City Equilibrium with Borrowing Constraints: Structural Estimation and General Equilibrium Effects.[*]

Amine Ouazad[†] Romain Rancière[‡]

October 2017

Abstract

This paper develops a general equilibrium model of location choice with social interactions where mortgage approval rates determine household-specific choice sets that differ across neighborhoods and years in observable and unobservable dimensions. Existence and local uniqueness of city equilibria enable comparative statics estimates of the impact of changes in borrowing constraints on neighborhood-level prices and demographics. Estimation the model using micro data on property transactions, household demographics, neighborhood amenities, mortgage applications, and bank liquidity for the San Francisco Bay area, reveals that the price sensitivity of borrowing constraints explains about two-thirds of the price elasticity of neighborhood demand. General equilibrium estimates of the impact of the relaxation of lending standards on prices and neighborhood demographics bring two out-of-sample predictions for the period 2000-2006: (i) an increase in house prices accompanied by a compression of the price distribution and (ii) a reduction in the isolation of Whites in line with evidence of gentrification in the San Francisco Bay. Both predictions are supported by empirical observation.

Keywords: Location choice, urban economics, borrowing constraints, segregation.

JELs: R21, R23, G21

[*]We thank Patrick Bayer, Jan Brueckner, Morris Davis, Gilles Duranton, Matthew Kahn, Scott Page, Stuart Rosenthal, Holger Sieg, the audiences of the CEPR Conference in Urban and Regional Economics conference in Philadelphia, INSEAD, Ecole polytechnique, University of Southern California, University of Virginia, and World Bank seminar series for fruitful comments. Marton Varga provided research assistance. The authors acknowledge financial support from INSEAD, the International Monetary Fund, the Chaire Banque de France.

[†]HEC Montreal, 3000, Chemin de la Côte Sainte Catherine, Montreal H3T 2A7, Canada. *amine.ouazad@hec.ca*.

[‡]University of Southern California, Kaprielian Hall, 3620 S. Vermont Ave, Los Angeles, CA 90089, *romain-ranciere@gmail.com*.

1 Introduction

The ability to secure mortgage credit is a key determinant of households' decision to purchase housing. Changes in lending standards contribute to the dynamics of cities by affecting housing decisions and house prices. This paper proposes and estimates a novel structural model of city equilibrium under borrowing constraints. The estimated model explains the role played by mortgage approval constraints in determining housing demand and its price elasticity. The model is then used to estimate how changes in lending standards affect the city equilibrium, the distribution of house prices, and the distribution of households across neighborhoods.

Structurally estimated models of consumer demand for differentiated goods have been widely used to characterize the market for durable goods such as cars (Berry, Levinsohn & Pakes 1995) or houses (Bayer & Timmins 2007, Bayer, Ferreira & McMillan 2007). While these papers are able to characterize the full set of price and demand elasticities on their respective markets, they do not explicitly account for financial frictions, and do not focus on measuring the general equilibrium effects associated with changes in such frictions. Credit constraints are important in housing markets – more than 80% of purchases involve mortgage financing.[1] Price changes impact households through both preferences and borrowing constraints; and as neighborhood-level housing supply is largely inelastic, changes in lending standards can have a sizable effect on the distribution of housing prices.

The paper estimates a structural housing choice model which accounts for the effect of endogenous borrowing constraints on household demand. The model structure stresses the equilibrium relationships between mortgage market conditions, location decisions, and house prices. Comparative statics of the estimated model allow to measure, in general equilibrium, the impact of a change in lending standards set by mortgage credit suppliers, on the distribution of house prices across the city, and on the spatial segregation of households.

Borrowing constraints are introduced in the location choice model by endogenizing housing choice sets, whose probabilities can be estimated using loan applications data. The model yields a novel decomposition of the own-price demand elasticity into two distinct terms: (i) an elasticity term that reflects the impact of a price increase on utility, and (ii) an elasticity term that reflects

[1] 81% in 2013 according to data provided by RealtyTrac.

the impact of a price increase on mortgage approval rates. Both own-price demand elasticity terms differ across households and neighborhoods. The second term reflects differences in the intensity of borrowing constraints across households with different characteristics, and financing purchases in different neighborhoods.

The model yields estimates of household's preferences and willingness to pay for housing amenities for the 4,416 neighborhoods (census blockgroups) of the San Francisco Bay area between 1990 and 2010. The estimation combines data on the entire universe of mortgage credit applications and transaction prices, with exhaustive census blockgroup characteristics and individual characteristics for a sample of 120,029 households (1% census). Individual mortgage approval probability estimates are based on the observation of lenders' discrete approval or denial choice, which results from screening borrowers' application based on information on loan amount and leverage (loan-to-income ratios) and a range of borrower and neighborhood characteristics such as property value, and observable and unobservable neighborhood prospects. When mortgage origination constraints are taken into account, households exhibit a higher willingness to pay for schools with higher test scores. Furthermore a large fraction of house price elasticity can be attributed to the sensitivity of mortgage approval rates to changes in prices.

A comparative statics analysis of the estimated model, based on a change in lending standards of the magnitude experienced between 2000 and 2006, shows that the model predicts well key moments of the actual distribution of house price changes in the San Francisco Bay area, and replicates the compression of the price distribution. The model also predicts, in line with the data, a reduction in the isolation of Whites, a pattern in line with the evidence of gentrifications in several areas of the San Francisco Bay.

Estimating the model and running the general equilibrium comparative statics analysis requires addressing a number of novel methodological and empirical challenges.

Berry et al. (1995) have shown how observed demand maps one-to-one with neighborhood utilities. This paper extends this key result to the case in which each household demand is conditional on the set of neighborhoods for which households' mortgage applications can be approved with some probability. In addition to proving the existence of a city equilibrium under borrowing constraints, we show that such equilibrium is locally unique under general conditions, and globally unique in the absence of social preferences. Existence and local uniqueness allow us to run comparative statics analyses of the effects of changes in lending standards.

