Idiosyncratic Risk of House Prices: Evidence from 26 Million Home Sales

Liang Peng[1] and Thomas G. Thibodeau[2]

September 28, 2013

Abstract

This paper uses about 26 million home sales to measure house price idiosyncratic risk for 7,580 U.S. zip codes during three periods: (1) when the U.S. housing market was stable (1996 to 2000), (2) booming (2001 to 2007), and (3) busting (2007 to 2012), and investigates the determinants of house price risk. We find very strong relationships between risk and some basic housing market characteristics. There is a U-shaped relationship between risk and zip-code level median household income; risk is higher in zip codes with more appreciation volatility; and risk is not compensated with higher appreciation.

Keywords: house prices; housing markets (R31)

1 Associate Professor, Finance Division, Leeds School of Business, 995 Regent Drive, University of Colorado-Boulder, Boulder, Colorado 80309-0419. Email: liang.peng@colorado.edu.

2 Professor, Finance Division, Leeds School of Business, 995 Regent Drive, University of Colorado-Boulder, Boulder, Colorado, 80309-0419. Email: tom.thibodeau@colorado.edu. Phone: 303.735.4021. Contact author.

I. Introduction

House prices play an extremely important role in the economy. Economists have substantiated the impact of house prices on consumption, saving, economic production, and asset pricing (see, e.g. Bhatia (1987), Engelhardt (1996), Lettau and Ludvigson (2004), Benjamin, Chinloy and Jud (2004), Case, Quigley and Shiller (2005), Yao and Zhang (2005), Lustig and Nieuwerburgh (2005), Campbell and Cocco (2007), Agarwal (2007), Bostic, Gabriel and Painter (2009), Lustig and Nieuwerburgh (2010), Gan (2010), Carroll, Otsuka and Slacalek (2011), Miller, Peng and Sklarz (2011a), Miller, Peng and Sklarz (2011b), Agarwal and Qian (2013), among many others). The subprime crisis and the subsequent great recession further highlight the importance of the interactions between house prices, the mortgage market, the credit market in general, and the economy (see, e.g. Deng and Quigley (2008), Demyanyk and Hemert (2011), An, Deng and Gabriel (2011), Piskorski and Tchistyi (2011), Brueckner, Calem and Nakamura (2012), Goetzmann, Peng and Yen (2012), among many others).

Given the central role that house prices play in the economy, it is surprising that the literature is virtually silent on a key attribute of house prices – their idiosyncratic risk. Homeowners own individual properties – not the average house in a city or in the U.S. market. Individual properties may sell for prices that are different not only from the average house but also from other properties with identical physical attributes and locations – heterogeneity in buyers' and sellers' valuation and frictions in the search process leads to uncertainty in transaction prices that is not driven by house attributes or market liquidity. We call this idiosyncratic risk. It is plausible that homeowners understand the existence of such risk. Therefore, the risk likely affects homeowners' behavior, such as their consumption, saving, and investment decisions. Moreover, an important empirical question is whether idiosyncratic risk is related to the performance of mortgages, such as default risk, prepayment, renegotiation, and the pricing of mortgage-backed securities.

To shed light on these questions, idiosyncratic risk needs to be measured and its determinants identified. The goal of this paper, therefore, is to measure idiosyncratic risk at the zip-code level and to investigate whether the risk is homogenous across markets. Specifically, we hypothesize that idiosyncratic risk is related to three market characteristics: median household income, the temporal volatility in house price appreciation, and the temporal average of house price appreciation. We expect the median household income and the temporal volatility in house price appreciation to be related to heterogeneity in market participants' valuation, and thus idiosyncratic risk of house prices. We relate idiosyncratic risk to the temporal average of house price appreciation to investigate whether idiosyncratic risk is compensated with higher average appreciation.

To test these hypotheses, we analyze 26 million home sales during the 1996 to 2012 period in 7,580 zip-code-delineated markets in the U.S. We split the sample into three periods with different housing market conditions: the stable period (1996:Q1 to 2000:Q4), the booming period (2001:Q1 to 2007:Q2), and the busting period (2007:Q3 to 2012:Q3). We assume that idiosyncratic risk is relatively homogeneous within each market and time period, so we are able to use the sufficiently large sample of home sales to test these hypotheses.

A key econometric challenge is that the idiosyncratic risk of house prices cannot be measured perfectly, as researchers are unable to observe all value-related attributes. Therefore, we adopt a two-step approach to measure the risk and test the hypotheses. First, we estimate a hedonic house price model, which relates the log of the transaction price to the same set of observed attributes and sale-time dummies, *for each of the 7,580 markets in each of the three periods separately*. Such a regression generates not only residuals but also a hedonic house price index for each market in each period, from which we calculate the temporal average and volatility of quarterly house price appreciation.

It is important to note that the regression residual is only partially driven by the idiosyncratic risk. We suggest that the residual consists of four parts, which are driven by the estimation errors in coefficients, unobserved attributes, the liquidity/thinness of the market, and the heterogeneity in valuation among market participants. The last part pertains to the idiosyncratic risk. This decomposition guides the second step of our analysis, in which we estimate the parameters of a cross-sectional regression *for each of the three time periods*. The dependent variable is the standard deviation of the regression residuals for each market; the explanatory variables consists of the three market characteristics plus variables that we use to control for variation in the first three components of the regression residuals. This regression allows us to test the hypotheses directly by looking at the significance level of the estimated coefficients for the three market characteristics. Moreover, the part of the residual standard deviation that is not explained by control variables constitutes a measure of the idiosyncratic risk for that zip code.

The results provide overwhelming evidence that idiosyncratic risk is not homogeneous across markets – it is significantly related to all three variables. Specifically, there is a very robust U-shaped relationship between idiosyncratic risk and market median household income: idiosyncratic risk is high in low-income markets, declines substantially for moderate-income markets, and then increases in high-income markets. Second, idiosyncratic risk is significantly and positively related to the temporal volatility in house price appreciation, which is also robust for all three periods and for alternative model specifications. Third, the idiosyncratic risk is significantly related to the temporal average house price appreciation in the market. However, the direction of the relationship varies across three time periods. This suggests the non-existence of a stable risk and return relationship for owner-occupied housing.

This paper aims to measure the idiosyncratic risk of house prices and understand its determinants. The strong evidence for house price risk heterogeneity across markets has important implications for many economic issues. Given the heterogeneity in idiosyncratic risk, it is important to analyze the possible impacts of house price risk on households' consumption, saving, and investment decisions, as well as the interactions between the risk and the performance of mortgages and the credit market in general. Further, house price risk is an important parameter in a large variety of urban economics and finance models, including housing tenure choice models (e.g. Ortalo-Magné and Rady (2002) and Davidoff (2006)), homeowners' portfolio allocation models (e.g. Goetzmann (1993), Brueckner (1997), and Englund, Hwang and Quigley (2002)), and homeowners' risk management models (Sinai and Souleles (2005) and Han (2008)). Incorporating heterogeneous house price risk into the above models can potentially help generate new insights into consumer behavior.

This paper compliments the literature on the determinants and implications of temporal volatility of house prices (e.g. Poterba (1991), Dolde and Tirtiroglue (1997), Ambrose, Buttimer and Thibodeau (2001), Capozza, Hendershott and Mack (2004), Miller and Peng (2006), Miles (2008), Zhou and Haurin (2010), among many others). Particularly, it corroborates the "risk segmentation" in metropolitan Denver reported by Peng and Thibodeau (2012), which shows that the temporal volatility of house price changes varies across markets and is significantly correlated with median household income. This paper shows that the idiosyncratic risk is related to income across the U.S.

The finding that low-income markets have much higher house price idiosyncratic risk has important policy implications. For decades, the US Government has been promoting homeownership, particularly among low-income neighborhoods (see Ambrose and Thibodeau (2004), Bostic and Gabriel (2006), and Jaffee, Quigley and Noll (2007) for information on the GSE's Affordable Housing Goals). In contrast to the government's efforts, economists have done relatively little in understanding welfare implications of homeownership of low-income households from the perspective of asset price risk. Findings in this paper raise the question whether homeownership is a desirable investment for low-income/poor households, as their homes would constitute a large portion of their wealth but be exposed to larger than average amounts of idiosyncratic house price risk.

The remainder of this paper is organized as follows. Next section discusses the research design. The third section introduces the data. The fourth section presents empirical evidence for house price risk segmentation. The final section concludes.

II. Research Design

II.1. Defining the Idiosyncratic Risk

Our analysis of idiosyncratic house price risk starts with the following standard hedonic model, which is assumed to be the true data generating process for properties located within a market,

$$\ln\left(V_{i,t}\right) = \alpha_t + \sum_{j=1}^{k} \beta_j X_{i,j} + u_{i,t}\ , \tag{1}$$

where $V_{i,t}$ is the transaction price of house i in period t; α_t is the intercept term; $X_{i,j}$ for $j = 1,\cdots,k$ are k *observed* attributes of the house; and $u_{i,t}$ is the pricing error that captures the part of price that is not explained by the observed house attributes. We further assume that the pricing error $u_{i,t}$ consists of three components,

$$u_{i,t} = \sum_{h=1}^{l} \rho_h U_{i,h} + \upsilon + \varepsilon, \tag{2}$$

where $U_{i,h}$ for $h = 1, \cdots, l$ are l *unobserved* attributes of the house; υ captures the pricing error related to the thinness of the housing market; and ε reflects the part of transaction price that is due to heterogeneity in market participants' valuation of the house. We call the standard deviation of ε the idiosyncratic risk.

To see the distinction between υ and ε, first consider a market of identical houses in which potential buyers have only two bids, \$3 and \$1. If the market is so thin that there is only one bid for each listed house and the seller always accepts the sole bid, the transaction price would be either \$3 or \$1 with the mean being \$2, which implies that the pricing error for each house is either \$1 or -\$1. If the market is liquid and there are numerous bids for each listed house and each seller always accepts the highest bid, the transaction price for each house would almost surely be \$3 and the mean would be \$3 as well, which means there is little pricing error. Therefore, holding constant the distribution of buyers' valuation, the more liquid the market, the smaller is the pricing error. We use υ to capture the pricing error due to the thinness of the market.

Now hold the thinness of the market constant, say there is only one bid and the seller always accept it, and increase the variance in buyers' bids, which are now \$3.5 or \$0.5. The transaction price would be either \$3.5 or \$0.5 and the mean is \$2, which implies that the pricing error for each house would be \$1.5 or -\$1.5. The pricing error is independent from the thinness of the market, and is driven by the heterogeneity in market participants' valuation, which is what we focus on in this paper. We use ε to capture this pricing error, and call its standard deviation idiosyncratic risk of the house price.

II.2. Testing Hypothesis

We test the following three hypotheses regarding the determinants of idiosyncratic risk of house prices.

Hypothesis 1*: Idiosyncratic risk of house prices is not related to the level of household income in the market.*

House prices are driven by sellers' ask prices and buyers' bids. We conjecture that both variables have greater heterogeneity in markets with lower household income, which leads to higher idiosyncratic risk of house prices. First, low-income homebuyers are exposed to greater uncertainty in income (see, e.g. Zhou and Haurin (2010) for evidence on the negative correlation between income level and income risk at the national level), and thus likely have a larger variation in their housing demand and bid prices. Second, low-income sellers' are also exposed to greater income uncertainty; therefore, their ability to absorb possible home equity loss varies more (see, e.g. Stein (1995), Genesove and Mayer (1997), and Lamont and Stein (1999) for evidence on equity constraints on home sellers' behavior), which leads to more variation in ask prices.

Hypothesis 2*: Idiosyncratic risk of house prices is not related to the temporal volatility of the house price appreciation in the market.*

Homebuyers' bid prices are affected not only by their income, but also their expectation of future house prices – all else equal, a house that is expected to have a higher future value will have a higher purchase price. The more uncertain future house prices, the more variation homebuyers likely have in their bid prices, which leads to greater variation in transaction prices. We conjecture that the temporal volatility in house price appreciation directly affects the variation in homebuyers' expected future house prices, and thus is related to idiosyncratic house price risk.

