators may be biased, due to function misspecification. To mitigate this concern, we apply the propensity score matching (PSM) method.

The matching estimation is obtained by simply comparing outcomes among units that received a particular treatment versus those that did not. Using terminology from the matching literature, we define immigrant households as the treatment group and native-born households as the comparison group, and we define the impact of immigrant status on the incidence of delinquency as the treatment effect. The basic purpose of matching is to find native-born households that are similar to the immigrant households in all relevant characteristics X, so that the systematic differences in delinquency behavior between these well-selected native-born and immigrant households can be attributed to the treatment effect.

One advantage of matching estimations (compared with regressions) is that the key identifying assumption is weaker: the effect of covariates on the outcome need not be linear, since the matching method estimates the effect by matching households with the same covariates instead of using a linear model for the effect of covariates. However, we should also note that matching is not a magic bullet to solve any unobservable variable bias. Similar to regression, matching is based on the assumption that the source of the selection bias is the set of observed covariates, X. That is, matching estimators would be biased if selection (into immigrant households) was based on unobservable variables.

Finding matches that are similar with respect to all relevant covariates, however, can be difficult if the number of covariates is large. Nevertheless, Rosenbaum and Rubin (1983) prove that matching on the (one-dimensional) propensity score (which is the estimated probability of a household being an immigrant household, given X) suffices to adjust for the differences in the observed covariates under two critical conditions: (i) the *unconfoundedness condition* that the only source of selection bias is the set of observed covariates; and (ii) the *common support assumption* that there is common support (or equivalently, overlapping support) for the propensity score distribution among the treated and untreated groups. Matching on the propensity score is called *propensity score matching*, which is the technique we will use for the following estimation. The key estimator is called the average treatment effect on the treated (ATT), which has a similar interpretation to the marginal effects in the probit and Heckman probit models: it measures the difference in mortgage delinquency rates between immigrant and native households. We test the quality of the matching in Appendix II.

There are various matching algorithms that differ in how the matched native-born (second-generation) households are selected. In this paper, we focus on kernel matching and nearest neighbor (NN) matching with caliper.[8] As in Smith and Todd (2005), we implement the trimming method to determine the region of common support: we drop 5 percent of the treatment observations (immigrant households) for which the propensity score density of the comparison observations (native-born households) is the lowest. The ATTs estimated by the kernel matching and nearest neighbor matching are reported in Table 7, with the kernel matching results in the left-hand panel: column 1 reports the delinquency gap between immigrant and native-born households. Column 2 reports the estimated delinquency gap between newly arrived immigrant households and native-born households, as well as that between long-standing immigrant households and native-born households. Column 3 reports the estimated delinquency gap between second-generation households and third-or-

[8]For the technical details of each matching algorithm, see Imbens (2004), Smith and Todd (2005), and Caliendo and Kopeinig (2008).

higher-generation households. Columns 4 to 6 in Table 7 report the corresponding results from the nearest neighbor matching.

Table 7: The Impact of Immigrant Status and the Duration Impact and Long-Term Impact of Immigration

(Estimated by the propensity score matching method)

	Kernel matching			Nearest neighbor matching		
	(1)	(2)	(3)	(4)	(5)	(6)
Immigrant	0.064**			0.051		
	(0.033)			(0.040)		
Duration of the immigrant's stay in the United States						
10-20 years		0.114*			0.083	
		(0.078)			(0.097)	
more than 20 years		0.046			0.005	
		(0.037)			(0.045)	
Second generation			-0.013			-0.025
			(0.016)			(0.024)
Other variables controlled						
demographic characteristics	Yes	Yes	Yes	Yes	Yes	Yes
socioeconomic status (non-financial factors)	Yes	Yes	Yes	Yes	Yes	Yes
socioeconomic status (financial factors)	Yes	Yes	Yes	Yes	Yes	Yes
mortgage characteristics	Yes	Yes	Yes	Yes	Yes	Yes

* Significant at 10% level, ** significant at 5% level, *** significant at 1% level

Consistent with our previous estimators, results from both matching algorithms show that immigrant households are likely to have a higher mortgage delinquency rate than native-born households. In addition, the immigrant-native delinquency gap is mainly driven by newly arrived immigrant households who have been in the United States for less than 20 years. Second-generation households are no different than third-or-higher-generation households in their delinquency behavior, suggesting that there is no long-term impact of immigration on mortgage delinquency.

