The Effect of Neighborhood Contagion on Mortgage Selection [*]

Rubén Hernández-Murillo [†]   Rajdeep Sengupta [‡]

October 2011

Abstract

In this paper we conduct an empirical investigation of how neighborhood mortgage adoption contagion affects mortgage product choice, with an emphasis on Hispanic borrowers.

We use loan-level mortgage data for metropolitan areas in California and Florida during 2004 and 2005, the peak years of the subprime mortgage boom.

We identify an important and statistically significant effect of contagion on consumer choice of hybrid mortgage products that were popular during this period, especially for Hispanic borrowers.

JEL Codes: G21, J15, R23, R31.
Keywords: Subprime mortgages, HMDA data, Hispanic population.

1   Introduction

In this paper we analyze the role of externalities in the choice of subprime mortgage products. In particular, we focus on whether high adoption rates of certain non-traditional mortgage (NTM) products at a given location affected the origination of similar products in subsequent periods.[1] We term this phenomenon as a *neighborhood contagion effect*. In addition, we examine the question of whether the race and ethnicity of the borrower is an important driver of such contagion effects.

The boom in subprime mortgages since the early 2000s has been the topic of much interest and research (see GAO, 2010 for an overview). Mayer, Pence, and Sherlund (2009) observe that an overwhelming majority—over 75 percent—of subprime mortgages originated over the 2003-2007 period were so-called short-term hybrids. Short-term hybrids are mortgage products that combine the (initial) payment stream of a fixed rate mortgage (FRM) that later resets into the payment stream of an adjustable rate mortgage (ARM). Subprime mortgage products 2/28 and 3/27 belong to this category. The 2/28 is a 30-year mortgage product that comprises an initial FRM for two years followed by an ARM for 28 years. The 3/27 is defined analogously. Notably, the ARM-leg of most of these mortgage products is scheduled to reset every 6 months.[2]

Naturally, short-term hybrid products are significantly different from traditional mortgage products. This adds to a widespread perception among academics and policymakers that most subprime mortgages were originations to borrowers with a lower degree of financial sophistication and therefore less likely to have full knowledge and understanding of the terms and conditions in their mortgage contracts. Several studies supporting this view have attributed these phenomena to the confusion about mortgage terms (Bucks and Pence, 2008),

---

[1]Conventional 30-year single-family mortgages which qualify under the underwriting criterion of Fannie Mae and Freddie Mac are generally classified as traditional mortgages.

[2]Arguably, this is a key difference between subprime hybrids and prime hybrids such as the 3/1, 5/1 and 7/1 mortgages, wherein the ARM-leg of the mortgage resets on an annual basis. Subprime hybrids reset at shorter maturities because they were designed to reduce lender exposure to long term contracts. Repeated resets were intended to *force* the borrower into refinancing the mortgage (see Gorton 2008; Bhardwaj and Sengupta, 2011, for details).

the lack of financial literacy among borrowers (Gerardi, Goette, and Stephan, 2010) and even the lack of proper disclosure of mortgage costs (Lacko and Pappalardo, 2007). These findings also show that numerical and cognitive ability does not always systematically predict riskier mortgage choices. Therefore, at least on the borrower side, it remains an open question as to why so many subprime originations included non-traditional mortgage products, which expanded at such a rapid pace.

Several explanations have been offered for the rapid expansion of NTMs or short term hybrids. It is important to note that hybrid mortgage products have been in existence since the 1980s and have been used in other segments of the mortgage market. But even in the jumbo and conventional mortgage segments, the rapid growth of hybrids has been a fairly recent phenomenon. Bhattacharya, Fabozzi, and Berliner (2005) note that the growth of hybrids is often attributed to a sharp increase in the steepness of the yield curve around 2002.[3] The demand for NTMs was also sustained by the strong growth in home prices since 2003 and the declining affordability of home seekers—some of whom used NTMs to expand their home-buying power (Agnell and Rowley, 2006). Studies of the subprime market have also noted that a segmentation of the mortgage market in terms of racial and ethnic differences resulted in part because lenders specialized across neighborhoods and marketed certain types of mortgage products more aggressively to minority borrowers, placing them at a higher risk of obtaining subprime loans (Bocian, Ernst, and Li 2006, 2008; Reid and Laderman 2009). While there is strong evidence in favor of each of the explanations described above, they do not reconcile the observed evidence of poor financial sophistication of subprime borrowers with the rapid expansion of riskier NTM mortgages.

In this paper, we seek to explain the rapid growth in short term hybrids among subprime originations in terms of a contagion effect.[4] Using loan-level data on subprime mortgage

---

[3]The 10-year/2-year spread that averaged -22 basis points in 2000 increased to an average in excess of 200 basis points around 2002. This raised the appeal of hybrid products making home buying more affordable. The borrowers attain a security of fixed payment rates for a short period of 2-3 years at rates significantly lower than those of 30-year fixed-rate mortgages.

[4]Again, we emphasize that this is perhaps not the only reason behind the expansion of NTMs in subprime mortgage. However, as we demonstrate below, there is strong evidence in support of this phenomenon.

originations we find that the probability of originating a hybrid product increases with the share of hybrids originated (as a proportion of total originations) at the same location in the previous period. Moreover, we find that this strong increase is significantly tied to the ethnicity of the borrower—especially if the borrower is Hispanic. Stated differently, we find that the probability of originating a hybrid product given that the borrower is Hispanic increases significantly with the share of hybrids originated to Hispanics (as proportion of total originations to Hispanics) at the same location in the previous period.

There are several strands of literature that describe the phenomenon observed in our study. The earliest mention of the spatial concentration of economic activity and related knowledge spillovers and information flows has been attributed to Marshall (1920). More recently, George and Waldfogel (2003), Waldfogel (2003, 2005) have shown how product choice and product diversification can be influenced by the size of the local minority population. The effects of agglomeration on consumer activity can be robust especially in environments where language and ethnic barriers form impediments. In particular, proximity to neighborhoods with a high concentration of home owners of the same ethnicity facilitates the diffusion of knowledge of the mortgage application process, likely through expanded word-of-mouth networks as well as local civic or religious organizations. This feature has been shown to be critical for Hispanic households facing mortgage and housing decisions (see Haurin and Rosenthal, 2009, and references therein). In addition, Haurin and Rosenthal (2009) also point to peer group and role model effects that might influence a household's financial decision. These effects bear relevance for our results although our study does not distinguish between the effects of agglomeration and role model effects.