Hypothesis 3*: Idiosyncratic risk of house prices is not related to the temporal average house price appreciation in the market.*

An important empirical question is whether homeowners are compensated for taking higher idiosyncratic risk with higher house price appreciation. Unlike stocks and bonds, houses cannot be bought in homogeneous units and transaction costs are significant. Therefore, households can only participate in markets where houses are affordable for them, and are excluded from markets with unaffordable houses, regardless the risk and return characteristics. Consequently, should there be houses that have high appreciation and low risk, the prices of such houses may not be driven up by increasing demand, as many households may not be able to afford such houses. We do not have a prior regarding the risk-return relationship.

We use a two-step approach to test the hypotheses. The first step uses the standard hedonic regression in (1) to relate transaction prices to a vector of observed house attributes for each market. Specifically, we estimate the following identical hedonic model for each market,

$$\begin{aligned}\ln\left(V_{i,t}\right) = \alpha_1 + \textstyle\sum_{s=2}^{T} \alpha_s Sold_{i,s} \\ &+ \beta_1 Size_{i,t} + \beta_2 Age_{i,t} + \beta_3 Age_{i,t}^2 + \beta_4 Age_{i,t}^3 + \beta_5 Old_{i,t} \\ &+ \beta_6 Bed1_{i,t} + \beta_7 Bed2_{i,t} + \beta_8 Bed4_{i,t} + \beta_9 Bed5_{i,t} + \beta_{10} Bath1_{i,t} \\ &+ \beta_{11} Bath1.5_{i,t} + \beta_{12} Bath2.5_{i,t} + \beta_{13} Bath3_{i,t} + \beta_{14} Bath3.5_{i,t} \\ &+ \beta_{15} Bath4_{i,t} + u_{i,t},\end{aligned} \tag{3}$$

where $Sold_{i,s}$ for $s = 1, \cdots, T$ are dummy variables for the 2nd to last periods, which equal 1 for the period when the house was sold and 0 otherwise; $Size_{i,t}$ is the square feet of the living area; $Age_{i,t}$ is the age of the house in decades since the house was built or 1940, which ever is later; $Age_{i,t}^2$ and $Age_{i,t}^3$ are squared and cubed $Age_{i,t}$; $Old_{i,t}$ is a dummy variable for houses built prior to 1940; $Bed1_{i,t}$ to $Bed5_{i,t}$ are dummy variables for having 1, 2, 4, and 5 (or more) bedrooms; $Bath1_{i,t}$ to $Bed4_{i,t}$ are dummy variables for having 1, 1.5, 2.5, 3, 3.5, and 4 (or more) bathrooms. We choose the above attributes because they are widely used in the literature *and* are observed in all zip codes in our sample. We use the same specification for all zip codes and each time period so that regression residuals are comparable across markets and over time.

Note that the specification of (3) suggests that the intercept measures the log of the transaction price of a house sold in the first period that has 3 bedrooms and 2 bathrooms. The coefficients of sale period dummies measure differences in log values in each period from the log value in the first period. We construct a hedonic price index from the intercept term and the estimated coefficients of quarterly sale dummy variables for each market. We use the standard deviation of the computed appreciation rates to measure the temporal volatility of house price appreciation, which is a variable hypothesized to be related to idiosyncratic risk.

Note the model in (1) and (2) suggests that the regression residuals from (3) contains four components.

$$\begin{aligned}\hat{u}_{i,t} &= \ln\left(V_{i,t}\right) - \left(\hat{\alpha}_t + \sum_{j=1}^{k}\hat{\beta}_j X_{i,j}\right) \\ &= \left(\alpha_t + \sum_{j=1}^{k}\beta_j X_{i,j} + u_{i,t}\right) - \left(\hat{\alpha}_t + \sum_{j=1}^{k}\hat{\beta}_j X_{i,j}\right) \\ &= \left(\alpha_t + \sum_{j=1}^{k}\beta_j X_{i,j}\right) - \left(\hat{\alpha}_t + \sum_{j=1}^{k}\hat{\beta}_j X_{i,j}\right) + \sum_{h=1}^{l}\rho_h U_{i,h} + \upsilon + \varepsilon\end{aligned} \tag{4}$$

The first component is essentially the estimation error, which we denote with e_i.

$$e_i = \left(\alpha_t + \sum_{j=1}^{k}\beta_j X_{i,j}\right) - \left(\hat{\alpha}_t + \sum_{j=1}^{k}\hat{\beta}_j X_{i,j}\right) \tag{5}$$

The second component is the pricing effect of the unobserved attributes, which we denote with τ_i .

$$\tau_i = \sum_{h=1}^{l}\rho_h U_{i,h} \tag{6}$$

The third and fourth components are the pricing effect of the market thinness, υ, and the impact of the heterogeneity in market participants' values on prices, ε. Rewriting equation (4) with the new notations leads to

$$\hat{u}_{i,t} = e_i + \tau_i + \upsilon + \varepsilon. \tag{7}$$

The variation of the regression residual is driven by the variation of the four factors.

To test the three hypotheses, we relate variation in ε to: (1) market median household income; (2) the temporal volatility of house price appreciation in the market; and (3) average house price appreciation. However, we only observe regression residuals $\hat{u}_i$, which comprises four different components, including ε. The variation in $\hat{u}_i$, therefore, measures the variation of ε with error, and the error is due to the variation in the other three components.

Taking this into account, we use the following model for hypothesis testing,

$$y_z = \alpha + \sum_{i=1}^{m}\beta_i^c C_{z,i} + \sum_{i=1}^{n}\beta_i^x X_{z,i} + \vartheta_z, \tag{8}$$

where y_z is the standard deviation of regression residuals in market z, which we call total risk; $C_{z,i}$ for $i=1,\cdots,m$ are m control variables that are expected to affect the variation in e_i, τ_i, and υ; $X_{z,i}$ for $i=1,\cdots n$ are variables that are hypothesized to be related to the variation in ε - idiosyncratic risk; and ϑ_z is an error term that is orthogonal to all explanatory variables. If a coefficient β_i^x is significantly different from 0, we reject the null hypothesis that the corresponding $X_{z,i}$ is not related to the idiosyncratic risk.

II.3. Specifying Regressions

II.3.A. Defining Markets and Sample Periods

We use zip-codes to delineate markets. There is a tradeoff between having more homogeneous houses/households and having more transactions in defining the market. Should houses be identical and households be homogeneous, the larger the market, the more transactions are observed, and the less is the estimation error in hedonic regressions. However, it is well documented that housing markets are segmented (see, e.g. Straszheim (1975), Schnare and Struyk (1976), Goodman (1978), Goodman (1981), and Goodman and Thibodeau (1998a)). Therefore, the larger the market, houses and households are likely more heterogeneous, coefficients of the hedonic regressions are less stable, and the measurement of household income in the market is less relevant for each household.

Zip-code-delineated markets are appropriate for this research. Goodman and Thibodeau (2003) demonstrate that zip-code districts performed about as well as census tracts and elementary school zones in delineating local neighborhoods with relatively homogeneous houses. Our dataset also provides a sufficient number of transactions for 7,580 zip-codes, which allows us to test our hypotheses in a large cross section. We use zip-codes to delineate markets also for practical reasons: the median household income is easily obtainable from the zip-code level census data.

We test the hypotheses for each of three time periods - (1) 1996:Q1 to 2000:Q4; (2) 2001:Q1 to 2007:Q2; and (3) 2007:3 and 2012Q3 - for two main reasons. First, coefficients in the hedonic regression are more likely constant in shorter sample periods. Therefore, by using short sample periods, our empirical results would be less affected by possible errors in the specification of the hedonic regression due to time-varying coefficients. Second, the three periods correspond to very different housing market conditions. U.S. house prices were relatively constant in the first period, increased significantly in the second period, and declined substantially in the third. Testing our hypotheses in the three periods helps shed light on the robustness of the results and possible time variation in the relationship between house price idiosyncratic risk and its determinants.

II.3.B. Control Variables

Equation (7) highlights the importance of controlling for the impact of estimation errors, unobserved attributes, and the thinness of the market on total risk in the zip-code level cross sectional regression (8). We use the following zip-code level variables to control for the magnitude of the estimation error. First, a larger sample likely leads to smaller estimation errors. Therefore, we include the number of transactions used in the hedonic regression as a control variable in (8). Second, the variance of the explanatory variables in the hedonic regression would affect the estimation error – the larger the variance, the more information the variables would provide to help identify their coefficients. As a result, we include the standard deviations of the two key continuous variables in the hedonic regression (3) – the size and the age of house – as control variables for estimation errors in (8).

It is important to control for but challenging to measure the variation of unobserved attributes – they are unobserved. However, it is plausible that older houses have more variation in amenities and attributes than newer homes – when houses age, variation in homeowners' financial conditions and their maintenance behavior would have more impact on the contemporaneous conditions and subsequent market values of the homes. Consistent with this, Goodman and Thibodeau (1995) and Goodman and Thibodeau (1997) report that the magnitude of residuals in hedonic house price regressions is systematically related to the age of the dwelling. Similar results of repeat sales regressions of house prices are reported by Goodman and Thibodeau (1998b). We include the average age (in decades) and the portion of traded houses that were built prior to 1940 in (8) to help capture the variation in unobserved house attributes and amenities. Further, larger homes tend to have more luxury features than smaller homes, which are also likely unobserved due to their nature of being non-standard. So we include the average size (in 1,000 square feet) to help control the variation in unobserved attributes.

The thinness of the market is related to the number of potential buyers relative to the total housing stock. A variable that helps measure thinness is housing market turnover, which is included in (8) and measured with the portion of the neighborhood owner-occupied housing stock that trades during the time period. In addition, the price level of a house affects the quantity of potential buyers. Expensive homes have fewer potential buyers than homes with modest prices. Further, the relationship between the market thinness and house prices is likely to be non-linear – homes with very low prices might have fewer amenities or be in less-desirable neighborhoods and thus have fewer potential buyers. Therefore, we include the median home value and its squared value in (8) to control for the impact of market thinness on total risk.

Two caveats are worth mentioning. First, our controls are not perfect. There is always the caveat that the hypothesized determinants of the idiosyncratic risk may be correlated with the magnitude of the estimation error, the variation in the unobserved house attributes, or the thinness of the market, which would lead to biased results. This caveat might be better addressed in future research that uses better data or are guided by new theories. Second, some control variables might be correlated with multiple factors in (7). For example, the average house age might be correlated with market thinness – older homes might have fewer potential buyers. Turnover might be correlated with not only the market thinness but also market size, which might be correlated with the variation in unobserved house attributes. Standard deviations of age and size, though indicating greater power in identifying coefficients, might be related to the variance in unobserved attributes. Consequently, we need to be cautious in interpreting the estimated coefficients for the control variables, which captures the aggregate impact of control variables on total risk. Nonetheless, the control variables are sensible and have been shown to correlate with the magnitude of hedonic regression residuals, through various mechanisms. For example, Thibodeau (2003) reports that the probability that a predicted house price is within ten (or twenty) percent of a property's observed transaction price is related to house age, dwelling size, the standard deviations of size and age of houses in the market, and the market turnover.

III. Data

This paper uses house transaction data from Zillow.com, which has a database of over 80 million owner-occupied homes sold between 1996:Q1 and 2012:Q3. Many of these properties are single-family attached homes or condominiums. This analysis restricts the sample to single-family detached properties, and excludes foreclosures and short sales. In addition, there is substantial variation in the reporting of housing attributes across property assessment districts (a primary source of Zillow.com information). Consequently, we exclude zip codes from the analysis if data on dwelling size, dwelling age, or number of bedrooms or bathrooms are not reported. Further, we restrict our sample to zip codes that had at least fifty owner-occupied homes in the 2000 Census and at least one hundred market-based transactions in each of the three time periods. Finally, we exclude zip codes with average quarterly appreciation rates greater than 20% or greater than three standard deviations away from the mean across zip codes in any of the three periods. The final sample for this analysis consists of 25,818,647 transactions of single-family detached homes sold over the 1996:Q1 through 2012:Q3 period. The properties are located in 7,580 zip codes across the United States. These zip codes are located in 207 of the largest MSAs in the U.S.