6 Conclusion

The United States has had a steady flow of immigration for many decades. The integration of immigrants in the United States, however, is not necessarily a smooth process. It entails uncomfortable adjustments among immigrants and American society. There is a large literature studying the impact of a large influx of immigrants on job markets, crime, social welfare and education. We extend the literature to study the effects of immigration on mortgage delinquency.

We find that immigrants are likely to have a higher delinquency rate than natives, even after controlling for a rich set of household demographic and socioeconomic status and mortgage characteristics. This finding is unlikely to be driven by unobservable financial constraints, local housing market conditions or unobservable characteristics of the metropolitan areas where immigrants tend to cluster. There is evidence of imperfect immigrant integration: the relatively high delinquency rate of immigrants is mainly driven by newly arrived immigrants who have been in the United States for less than 20 years. Immigrants who have resided in the United States for more than 20 years are no different than natives in mortgage delinquency rates. In addition, there is no evidence that the second generation has a higher delinquency rate than the third-or-higher generation. Additional analysis reaffirms that our main results are robust to potential sample selection biases and functional misspecification. Our findings indicate that immigrants as a whole did not play an important role in causing the recent foreclosure crisis. It should be noted that, in this paper, we have defined mortgage delinquency as being behind in the mortgage payment, for at least one month. Our main results are also robust to the definition of more serious mortgage delinquency, as shown in Table A-2.

As we pointed out, mortgage-indebted households were asked about their delinquency status for the first time in 2009 in the PSID survey. This data limitation prevents us from using data from earlier years. That being said, we can still make some predictions for the immigrant-native difference in the early years of the financial crisis by using the PSID data from these years. For example, the 2007 PSID data show that immigrant households tend to have worse financial status and riskier mortgage products than natives. We can deduce that the empirical magnitudes documented in this paper are conservative estimates of the overall effects of immigrants on mortgage delinquency during the recent housing crisis.

Appendix I. Definition of Variables

Head: The head of the household. Relationship to Head is "Head" (code 10) and values of Sequence Number are in the range of 1-20.

Household demographic characteristics

Immigrant: An indicator variable of immigrant households. A sample of 441 post-1968 immigrant families was added in 1997. In 1999, an additional 70 families were added, for a total of 511 immigrant families as of 1999. We use the 1968 family interview number to identify the immigrant households. Immigrant sample families have values greater than 3000 and less than 5000. (Values from 3001 to 3441 indicate that the immigrant family was first interviewed in 1997; values from 3442 to 3511 indicate that the immigrant family was first interviewed in 1999.)

duration1020: An indicator variable of immigrant households who have been in the United States for less than 20 years. This variable equals 0 for other immigrant households and native-born households.

duration2030: An indicator variable of immigrant households who have been in the United States for 21 to 30 years. This variable equals 0 for other immigrant households and native-born households.

duration3040: An indicator variable of immigrant households who have been in the United States for more than 31 years. This variable equals 0 for other immigrant households and native-born households.

Second generation: An indicator variable of second-generation households. Second-generation households are native-born households where at least one of the parents of the head is foreign born.

Age: Age of the head.

Male: An indicator variable of the male head.

Married: An indicator variable of a married head.

White: An indicator variable of the head being white.

Black: An indicator variable of the head being black.

Other race: An indicator variable of the head being any other race.

Region: A categorical variable of the region where the household lived in 2009.

Household socioeconomic characteristics

Less than high school: An indicator variable of the head having less than high school education.

High school: An indicator variable of the head having high school degree but not college degree.

College: An indicator variable of the head having a college degree or higher.

Unemployed: An indicator variable of the head being unemployed.

Employed: An indicator variable of the head being employed.

Not in labor force: An indicator variable of the head not being in the labor force.

Yearly family income: The total family income in the past year.

Income growth rate: The income growth rate. It equals to yearly family income divided by average family income in the past four waves.

Total assets: The dollar amount of a household's assets.