The structural parameters of the model, which include household preferences for a vector of key observable amenities and location-specific unobservable fixed effects are estimated using a simulated method of moments estimator with panel data.[2] The probability of approval for a loan application is estimated using Home Mortgage Disclosure Act (HMDA) data, which includes both the income and race of the applicant, and is matched to longitudinal data on neighborhood time-varying characteristics – including geocoded transaction prices. In estimating approval probabilities, we exploit the very unique characteristics of HMDA data which is to provide information on the universe of mortgage *applications* and their associated approval decisions, rather that inferring the intensity of borrowing constraints from the information about successful applicants only.[3]

The model's mortgage approval specification flexibly estimates how banks' lending standards depend on the rich heterogeneity in observable and unobservable individual, property, and neighborhood dimensions. Lending standards constrain households' neighborhood choice set, likely affecting household demand. The estimation of mortgage approval probabilities is subject to potential endogeneity biases due to unobservable borrowers' characteristics. We estimate the impact of borrower and mortgage characteristics using a set of instruments combining information on banks' balance sheet liquidity, at the national level, with information on the location of their branches across neighborhoods. Bank liquidity, derived from national balance sheets, predicts local loan characteristics, and is unlikely to be confounded by neighborhood-level unobservable demand factors. Finally, house prices are typically endogenous with respect to local unobservables, an issue shared with prior consumer choice models; this leads to an upward bias in the estimated impact of prices on neighborhood utility. We use the set of characteristics of two-step adjacent neighborhoods as instruments for neighborhood price, in a similar way as Bayer & Timmins (2007).

The model yields estimates of household preferences under borrowing constraints along seven broad dimensions: house price, neighbors' observables, housing quality, distance to the central business district, and school test scores as measured by California's Academic Performance Index. The model also estimates household preference heterogeneity using data from the 1% Census'

[2] The total number of potential choice sets leads to our use of Simulated GMM: 2 to the power of the number of neighborhoods is a very large number, leading to computationally intractable estimation when using non-simulated demand.

[3] Other datasets (e.g. loan performance data, corelogic) include additional informations on *originated loans* but they do not include approval decisions on *applications* or the characteristics of denied applications.

120,029 households: preferences for neighborhoods vary according to the household's race, income, and unobserved heterogeneity. In contrast with a model ignoring borrowing constraints, our model estimates stronger preferences for neighborhoods with high-performing public schools. Importantly, when the effect of a price change on mortgage approval rate is controlled for, the estimated price sensitivity of base utility is three times lower for a median income household.

The empirical decomposition of the estimation of total price elasticity between the conditional demand elasticity (assuming constant mortgage approval probabilities) and the borrowing elasticity (capturing the effect of a price change approval probabilities) reveals that borrowing elasticity accounts on average for about 70 percent of total price elasticity (the mean ratio of borrowing elasticity over total elasticity is 0.69). Borrowing elasticity exhibits a large degree of heterogeneity reflecting considerable differences in the intensity of the borrowing constraints across households, and displays a correlation with house prices that is more than twice that of conditional demand elasticity (-0.54 vs. -0.24).

Using the estimated structural model, we can measure the effect of lending standards on house prices and spatial segregation in general equilibrium. Our first comparative statics result shows that a relaxation in lending standards, as observed in the Bay area between 2000 and 2006, leads to a compression of the price distribution: both in the model and in the data, neighborhoods that were less expensive in 2000 saw a larger increase between 2000 and 2006 than initially more expensive neighborhoods.[4] Without using any information on the period 2001-2006 other than the realized change in lending standards, the model is able to approximate key moments of the 2000-2006 price change distribution, and yields a negative correlation between price changes and initial price change that is about 40 percent of the actual correlation (0.29 vs. 0.75). We also analyze the effect of the same change in lending standards on spatial segregation. Regarding changes in segregation, the model replicate qualitatively well several features of the data such as the reduction in the isolation of Whites, in line in line with evidence of gentrification in the San Francisco Bay.

We test the robustness of our comparative statics results to the introduction of an option to rent, to the introduction of a neighborhood-specific supply elasticity (constructed using satellite data from the U.S. Geological Survey on land development combined with local measures of land

[4] A similar cross-sectional dispersion in house price changes has been documented for San Diego by Landvoigt, Piazzesi & Schneider (2015).

slope and ruggedness), and to changes in population size and racial composition due to migration flows in and out of the San Francisco Bay. The robustness tests yield results that are similar to our baseline findings.

This paper contributes to the literature along several dimensions. Our general equilibrium comparative statics analysis is, to our knowledge, the first attempt in the literature to measure the effects of changes in lending standards on the distribution of prices and households within an estimated structural framework of a city equilibrium. Such analysis is particularly useful in the case of housing (as opposed to consumer products) because housing supply elasticity is very far from perfectly elastic, and therefore demand shocks typically lead to substantial changes in the price distribution. By using an estimated general equilibrium approach, this paper complements the literature linking changes in credit conditions to changes in house prices based on regression discontinuity design frameworks (Favara & Imbs 2015, Adelino, Schoar & Severino 2012) or on calibrated models (Glaeser, Gottlieb & Gyourko 2012, Landvoigt et al. 2015).

We introduce borrowing constraints, modeled as household's choice sets, in the workhorse model of the consumer choice literature developed in Berry, Levinsohn & Pakes (1995) and Petrin (2002) for consumer products, in Bayer & Timmins (2007) and Bayer, Ferreira & McMillan (2007) for housing.[5] While the approach bears similarity with the modeling of incomplete product availability (Conlon & Mortimer 2013) or limited information on products (Goeree 2008), it differs from it in two very important ways. First, the reduction of the choice set in our model is due to the endogenous decision of banks to approve mortgage loan applications. Second, the reduction of the choice set in our model comes from the (housing) demand side rather than from the supply side.