At the heart of these explanations is the phenomenon whereby the actions of one set of economic decision makers influence the purchases and financial decisions of others. This form of observational or *social learning* can dominate private information about products and contracts leading consumers to act as if they are in an *informational cascade* (Bikhchandani, Hirshleifer, and Welch, 1998). Since empirical evidence on observational learning is

significantly difficult to obtain, efforts in the literature have concentrated on randomized natural field experiments as in Cai, Chen, Fang, and Zhou (2009). While our study alludes to the fact that individual borrower decisions could be affected by observing others' choices, it cannot establish the existence of observational learning.

Our study uses a discrete choice model to determine how individual and neighborhood characteristics determine the choice of mortgage contracts. There are several studies that analyze mortgage choice in this framework. Using micro data from the 1983 Survey of Consumer Finances, Gabriel and Rosenthal (1991) suggest that minority households are significantly less likely to obtain conventional home loans than white borrowers, even after controlling for proxies of default risk. Hendershott, LaFayette, and Haurin (1997) in turn analyzed the joint choice of loan-to-value ratios and mortgage type using data from the 1984 American Housing Survey. They focused on the role of Federal Housing Administration (FHA) financing in easing mortgage qualification constraints. The authors did not analyze credit or demographic factors. Pennington-Cross and Nichols (2000) introduced borrower credit history in the analysis, matching data reported under the Home Mortgage Disclosure Act (HMDA) with data from a credit bureau. They found that credit history plays an important role in the FHA vs. conventional choice, and while they did not find that black borrowers were less likely to select FHA-mortgages than other borrowers, the authors found that Hispanics were more likely to use FHA than white borrowers.

More recent studies of mortgage choice have focused on the choice of prime vs. nonprime mortgages. Courchane, Surette, and Zorn (2004) used a survey of prime and *subprime borrowers* (defined as those with higher credit risk) to study whether subprime borrowers were being channeled to the subprime segment. They found that subprime borrowers were less knowledgable about the mortgage process and were less likely to search for the best mortgage rates. Their results indicate that in addition to typical underwriting factors, Hispanic ethnicity and age are important determinants of market segmentation, and that Hispanic borrowers and older borrowers may disproportionately find themselves in the subprime market.

To the best of our knowledge, ours is one of the first studies to analyze mortgage product choice during the recent subprime lending boom. The study closest in spirit to ours is Reid and Laderman (2009), which also analyzes mortgage choice during the same period, but their focus is foreclosure outcomes across racial or ethnic groups and they do not consider neighborhood and contagion effects on the choice of mortgage products. Reid and Laderman (2009) use HMDA data merged with a proprietary dataset on loan performance collect by LPS Applied Analytics, Inc., for the period 2004 to 2006 for California.[5] They examine the choice between four mortgage products: prime fixed rate mortgages, prime adjustable rate mortgages, subprime fixed rate mortgages, and subprime adjustable rate mortgages. In the choice decision, the authors emphasize the role of different market channels. Their results indicate that black and Hispanic borrowers in California had access to markedly different markets than comparable white borrowers and this market segmentation played an important role in their likelihood of receiving a higher-priced loan. The authors also study the differences in foreclosure rates across racial or ethnic groups and find that after controlling for the choice of mortgage product, the differences in foreclosures fall considerably, suggesting that the disparities primarily result from differential access to lending markets.

In the next section 2, we describe the data sets used in the analysis. In section 3 we present the methodology. We discuss the results in section 5, and we provide concluding comments in section 6.

2   Data

In this study, we use two sources of mortgage data. We combine data reported under HMDA with data on private-label asset backed securities from Core Logic Information Solutions, Inc. (CL), for California and Florida during 2004 and 2005.

The HMDA data contain limited information on the characteristics of the loan, including

---

[5]The LPS dataset is also referred to as the *McDash data*, since McDash Analytics was the original compiler.

the amount, type (conventional or government-backed), purpose (purchase, home improvement, or refinancing), and lien status (first or subordinate). Most of the observations correspond to loans originated in metropolitan areas, as only lenders with assets over a certain threshold are required to report data on the loans they originate or purchase. The data set also contains information on the race and ethnicity of the applicants. In this study we focus on Hispanics, who can be of any race. HMDA data do not include information about the credit-risk profile of the borrowers other than their income. Finally, the data also contain information on the property associated with the mortgage, such as whether it is owner occupied and the Census tract and metropolitan area of its location.

The CL data contain extensive information on the characteristics of loans that were securitized in private-label subprime pools. The information on loan characteristics include, among other variables, the interest rate, the mortgage product type, the terms of the loan (including information on amortization terms and the reset periods of adjustable rate mortgages), and whether the loan has a prepayment penalty or the borrower has been required to acquire private mortgage insurance. The CL data also provide information on the borrower, including the FICO credit score, the borrower's reported debt-to-income ratio, and the extent to which the borrower's income was documented. The information on the property includes the sale or appraised price, the type of property, and the state and ZIP code of its location.

The HMDA and CL data sets were matched following the algorithm described by Haughwout, Mayer, and Tracy (2009). The procedure uses the loan originator names, the loan amounts, as well as the dates of originations, and also involves associating ZIP codes with Census tracts to establish search areas of HMDA loans for each loan in CL.

3   Methodology

3.1   Mortgage choice

In this section we describe our methodology for determining the borrowers' choice of mortgage product. Our motivation here is to examine the determinants of the choice of hybrid mortgage products, defined above as 2/28 and 3/27 mortgages. As mentioned earlier, although hybrids (such as the 3/1, 5/1 and 7/1) have been in existence in other segments of the U.S. mortgage market, the 2/28 and 3/27 products in particular were unique to the subprime product universe.

We model the mortgage choice problem with a standard probit regression. We assume that the difference between the benefit and cost of adopting a hybrid product over all other product types can be denoted in terms of an unobserved variable y^*. Therefore the net utility of adopting a hybrid product for a borrower i in period t in Census track k is described as

$$y^*_{ikt} = \mathbf{x}'_{it}\beta + \mathbf{w}'_{kt}\delta + \gamma z_{kt-1} + u_{ikt} \tag{1}$$

where $\mathbf{x}_{it}$ is a set of contemporaneous borrower and loan characteristics, $\mathbf{w}_{kt}$ is a set of geographic characteristics, and z_{kt-1} is a scalar measure of contagion defined at the Census tract level. Following standard conventions of the index function model, we define our probit estimation procedure in terms of observable covariates as

$$\Pr[y_{ikt} = 1 | \mathbf{x}_{it}, \mathbf{w}_{kt}, z_{kt-1}] = \Phi(\mathbf{x}'_{it}\beta + \mathbf{w}'_{kt}\delta + \gamma z_{kt-1}), \tag{2}$$

where Φ denotes the standard normal cdf.