Total debts other than mortgages: The total dollar amount of other debts besides the home mortgages.

Monetary transfers to relatives: The yearly dollar amount of monetary transfers to relatives not living in the households.

Mortgage characteristics

Main mortgage: All the households in our data have at most two mortgages on their homes. The main mortgage is defined to be the delinquent mortgage if there is only one delinquent mortgage. Otherwise, it is the mortgage that has the higher remaining principal.

CLTV: The current combined loan-to-value ratio. It equals the total principals of all mortgages on the household's primary residence divided by the current value of the household's primary residence.

DSR: The current debt-service ratio. The DSR is calculated as the ratio of the sum of the current mortgage principal plus interest payments, property taxes and insurance to current family income.

Mortgage term: The length of time it takes to pay off the entire mortgage. Most mortgages have terms of 15 years or 30 years.

Mortgage age: The length of time the household has been paying the mortgage.

Two mortgages: An indicator variable of two mortgages. This variable equals 0 for households having one mortgage, and 1 for households having two mortgages on their home.

ARM: An indicator variable of adjustable-rate mortgage. It equals 1 if the main mortgage has an adjustable interest rate and 0 if the main mortgage has a fixed interest rate.

Interest rate: The interest rate on the main mortgage.

Refinance: An indicator variable for a refinanced mortgage. It equals 1 if the main mortgage is refinanced and 0 if the main mortgage is the original.

Recourse: An indicator variable for a recourse mortgage.

Appendix II. Matching Quality

One important step in the implementation of propensity score matching is to assess the matching quality. In particular, we need to check the unconfoundedness condition and the common support condition.

II.1 Unconfoundedness Condition

Since we match on the propensity score instead of all covariates, we must ensure that the matching procedure is able to balance the distribution of the relevant variables in both the treatment and comparison groups. A helpful theorem suggested by Rosenbaum and Rubin (1983) states that, after matching on the propensity score, the relevant covariate X should be independent of the immigrant (second-generation) status.

Table A-1 provides three measures of the quality of the matching. First, the upper part of Table A-1 shows that, after matching, the mean differences in household characteristics between immigrant (second-generation) households and native-born (third-or-higher-generation) households become statistically insignificant and economically negligible. For example, before matching, the immigrant-native difference in the proportion of a white head is 0.252 and significant at the 1% significance level (see Table 1); after matching, the difference is much smaller (0.092) and becomes statistically insignificant. This implies that, after matching, the treatment group and comparison group are comparable with respect to the relevant covariates.

Second, the third row from the bottom in Table A-1 reports the pseudo R-squared of a probit regression of the immigrant (second-generation) status on the covariates X. The reported R-squared tells us to what extent the immigrant (second-generation) status is correlated with all of the covariates. After matching, the pseudo R-squared decreases to 3.4 percent (3.2 percent) from 35.4 percent (12.2 percent), as shown in Table 1. This result also suggests that the unconfoundedness condition is likely to hold.

Third, the second row from the bottom reports the p-value of the F-test of the joint significance of all the covariates X in the probit regression of the immigrant (second-generation) status on X. The reported p-value shows that, after matching, the immigrant (second-generation) status is unlikely to have a linear relationship with the covariates. This result also provides evidence in favor of the unconfoundedness condition.