The paper focuses on estimating the effect of a change in lending standards on city equilibrium over a medium run horizon (e.g. between 2000 and 2006). Running comparative statics in a static framework keeps the general equilibrium analysis tractable and computationally manageable with a large number of neighborhoods, and a rich structure of heterogeneity in preferences as well as in mortgage approval decisions. In doing so we leave aside the short-run dynamics of housing markets (e.g. the role of vacancies and moving costs). The dynamic aspects of housing demand decision are treated by Bayer, McMillan, Murphy & Timmins (2016) using a structural framework estimated at annual frequency.

[5]See Holmes & Sieg (2015) for a comprehensive survey of structural estimation in urban economics.

The paper proceeds as follows. Section 2 introduces the economic model of location choice with borrowing constraints and its equilibrium properties. Section 3 introduces the empirical approach for identifying (section 3.1) and estimating (section 3.2) the model by simulated generalized method of moments. Empirical findings are described in Section 4. Section 5 presents the general equilibrium analysis of a change in lending standards. Section 6 extends to the general equilibrium analysis to incorporate the impact of local housing supply elasticity and the role of tenure choice. Section 7 concludes.

2 The Model

We model a metropolitan area as made of a finite number of neighborhoods, each composed of housing units, and inhabited by a corresponding population of households. Time is discrete and there is a finite number of time periods $t = 1, 2, \ldots, T$. In year t, the metropolitan area has a number N_t of households indexed by $i = 1, 2, \ldots, N_t$. There is a fixed number of $J \geq 2$ neighborhoods and each neighborhood $j = 1, 2, \ldots, J$ in year $t = 1, 2, \ldots, T$ has a number s_{jt} of housing units. Household i in year t chooses a neighborhood $j(i,t) \in \{1, 2, \ldots, J\} \equiv \mathbb{J}$, where $j(i,t)$ is constrained to belong to the *choice set* $C_{it} \subset \mathbb{J}$ of household i in year t. The choice set C_{it} is a subset of $\mathbb{J}$ and is equal to $\mathbb{J}$ if and only if household i is unconstrained in year t.

(i) Neighborhood Choice $j(i,t)$ Conditional on Credit Availability

Household i derives utility U_{ijt} from neighborhood j in year t. Household i's neighborhood choice $j(i,t)$ is based on maximizing utility within the choice set C_{it}, i.e. $j(i,t) = \text{argmax}_{j \in C_{it}} U_{ijt}$. Utility U_{ijt} is allowed to vary across neighborhoods, across individuals, and to vary across neighborhoods according to individual characteristics. Thus:

$$U_{ijt} = \underbrace{\delta_{jt} + \mathbf{x}'_{it}\Omega\mathbf{z}_{jt} + \tilde{\boldsymbol{\beta}}'_{it}\mathbf{z}_{jt}}_{u_{ijt}} + \varepsilon_{ijt} \tag{1}$$

where u_{ijt} is the explained part of household utility. All throughout the paper, bold symbols denote vectors. Household i's observable characteristics are stacked in a vector $\mathbf{x}_{it}$ of size $K_1 \geq 1$. Similarly, neighborhood j's observable characteristics are summarized by a vector $\mathbf{z}_{jt}$ of size $K_2 \geq 1$.[6] In utility specification (1), δ_{jt} is the base utility that all households derive from living in neighborhood j in year t. The vector $\boldsymbol{\delta}_t$ thus captures variation in utility across neighborhoods independently of individual characteristics. The term $\mathbf{x}'_{it}\Omega\mathbf{z}_{jt}$ captures the observable heterogeneity of neighborhood utility across households. It is a series of interaction terms between the K_1 household characteristics $\mathbf{x}_{it}$ and the K_2 neighborhood characteristics $\mathbf{z}_{jt}$, where Ω is the matrix of K_1K_2 interaction coefficients. Therefore $\mathbf{x}'_{it}\Omega\mathbf{z}_{jt}$ captures how neighborhood utility varies according to households' observable characteristics. Additionally, the term $\tilde{\boldsymbol{\beta}}'_{it}\boldsymbol{z}_{jt}$ captures the unobservable heterogeneity of households' preferences for neighborhood amenities. $\tilde{\boldsymbol{\beta}}_{it}$ is a normally-distributed vector of household-level heterogeneity with mean 0 and with variance-covariance matrix Σ. The standard deviation of $\tilde{\beta}_{itk}$, i.e. the k-th diagonal element of Σ, is noted σ_k, $k = 1, 2, \ldots, K_2$. Normalizing each random coefficient $\tilde{\beta}_{itk}$ by the corresponding standard deviation σ_k allows utility to be written as a function of a standard normal random term $b_{itk} = \tilde{\beta}_{itk}/\sigma_k$. Following Berry et al. (1995), the explained utility u_{ijt} of one neighborhood, say neighborhood 1, is set to zero by convention, so that $U_{i1t} \equiv \varepsilon_{i1t}$.

In turn, base utility δ_{jt} is decomposed into the impact of the log price, of observable local amenities $\mathbf{z}'_{jt}\boldsymbol{\beta}$, and of unobservable neighborhood heterogeneity ξ_j:

$$\delta_{jt} = -\alpha \log(p_{jt}) + \mathbf{z}'_{jt}\boldsymbol{\beta} + \xi_j + \zeta_{jt} \tag{2}$$