Zip code level household income, median house value, and the number of owner-occupied housing units are obtained from the 2000 Census' STF3 file. The median household income for the 7,580 zip codes ranges from \$14,000 to \$196,000, with the mean being \$50,000. The median homeowner's estimate of home value has an average of \$156,000, with a range from \$27,000 to \$1,000,000 (the top coded value category).

Table 1 reports summary statistics, including the mean, the standard deviation, the minimum, and the maximum across 7,580 zip codes, for 12 zip code level variables. The first two are the median household income and the median value of owner-occupied housing units in the 2000 Census data. The remaining ten variables are reported for each of the three time periods: 1996:Q1-2000:Q4, 2001:Q1-2007:Q2, and 2007:Q3-2012:Q3. The variables include the number of transactions, the market turnover (the number of transactions as a percentage of owner-occupied housing units in 2000 Census data), the average size, the standard deviation of size, the average age, the standard deviation of age of traded houses, the portion of the traded houses that were built prior to 1940, the arithmetic mean and the standard deviation of quarterly appreciation rates of the zip code house price index, and the standard deviation of the hedonic regression residuals.

The statistics in Table 1 are consistent with some stylized facts. First, the statistics support the notion that house prices were relatively stable in the first period (the average quarterly appreciation rate is 1.9%), increased significantly in the second period (the average appreciation rate is 2.3%), and declined substantially in the third period (the average appreciation rate is -0.2%). Second, the table shows no or negative risk-return relationship at the aggregate level. The standard deviation of the quarterly house price appreciation rate is the lowest in the "booming" period and the highest in the "busting" period. This corroborates the findings of Peng and Thibodeau (2012) and Han (2013). Third, the table suggests a positive price volume relationship across time, which is often documented in the literature (see, e.g. Clayton, Miller and Peng (2010)). Specifically, there are more transactions (1,722 on average) and higher turnover (33% on average) in the "booming" period than in the "normal" period (936 transactions and 17% turnover on average) or the "busting" period (748 transactions and 15% turnover on average) even after we make adjustments for the duration of the time periods.

Figures 1 to 3 provide the histograms of the dependent variable in the regression (8) – total risk – for the three time periods, which is the standard deviation of hedonic regression residuals in each zip code. The distribution of total risk is similar in the first two time periods, and shifts to the right in the third period when house prices declined. Figure 4 plots the histogram of the 2000 Census median household income. All four figures suggest skewed distributions.

Table 2 reports correlations between 11 zip code level variables that describe the local economic and housing market conditions for each of the three time periods. The variables are the 2000 Census median household income and median house value, market turnover, the average and standard deviation of the size and age of traded houses, the portion of the traded houses that were built before 1940, the average and standard deviation of quarterly home price appreciation rates and the standard deviation of the hedonic regression residuals (i.e. the total risk). This table establishes a few bivariate relationships for total risk that help shed light on the tested hypotheses. The correlation between total risk and income is strongly negative in each period (-0.24, -0.33, and -0.43, respectively). The correlations between total risk and average house price appreciation (0.39, 0.24, and 0.37) and between total risk and the standard deviation of appreciation (0.51, 0.62, 0.55) are strongly positive. Further, it is worth noting that the total risk is correlated with the age of traded houses. The correlations with the average age are 0.35, 0.29, and 0.27, and the correlations with the portion of old houses are 0.37, 0.35, and 0.33. This highlights the importance of controlling for pricing error due to variation in unobserved house attributes, which house age captures.

Figures 5 to 7 display the relationship between total risk and market median household income across zip codes for each of the three time periods by showing the 25th percentile, the median, and the 75th percentile of total risk by household income percentile (with roughly 76 zip codes in each income percentile). The three figures show very similar patterns: total risk declines as household income increases. This is consistent with the negative correlation between the total risk and the household income reported in Table 2.

IV. Results

We estimate four specifications of equation (8) and report the results in Tables 3 to 6. The specifications differ in two respects: the way we calculate the standard deviation of zip code residuals and whether we include MSA fixed effects. Table 3 reports the results of the "baseline" specification, which does not include MSA fixed effects. The standard deviation of residuals is calculated using all traded houses in each zip code/period.

Results in Table 3 show that the coefficients of household income are statistically significant at the 1% level and negative in all three periods (-0.209, -0.268, and -0.271), which strongly rejects Hypothesis I. Further, the relationship between idiosyncratic risk and market median household income is nonlinear – the coefficients of squared household income are statistically significant in all three periods (0.064, 0.100, and 0.108). To visualize the relationship, we plot the idiosyncratic risk level as the quadratic function of household income implied by the estimated coefficients, with the values of control variables set to their across zip code mean. Figure 8 shows a clear U-shape relationship: idiosyncratic risk is high for low-income markets, declines substantially for moderate-income markets, and then increases for high-income markets. Note that the median household income for the 7,580 zip codes is about $50,000, so about half the observations are depicted in the left quarter of Figure 8.

Table 3 also provides results that strongly reject Hypotheses II and III. The estimated coefficients for temporal volatility in house price appreciation are statistically significant at the 1% level for all three periods (0.429, 0.856, and 0.492). The higher the temporal volatility in home value appreciation, the greater the idiosyncratic risk in the local housing market. The estimated coefficients for the average house price appreciation rate are also statistically significant at the 1% level; however, they change signs across periods: positive (0.186) in the first period and negative (-0.675 and -0.329) in the last two periods. The alternating sign of the average house price appreciation rate suggests the non-existence of a robust positive risk return relationship and that homeowners are not always compensated for taking higher idiosyncratic house price risk. Note that some control variables have unexpected signs. For example, market turnover turns out to be negatively related to total risk, though we expected lower total risk in more liquid markets. We suspect that such counterintuitive results might be due to unobserved market specific conditions that are related to control variables.

The second specification controls for unobserved metropolitan area market specific conditions by including Metropolitan Statistical Area (MSA) fixed effects. To have sufficient degrees of freedom to estimate the fixed effects, we restricted the sample to zip codes in MSAs that have at least 5 zip codes in our sample, which leads to 6,618 zip codes. The results are in Table 4. A noticeable difference between Table 4 and Table 3 is that the Adjusted R-squares are substantially higher in Table 4 (0.534, 0.599, and 0.585) than in Table 3 (0.426, 0.494, and 0.493). This substantiates the importance of controlling for metropolitan area specific house price risk.

Table 4 provides results that are consistent with Table 3 in rejecting Hypothesis I. The coefficients of household income are negative and significant (-0.163, -0.180, and -0.199 in three periods) and the coefficients of squared income are positive and significant. Figure 9 plots the relationships between total risk and household income in the three periods indicated by Table 4, which are very similar to the relationships displayed in Figure 8. Further, Hypothesis II is also rejected with statistically significant and positive coefficients of temporal volatility in house price appreciation in all three periods. The coefficients of average house price appreciation are positive for the first two periods, and negative for the third – not robust across time periods or across alternative specifications.

Another possible econometric issue that might bias the results pertains to the way we calculate the standard deviation of residuals in Tables 3 and 4. Some zip codes have many more transactions than others; therefore, standard deviations calculated from larger samples might be more accurate than those calculated from smaller samples. During the first time period, for example, the number of transactions in a zip code ranges from 100 to 7,928 (Table 1). Including the number of transactions might not completely eliminate this problem: as Tables 3 and 4 show, the number of transactions is not always negatively related to total risk, possibly due to the fact that the number of transactions may be correlated with unknown variables that affect total risk. To mitigate the impact of this statistical attribute on our analyses, we randomly draw 100 residuals from each zip code in each period, calculate the standard deviation using the 100 residuals, and use this bootstrapped standard deviation as the dependent variable. Because all standard deviations are now estimated using 100 observations, we delete the number of observations as an explanatory variable. Table 5 reports results of regressions without the MSA fixed effects and Table 6 includes MSA fixed effects.

Results in Table 5 and 6 are very robust in rejecting Hypothesis I. The coefficients of household income are always negative and statistically significant, and the coefficients of squared income are always positive and statistically significant. Figures 10 and 11 graph the relationships, which are very similar to the relationships shown in Figures 8 and 9. Hypothesis II is once again strongly rejected in Tables 5 and 6: the coefficient of temporal volatility in house price appreciation is positive and significant in each period. Hypothesis III is rejected, but as in Tables 5 and 6, the sign of the coefficient of average house price appreciation is not stable over time.

V. Conclusion

This paper empirically measures the idiosyncratic risk of house prices in 7,580 zip-code-delineated markets in the U.S. and investigates the determinants of risk. Using 26 million home sales during the 1996 to 2012 period, we find strong evidence that the idiosyncratic risk is not identical across markets: it is significantly related to median household income and to the temporal volatility and average in house price appreciation in each market. Particularly, there is a U-shaped relationship between risk and median household income: idiosyncratic risk is high for low-income markets, declines substantially for moderate-income markets, and then increases for high-income markets. The results are robust across three periods: when the U.S. housing market was stable (1996:Q1 to 2000:Q4), booming (2001:Q1 to 2007:Q2), and the busting (2007:Q3 to 2012:Q3). Further, idiosyncratic risk is significantly and positively correlated with the temporal volatility in house price appreciation. This is consistent with the notion that temporal volatility in house price appreciation increases heterogeneity in market participants' valuation of houses. Finally, while idiosyncratic risk is significantly correlated with the average house price appreciation rates in each period, the relationship changes direction across periods. Therefore, there is no evidence that idiosyncratic risk is compensated with higher house price appreciation. The heterogeneity in the idiosyncratic risk and its correlation with market characteristics has important implications for many housing models as well as for U.S. government policies that promote homeownership among low-income households.

Acknowledgements: The authors thank Stan Humphries and Svenja Gudell for providing access to and assistance with Zillow.com data. We also thank Stephen Billings, Allen Goodman, Mark Shroder and Todd Sinai for their comments and seminar participants at Johns Hopkins' Carey School of Business. The authors take responsibility for any errors in this paper.

REFERENCES

Agarwal, Sumit. 2007. "The Impact of Homeowners' Housing Wealth Misestimation on Consumption and Saving Decisions." *Real Estate Economics* 35, 135-154.

Agarwal, Sumit, and Wenlan Qian. 2013. "Access to Housing Wealth and Consumption: Evidence from a Natural Experiment." *National University of Singapore working paper*.

Ambrose, Brent, Richard Buttimer, and Thomas G. Thibodeau. 2001. "A New Spin on the Jumbo/Conforming Loan Rate Differential." *Journal of Real Estate Finance and Economics* 23, 309-335.

Ambrose, Brent, and Thomas G. Thibodeau. 2004. "Have the GSE Affordable Housing Goals Increased the Supply of Mortgage Credit?" *Regional Science and Urban Economics* 34, 263-273.

An, Xudong, Yongheng Deng, and Stuart Gabriel. 2011. "Asymmetric Information, Adverse Selection, and the Pricing of CMBS." *Journal of Financial Economics* 100, 304-325.

Benjamin, John D., Peter Chinloy, and G. Donald Jud. 2004. "Real Estate versus Financial Wealth in Consumption." *Journal of Real Estate Finance and Economics* 29, 341-354.

Bhatia, Kul B. 1987. "Real Estate Assets and Consumer Spending." *Quarterly Journal of Economics* 102, 437-444.

Bostic, Raphael, Stuart Gabriel, and Gary Painter. 2009. "Housing Wealth, Financial Wealth, and Consumption: New Evidence from Micro Data." *Regional Science and Urban Economics* 39, 79-89.

Bostic, Raphael W., and Stuart A. Gabriel. 2006. "Do the GSEs Matter to the Low-income Housing Market? An Assessment of the Effect of GSE Loan Purchase Goals on California Housing Outcomes." *Journal of Urban Economics* 59, 458-475.

Brueckner, Jan K. 1997. "Consumption and Investment Motives and the Portfolio Choices of Homeowners." *Journal of Real Estate Finance and Economics* 15, 159-180.

Brueckner, Jan K., Paul S. Calem, and Leonard I. Nakamura. 2012. "Subprime Mortgages and the Housing Bubble." *Journal of Urban Economics* 71, 230–243.