The vector $\mathbf{x}$ includes borrower and loan characteristics, including the FICO credit score, the loan-to-value ratio, the debt-to-income ratio, the loan amount, an indicator variable for whether the loan is a purchase or refinance, and an indicator variable for whether the borrower's income is fully documented. Additionally, the vector includes an indicator variable

x^H that takes a value of 1 if the borrower is Hispanic and 0 otherwise.

The vector **w** defines a set of geographic and economic characteristics, which includes an indicator variable for loans originated in Florida, the median family income of the Census tract relative to the median income of the corresponding metropolitan area, the share of the minority population in the Census tract, and variables summarizing local economic conditions such as the county-level unemployment rate and the county-level growth rate in house prices since the previous quarter.

Our focus of interest centers on the contagion measure, reflecting the adoption rate of hybrid products at the same location in the previous period. A positive effect of this measure on subsequent adoption rates of hybrid mortgage products is viewed as evidence of contagion.

Let Q_{kt} denote the set of all subprime originations in our sample that were originated in Census tract k during period t. We denote q_{kt} as an element of this set. Next, we define the subsets $Q_{kt}^{hybrid} \subset Q_{kt}$ as the set of hybrid mortgages among Q_{kt} and $Q_{kt}^H \subset Q_{kt}$ as the subset of mortgages originated to Hispanics. Using the above notation we define two measures of the contagion as follows

$$z_{k,t-1}^1 = \frac{\#(q_{kt-1} \in Q_{kt-1}^{hybrid})}{\#(q_{kt-1} \in Q_{kt-1})} \tag{3}$$

$$z_{k,t-1}^2 = \frac{\#[(q_{kt-1} \in Q_{kt-1}^{hybrid}) \cap (q_{kt-1} \in Q_{kt-1}^H)]}{\#(q_{kt-1} \in Q_{kt-1}^H)} \tag{4}$$

The first contagion measure, $z_{k,t-1}^1$, is the fraction of hybrid originations as a proportion of total originations in a Census tract in the previous quarter. A positive effect of the first measure on the probability of choosing a hybrid mortgage product in the next quarter is viewed as evidence of contagion in product choice. Notably, the first measure of contagion is independent of the race or ethnicity of the borrower.

The second contagion measure, $z_{k,t-1}^2$, is the fraction of hybrid originations to Hispanic borrowers as a proportion of total originations to Hispanic borrowers in a Census tract in the previous quarter. Again, a positive effect of this measure on the probability of adopting

a hybrid product in the subsequent period is viewed as evidence of contagion. Notably, the second measure is particularly relevant for Hispanic borrowers.

We are interested in capturing the idea that the measures of contagion are reinforced by the ethnicity of the borrower especially if the borrower is Hispanic. To capture this effect, we introduce an interaction term to the regression as follows. In the model of discrete choice we allow for the interaction of the contagion measure z_{kt-1} with the ethnicity indicator, x^H.

$$\Pr[y_{ikt} = 1 | \mathbf{x}_{it}, \mathbf{w}_{kt}, z_{kt-1}] = \Phi(\mathbf{x}'_{it}\beta + \mathbf{w}'_{kt}\delta + \gamma z_{kt-1} + \eta z_{kt-1} \times x^H). \tag{5}$$

Our object of interest is the marginal effect of the interaction term, which we compute according Ai and Norton (2003) and Norton, Wang, and Ai (2004). These studies show that the sign of the coefficient for the interaction term, η, does not necessarily reflect the sign of the interaction effect. Furthermore, standard summary measures of marginal effects are inappropriate and instead, the marginal effect of the interaction should be computed for each observation in the sample.[6]

4   Summary statistics

In this section, we provide an overview and summary statistics of the merged data set. Our sample comprises of a little over 1.2 million individual first-lien mortgages originated in California and Florida from 2004:Q2 to 2005:Q4 (the initial quarter, 2004:Q1, is used in the calculation of the contagion measures). Table 1 presents summary statistics for the entire sample of loans, table 2 presents similar statistics for the sample of hybrid loans, and, finally, table 3 presents summary statistics for the sample of non-hybrid loans. All tables

---

[6] Standard marginal effects with respect to variable x are computed according to the formula, $\frac{\partial \Pr[y=1|X]}{\partial x} = \frac{\partial \Phi(X'\beta)}{\partial x}$, if variable x is continuous, or as the discrete change in probability, $\frac{\Delta \Pr[y=1|X]}{\Delta x} = \{\Phi(X'\beta)|_{x=1} - \Phi(X'\beta)|_{x=0}\}$, if variable x is binary. The marginal effect of the interaction of a discrete variable x^H with a continuous variable x, however, is computed as the following cross-partial effect $\frac{\Delta \partial \Pr[y=1|X]}{\Delta x^H \partial x} = \frac{\Phi(X'\beta)}{\partial x}\Big|_{x^H=1} - \frac{\partial \Phi(X'\beta)}{\partial x}\Big|_{x^H=0}$. Standard errors of all these effects are obtained with the delta method, as usual.

additionally report summary statistics for the subsamples of loans originated to Hispanic and non-Hispanic borrowers.

The last three columns of table 1 provide us with the details of the entire sample of originations. The average loan amount is around $300,000 dollars. The mean FICO score is around 650 and the mean loan-to-value ratio at 76% is less than the 80% required on conventional mortgages without private mortgage insurance. More than half (58%) of the originations are refinances and little less than half (46%) of the loans provide full documentation on the loan application. For the purposes of this study it is important to note that about 57% of originations in our sample are hybrid products and 29% of originations are to Hispanic households.

From table 1, we observe that, on average, originations of hybrid loans to Hispanics (65%) make for a higher share of the hybrid products than in the overall sample (57%). On average originations to Hispanics in our sample appear to have marginally lower FICO scores, marginally higher loan-to-value ratios, higher debt-to-income ratios and lower levels of full documentation on the mortgage. This suggests that originations to Hispanics have ex-ante riskier characteristics than other originations.

Next, we compare originations of hybrid products in table 2 with originations of other (non-hybrid) products in table 3. Notably, 33% of hybrid products are originated to Hispanic borrowers whereas the share of non-hybrids among Hispanics is only 23%. Originations with hybrid products have a significantly lower FICO scores on average—there is almost a 60 point difference in the mean scores of hybrid originations vis-à-vis non-hybrid originations. In comparison with non-hybrid originations, hybrid originations have higher loan-to-value ratios and higher debt-to-income ratios. At the same time, hybrid originations also have a higher percentage of loans that are refinances or have full documentation on the mortgage. The loan amount is typically lower on hybrid originations.