In utility specification (1), the random term ε_{ijt} is extreme-value distributed. Following McFadden (1973), household i's probability of choosing neighborhood j in year t, within his choice set C_{it}, is:

$$P\left(U_{ijt} \geq U_{ikt},\, \forall k \in C_{it} \,\middle|\, \boldsymbol{\delta}_{\cdot t}, \mathbf{z}_{\cdot t}, \mathbf{x}_{it}, \tilde{\boldsymbol{\beta}}_{it}, C_{it}\right) = \mathbf{1}(j \in C_{it}) \cdot \frac{\exp(\delta_{jt} + \boldsymbol{x}'_{it}\Omega\mathbf{z}_{jt} + \tilde{\boldsymbol{\beta}}'_{it}\mathbf{z}_{jt})}{\sum_{k \in C_{it}} \exp(\delta_{kt} + \boldsymbol{x}'_{it}\Omega\mathbf{z}_{kt} + \tilde{\boldsymbol{\beta}}'_{it}\mathbf{z}_{kt})} \tag{3}$$

where $\mathbf{1}(j \in C_{it}) = 1$ whenever $j \in C_{it}$, and 0 otherwise. Integrating over the vector of random coefficients $\tilde{\boldsymbol{\beta}}_{it}$ obtains the probability conditional on observables and on base utility:

$$P(j, t | \boldsymbol{\delta}_{\cdot t}, \mathbf{z}_{\cdot t}, \mathbf{x}_{it}, C_{it}) = \int_{\tilde{\boldsymbol{\beta}}} P(j, t | \boldsymbol{\delta}_{\cdot t}, \mathbf{z}_{\cdot t}, \boldsymbol{x}_{it}, \tilde{\boldsymbol{\beta}}, C_{it}) \cdot f(\Sigma^{-1/2}\tilde{\beta}) \cdot d\tilde{\beta} \tag{4}$$

[6] $\mathbf{z}_{jt}$ can also include the characteristics of loans for location j for household i in year t. This is discussed in Section 4.3.

where f is the density function of the standard i.i.d. multivariate normal of dimension K_2.

(ii) Endogenizing the Choice Set $C_{it} \subset \mathbb{J}$

This paper focuses on constraints on the choice set driven by mortgage credit approval decisions. Lending standards vary flexibly across individuals, locations, and years, and thus there does not typically exist a ranking of neighborhoods by increasing approval rates independently of household characteristics. A neighborhood j belongs to the choice set C_{it} of household i in year t if a mortgage lender approves household i's application in neighborhood j.[7] The probability that a neighborhood j belongs to household i's choice set is thus the probability that a mortgage lender approves household i's application.

The probability of a choice set $C_{it} \subset \mathbb{J} = \{1, 2, \ldots, J\}$ is the product of banks' approval probabilities for the applications of household i. ϕ_{ijt} is the probability that household i is approved for a mortgage in neighborhood j in year t. With independent draws, the probability that household i's choice set C_{it} is exactly C is the binomial probability:

$$P(C_{it} = C) = \Pi_{j \in C} P(j \in C) \Pi_{j \notin C} \left[1 - P(j \in C)\right] \equiv \Pi_{j \in C} \phi_{ijt} \Pi_{j \notin C} (1 - \phi_{ijt}) \tag{5}$$

Such formula assumes independent and identically distributed decisions, and an extension of the model allows for non-zero correlation in banks' decisions across neighborhoods for a given household i.[8]

The probability ϕ_{ijt} of approval is the outcome of banks' benefit-cost analysis of originating a mortgage loan for household i for a house in neighborhood j. Such benefit-cost analysis is summarized in a latent variable $approves^*_{ijt}$. Household i's application in neighborhood j is approved if and only if $approves^*_{ijt} \geq 0$. The benefit-cost $approves^*_{ijt}$ depends on household observables $\mathbf{x}_{it}$ and neighborhood characteristics $\mathbf{z}_{jt}$.

$$j \in C \quad \text{whenever} \quad approves^*_{ijt} = \mathbf{x}'_{it}\boldsymbol{\psi} + \mathbf{z}'_{jt}\boldsymbol{\gamma} + \mathbf{x}'_{it}\Psi\mathbf{z}_{jt} - e_{ijt} > 0 \tag{6}$$

[7] The endogeneity of households' choice of mortgage lender when applying is addressed in Section 4.2.

[8] Results available from the authors. An updated formula 5 features the correlation of binomial draws across neighborhoods. Assuming independent approval probabilities gives a lower bound of the impact of borrowing constraints: with uncorrelated outcomes, an individual who applies to multiple neighborhoods obtains a higher approval rate for at least one application.

Neighborhood characteristics $\mathbf{z}_{jt}$ include a neighborhood fixed effect that captures location-specific underwriting unobservables. There will be spatial variation in lending standards if, for instance, expected trends in the value of the housing collateral differ across locations. The neighborhood's log price either enters directly in $\mathbf{z}_{jt}$, or enters as the log of the loan-to-income ratio in both $\mathbf{x}_{it}$ (log household income) and $\mathbf{z}_{jt}$ (log house price). In both cases, a change in log price affects the probability of the choice set. Ψ is the matrix of interaction terms: neighborhood characteristics affect the probability of approval non-monotonically as the sign of $\partial approves^*_{ijt}/\partial z_{jtl} = \gamma_l + \sum_{k=1}^{K} \psi_{kl} x_{itk}$ depends on household characteristics.

e_{ijt} is an i.i.d. logistically distributed random heterogeneity with cumulative distribution function $\Lambda(\cdot)$.[9] The probability of approval in neighborhood j for household i is therefore simply a logistic function of individual and household observables $P(j \in C_{it}) = \Lambda\left(\mathbf{x}'_{it}\boldsymbol{\psi} + \mathbf{z}'_{jt}\boldsymbol{\gamma} + \mathbf{x}'_{it}\Psi\mathbf{z}_{jt}\right)$. Combining (5) and such logit specification, the probability of the choice set depends on the vector of all neighborhood and household covariates:

$$P(C_{it} = C|\mathbf{x}_{it}, \mathbf{z}_{\cdot t}) = \Pi_{j\in C}\Lambda\left(\mathbf{x}'_{it}\boldsymbol{\psi} + \mathbf{z}'_{jt}\boldsymbol{\gamma} + \mathbf{x}'_{it}\Psi\mathbf{z}_{jt}\right)\Pi_{j\notin C}\left\{1 - \Lambda\left(\mathbf{x}'_{it}\boldsymbol{\psi} + \mathbf{z}'_{jt}\boldsymbol{\gamma} + \mathbf{x}'_{it}\Psi\mathbf{z}_{jt}\right)\right\} \tag{7}$$