Campbell, John Y., and Joao F. Cocco. 2007. "How do House Prices Affect Consumption? Evidence from Micro Data" *Journal of Monetary Economics* 54, 591-621.

Capozza, Dennis R., Patric H. Hendershott, and Charlotte Mack. 2004. "An Anatomy of Price Dynamics in Illiquid Markets: Analysis and Evidence from Local Housing Markets." *Real Estate Economics* 32, 1-32.

Carroll, Christopher D., Misuzu Otsuka, and Jirka Slacalek. 2011. "How Large is the Housing Wealth Effect? A New Approach." *Journal of Money, Credit and Banking* 43, 55-79.

Case, Karl E., John M. Quigley, and Robert J. Shiller. 2005. "Comparing Wealth Effects: The Stock Market versus the Housing Market." *The B.E. Journal of Macroeconomics* 5, 1-34.

Clayton, Jim, Norman Miller, and Liang Peng. 2010. "Price-Volume Correlation in the Housing Market: Causality and Co-Movements." *Journal of Real Estate Finance and Economics* 40, 14-40.

Davidoff, Thomas. 2006. "Labor Income, Housing Prices, and Homeownership." *Journal of Urban Economics* 59, 209-235.

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

Deng, Yongheng, and John M. Quigley. 2008. "Index Revision, House Price Risk, and the Market for Home Price Derivatives." *Journal of Real Estate Finance and Economics* 37, 191-210.

Dolde, Walter, and Dogan Tirtiroglue. 1997. "Temporal and Spatial Information Diffusion in Real Estate Price Changes and Variances." *Real Estate Economics* 25, 539–565.

Engelhardt, Gary V. 1996. "House Prices and Home Owner Saving Behavior." *Regional Science and Urban Economics* 26, 313-336.

Englund, Peter, Min Hwang, and John M. Quigley. 2002. "Hedging Housing Risk." *Journal of Real Estate Finance and Economics* 24, 167-200.

Gan, Jie. 2010. "Housing Wealth and Consumption Growth: Evidence from a Large Panel of Households." *Review of Financial Studies* 23, 2229-2267.

Genesove, David, and Christopher J. Mayer. 1997. "Equity and Time to Sale in the Real Estate Market." *American Economic Review* 87, 255-269.

Goetzmann, William N. 1993. "The Single Family Home in the Investment Portfolio." *Journal of Real Estate Finance and Economics* 6, 201-222.

Goetzmann, William N., Liang Peng, and Jacqueline Yen. 2012. "The Subprime Crisis and House Price Appreciation." *Journal of Real Estate Finance and Economics* 44, 36-66.

Goodman, Allen. 1978. "Hedonic Prices, Price Indices, and Housing Markets." *Journal of Urban Economics* 5, 471-484.

Goodman, Allen. 1981. "Housing Submarkets within Urban Areas: Definitions and Evidence." *Journal of Regional Science* 21, 175–185.

Goodman, Allen, and Thomas G Thibodeau. 2003. "Housing Market Segmentation and Hedonic Prediction Accuracy." *Journal of Housing Economics* 12, 181-201.

Goodman, Allen, and Thomas G. Thibodeau. 1995. "Age-Related Heteroskedasticity in Hedonic House Price Equations." *Journal of Housing Research* 6, 25-42.

Goodman, Allen, and Thomas G. Thibodeau. 1997. "Dwelling Age Heteroskedasticity in Hedonic House Price Equations: An Extension." *Journal of Housing Research* 8, 299-317.

Goodman, Allen, and Thomas G. Thibodeau. 1998a. "Housing Market Segmentation." *Journal of Housing Economics* 7, 121-143.

Goodman, Allen, and Thomas G. Thibodeau. 1998b. "Dwelling Age Heteroskedasticity in Repeat Sales House Price Equations." *Real Estate Economics* 26, 149-169.

Han, Lu. 2008. "Hedge House Price Risk in the Presence of Lumpy Transaction Costs." *Journal of Urban Economics* 64, 270-287.

Han, Lu. 2013. "Understanding the Puzzling Risk-Return Relationship for Housing." *Review of Financial Studies* 26, 877-928.

Jaffee, Dwight M., John M. Quigley, and Roger G. Noll. 2007. "Housing Subsidies and Homeowners: What Role for Government-Sponsored Enterprises?" *Brookings-Wharton Papers on Urban Affairs*.

Lamont, Owen, and Jeremy C. Stein. 1999. "Leverage and House-Price Dynamics in US Cities." *Rand Journal of Economics* 30, 498-514.

Lettau, Martin, and Sydney Ludvigson. 2004. "Understanding Trend and Cycle in Asset Values: Reevaluating the Wealth Effect on Consumption." *American Economic Review* 94, 276–299.

Lustig, Hanno N., and Stijn G. Van Nieuwerburgh. 2005. "Housing Collateral, Consumption Insurance, and Risk Premia: An Empirical Perspective." *Journal of Finance* 60, 1167-1219.

Lustig, Hanno, and Stijn Van Nieuwerburgh. 2010. "How Much does Household Collateral Constrain Regional Risk Sharing?" *Review of Economic Dynamics* 13, 265–294.

Miles, William. 2008. "Volatility Clustering in U.S. Home Prices." *Journal of Real Estate Research* 30, 74-90.

Miller, Norman, and Liang Peng. 2006. "Exploring Metropolitan Housing Price Volatility." *Journal of Real Estate Finance and Economics* 33, 5-18.

Miller, Norman, Liang Peng, and Michael Sklarz. 2011a. "Economic Impact of Anticipated House Price Changes - Evidence from Home Sales." *Real Estate Economics* 39, 345-378.

Miller, Norman, Liang Peng, and Michael Sklarz. 2011b. "House Prices and Economic Growth." *Journal of Real Estate Finance and Economics* 42, 522-541.

Ortalo-Magné, François, and Sven Rady. 2002. "Tenure Choice and the Riskiness of Non-Housing Consumption." *Journal of Housing Economics* 11, 226-279.

Peng, Liang, and Thomas Thibodeau. 2012. "Risk Segmentation of American Homes: Evidence from Denver." *Real Estate Economics* forthcoming.

Piskorski, Tomasz, and Alexei Tchistyi. 2011. "Stochastic House Appreciation and Optimal Mortgage Lending." *Review of Financial Studies* 24, 1407-1446.

Poterba, James M. 1991. House Price Dynamics: The Role of Taxes and Demography." *Brookings Papers on Economic Activity* 2, 143-183.

Schnare, Anne, and Raymond Struyk. 1976. "Segmentation in Urban Housing Markets." *Journal of Urban Economics* 3, 146-166.

Sinai, Todd, and Nicholas S. Souleles. 2005. "Owner-Occupied Housing as a Hedge against Rent Risk." *Quarterly Journal of Economics* 120, 763-789.

Stein, Jeremy C. 1995. "Prices and Trading Volume in the Housing Market: A Model with Down-Payment Effects." *Quarterly Journal of Economics* 110, 379-406.

Straszheim, Mahlon R. 1975. *An econometric analysis of the urban housing market* (National Bureau of Economic Research, New York).

Thibodeau, Thomas G. 2003. "Marking Single-Family Property Values to Market." *Real Estate Economics* 31, 1-22.

Yao, Rui, and Harold Zhang. 2005. "Optimal Consumption and Portfolio Choices with Risky Housing and Borrowing Constraints." *Review of Financial Studies* 18, 197-239.

Zhou, Yu, and Donald R. Haurin. 2010. "On the Determinants of Housing Value Volatility." *Journal of Real Estate Research* 32, 377-396.

Table 1 Data Summary

This table reports the mean, the standard deviation, the minimum, and the maximum across 7,580 zip codes of the median household income and the median value of owner-occupied housing units in the 2000 census data. It also reports the same statistics for 10 zip code level variables in each of three consecutive sample periods (1996:Q1-2000:Q4, 2001:Q1-2007:Q2, and 2007:Q3-2012:Q3). The variables include the number of transactions of single-family houses in the sample and the housing market turnover (the number of transactions as a percentage of owner-occupied single family housing units as reported in 2000 census data). They further include the average size (square feet), the standard deviation of size (square feet), the average age (years), the standard deviation of age (years) of the traded houses in the sample, and the portion of the traded houses that were built before 1940. The last three variables are generated from hedonic regressions of house transaction prices for each of the 7,580 zip codes, which provide a hedonic house price index and regression residuals for each zip code. The variables are the arithmetic mean and the standard deviation of quarterly appreciation rates of the zip code house price index (percentage changes in index price levels between consecutive quarters) and the standard deviation of the hedonic regression residuals.

Variables	Mean	Std. Dev.	Min	Max
2000 Census				
Median Household Income ($1,000)	50	18	14	196
Median Housing Unit Value ($1,000)	156	102	27	1,000
1996:Q1 - 2000:Q4				
Number of Transactions	936	907	100	7,928
Turnover	17%	12%	1%	152%
Size (square feet): Average	1,731	392	912	4,331
Size (square feet): Standard Dev.	667	275	141	2,918
Age (years): Average	42	14	13	72
Age (years): Standard Dev.	17	7	6	89
Portion of houses built before 1940	16%	21%	0%	99%
Quarterly zip code index appreciation: Average	1.9%	1.4%	-1.2%	9.6%
Quarterly zip code index appreciation: Standard Dev.	9.4%	7.2%	0.7%	57.2%
Zip code hedonic regression residuals: Standard Dev.	0.22	0.07	0.04	0.61
2001:Q1 - 2007:Q2				
Number of Transactions	1,722	1,635	105	14,541
Turnover	33%	27%	1%	297%
Size (square feet): Average	1,764	400	923	3,984
Size (square feet): Standard Dev.	796	412	170	4,878
Age (years): Average	39	15	8	72
Age (years): Standard Dev.	21	11	0	149
Portion of houses built before 1940	15%	20%	0%	99%
Quarterly zip code index appreciation: Average	2.3%	1.3%	-0.1%	8.8%
Quarterly zip code index appreciation: Standard Dev.	7.9%	5.5%	1.1%	50.4%
Zip code hedonic regression residuals: Standard Dev.	0.22	0.08	0.06	0.66
2007:Q3 - 2012:Q3				
Number of Transactions	748	601	100	6,709
Turnover	15%	11%	1%	221%
Size (square feet): Average	1,811	420	962	4,077
Size (square feet): Standard Dev.	722	294	164	6,667
Age (years): Average	38	16	6	72
Age (years): Standard Dev.	18	7	0	90
Portion of houses built before 1940	14%	20%	0%	99%
Quarterly zip code index appreciation: Average	-0.2%	2.0%	-5.8%	11%
Quarterly zip code index appreciation: Standard Dev.	11.5%	8.3%	1.4%	68.4%
Zip code hedonic regression residuals: Standard Dev.	0.24	0.08	0.05	0.68

Table 2.A. Variable Correlations: 1996:Q1 - 2000:Q4

This table reports the correlations between 11 zip code level variables, which are calculated from 7,580 zip codes, in the 1996 - 2000 period. The variables are the 2000 census median household income (V1, in $100,000) and median house value (V2, in $100,000), housing market turnover (V3, the number of transactions as a percentage of owner-occupied single family housing units as reported in 2000 census data), the average (V4) and standard deviation (V5) of the size (1,000 square feet) of traded houses, the average (V6) and standard deviation (V7) of the age (10 years) of the traded houses, and the portion of the traded houses that were built before 1940 (V8), the arithmetic mean (V9) and the standard deviation (V10) of quarterly appreciation rates of the zip code house price index (percentage changes in index price levels between consecutive quarters) and the standard deviation of the hedonic regression residuals (V11).