Table A-1: Kernel Matching Quality

	Immigrant	Native	p-value	Second-generation	Third-or-higher generation	p-value
Demographic information						
age	45.612	44.861	0.594	49.958	47.376	0.143
male	0.843	0.824	0.679	0.849	0.848	0.986
married	0.724	0.716	0.888	0.790	0.762	0.606
race						
black	0.090	0.126	0.341	0.034	0.139	0.004
white	0.582	0.549	0.585	0.874	0.798	0.114
others	0.328	0.325	0.958	0.092	0.064	0.410
region						
northeast	0.149	0.124	0.545	0.151	0.149	0.963
north central	0.097	0.162	0.116	0.235	0.260	0.666
south	0.276	0.297	0.711	0.311	0.351	0.510
west	0.478	0.418	0.329	0.294	0.233	0.283
Socioeconomic status (non-financial factors)						
education						
less than high school	0.284	0.208	0.151	0.084	0.072	0.731
high school degree	0.463	0.481	0.770	0.462	0.532	0.287
college degree or higher	0.254	0.312	0.295	0.454	0.396	0.373
employment status						
unemployed	0.082	0.058	0.439	0.050	0.047	0.917
not in labor force	0.060	0.113	0.121	0.134	0.137	0.952
employed	0.858	0.829	0.514	0.815	0.815	0.996
Socioeconomic status (financial factors)						
yearly family income	94994	98334	0.822	1.40E+05	1.40E+05	0.937
income growth rate	0.477	0.667	0.590	0.284	0.326	0.711
total assets	171.820	318.050	0.261	510.600	587.210	0.761
total debts other than mortgages	12654	19417	0.208	12433	15310	0.501
monetary transfers to relatives	1836	985.96	0.484	1865.5	1225.7	0.541
Mortgage information						
CLTV	0.758	0.789	0.581	0.616	0.642	0.527
DSR	0.230	0.263	0.407	0.156	0.156	0.999
VRM	0.134	0.118	0.692	0.118	0.095	0.581
interest rate	6.191	6.190	0.995	5.983	6.023	0.831
mortgage term	24.970	25.062	0.920	23.345	23.572	0.827
mortgage age	4.179	4.024	0.736	4.538	4.542	0.993
refinance	0.485	0.464	0.729	0.437	0.438	0.985
two mortgages	0.134	0.168	0.444	0.185	0.163	0.660
recourse	0.485	0.398	0.153	0.261	0.248	0.830
Pseudo R-squared	6%			5%		
P-value	0.785			0.934		
Number of observations	134	2219		119	2097	

II.2 Common Support Condition

The PSM estimator is only defined in the region of common support. Heckman, Ichimura and Todd (1997) note that matching an incomparable comparison group to the treated group is a major source of evaluation bias as conventionally measured. Only the subset of the native households that is comparable with immigrant households should be used in the analysis (Dehejia and Wahba, 1999.) Hence, an important step is to check the overlap and the region of common support between the immigrant and native-born households as well as between the second-generation and third-or-higher-generation households.

The most straightforward way to check the common support defined by the trimming method is a visual analysis of the density distribution of the propensity score for both the treatment and comparison groups.

Figure 1 displays the propensity score distribution of the immigrant households on the top, and that of the native-born households on the bottom. Similarly, Figure 2 displays the propensity score distribution of the second generation on the top and that of the third-or-higher generation on the bottom. Both figures show evidence of overlapping of the propensity score distributions, especially for the matching between second-generation and third-or-higher-generation households, as shown in Figure 2.

In summary, various measures of matching quality show that our PSM estimation with kernel matching and the trimming method is reliable. In addition, our empirical results are also robust to other matching algorithms, such as nearest neighbor matching.