Table 2 also demonstrates that among hybrid mortgage originations, there is not much to distinguish between borrowers that are Hispanic with those who are non-Hispanic. Most

mortgage and borrower characteristic such as FICO scores, loan-to-value ratios, debt-to-income ratios are comparable. However, we do observe that Hispanic borrowers have a smaller proportion of loans that are refinances and have full documentation on the mortgage. Almost in the same fashion, table 3 shows that within non-hybrid mortgage originations, there is not much distinguishing borrowers who are Hispanic from those who are non-Hispanic.

5 Results

Tables 4 and 5 present the results of the estimation. Table 4 presents the results for the specifications with the first contagion measure, defined as the count of hybrid products as a share of all mortgage originations in the Census tract. Table 5 presents the results with the second contagion measure, defined as the count of hybrid products originated for Hispanics as a share of all mortgage originations for Hispanics in the Census tract. The tables report the estimated coefficients and the average marginal effects (calculated for all variables including the interaction term averaging the measures defined in footnote 6 over all observations) for various specifications with different sets of regressors.[7]

In addition to the contagion measure, the Hispanic indicator, and the interaction of the two, we consider additional regressors according to 3 specifications. Model 1 includes borrower and loan characteristics and the following geographic characteristics: an indicator for Florida, Census tract income, and Census tract minority share. Model 2 includes additionally house price growth and unemployment rates measured at the county level. Finally, model 3 includes also quarterly dummies.

The estimated coefficients in tables 4 and 5 indicate that a standard set of borrower- and loan-level characteristics have the expected sign on the probability of adoption of hybrid products. For example, a higher FICO score, a higher origination amount, and whether the

---

[7]In the computation of the average marginal effects, the derivatives take into account the interaction term.

Table 1: Summary statistics. All mortgage products

	Hispanic=1			Hispanic=0			Total		
	N	mean	sd	N	mean	sd	N	mean	sd
Hybrid	347,963	0.65	0.48	870,879	0.54	0.50	1,218,842	0.57	0.50
Hispanic	347,963	1	0	870,879	0	0	1,218,842	0.29	0.45
Z_k	320,964	0.63	0.16	788,380	0.56	0.19	1,109,344	0.58	0.19
Z_h	308,019	0.67	0.22	643,779	0.64	0.30	951,798	0.65	0.28
FICO	347,963	649.61	65.39	870,879	652.82	72.01	1,218,842	651.90	70.20
Loan-to-Value (%)	347,963	77.55	12.16	870,879	76.08	13.23	1,218,842	76.50	12.95
Debt-to-Income (%)	347,963	29.15	19.71	870,879	26.11	20.08	1,218,842	26.98	20.02
Loan amount (thou.)	347,963	270.21	140.63	870,879	311.92	210.52	1,218,842	300.01	194.08
Full Doc	347,963	0.40	0.49	870,879	0.48	0.50	1,218,842	0.46	0.50
Refinance	347,963	0.51	0.50	870,879	0.61	0.49	1,218,842	0.58	0.49
Florida	347,963	0.27	0.44	870,879	0.32	0.47	1,218,842	0.31	0.46
Tract income (%)	347,800	85.03	31.01	870,617	98.53	36.39	1,218,417	94.67	35.47
Tract minority (%)	347,925	61.86	26.03	870,657	40.57	26.36	1,218,582	46.65	27.97
HPI growth (%)	321,258	6.07	1.97	787,626	5.95	2.00	1,108,884	5.99	1.99
Unemployment (%)	321,328	5.81	1.86	788,998	5.40	1.58	1,110,326	5.52	1.67

Hybrid is an indicator that takes the value of 1 if the origination is a 2/28 or 3/27 mortgage.
Z_k is the count of hybrid products as share of all mortgages at the Census tract level.
Z_h is the count of hybrid products for Hispanics as share of all mortgages for Hispanics at the Census tract level.
Tract Income is the Census tract median family income relative to metropolitan median family income.
Tract Minority is the Census tract share of nonwhite population.
HPI growth is the period-to-period growth in the CoreLogic county-level house price index computed at the end-of-period quarter of origination.
Unemployment is the county-level unemployment rate for the end-of-period quarter of origination.

13

Table 2: Summary statistics. Hybrid mortgage products

	Hispanic=1			Hispanic=0			Total		
	N	mean	sd	N	mean	sd	N	mean	sd
Hybrid	225,836	1	0	468,363	1	0	694,199	1	0
Hispanic	225,836	1	0	468,363	0	0	694,199	0.33	0.47
Z_k	208,273	0.64	0.15	420,288	0.60	0.17	628,561	0.62	0.17
Z_h	200,982	0.68	0.21	355,653	0.66	0.28	556,635	0.67	0.26
FICO	225,836	635.70	63.39	468,363	623.42	64.92	694,199	627.42	64.68
Loan-to-value (%)	225,836	79.74	10.48	468,363	79.21	11.44	694,199	79.39	11.14
Debt-to-income (%)	225,836	32.25	18.76	468,363	30.83	19.22	694,199	31.29	19.08
Loan amount (thou.)	225,836	260.72	122.44	468,363	270.76	152.61	694,199	267.49	143.57
Full Doc	225,836	0.44	0.50	468,363	0.55	0.50	694,199	0.51	0.50
Refinance	225,836	0.45	0.50	468,363	0.58	0.49	694,199	0.54	0.50
Florida	225,836	0.28	0.45	468,363	0.35	0.48	694,199	0.33	0.47
Tract income (%)	225,751	83.38	29.11	468,248	93.49	31.99	693,999	90.20	31.44
Tract minority (%)	225,817	62.43	25.80	468,273	43.15	26.99	694,090	49.43	28.10
HPI growth (%)	208,444	6.11	1.95	419,747	6.05	1.98	628,191	6.07	1.97
Unemployment (%)	208,486	5.85	1.89	420,510	5.49	1.65	628,996	5.60	1.74

Hybrid is an indicator that takes the value of 1 if the origination is a 2/28 or 3/27 mortgage.
Z_k is the count of hybrid products as share of all mortgages at the Census tract level.
Z_h is the count of hybrid products for Hispanics as share of all mortgages for Hispanics at the Census tract level.
Tract Income is the Census tract median family income relative to metropolitan median family income.
Tract Minority is the Census tract share of nonwhite population.
HPI growth is the period-to-period growth in the CoreLogic county-level house price index computed at the end-of-period quarter of origination.
Unemployment is the county-level unemployment rate for the end-of-period quarter of origination.