	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11
V1: Household Income	0.71	0.14	0.69	0.46	-0.31	-0.02	-0.26	0.08	-0.13	-0.24
V2: Home Value	1	0.14	0.57	0.42	-0.01	-0.01	-0.03	0.27	-0.07	-0.01
V3: Turnover		1	0.09	0.18	-0.21	0.11	-0.13	-0.07	-0.37	-0.09
V4: Size: Average			1	0.68	-0.37	-0.03	-0.22	0.05	-0.05	-0.08
V5: Size: Std.				1	-0.24	0.21	-0.13	0.09	-0.03	0.04
V6: Age: Average					1	-0.11	0.77	0.16	0.14	0.35
V7: Age: Std.						1	-0.06	0.42	0.01	0.09
V8: Portion before 1940							1	0.17	0.21	0.37
V9: Appreciation: Mean								1	0.65	0.39
V10: Appreciation: Std.									1	0.51
V11: Residuals: Std.										1

Table 2.B. Variable Correlations: 2001:Q1 - 2007:Q2

This table reports the correlations between 11 zip code level variables, which are calculated from 7,580 zip codes, in the 2001 - 2007 period. The variables are the 2000 census median household income (V1, in $100,000) and median house value (V2, in $100,000), housing market turnover (V3, the number of transactions as a percentage of owner-occupied single family housing units as reported in 2000 census data), the average (V4) and standard deviation (V5) of the size (1,000 square feet) of traded houses, the average (V6) and standard deviation (V7) of the age (10 years) of the traded houses, and the portion of the traded houses that were built before 1940 (V8), the arithmetic mean (V9) and the standard deviation (V10) of quarterly appreciation rates of the zip code house price index (percentage changes in index price levels between consecutive quarters) and the standard deviation of the hedonic regression residuals (V11).

	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11
V1: Household Income	0.71	0.02	0.67	0.32	-0.26	-0.13	-0.25	-0.11	-0.14	-0.33
V2: Home Value	1	-0.06	0.50	0.21	0.03	-0.15	-0.03	0.09	-0.05	-0.15
V3: Turnover		1	0.12	0.40	-0.35	0.29	-0.17	-0.03	-0.22	-0.12
V4: Size: Average			1	0.57	-0.43	-0.03	-0.27	-0.07	-0.08	-0.16
V5: Size: Std.				1	-0.37	0.41	-0.20	-0.00	-0.10	-0.10
V6: Age: Average					1	-0.16	0.76	-0.11	-0.20	0.29
V7: Age: Std.						1	-0.08	0.02	0.02	0.08
V8: Portion before 1940							1	0.10	0.27	0.35
V9: Appreciation: Mean								1	0.50	0.24
V10: Appreciation: Std.									1	0.62
V11: Residuals: Std.										1

Table 2.C. Variable Correlations: 2007:Q3 - 2012:Q3

This table reports the correlations between 11 zip code level variables, which are calculated from 7,580 zip codes, in the 2007 - 2012 period. The variables are the 2000 census median household income (V1, in $100,000) and median house value (V2, in $100,000), housing market turnover (V3, the number of transactions as a percentage of owner-occupied single family housing units as reported in 2000 census data), the average (V4) and standard deviation (V5) of the size (1,000 square feet) of traded houses, the average (V6) and standard deviation (V7) of the age (10 years) of the traded houses, and the portion of the traded houses that were built before 1940 (V8), the arithmetic mean (V9) and the standard deviation (V10) of quarterly appreciation rates of the zip code house price index (percentage changes in index price levels between consecutive quarters) and the standard deviation of the hedonic regression residuals (V11).

	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11
V1: Household Income	0.71	0.11	0.65	0.39	-0.24	-0.16	-0.25	-0.10	-0.28	-0.43
V2: Home Value	1	-0.01	0.48	0.27	0.04	-0.12	-0.02	-0.11	-0.19	-0.26
V3: Turnover		1	0.24	0.20	-0.41	-0.03	-0.17	-0.11	-0.20	-0.10
V4: Size: Average			1	0.69	-0.46	-0.10	-0.28	-0.09	-0.24	-0.25
V5: Size: Std.				1	-0.31	0.16	-0.19	-0.04	-0.15	-0.06
V6: Age: Average					1	-0.04	0.74	0.19	0.20	0.27
V7: Age: Std.						1	0.01	0.12	0.12	0.17
V8: Portion before 1940							1	0.26	0.26	0.33
V9: Appreciation: Mean								1	0.75	0.37
V10: Appreciation: Std.									1	0.55
V11: Residuals: Std.										1

Table 3 Determinants of Idiosyncratic House Price Risk: Baseline Model

This table reports results of the zip code level cross-sectional regressions regarding the determinants of idiosyncratic house price risk. The sample consists of 7,580 zip codes. The dependent variable is the standard deviation of residuals from the hedonic regression for each zip code. The explanatory variables are for the same zip code, including the 2000 census median household income (in $100,000) and its squared value, the median house value (in $100,000) and its squared value, the number of transactions (1,000), the housing market turnover (the ratio of the number of transactions to the owner-occupied single family housing units in 2000 census), the average and standard deviation of the size (1,000 square feet) of traded houses, the average and standard deviation of the age (10 years) of traded houses, the portion of the traded houses that were built before 1940, and the arithmetic mean and the standard deviation of quarterly appreciation rates of the zip code house price index. The three regressions have identical specifications but different sample periods. Coefficients with ***, **, and * are significant at 1%, 5%, and 10% level.

	I	II	III
Variables	1996:Q1 - 2000:Q4	2001:Q1 - 2007:Q2	2007:Q3 - 2012:Q3
Intercept	0.125***	0.183***	0.194***
Household Income	-0.209***	-0.268***	-0.271***
Household Income Squared	0.064***	0.100***	0.108***
Index Appreciation: Std. Dev.	0.429***	0.856***	0.492***
Index Appreciation: Average	0.186***	-0.675***	-0.329***
Home Value	0.007***	0.003	-0.016***
Home Value Squared	0.000	0.000	0.002***
Number of Transactions	-0.004***	0.003***	0.012***
Turnover	0.097***	0.008***	0.017***
Size: Average	0.021***	0.029***	0.017***
Size: Std. Dev.	0.025***	-0.002	0.034***
Age: Average	0.010***	0.005***	0.007***
Age: Std. Dev.	0.008***	0.003***	0.006***
Portion before 1940	0.028***	0.026***	0.028***
Adjusted R-square	0.426	0.494	0.493
F Value	432.7	569.8***	501.0***

Table 4 Determinants of Idiosyncratic House Price Risk: MSA Fixed Effects

This table reports results of the zip code level cross-sectional regressions regarding the determinants of idiosyncratic house price risk. The sample consists of 6,618 zip codes. The dependent variable is the standard deviation of residuals from the hedonic regression for each zip code. The explanatory variables include MSA dummy variables, the 2000 census median household income (in $100,000) and its squared value, the median house value (in $100,000) and its squared value, the number of transactions (1,000), the housing market turnover (the ratio of the number of transactions to the owner-occupied single family housing units in 2000 census), the average and standard deviation of the size (1,000 square feet) of traded houses, the average and standard deviation of the age (10 years) of the traded houses, the portion of the traded houses that were built before 1940, and the arithmetic mean and the standard deviation of quarterly appreciation rates of the zip code house price index. The three regressions have identical specifications but different sample periods. Coefficients with ***, **, and * are significant at 1%, 5%, and 10% level.

	I	II	III
Variables	1996:Q1 - 2000:Q4	2001:Q1 - 2007:Q2	2007:Q3 - 2012:Q3
MSA Fixed Effect	Yes	Yes	Yes
Intercept	0.056***	0.057***	0.072***
Household Income	-0.163***	-0.180***	-0.199***
Household Income Squared	0.044***	0.062***	0.075***
Index Appreciation: Std. Dev.	0.358***	0.668***	0.486***
Index Appreciation: Average	0.190***	0.277***	-0.790***
Home Value	0.017***	0.014***	0.004
Home Value Squared	-0.001	-0.001**	0.000
Number of Transactions	-0.003***	0.002***	0.008***
Turnover	0.077***	0.003	0.013*
Size: Average	0.018***	0.020***	0.012***
Size: Std. Dev.	0.023***	0.007***	0.029***
Age: Average	0.013***	0.011***	0.012***
Age: Std. Dev.	0.009***	0.004***	0.010***
Portion before 1940	0.032***	0.027***	0.036***
Adjusted R-square	0.534	0.599	0.585
F Value	44.3***	57.4***	54.2***

Table 5 Determinants of Idiosyncratic House Price Risk: Baseline Model with Bootstrapped Dependent Variable

This table reports results of the zip code level cross-sectional regressions regarding the determinants of idiosyncratic house price risk. The sample consists of 7,580 zip codes. The dependent variable is the standard deviation of *100 randomly selected residuals* from the hedonic regression for each zip code. The explanatory variables are for the same zip code, including the 2000 census median household income (in $100,000) and its squared value, the median house value (in $100,000) and its squared value, the number of transactions (1,000), the housing market turnover (the ratio of the number of transactions to the owner-occupied single family housing units in 2000 census), the average and standard deviation of the size (1,000 square feet) of traded houses, the average and standard deviation of the age (10 years) of traded houses, the portion of the traded houses that were built before 1940, and the arithmetic mean and the standard deviation of quarterly appreciation rates of the zip code house price index. The three regressions have identical specifications but different sample periods. Coefficients with ***, **, and * are significant at 1%, 5%, and 10% level.

	I	II	III
Variables	1996:Q1 - 2000:Q4	2001:Q1 - 2007:Q2	2007:Q3 - 2012:Q3
Intercept	0.125***	0.182***	0.202***
Household Income	-0.212***	-0.257***	-0.258***
Household Income Squared	0.064***	0.095***	0.102***
Index Appreciation: Std. Dev.	0.456***	0.830***	0.447***
Index Appreciation: Average	0.11	-0.652***	-0.283***
Home Value	0.008***	0.002	-0.018***
Home Value Squared	0.000	0.000	0.002***
Turnover	0.085***	0.014***	0.042***
Size: Average	0.022***	0.027***	0.016***
Size: Std. Dev.	0.022***	0.001	0.037***
Age: Average	0.009***	0.005***	0.008***
Age: Std. Dev.	0.008***	0.003***	0.006***
Portion before 1940	0.029***	0.027***	0.024***
Adjusted R-square	0.388	0.462	0.425
F Value	401.7***	543.8***	467.3***

Table 6 Determinants of Idiosyncratic House Price Risk: MSA Fixed Effects with Bootstrapped Dependent Variable

This table reports results of the zip code level cross-sectional regressions regarding the determinants of idiosyncratic house price risk. The sample consists of 6,618 zip codes. The dependent variable is the standard deviation of *100 randomly selected residuals* from the hedonic regression for each zip code. The explanatory variables include MSA dummy variables, the 2000 census median household income (in $100,000) and its squared value, the median house value (in $100,000) and its squared value, the number of transactions (1,000), the housing market turnover (the ratio of the number of transactions to the owner-occupied single family housing units in 2000 census), the average and standard deviation of the size (1,000 square feet) of traded houses, the average and standard deviation of the age (10 years) of the traded houses, the portion of the traded houses that were built before 1940, and the arithmetic mean and the standard deviation of quarterly appreciation rates of the zip code house price index. The three regressions have identical specifications but different sample periods. Coefficients with ***, **, and * are significant at 1%, 5%, and 10% level.

	I	II	III
Variables	1996:Q1 - 2000:Q4	2001:Q1 - 2007:Q2	2007:Q3 - 2012:Q3
MSA Fixed Effect	Yes	Yes	Yes
Intercept	0.060***	0.051***	0.078***
Household Income	-0.171***	-0.161***	-0.184***
Household Income Squared	0.047***	0.053***	0.069***
Index Appreciation: Std. Dev.	0.379***	0.644***	0.444***
Index Appreciation: Average	0.151*	0.391***	-0.715***
Home Value	0.019***	0.013***	0.002
Home Value Squared	-0.001***	-0.001	0.000
Turnover	0.070***	0.006***	0.025***
Size: Average	0.017***	0.018***	0.010***
Size: Std. Dev.	0.021***	0.009***	0.032***
Age: Average	0.012***	0.011***	0.013***
Age: Std. Dev.	0.009***	0.004***	0.009***
Portion before 1940	0.031***	0.030***	0.034***
Adjusted R-square	0.493	0.558	0.539
F Value	38.0***	49.1***	45.4***

Figure 1. Histogram of Total Risk in Zip Codes (1996:Q1 - 2000:Q4)
The vertical axis is the number of zip codes, and the horizontal axis is total risk.