(iii) Neighborhood Demand

Total demand for neighborhood j in year t is the weighted average of the conditional demand for neighborhood j, conditional on each possible choice set. The set of choice sets $\mathbb{C}$ is the set of all possible subsets of $\mathbb{J} = \{1, 2, \ldots, J\}$. The probability of a choice set C has been derived above.

$$P(j, t|\boldsymbol{\delta}_{\cdot t}, \mathbf{z}_{\cdot t}, \mathbf{x}_{it}) = \sum_{C\in\mathbb{C}} \underbrace{P(C|\mathbf{x}_{it}, \mathbf{z}_{\cdot t}, j, t)}_{probability\,of\,choice\,set\,C} \cdot \underbrace{P(j, t|\boldsymbol{\delta}_{\cdot t}, \mathbf{z}_{\cdot t}, x_{it}, C)}_{demand\,conditional\,on\,set\,C} \tag{8}$$

As the set of all possible choice sets $\mathbb{C}$ has 2^J elements, total demand for neighborhood j in year t is the sum of 2^J conditional demands. Equation (8) is the total demand of households with characteristics $\mathbf{x}_{it}$ with a given income, race, education, age, and other characteristics. The last step in obtaining total demand is to integrate over the distribution of individual demographics. We

[9]The paper's section 4.2 also considers a normal distribution for e_{ijt} when instrumenting the covariates. Results are similar.

note $D(j,t)$ such total demand for neighborhood j in year t:

$$D(j,t) \equiv P\left(j,t\,|\boldsymbol{\delta},\mathbf{z}_{\cdot t},f(\cdot)\right) = \sum_{C\in\mathbb{C}}\int_{\mathbf{x}} P(C|\mathbf{z}_{\cdot t},x_{it})\cdot P(j,t|\boldsymbol{\delta},\mathbf{z}_{\cdot t},x_{it},C_{it})\cdot f(\mathbf{x}_{it})d\mathbf{x}_{it} \tag{9}$$

And $f(\mathbf{x}_{it})$ is the city-level distribution of household characteristics. Such distribution depends on time t and thus allows for changes in the demographic make-up of the city between t and $t+1$, e.g. driven by migrations in and out of the metropolitan area.

(iv) Demand Elasticities

Changes in prices affect both household utility and choice set probabilities. Hence demand elasticities conditional on the choice set differ from total demand elasticities. The first kind of (own-price) demand elasticity $\eta_{|C}$ is equal to the effect of a change in the log price on demand while keeping the choice set probabilities $P(C_{it})$, $i=1,2,\ldots,N$, $t=1,2,\ldots,T$ constant. For a fixed probability distribution $P(C)$ over choice sets, we note $\eta_{jt|C}$ the own-price demand elasticity for neighborhood j in year t conditional on $P(C)$.

$$\begin{aligned}\eta_{jt|C} &= \frac{1}{D(j,t)}\left.\frac{\partial D(j,t)}{\partial\log(p_{jt})}\right|_{P(C)\,fixed} \qquad (10)\\ &= -\frac{1}{D(j,t)}\sum_{C\in\mathbb{C}}\int_X \alpha(\mathbf{x}_{it},\tilde{\boldsymbol{\beta}}_{it})\cdot P(C|\mathbf{z}_{\cdot t},\mathbf{x}_{it})P(j,t|\boldsymbol{\delta},\mathbf{z}_{\cdot t},\mathbf{x}_{it},C)\left(1-P(j,t|\boldsymbol{\delta},\mathbf{z}_{\cdot t},\mathbf{x}_{it},C)\right)f(\mathbf{x}_{it})d\mathbf{x}_{it}\end{aligned}$$

where $\alpha(\mathbf{x}_{it},\tilde{\boldsymbol{\beta}}_{it})$ is the impact of log price on base utility for a household with characteristics $\mathbf{x}_{it}$ and random coefficient $\tilde{\boldsymbol{\beta}}_{it}$ (Equation (2)). In the simplified case where the model does not feature random coefficients, and when there is a large number of neighborhoods so that $P(j)\ll 1$, such conditional demand elasticity is well approximated by $-\alpha$.

Prices also affect the probabilities $P(C|\mathbf{z}_{\cdot t},\mathbf{x}_{it})$ over choice sets $C\in\mathbb{C}$. Total demand elasticity η is equal to the sum of two effects: the price effect on demand due to utility changes, and the price effect on demand due to changes in choice set probabilities,

$$\eta_{jt} = \eta_{jt|C} + \sum_{C\in\mathbb{C}}\int_{\mathbf{x}}\left[\frac{\partial}{\partial\log(p_{jt})}P(C|\mathbf{z}_{\cdot t},\mathbf{x}_{it})\right]\cdot P(j,t|\boldsymbol{\delta}_{\cdot t},\mathbf{z}_{\cdot t},\mathbf{x}_{it},C)\cdot f(\mathbf{x}_{it})d\mathbf{x}_{it} \tag{11}$$

The second term of this expression (11) is the borrowing elasticity. An increase in house price in neighborhood j causes a decline in demand because of (i) the lower utility value of the neighborhood and (ii) the lower approval rate in neighborhood j.

The impact of a *log* price change on the choice-set probability is simply derived from the individual approval probabilities.

$$\begin{aligned} \frac{\partial}{\partial \log(p_{jt})} P(C|\mathbf{z}_{\cdot t}, \mathbf{x}_{it}) &= \frac{\partial}{\partial \log(p_{jt})} \{\Pi_{k \in C} \phi_{ijt} \Pi_{k \notin C} (1 - \phi_{ijt})\} \\ &= -a \cdot \underbrace{(\mathbf{1}(j \in C) - \phi_{ijt})}_{\equiv \Phi_{ijt}(C)} P(C|\mathbf{z}_{\cdot t}, \mathbf{x}_{it}) \end{aligned} \tag{12}$$

where $-a$ is the coefficient of *log* price in the approval specification (6). The closed-form expression of the borrowing elasticity follows by plugging (12) into (11).