Table 3: Summary statistics. Non-hybrid mortgage products

	Hispanic=1			Hispanic=0			Total		
	N	mean	sd	N	mean	sd	N	mean	sd
Hybrid	122,127	0	0	402,516	0	0	524,643	0	0
Hispanic	122,127	1	0	402,516	0	0	524,643	0.23	0.42
Z_k	112,691	0.60	0.17	368,092	0.52	0.20	480,783	0.54	0.20
Z_h	107,037	0.64	0.23	288,126	0.61	0.32	395,163	0.62	0.30
FICO	122,127	675.32	61.11	402,516	687.02	64.36	524,643	684.30	63.81
Loan-to-value (%)	122,127	73.49	13.89	402,516	72.44	14.20	524,643	72.69	14.13
Debt-to-income (%)	122,127	23.42	20.14	402,516	20.63	19.66	524,643	21.28	19.80
Loan amount (thou.)	122,127	287.77	167.78	402,516	359.81	254.01	524,643	343.04	238.71
Full Doc	122,127	0.35	0.48	402,516	0.39	0.49	524,643	0.38	0.49
Refinance	122,127	0.62	0.49	402,516	0.65	0.48	524,643	0.64	0.48
Florida	122,127	0.26	0.44	402,516	0.29	0.45	524,643	0.28	0.45
Tract income (%)	122,049	88.08	34.05	402,369	104.38	40.14	524,418	100.59	39.41
Tract minority (%)	122,108	60.80	26.43	402,384	37.55	25.28	524,492	42.97	27.38
HPI growth (%)	112,814	5.98	2.00	367,879	5.84	2.03	480,693	5.87	2.02
Unemployment (%)	112,842	5.73	1.80	368,488	5.31	1.49	481,330	5.41	1.58

Hybrid is an indicator that takes the value of 1 if the origination is a 2/28 or 3/27 mortgage.
Z_k is the count of hybrid products as share of all mortgages at the Census tract level.
Z_h is the count of hybrid products for Hispanics as share of all mortgages for Hispanics at the Census tract level.
Tract Income is the Census tract median family income relative to metropolitan median family income.
Tract Minority is the Census tract share of nonwhite population.
HPI growth is the period-to-period growth in the CoreLogic county-level house price index computed at the end-of-period quarter of origination.
Unemployment is the county-level unemployment rate for the end-of-period quarter of origination.

origination is a refinance reduce the likelihood of acquiring a hybrid product. On the other hand, a higher loan-to-value ratio, a higher debt-to-income ratio, and whether the origination has full documentation increases the probability of acquiring a hybrid. The average marginal effects indicate that the magnitude of these effects is economically important. An increase of one standard deviation (about 70 points) in FICO scores relative to the mean (about 650 points) reduces the likelihood of acquiring a hybrid by about 20 percentage points, on average. Similarly, a one-standard deviation increase in the loan-to-value ratio (about 13 percent) relative to the mean (about 76.5 percent) increases the probability of acquiring a hybrid product by about 10 percentage points, on average. For any given specification, the effect of these covariates is similar across the two contagion measures.

Tables 4 and 5 indicate that both contagion measures have a positive effect on the adoption of hybrid products. In other words, on average, the origination of hybrid products responds to past choices of the average population (the first contagion measure), and to the past choices of the Hispanic population (the second contagion measure). The results indicate, however, that the two alternative contagion measures produce marginal effects on the adoption probabilities of different magnitude. In table 4, a 10 percentage-point increase in the first contagion measure, namely the share of hybrid products in a given Census tract in the previous quarter, leads to an increase in the likelihood of acquiring a hybrid product of about 80 to 100 basis points. (The contagion variable varies between 0 and 1. Therefore, a coefficient of 0.0841 for the average marginal effect under Model 3 corresponds to an increase of about 80 basis points in the probability of hybrid origination in response to a 10 percentage point increase in the contagion measure: $0.0841 \times 0.10 = 0.0084$.) In table 5, a 10 percentage-point increase in the second contagion measure, namely the share of hybrids acquired by Hispanics in the previous quarter, leads to an increase in the likelihood of acquiring a hybrid product of about 30 basis points. (A coefficient of 0.0267 for the average marginal effect under Model 3 corresponds to an increase of about 30 basis points in the probability of hybrid origination in response to a 10 percentage point increase in the

contagion measure: $0.0267 \times 0.10 = 0.0027$.) Note that the marginal effects described here refer to the partial effect on the probability of acquiring a hybrid product averaged across all borrowers irrespective of ethnicity. Since the first contagion measure is defined independent of ethnicity, whereas the second contagion measure includes it, it is not surprising that the magnitude of the first is larger.

The results from both sets of models also indicate that Hispanics are more likely to acquire a hybrid product than other borrowers by about 3 percentage points, and these effects are statistically significant in both sets of models.

Considering the marginal effect of the interaction term we find that according to the average effect reported in table 4, the average marginal effect of the interaction is not statistically significant in any of the specifications of the regressors. In contrast, the average marginal effect of the interaction is statistically significant in all the specifications of table 5. The coefficients indicate that, on average, the differential marginal effect of the contagion measure for Hispanics relative to other borrowers is slightly larger than the magnitude of the average contagion effect for the entire sample of about 30 basis points for a 10 percentage point increase in the contagion measure.

As suggested by Ai and Norton (2003) and Norton, Wang, and Ai (2004), computing the average of the marginal effect of the interaction term may obscure the effects across different observations. Hence, a plot of the marginal effects of the interaction against the predicted probability of each observation helps illustrate better the impact of the interaction term. Figures 1 and 2 illustrate the marginal effect of the interaction, along with the corresponding z-statistics that illustrate their statistical significance across all observations, and are also plotted against the predicted probability of hybrid origination. In the first case, the interaction effect has an inverted S-shape and varies from positive to negative values. However, except at the extremes of the predicted probability, the interaction effect is not statistically significant. In the second case, the interaction effect has an inverted U-shape and is always positive, reaching a peak of about a 45 basis points change in the hybrid origi-

nation rate to a 10 percentage point increase in the contagion measure. The peak is reached near the region of a predicted probability of hybrid adoption of about 50 percent.

It is interesting to note that the narrower definition of contagion which conditions not only on product type but also on ethnicity generates an interaction effect that is positive and statistically significant for all the observations in the sample in figure 2, while the interaction effect of the broader definition of contagion is statistically significant only for observations with very low or very high values of the predicted probability of hybrid origination in figure 1.

We interpret these results as evidence of differential effects of contagion among Hispanic borrowers relative to non-Hispanic borrowers. In other words, there seems to be an externality in the mortgage choices of Hispanics, in that they respond to the previous choices of Hispanics. In contrast, Hispanics do not seem to respond to the previous choices of the average population differentially from other borrowers.