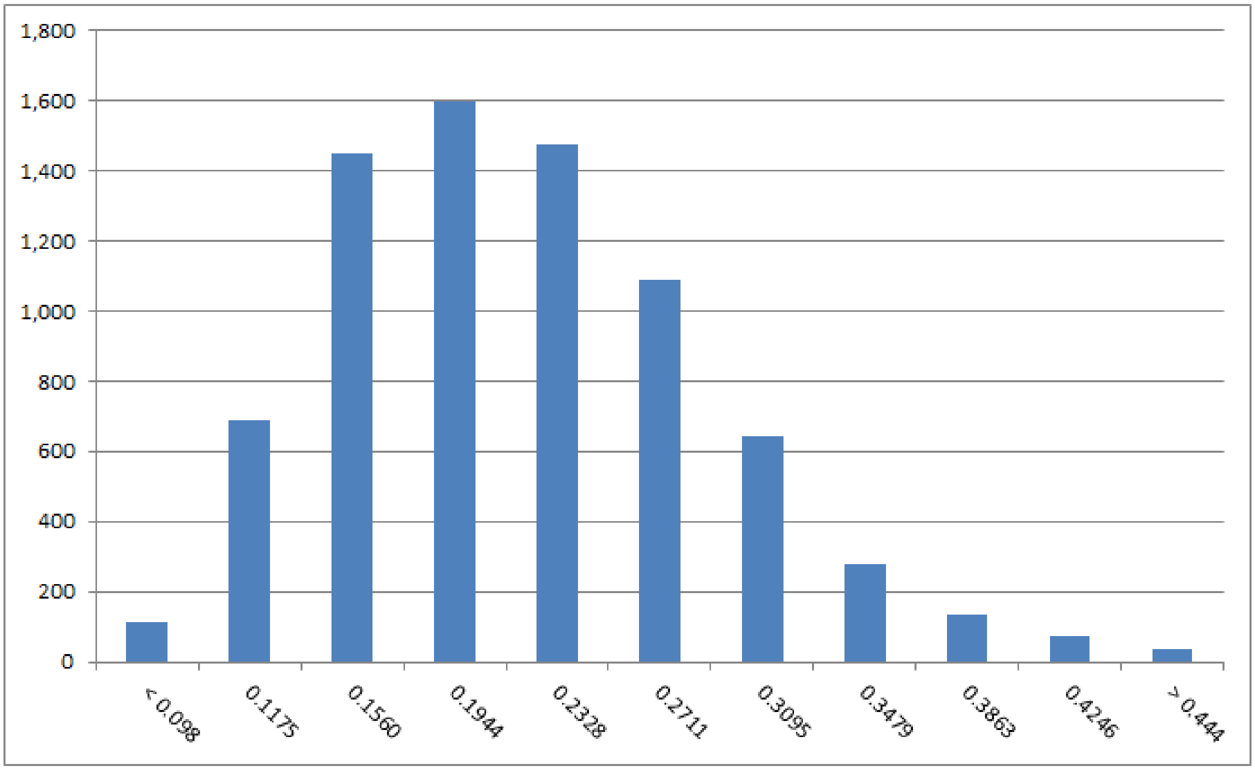

Figure 2. Histogram of Total Risk in Zip Codes (2001:Q1 - 2007:Q2)
The vertical axis is the number of zip codes, and the horizontal axis is total risk.

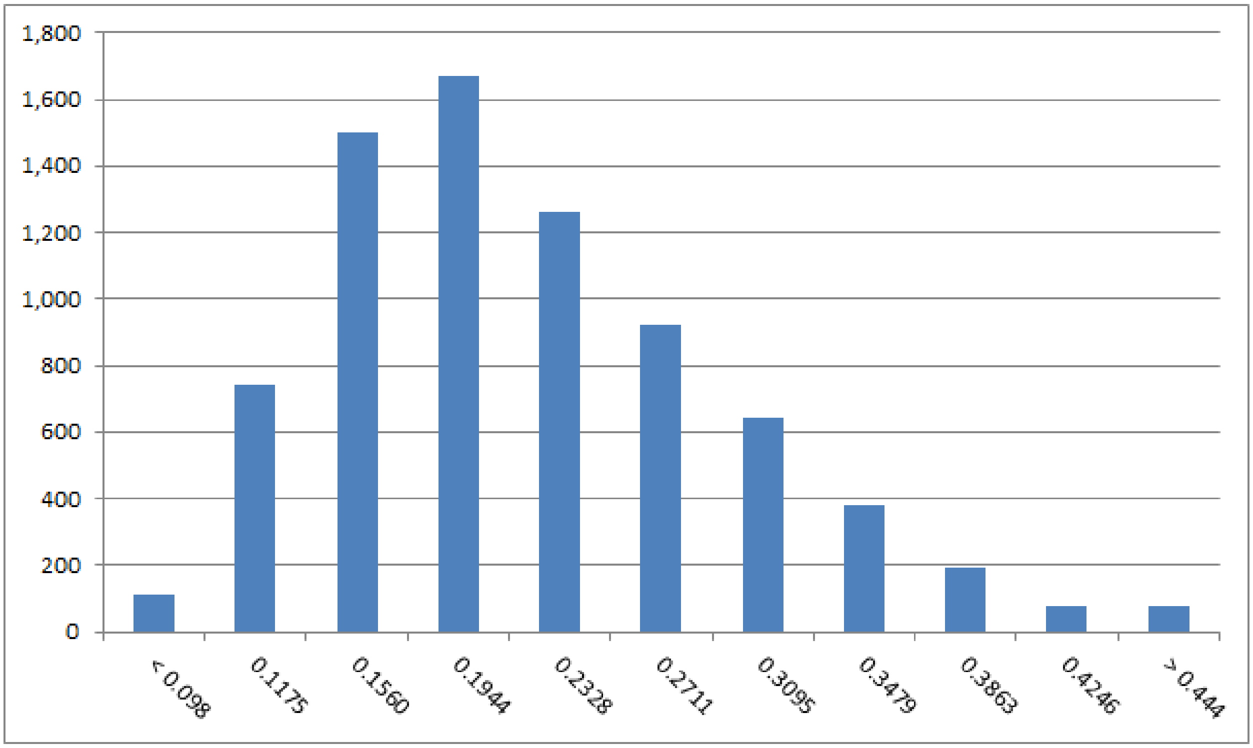

Figure 3. Histogram of Total Risk in Zip Codes (2007:Q3 - 2012:Q3)
The vertical axis is the number of zip codes, and the horizontal axis is total risk.

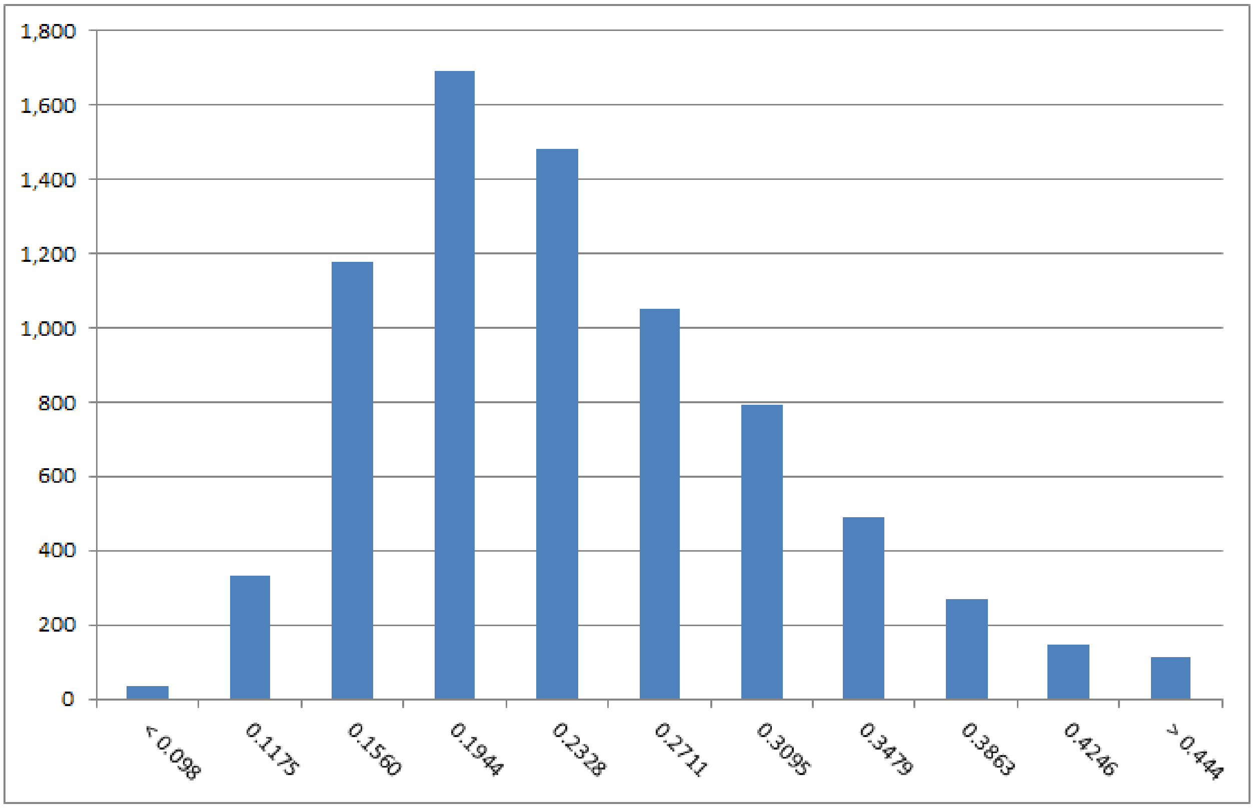

Figure 4. Histogram of 2000 Census Median Household Income

The vertical axis is the number of zip codes, and the horizontal axis is zip code median household income in 2000.

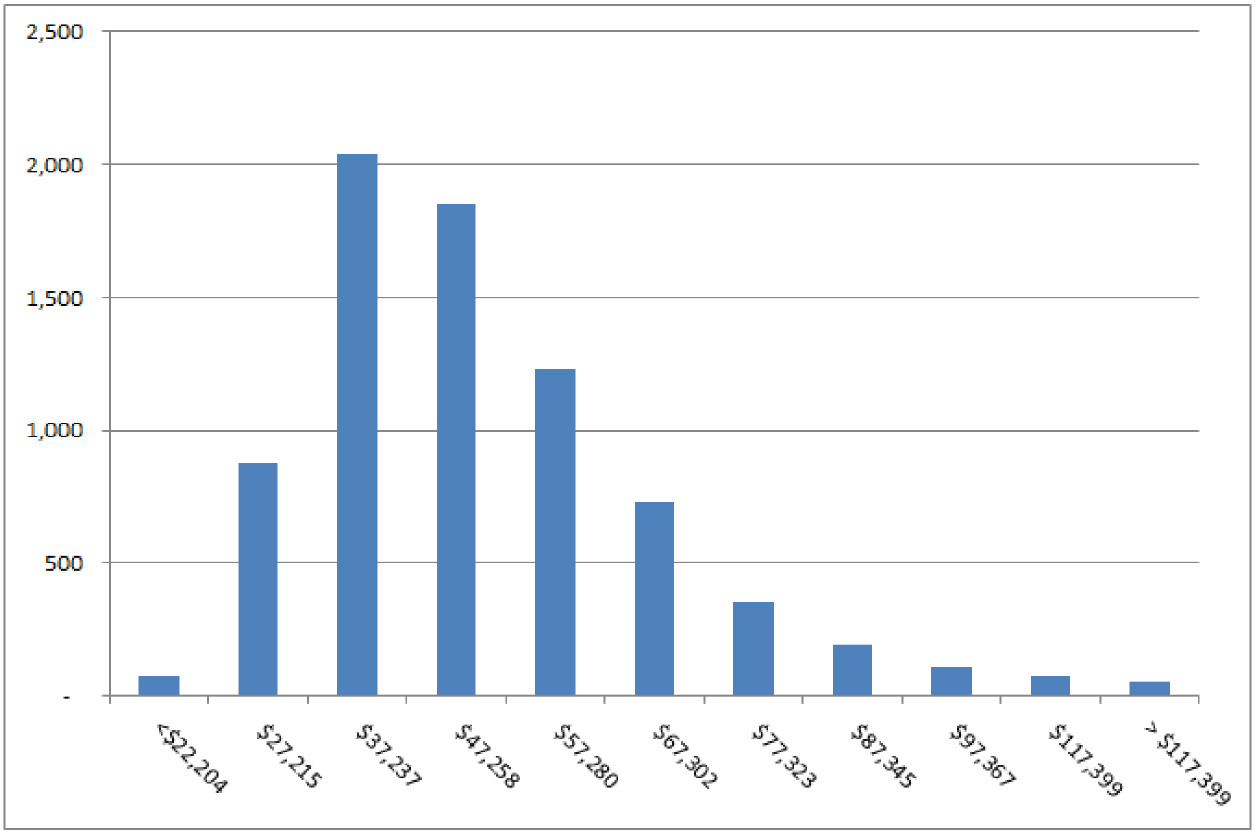

Figure 5. Total Risk and Income Percentiles (1996:Q1 - 2000:Q4)
The vertical axis is total risk, and the horizontal axis is the percentile of median household income in 2000.