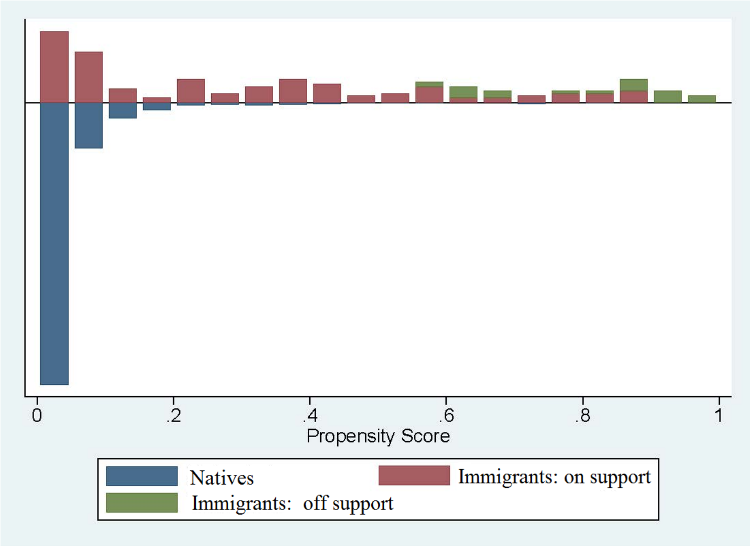

Figure 1: The propensity score distributions of immigrant and native households

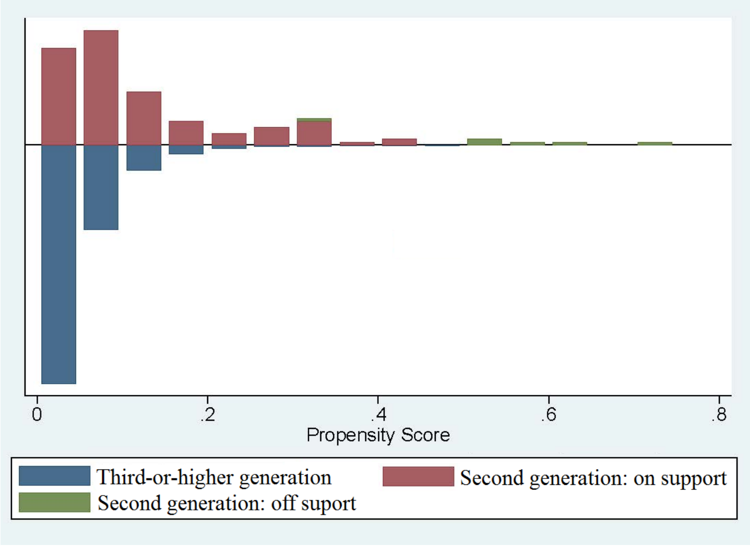

Figure 2: The propensity score distributions of second-generation and third-or-higher-generation households

Table A-2: Results by Using the Definition of More Serious Delinquency

	Probit						Heckman probit						PSM (kernel matching)		
	(1)		(2)		(3)		(4)		(5)		(6)		(7)	(8)	(9)
	mfx	coef.	mfx	coef.	mfx	coef.	mfx	coef.	mfx	coef.	mfx	coef.	mfx	mfx	mfx
Immigrant	0.017	0.410**					0.012	0.369**					0.038		
		(0.164)						(0.180)					(0.027)		
Duration of the immigrant's stay															
10-20 years			0.043	0.706***					0.034	0.681**				0.101*	
				(0.214)						(0.267)				(0.068)	
21-30 years			0.008	0.215					0.005	0.188				0.026	
				(0.249)						(0.240)				(0.030)	
31-40 years			0.016	0.383					0.010	0.306					
				(0.349)						(0.293)					
Second Generation					-0.008	-0.624**					-0.007	-0.644*			-0.015*
						(0.242)						(0.356)			(0.009)
Other variables controlled															
demographic characteristics		Yes		Yes		Yes		Yes		Yes		Yes	Yes	Yes	Yes
socioeconomic status		Yes		Yes		Yes		Yes		Yes		Yes	Yes	Yes	Yes
mortgage characteristics		Yes		Yes		Yes		Yes		Yes		Yes	Yes	Yes	Yes

* Significant at 10% level, ** significant at 5% level, *** significant at 1% level. The marginal effect of a dummy variable measures the impact of a discrete change in the dummy variable from 0 to 1.

7 References

Borjas, G. (2002), "Homeownership in the Immigrant Population," *Journal of Urban Economics*, 52(3), 448-476.

Caliendo, M. and Kopeinig, S. (2008), "Some Practical Guidance for the Implementation of Propensity Score Matching," *Journal of Economic Surveys*, 22, 31-72.

Callis, R. (1997), "Moving to America—Moving to Homeownership," *Current Housing Reports*, H121/97-2: United States Census Bureau.

Cameron, A. and Trivedi, P. (2005), "Microeconometrics: Methods and Applications," New York, Cambridge University Press.

Campbell, J. and Cocco, J. (2012), "A Model of Mortgage Default," *NBER working paper* No. 17516.

Corwin, R. (2012), "How Immigration Is Impacting the U.S. Social Economy," available at http://www.socialissues.us/16301.html

Coulson, E. (2011), "Immigration and Housing," *Encyclopedia of Housing*, Second Edition, Sage Publishers.

Dail, P. (2009), "Immigration and Migration in America: Social Impact and Social Response," *International Migration*, 26(4), 441-452.