6   Conclusion

Since the subprime mortgage crisis, a significant volume of literature has emerged examining the causes and consequences of the subprime debacle. Fewer studies have examined the emergence, growth, and rapid expansion of subprime mortgages since 2000. In this respect, the significant growth of short-term subprime hybrid products from less than 5% of all subprime mortgages originated in 2000 to almost 75% of mortgages originated in 2006 has remained a mystery to most researchers. Considering the significant evidence showing that most subprime borrowers lacked significant financial sophistication, this rapid expansion of hybrid products is even more puzzling.

In this paper we attempt to shed light on this puzzle from the perspective of the borrower. We argue that neighborhood contagion may have a significant role to play in the dissemination of hybrid products. We do not argue that this was the sole cause of expan-

Table 4: Hybrid probit regression. Coefficients and average marginal effects. Contagion Z_k

Variables	Coefficients			Average Marginal Effects		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Z_k	0.3270	0.3424	0.2855	0.0985	0.1030	0.0841
Hispanic	0.0847	0.0862	0.1068	0.0313	0.0318	0.0335
Hispanic × Z_k	0.0389	0.0397	0.0145	0.0092	0.0093	0.0023
FICO	-0.6702	-0.6715	-0.6728	-0.1952	-0.1955	-0.1953
Loan-to-value	0.3260	0.3303	0.3304	0.0950	0.0961	0.0959
Debt-to-income	0.2094	0.2090	0.2086	0.0610	0.0608	0.0606
Loan amount	-0.1426	-0.1566	-0.1537	-0.0416	-0.0456	-0.0446
Full Doc	0.0155	0.0187	0.0173	0.0045	0.0055	0.0050
Refinance	-0.6004	-0.6012	-0.6038	-0.1727	-0.1727	-0.1731
Florida	-0.2890	-0.3384	-0.3383	-0.0838	-0.0978	-0.0976
Tract income	-0.0098	-0.0013	-0.0061	-0.0029	-0.0004	-0.0018
Tract minority	0.0428	0.0465	0.0484	0.0125	0.0135	0.0141
HPI growth		0.0023	0.0079		0.0007	0.0023
Unemployment		-0.0222	-0.0280		-0.0065	-0.0081
Constant	0.4134	0.5275	0.5161			
Quarterly dummies	No	No	Yes	No	No	Yes
No. Obs.	1108962	1105836	1105836	1108962	1105836	1105836

The variables *FICO, Loan-to-Value, Debt-to-Income, Loan Amount, Tract Income,* and *Tract Minority* were standardized before entering them in the probit estimation.

Hispanic is an indicator that takes the value of 1 if the loan's borrower ethnicity is Hispanic and 0 otherwise.

Z_k is the count of hybrid products as a share of all mortgages at the Census tract level.

The average marginal effects for all variables x^k, except the interaction term, correspond to $\sum_{j=1}^{N} \frac{\partial \Phi(x'_j \beta)}{\partial x^k}$, where Φ is the standard normal cdf.

The average marginal effect of the interaction term, *Hispanic* × Z_k is computed as the average of the cross-partial derivative $\sum_{j=1}^{N} \frac{\Delta \partial \Phi(x'_j \beta)/\partial Z_k}{\Delta Hispanic}$.

Table 5: Hybrid probit regression. Coefficients and average marginal effects. Contagion Z_h

Variables	Coefficients			Average Marginal Effects		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Z_h	0.0771	0.0802	0.0564	0.0347	0.0356	0.0267
Hispanic	0.0139	0.0162	0.0340	0.0285	0.0292	0.0305
Hispanic $\times Z_h$	0.1294	0.1294	0.1096	0.0368	0.0367	0.0311
FICO	-0.6721	-0.6734	-0.6745	-0.1965	-0.1968	-0.1965
Loan-to-value	0.3428	0.3476	0.3464	0.1002	0.1016	0.1009
Debt-to-income	0.2087	0.2083	0.2077	0.0610	0.0608	0.0605
Loan amount	-0.1655	-0.1822	-0.1750	-0.0484	-0.0532	-0.0510
Full Doc	0.0226	0.0253	0.0236	0.0066	0.0074	0.0069
Florida	-0.6176	-0.6182	-0.6213	-0.1782	-0.1783	-0.1787
Florida	-0.2826	-0.3332	-0.3312	-0.0826	-0.0972	-0.0963
Tract income	-0.0214	-0.0118	-0.0165	-0.0063	-0.0035	-0.0048
Tract minority	0.0516	0.0566	0.0558	0.0151	0.0165	0.0163
HPI growth		0.0016	0.0066		0.0005	0.0019
Unemployment		-0.0220	-0.0278		-0.0064	-0.0081
Constant	0.5719	0.6957	0.6608			
Quarterly dummies	No	No	Yes	No	No	Yes
No. Obs.	951456	950543	950543	951456	950543	950543

The variables *FICO*, *Loan-to-Value*, *Debt-to-Income*, *Loan Amount*, *Tract Income*, and *Tract Minority* were standardized before entering them in the probit estimation.
Hispanic is an indicator that takes the value of 1 if the loan's borrower ethnicity is Hispanic and 0 otherwise.
Z_h is the count of hybrid products for Hispanics as a share of all mortgages for Hispanics at the Census tract level.