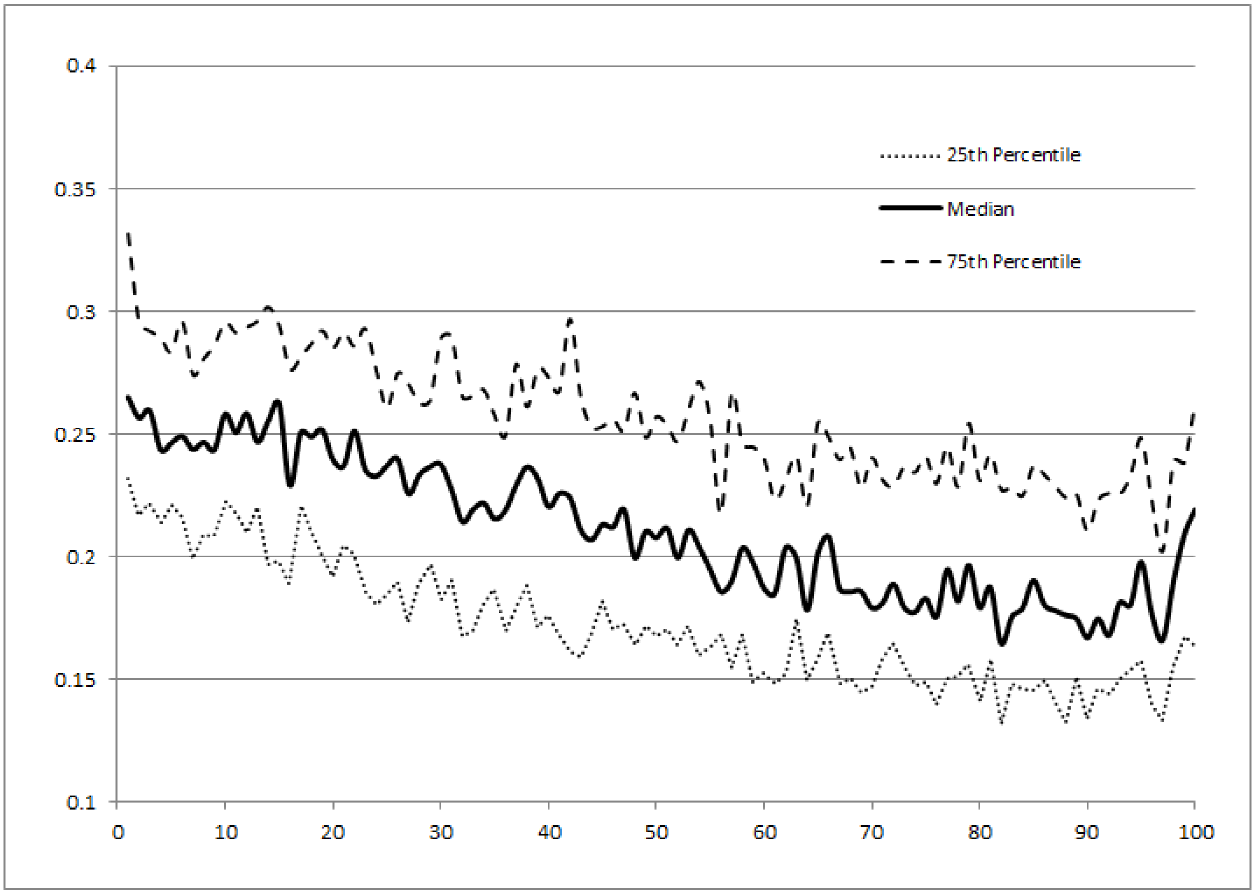

Figure 6. Total Risk and Income Percentiles (2001:Q1 - 2007:Q2)
The vertical axis is total risk, and the horizontal axis is the percentile of median household income in 2000.

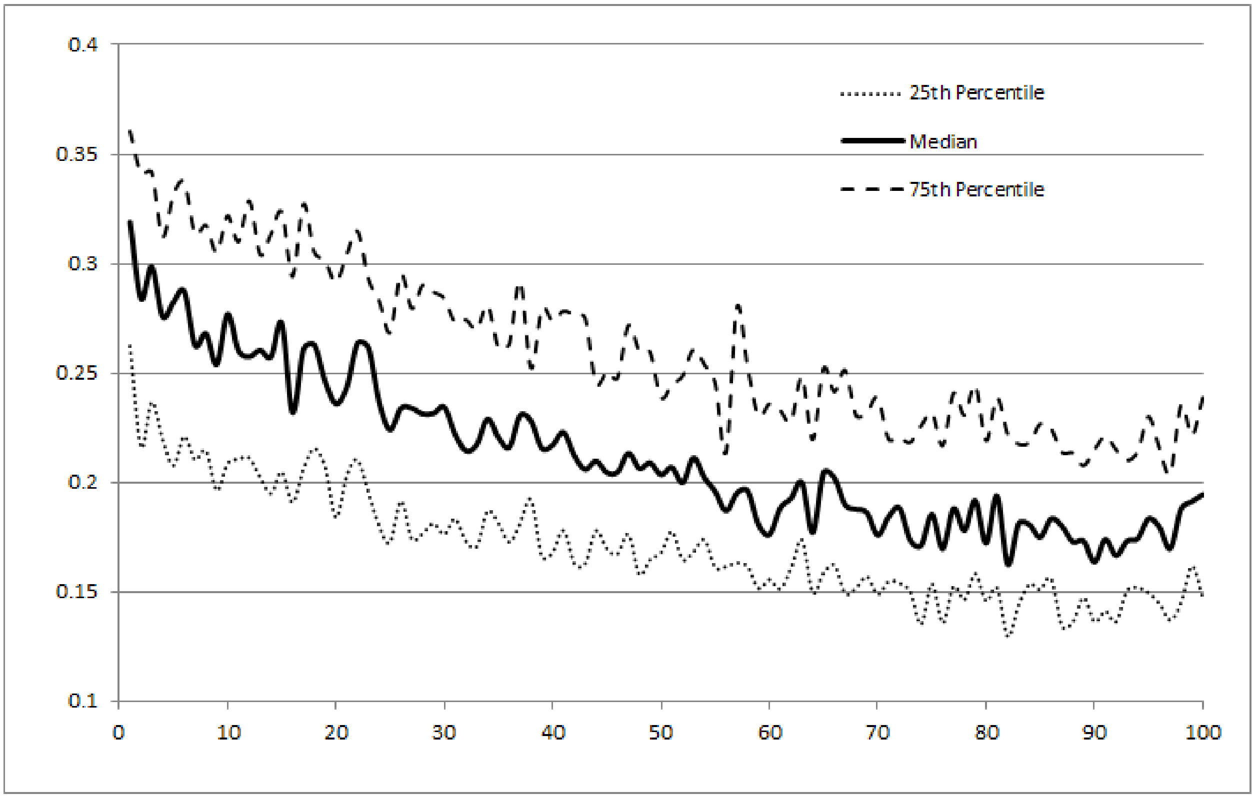

Figure 7. Total Risk and Income Percentiles (2007:Q3 - 2012:Q3)
The vertical axis is total risk, and the horizontal axis is the percentile of median household income in 2000.

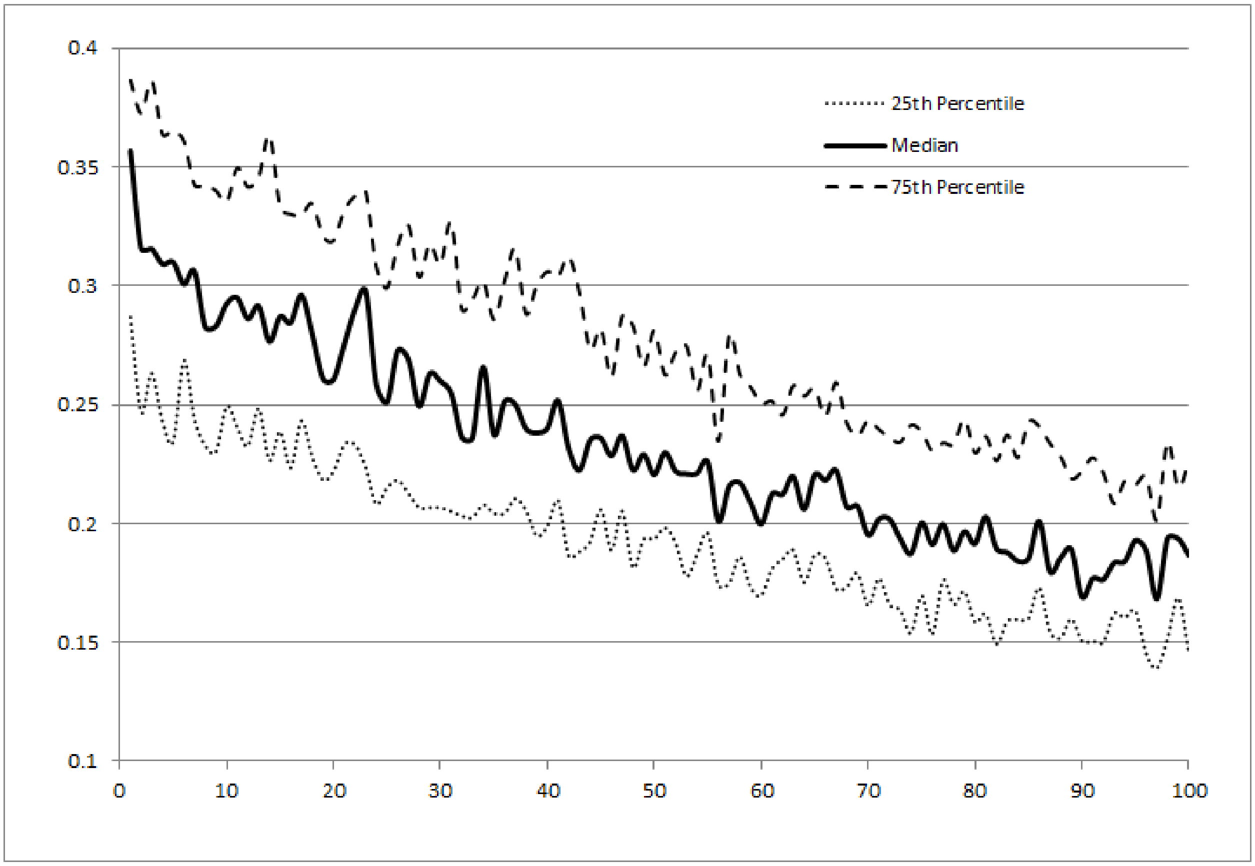

Figure 8. Risk-Income Relationship Suggested by Table 3
The vertical axis is idiosyncratic risk, and the horizontal axis is the median household income in 2000.