(v) City Equilibrium

Households choose their location based on the set of amenities that includes neighbors' demographics – households may exhibit preferences for neighbors' education, income, and race. Therefore, two sets of conditions need to be satisfied at equilibrium.

The first set of equilibrium conditions is the equality of the demand and supply of housing units for each neighborhood. Subsections (i)–(iv) above have introduced household demand and its properties. The supply of housing in neighborhood j depends on the log price of housing in neighborhood j such that the supply of housing $S_j(p_{jt})$ in neighborhood j is an increasing and concave function of p_{jt}. In this paper's empirical analysis, Sections 4–5, the supply of housing units is initially assumed perfectly inelastic, so that $S_j(p_{jt}) = s_{jt}$ for every j, t.[10]

Households have preferences for neighborhood demographics whenever the vector $\mathbf{z}_{jt}$ of neighborhood characteristics includes other households' demand for the same neighborhood. Neighborhood characteristics $\mathbf{z}_{jt} = (\mathbf{w}_{jt}, \mathbf{v}_{jt})$ thus comprise a set $\mathbf{w}_{jt}$ of exogenous characteristics (e.g. the age of structures in the neighborhood) and a set of endogenous neighbor demographics $\mathbf{v}_{jt}$ (e.g. fraction of households with a college degree in neighborhood j). Note $v_{jt}(x)$ the fraction of households with characteristic x in neighborhood j in year t. Write X^+ the finite set of neighbors' characteristics (e.g. neighbors' income and race) that affect household utility, then $\mathbf{v}_{jt} = (v_{jt}(x),\, x \in X^+)$.

[10] Supply elasticity is introduced and estimated in Section 6.2.

Definition 1. *Given the distribution of household observables $f(\mathbf{x}_{it})$ in the city, given the vector of neighborhood observables $\mathbf{z}_{jt}$, excluding log price, and neighborhood-level unobservables $\boldsymbol{\xi}_j$, $\boldsymbol{\zeta}_{jt}$, for each neighborhood $j = 1, 2, \ldots, J$, in each year $t = 1, 2, \ldots, T$, a* city equilibrium *is a vector of prices and neighborhood demographics $(\mathbf{p}^*, \mathbf{v}^*) = \left(p^*_{jt}, \mathbf{v}^*_{jt}\right)_{j=2,\ldots,J;t=1,2,\ldots,T}$ such that:*

$$\forall j \in \mathbb{J} \setminus \{1\}, \forall t \in \{1, 2, \ldots, T\}, \quad D(j, t; \mathbf{p}^*, \mathbf{v}^*) = S_j\left(p^*_{jt}\right) \tag{13}$$

$$\forall x \in X^+, \quad v_{jt}(x) = \frac{N_x}{N} \cdot \frac{D(j, t|x; \mathbf{p}^*, \mathbf{v}^*)}{D(j, t; \mathbf{p}^*, \mathbf{v}^*)} \tag{14}$$

where N_x is the number of households with characteristics x. Condition (13) *expresses the equality of the demand and supply of housing. Condition* (14) *expresses the equality of households' neighborhood demographics and the demand of each demographic group. $D(j, t|x; \mathbf{p}^*, \mathbf{v}^*)$ is the demand for neighborhood j in year t of a household with characteristic x. $\frac{D(j,t|x;\mathbf{p}^*,\mathbf{v}^*)}{D(j,t;\mathbf{p}^*,\mathbf{v}^*)}$ is the fraction of such households as derived from equilibrium neighborhood demand.*

The Appendix's lemma 1 proves equilibrium existence using Brouwer's theorem. An equilibrium is locally unique as in Debreu (1970), proven in Lemma 2. Local uniqueness enables comparative statics analysis of lending standards around the observed equilibrium (Section 5). A stronger result of global uniqueness is obtained in Lemma 3 when households do not exhibit preferences for peers' demographics, using an argument of gross substitutability. The three equilibrium properties are grouped in the following proposition:

Proposition 1. (Equilibrium properties) *The city equilibrium exists, i.e. there exists an equilibrium vector of prices and neighborhood demographics $(\mathbf{p}^*, \mathbf{v}^*)$. The city equilibrium is locally unique almost surely: with probability 1, the Jacobian $\frac{\partial \mathbf{D}}{\partial(\mathbf{p},\mathbf{v})}$ of the demand vector w.r.t. $(\mathbf{p}, \mathbf{v})$ is of full rank. Equivalently, with probability 1, there exists an open set ω that contains $(\mathbf{p}^*, \mathbf{v}^*)$ such that, $(\mathbf{p}^*, \mathbf{v}^*)$ is the unique equilibrium vector in ω. The city equilibrium is globally unique if households do not exhibit preferences for peers' demographics.*

While the city typically exhibits multiple, locally unique, equilibria, the next section shows that observations of neighborhood characteristics and mortgage approvals leads to the identification of a unique set of structural preference parameters and lending standards.

2.1 Discussion of the set-up

The introduction of mortgage approval constraints through endogenous choice set probabilities is the key innovation of the model presented above. The rest of the model is a general equilibrium framework of location choice with social interactions. In the baseline model, the utility that a household i derives from living in neighborhood j depends on the log price. An alternative, following Berry et al. (1995), is to make it depend on of the log non-housing consumption, i.e. the log of the difference between household income and mortgage payments. We derive and estimate this alternative set-up in Section 6 of the online appendix. The main additional advantage of this framework is to introduce a role for mortgage payments, thus capturing another dimension of lending standards. In the baseline model, we also choose ignore the option to rent and focus instead on alternative choices for buying a house. Such rental option is introduced in an extension of the model discussed in Section 6.1. In another extension of the model, presented in Section 6.2, we introduce housing supply elasticity (in the basic model housing supply is inelastic). The goal of introducing these extensions is both to show the flexibility of the model and of its structural estimation procedure, and to assess the robustness of the estimation and general equilibrium results.