The average marginal effects for all variables x^k, except the interaction term, correspond to $\sum_{j=1}^{N} \frac{\partial \Phi(x_j'\beta)}{\partial x^k}$, where Φ is the standard normal cdf.
The average marginal effect of the interaction term, *Hispanic* $\times Z_h$, is computed as the average of the cross-partial derivative $\sum_{j=1}^{N} \frac{\Delta \partial \Phi(x_j'\beta)/\partial Z_h}{\Delta Hispanic}$.
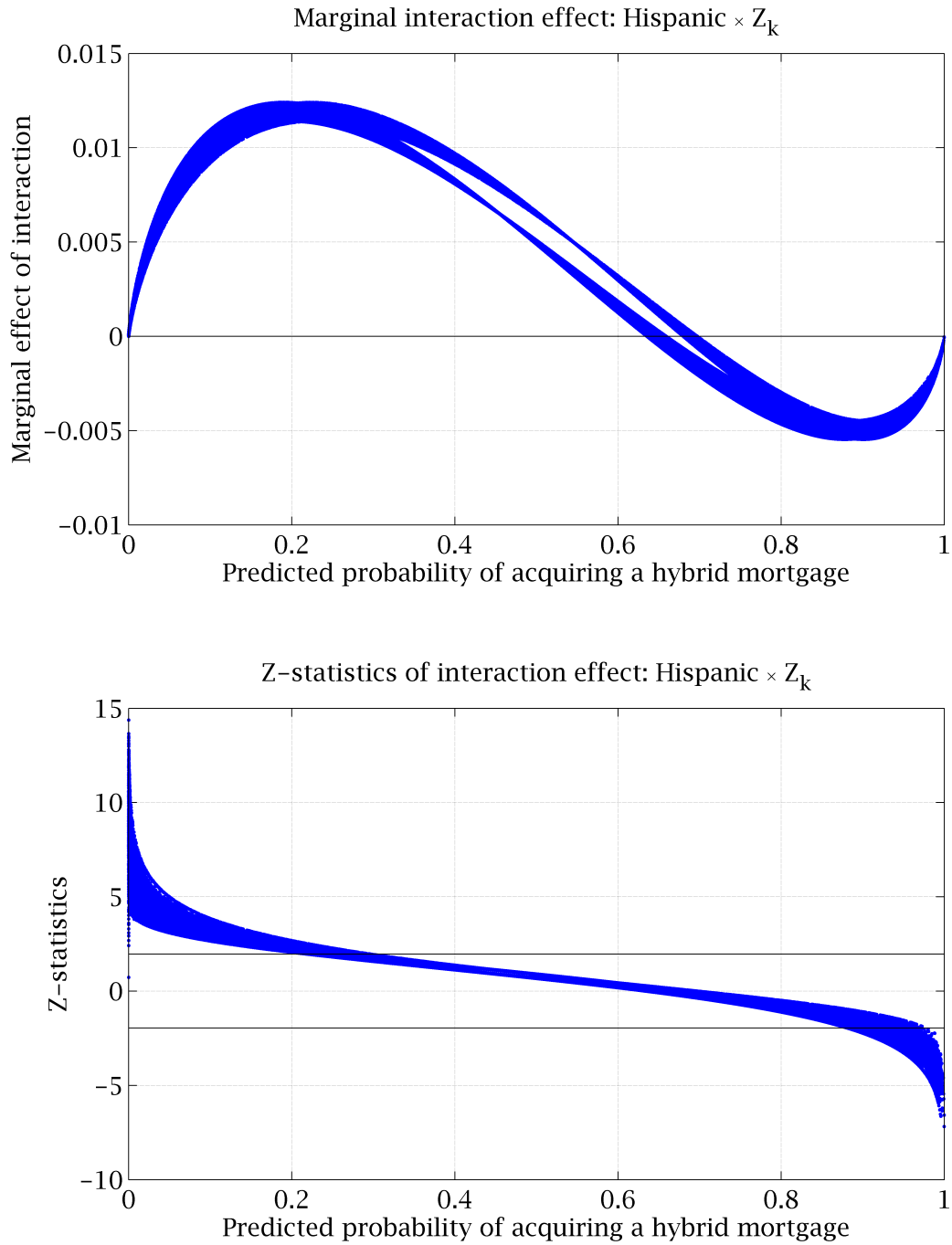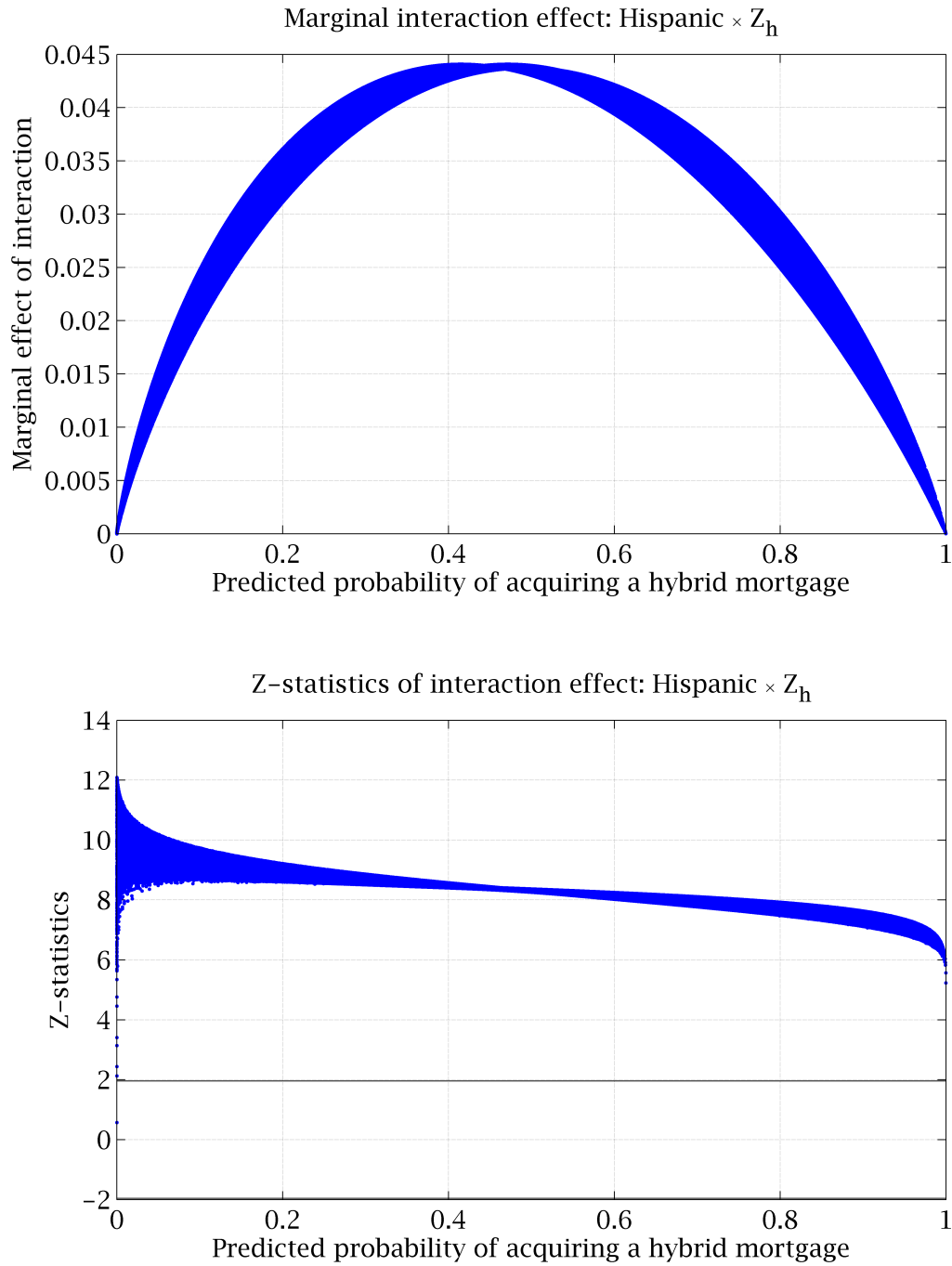
20

Figure 1: Marginal interaction effect. Contagion Z_k

Figure 2: Marginal interaction effect. Contagion Z_h

sion of subprime hybrid products during this period. Several other proximate causes—such as steepness of the yield curve, the affordability of hybrid products in an environment of rapid house price growth, aggressive lending efforts by brokers and banks in an effort to capture market share—may also be responsible in the widespread use of hybrid products in the subprime universe.