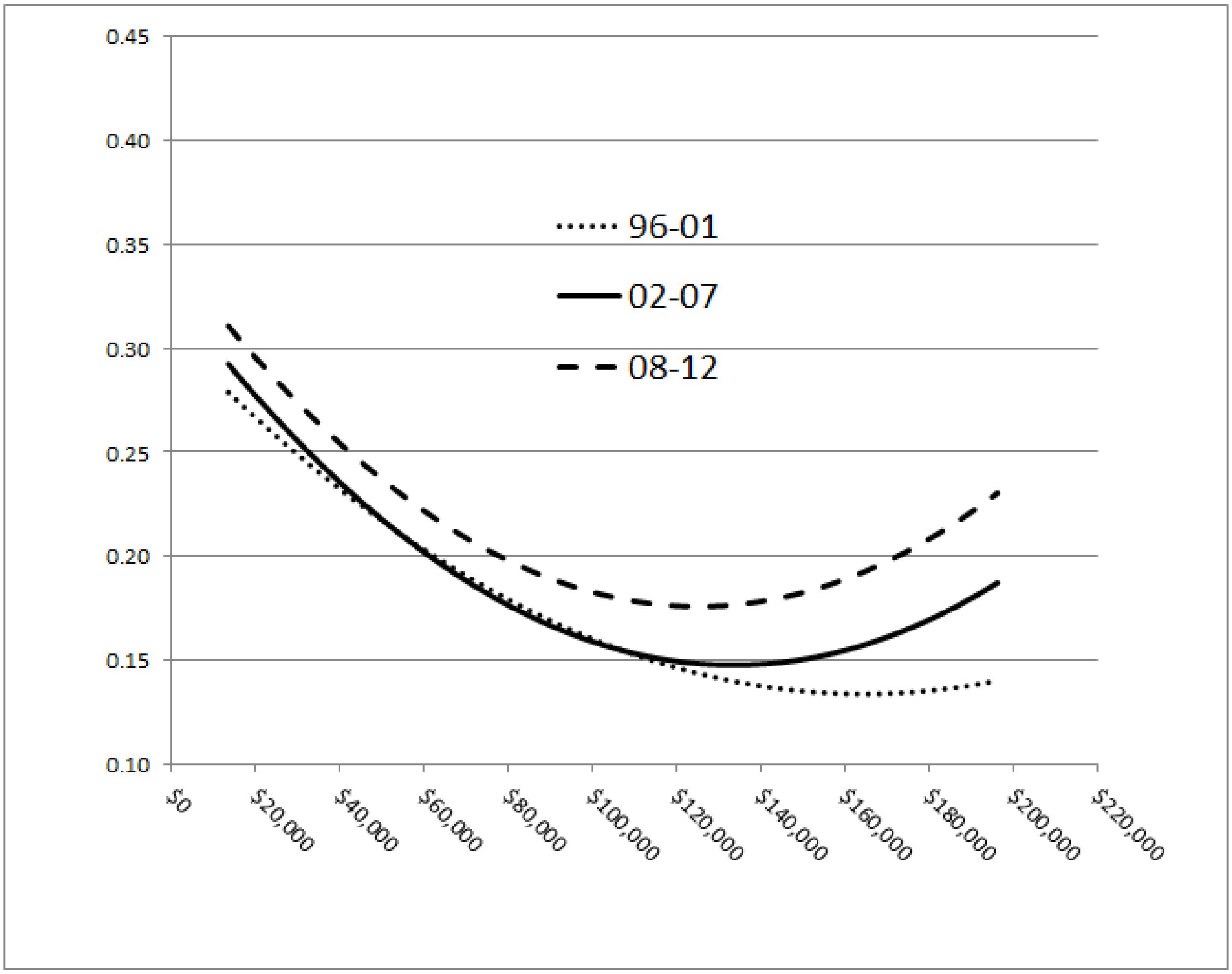
0.45
0.40
0.35
0.30
0.25
0.20
0.15
0.10
96-01
02-07
08-12
$0
$20,000
$40,000
$60,000
$80,000
$100,000
$120,000
$140,000
$160,000
$180,000
$200,000
$220,000

Figure 9. Risk-Income Relationship Suggested by Table 4
The vertical axis is idiosyncratic risk, and the horizontal axis is the median household income in 2000.

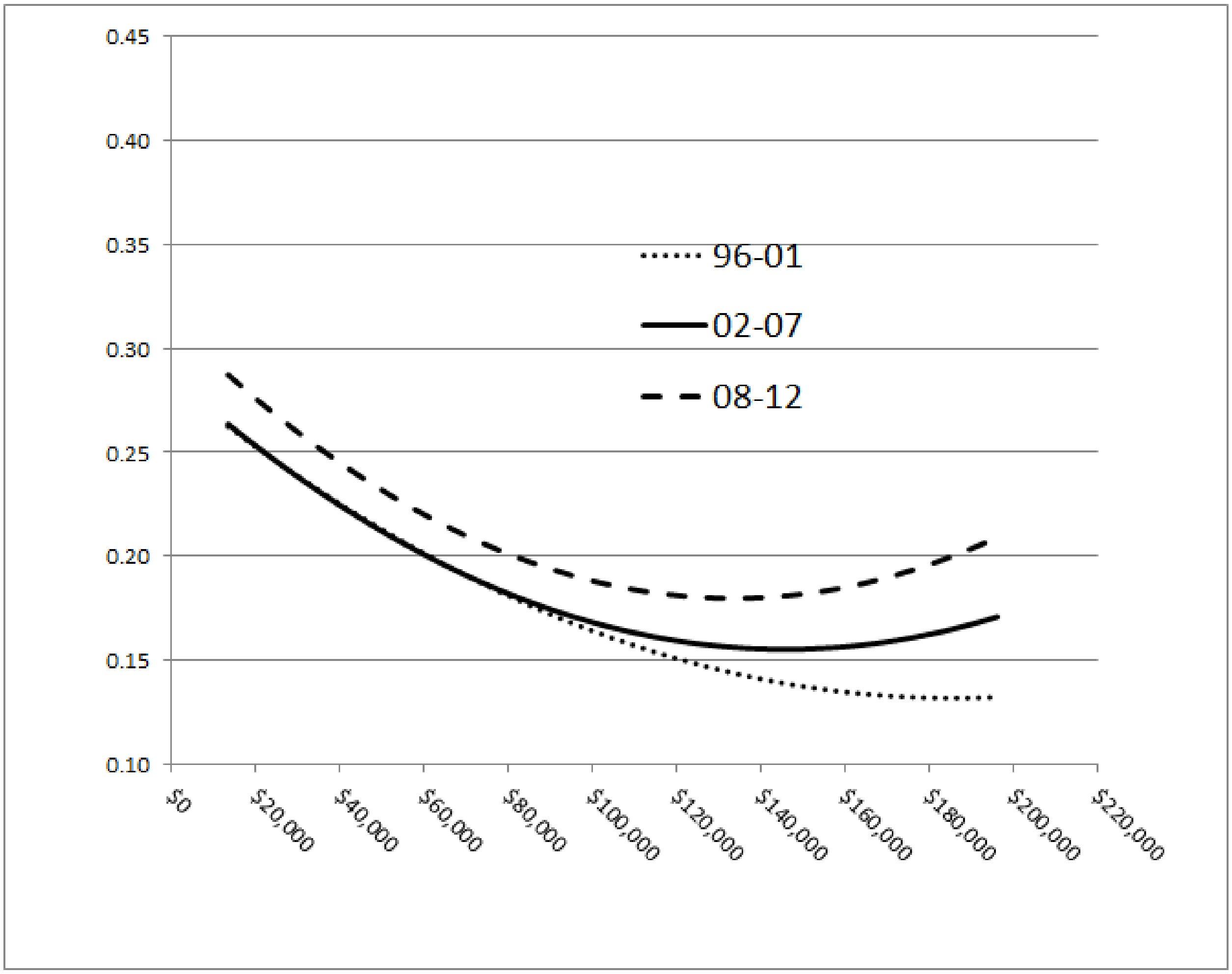

Figure 10. Risk-Income Relationship Suggested by Table 5
The vertical axis is idiosyncratic risk, and the horizontal axis is the median household income in 2000.

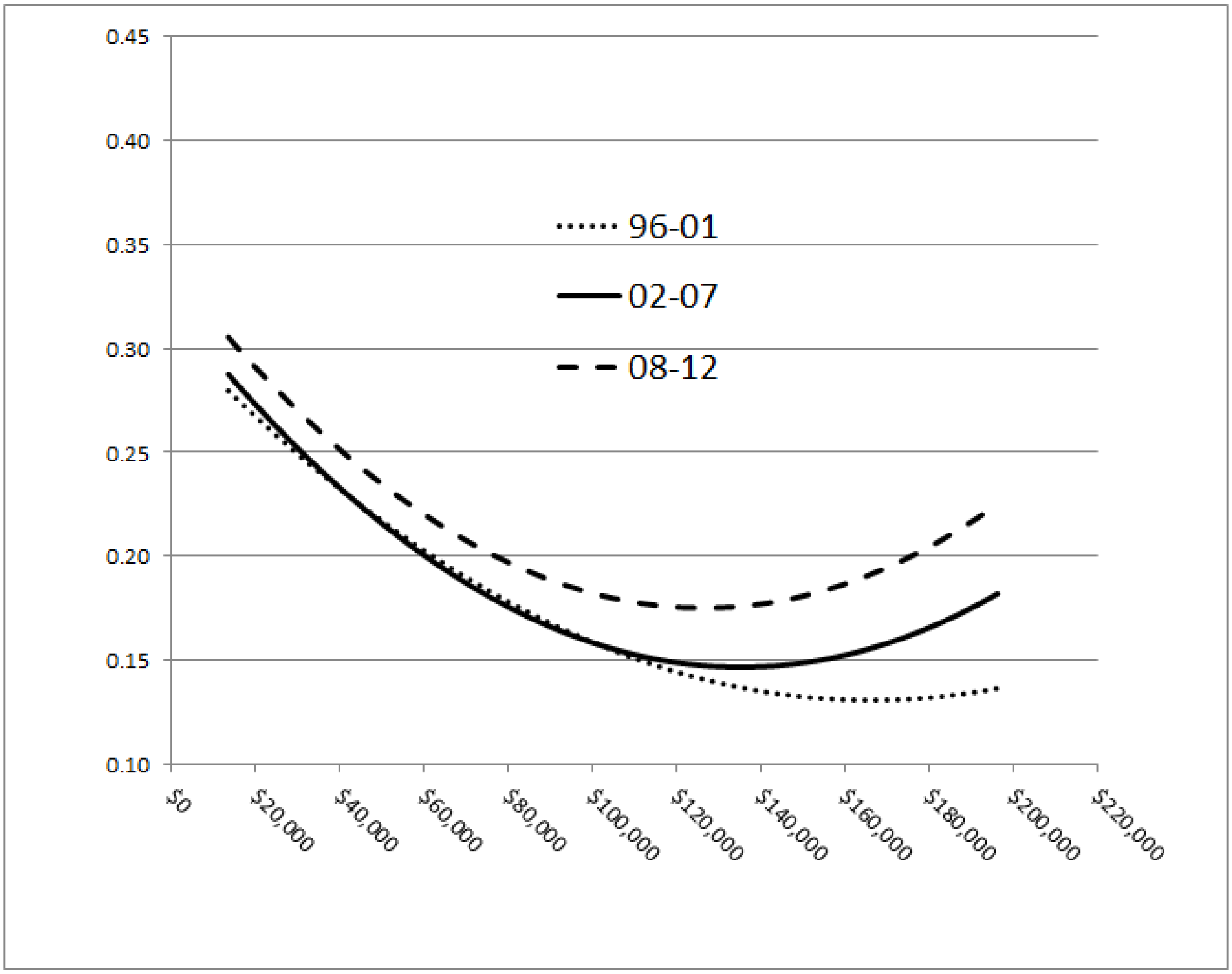

Figure 11. Risk-Income Relationship Suggested by Table 6
The vertical axis is idiosyncratic risk, and the horizontal axis is the median household income in 2000.