3 Identification and Estimation

This section presents the empirical approach for the identification (section 3.1) and estimation (section 3.2) of borrowing constraints and of households' preferences for neighborhood amenities and neighborhood demographics.

3.1 Identification Strategy

The model includes three sets of parameters to estimate, which we identify using three sets of corresponding moments. The first set of parameters pins down the probability of approval for a mortgage application, given the characteristics of borrower, lender, and houses, as in the lending standards' specification (6). This paper identifies such parameters using an instrumental variable strategy. The second set of parameters measures the impact of neighborhood observable and unobservable characteristics on base utility (specification (2)). Log house prices are typically positively correlated with unobservable amenities. The paper presents a set of IV moments as in by Berry et al. (1995)

and Bayer & Timmins (2007) that use both neighborhood-specific fixed effects and the exogenous characteristics of second-degree adjacent neighborhoods. Finally, the third set of estimated parameters measures how the valuation of neighborhood amenities depends on household characteristics. We identify such preference heterogeneity using micromoments as in Imbens & Lancaster (1994) and Petrin (2002).

This paper's identification strategy relies on the use of mortgage approval data for the estimation of choice set probabilities. Abaluck & Adams (2017) proposes an alternative approach: endogenous choice sets are estimated without the requirement of auxiliary data such as mortgage approval data. Such methodology would yield estimates of the sensitivity of choice sets to price changes. It would not allow however identify the full relationship between choice sets and observables, that is, in our case, the mortgage approval rule. By contrast, this paper's methodology, while staying computationally tractable, builds on direct evidence of mortgage approval constraints, and thus allows for policy experiments of changes in lending standards.

Note that the potential multiplicity of equilibrium in the model does not prevent point identification. Following the same logic of Brock & Durlauf (2001) we show that the linearity (or polynomial nature) of social preferences, together with logit demand, implies identification of the model. The detailed, step by step proof of identification of this paper's model is provided in Section 2.1 of the online appendix.

(i) Lending Standards

The approval specification (6) predicts the probability of approval based on the characteristics of the household, the time-varying characteristics of the neighborhood, neighborhood fixed effects, and interactions between individual and neighborhood characteristics. The neighborhood fixed effect captures non time-varying unobservables that are correlated with the unobservable collateral (house) value for the loan and have an impact on the probability of approval.

The identification challenge lies at least (i) in the lack of observability of the full range of time-varying household and house covariates that mortgage originators consider in the underwriting process, (ii) in the endogeneity of households' decision to apply for a mortgage, and (iii) in the endogeneity of the choice of a specific mortgage lender.[11] Empirical Section 4.2 below presents

[11] Such endogeneity concerns are formalized in Section 2.5 of the Online Appendix.

a set of instrumental variables $\mathbf{\Theta}_{ijt}$ based on banks' national balance sheets that are arguably uncorrelated with the unobservables $\mathbf{e}_{ijt}$ of specification (6); determine banks' lending standards and thus the probability of approval; and predict the characteristics $\mathbf{x}_{it}$ of mortgage applicants. Such instrumental variables, used in the context of a logit, provide the first set of moment conditions:

$$E\left[G_0\left(\psi_0,\gamma_0,\omega_0\right)\right]=E\left[\Lambda\left(\mathbf{x}_{it}'\psi+\mathbf{z}_{jt}'\gamma+\mathbf{x}_{it}'\Psi\mathbf{z}_{jt}\right)\mathbf{\Theta}_{ijt}\right]=0 \tag{15}$$

(ii) Base Utility

The second set of moments aims at identifying the impact of neighborhood amenities and neighborhood demographics on neighborhood base utility (Specification (2)). Identification and estimation of such parameters $(\alpha,\boldsymbol{\beta},\boldsymbol{\xi})$ encounters at least three challenges (i) neighborhood utility U_{ijt} and base utility δ_{jt} in specification (1) are not directly observable; (ii) demand, defined in equation (9), is the sum of a large number of 2^J terms, one for each potential choice set $C_{it}\in 2^{\{1,2,\dots,J\}}$, (iii) unobservable neighborhood amenities ζ_{jt} may be correlated with the price of housing and with observable neighborhood amenities (e.g. school quality).

On the first issue, Berry et al. (1995) has shown that, given the interaction parameters Ω, and the variance of the random coefficients Σ, there is a one-to-one mapping between the vector of observed demands $\mathbf{D}=(D_{jt})_{j\in\mathbb{J}\setminus\{1\},t\in\{1,2,\dots,T\}}$ and the vector of base utilities $\boldsymbol{\delta}=(\delta_{jt})_{j\in\mathbb{J}\setminus\{1\},t\in\{1,2,\dots,T\}}$, whenever households are unconstrained, i.e. $\phi_{ijt}=1$ for all households, all neighborhoods and all years. The Appendix's Proposition 3 shows that a similar one-to-one mapping exists when demand is the weighted average of demands conditional on each choice set. The unique vector $\boldsymbol{\delta}$ is obtained by iterating the following sequence:

$$\hat{\boldsymbol{\delta}}^k=\hat{\boldsymbol{\delta}}^{k-1}+\log\left(D_{jt}\right)-\log\left(D\left(j,t|\hat{\boldsymbol{\delta}}^{k-1}\right)\right) \tag{16}$$

and at the limit $\hat{\boldsymbol{\delta}}=\lim_{k\to\infty}\hat{\boldsymbol{\delta}}^k$ of such sequence, predicted demand $D\left(j,t|\lim_{k\to\infty}\hat{\boldsymbol{\delta}}^k\right)$ is equal to observed demand D_{jt}.