However, this study adds to this list of proximate causes, arguing that neighborhood contagion effects of product adoption may have a significant role to play as well. First, the proportion of hybrid subprime products in a given location increases the likelihood that a subprime borrower at the location will adopt a hybrid product subsequently. Notably, this effect of product adoption is independent of the race or ethnicity of the borrower. Further analysis reveals a second neighborhood contagion effect that depends on whether the borrower is Hispanic. We observe a strong contagion effect on the adoption of hybrid mortgages for Hispanics in the following way: An increase in the origination of hybrid products among Hispanic subprime borrowers as a proportion of total hybrid products at a given location increases the likelihood that a subprime borrower who is Hispanic will adopt this product subsequently. This second effect is similar to the first, but significant only for Hispanic borrowers. Our interpretation of this result is that, while the contagion effect holds in general for all products, its effect is significantly more pronounced when one considers the adoption rates of Hispanic borrowers.

While this study is a first step towards understanding the role of neighborhood contagion effects in increasing the appeal and widespread dissemination of subprime hybrid products, further analysis is needed to determine the mechanism by which this growth occurred. In this respect, we are agnostic about the exact mechanism by which these neighborhood effects determined mortgage choice. We hope that this phenomenon will merit further research in light of the evidence presented here.

References

C. Agnell and C. D. Rowley. Breaking New Ground in U.S. Mortgage Lending. Outlook, FDIC, Summer 2006.

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

G. Bhardwaj and R. Sengupta. Subprime Mortgage Design. Working Paper 2008-039D, Federal Reserve Bank of St. Louis, August 2011.

A. K. Bhattacharya, F. J. Fabozzi, and W. S. Berliner. An Overview of Mortgages and the Mortgage Market. In F. J. Fabozzi and S. V. Mann, editors, *The Handbook of Fixed Income Securities*, pages 487–512. McGraw-Hill, 2005.

S. Bikhchandani, D. Hirshleifer, and I. Welch. Learning from the Behavior of Others: Conformity, Fads, and Informational Cascades. *Journal of Economic Perspectives*, 12(3):151–170, Summer 1998.

D. G. Bocian, K. S. Ernst, and W. Li. Unfair Lending: The Effect of Race and Ethnicity on the Price of Subprime Mortgages. Technical report, Center for Responsible Lending, 2006.

D. G. Bocian, K. S. Ernst, and W. Li. Race, Ethnicity and Subprime Home Loan Pricing. *Journal of Economics and Business*, 60:110–124, 2008.

B. Bucks and K. Pence. Do borrowers know their mortgage terms? *Journal of Urban Economics*, 64(2):218–233, 2008.

H. Cai, Y. Chen, H. Fang, and L.-A. Zhou. Microinsurance, Trust and Economic Development: Evidence from a Randomized Natural Field Experiment. Working Paper 15396, National Bureau of Economic Research, October 2009.

M. J. Courchane, B. J. Surette, and P. M. Zorn. Subprime Borrowers: Mortgage Transitions and Outcomes. *Journal of Real Estate Finance and Economics*, 29(4):365–392, 2004.

S. A. Gabriel and S. S. Rosenthal. Credit Rationing, Race, and the Mortgage Market. *Journal of Urban Economics*, 29(3):371 – 379, 1991. ISSN 0094-1190.

GAO. Nonprime Mortgages: Analysis of Loan Performance, Factors Associated with Defaults, and Data Sources. Report GAO-10-805, U.S. Government Accountability Office, August 2010.

L. George and J. Waldfogel. Who Affects Whom in Daily Newspaper Markets? *Journal of Political Economy*, 111(4):765–84, August 2003.

K. Gerardi, L. Goette, and M. Stephan. Financial Literacy and Subprime Mortgage Delinquency: Evidence from a Survey Matched to Administrative Data. Working Paper Working Paper, Federal Reserve Bank of Atlanta, April 2010.

G. Gorton. The Panic of 2007. In *Maintaining Stability in a Changing Financial System, Proceedings of the 2008 Jackson Hole Conference*. Federal Reserve Bank of Kansas City, August 2008.

A. Haughwout, C. Mayer, and J. Tracy. Subprime Mortgage Pricing: The Impact of Race, Ethnicity, and Gender on the Cost of Borrowing. Federal Reserve Bank of New York. Staff Reports Staff Report no. 368., Federal Reserve Bank of New York, apr 2009.

D. R. Haurin and S. S. Rosenthal. Language, Agglomeration and Hispanic Homeownership. *Real Estate Economics*, 37(2):155–183, Apr 2009.

P. H. Hendershott, W. C. LaFayette, and D. R. Haurin. Debt Usage and Mortgage Choice: The FHA-Conventional Decision. *Journal of Urban Economics*, 41(2):202–217, 1997. ISSN 0094-1190.

J. M. Lacko and J. K. Pappalardo. Improving Consumer Mortgage Disclosures: An Empirical Assessment of Current and Prototype Disclosure Forms. Staff report, Federal Trade Commission Bureau of Economics, June 2007.

A. Marshall. *Principles of economics*. Macmillan, eighth edition, 1920.

C. Mayer, K. Pence, and S. M. Sherlund. The Rise in Mortgage Default. *Journal of Economic Perspectives*, 23(1):27–50, Winter 2009.

E. C. Norton, H. Wang, and C. Ai. Computing Interaction Effects and Standard Errors in Logit and Probit Models. *The Stata Journal*, 4(2):154–167, 2004.

A. Pennington-Cross and J. Nichols. Credit History and the FHA-Conventional Choice. *Real Estate Economics*, 28(2):307–336, 2000.

C. Reid and E. Laderman. The Untold Costs of Subprime Lending: Examining the Links among Higher Priced Lending, Foreclosures and Race in California. Technical report, Federal Reserve Bank of San Francisco, 2009.

J. Waldfogel. Preference Externalities: An Empirical Study of Who Benefits Whom in Differentiated Product Markets. *The RAND Journal of Economics*, 34(3):557–568, Autumn 2003.

J. Waldfogel. Does Consumer Irrationality Trump Consumer Sovereignty? *The Review of Economics and Statistics*, 87(87):691–696, November 2005.