Day, J. (1996), "Population Projections of the United States by Age, Sex, Race, and Hispanic Origin: 1995 to 2050," United States. Bureau of the Census, Current Population Reports, 25-1130, U.S. Government Printing Office, Washington, D.C.

Dehejia, R. and Wahba, S. (1999) "Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs," *Journal of the American Statistical Association*, 94(448), 1053-1062.

Deng, Y., Quigley, J. and Van Order, R. (2000), "Mortgage Terminations, Heterogeneity and the Exercise of Mortgage Options," *Econometrica*, 68(2), 275-307.

Foote, C., Gerardi, K. and Willen, P. (2008), "Negative Equity and Foreclosure: Theory and Evidence," Public Policy Discussion Paper from Federal Reserve Bank of Boston.

Friedman, S., Rosenbaum, E. and Schill, M. (1998), "The Housing Conditions of Immigrants in New York City," Working paper, Fannie Mae Foundation.

Ghent, A. and Kudlyak, M. (2011), "Recourse and Residential Mortgage Default: Evidence from US States," *Review of Financial Studies*, 24, pp. 3139-3186.

Haughwout, A., Peach, R., and Tracy, J. (2008), "Juvenile Delinquent Mortgages: Bad Credit or Bad Economy?" *Journal of Urban Economics*, 64(2), 246-257.

Heckman, J., (1979), "Sample Selection Bias as a Specification Error," *Econometrica*, 47(1), pp. 153-161.

Heckman, J., Ichimura, H. and Todd, P. (1997), "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program," *Review of Economic Studies*, 64(4), pp. 605-654.

Hu, W. (2000), "Immigrant Earnings Assimilation: Estimates from Longitudinal Data," *The American Economic Review*, 90(2), 368-372.

Imbens, G. (2004), "Nonparametric Estimation of Average Treatment Effects under Exogeneity: A Review," *Review of Economics and Statistics*, 86(1), 4-29.

Jimenez, T.R. (2011), "Immigrants in the United States: How Well Are They Integrating into Society?" Washington, DC: Migration Policy Institute.

Joint Center for Housing Studies of Harvard University (JCHS) (2000), "The State of the Nation's Housing 2000," Cambridge: Harvard University.

Kochhar, R., Gonzalez-Barrera, A. and Dockterman, D. (2009), "Through Boom and Bust: Minorities, Immigrants and Homeownership," Pew Hispanic Center Working Paper.

Lin, Z., Rosenblatt, E. and Yao, V. (2009), "Spillover Effects of Foreclosures on Neighborhood Property Values," *Journal of Real Estate Finance and Economics*, 38, 387-407.

Malkin, M. (2008), "Kill the Bailout: Illegal Immigration and the Mortgage Mess," available at http://michellemalkin.com/2008/09/24/illegal-immigration-and-the-mortgage-mess/

Mayer, C., Pence, K. and Sherlund, S. (2009), "The Rise in Mortgage Defaults," *Journal of Economic Perspectives*, 23 (1), 27-50.

Myers, D., Baer, W. and Choi, C. (1996), "The Changing Problem of Overcrowded Housing," *Journal of the American Planning Association*, 62, 66-84.

Myers, D. and Park, J. (1999), "The Role of Occupational Achievement in Homeownership Attainment by Immigrants and Native Borns in Five Metropolitan Areas," *Journal of Housing Research*, 10 (1), pp. 61-93.

Osili, U. and Xie, J. (2009), "Do Immigrants and Their Children Free Ride More Than Natives?" *The American Economic Review*, 99(2), 28-34.

Rosenbaum, P. and Rubin, D. (1983), “The Central Role of the Propensity Score in Observational Studies For Causal Effects,” *Biometrika*, 70 (1), pp. 41-50.

Smith, J. and Todd, P. (2005), “Does Matching Overcome Lalonde’s Critique of Nonexperimental Estimators?” *Journal of Econometrics*, 125, pp. 305-353.

Vandell, K. (1978), “Default Risk Under Alternative Mortgage Instruments,” *The Journal of Finance*, 33 (5), 1279-1296.

Zandi, M. (2009), “Financial Shock: Global Panic and Government Bailouts: How We Got Here and What Must Be Done to Fix It,” Upper Saddle River, NJ: Pearson Education.