In the contraction mapping (16) , demand $D\left(j,t|\hat{\boldsymbol{\delta}}\right)$ is the sum of demand for all possible choice sets $C\subset\{1,2,\dots,J\}$, weighted by the probability of that choice set $P(C)$. With J neighborhoods, there are 2^J choice sets. In our empirical application where households choose among more than

4000 neighborhoods, computing demand over a sum of 2^{4000} choice sets is unfeasible. We proceed by simulating $D\left(j,t|\hat{\boldsymbol{\delta}}\right)$. A set of $\mathcal{S}$ sets noted C_{it}^s, $s=1,2,\ldots,\mathcal{S}$, is drawn for each household i in each year t, where the probability of drawing the set $C \subset \mathbb{J}$ of neighborhoods is $\Pi_{j\in C}\hat{\phi}_{ijt}\Pi_{j\notin C}\left(1-\hat{\phi}_{ijt}\right)$. Estimation of demand $D\left(j,t|\hat{\boldsymbol{\delta}}\right)$ given base utilities, and observables thus is the average of the demands based on the simulated choice sets C_{it}^s:

$$\hat{D}\left(j,t|\hat{\boldsymbol{\delta}}\right)=\sum_{i=1}^{N}\frac{1}{\mathcal{S}}\sum_{s=1}^{\mathcal{S}}D\left(j,t|\hat{\boldsymbol{\delta}},C_{it}^s\right) \tag{17}$$

with $D\left(j,t|\hat{\boldsymbol{\delta}},C_{it}^s\right)$ the demand conditional on the choice set C_{it}^s. We apply Train's (2009) simulated method to obtain a consistent estimate of demand. Furthermore, we rely on Monte Carlo integration following Geweke (1996) to integrate over household characteristics $\int_{\mathbf{x}} f(\cdot)dx$ and over the random coefficients $\int_{\tilde{\boldsymbol{\beta}}} f(\cdot)d\tilde{\boldsymbol{\beta}}$. The simulated demand $\hat{D}$ thus replaces the actual demand D in the sequence $(\boldsymbol{\delta}^k)$ specified in (16).

The contraction mapping method provides an estimate $\hat{\delta}_{jt}$ for each neighborhood in each year. The next step is to identify the impact of neighborhood amenities on base utility. Least squares applied to regression (2) typically yields a positive coefficient on log price, while consumer theory suggests a negative coefficient $-\alpha$. Higher prices are positively correlated with unobservable amenities, leading to an upward bias on the log(price) coefficient in households' utility function. Empirical section 4.3 introduces L' time-varying instruments $\boldsymbol{\Xi}=\Xi_{jt}$ arguably orthogonal to the unobservable neighborhood amenities ζ_{jt}, providing a set of moment conditions for the estimation of the impact of observable neighborhood amenities, house value, neighbors' demographics $\mathbf{z}_{jt}$ on base utility δ_{jt}.

As the specification is estimated in panel, specification (2) relating base-utility to house and neighborhood characteristics $\mathbf{z}_{jt}$ is first-differenced. The first moment condition expresses the orthogonality of $\Delta\boldsymbol{\Xi}$ with $\Delta\boldsymbol{\zeta}$:

$$E(\Delta\boldsymbol{\Xi}\Delta\boldsymbol{\zeta})=0 \tag{18}$$

where the change in unobservable amenities $\Delta\zeta_{jt}$ depends on observable neighborhood characteristics: $\Delta\zeta_{jt}=\Delta\delta_{jt}-\Delta\mathbf{z}_{jt}'\boldsymbol{\beta}$, whose estimator $\Delta\hat{\zeta}_{jt}=\Delta\hat{\delta}_{jt}-\Delta\mathbf{z}_{jt}'\boldsymbol{\beta}$ follows from the previously estimated vector of base utilities $\hat{\boldsymbol{\delta}}$.

$\Delta\boldsymbol{\zeta}$ depends on lending standards $(\boldsymbol{\psi}, \boldsymbol{\gamma}, \Psi)$ and on preference parameters $(\alpha, \boldsymbol{\beta}, \boldsymbol{\xi}, \Omega, \Sigma)$ as the contraction mapping depends on predicted approval probabilities $\hat{\phi}_{ijt}$ and on the non linear parameters (Ω, Σ). We write the moment condition (18) as a function $E\left[G_1(\theta)\right] = E\left[\Delta\Xi\Delta\boldsymbol{\zeta}(\boldsymbol{\theta})\right] = 0$ of the vector $\boldsymbol{\theta} = (\boldsymbol{\psi}, \boldsymbol{\gamma}, \Psi, \alpha, \boldsymbol{\beta}, \boldsymbol{\xi}, \Omega, \Sigma)$ of lending standards and preference parameters.

Moment condition (18) together with demand simulation (17) and the contraction mapping (16) allows for the estimation of common households' preferences $\boldsymbol{\beta}$ for neighborhood amenities; households' preference heterogeneity is estimated in the next subsection.

(iii) Preference Heterogeneity

The matrix of interaction coefficients Ω in utility specification (1) measures observable heterogeneity in household preferences. The variance covariance matrix $\Sigma = E\left[\tilde{\boldsymbol{\beta}}_{it}\tilde{\boldsymbol{\beta}}_{it}'\right]$ captures unobservable heterogeneity in household preferences. Identification of the parameters Ω and Σ in (1) is provided by a third set of micro moment conditions following Imbens & Lancaster (1994) and Petrin (2002). Specifically, the third set of micro moments ensures that the predicted spatial distribution of households, e.g. by race and income, matches the observed distributions of such households across neighborhoods.

Two data sources provide moment conditions useful for such identification. First, Census micro data provides the characteristics $\mathbf{x}_{it}$ of a sample of size N_t where $i = 1, 2, \ldots, N_t$. Second, neighborhood-level data provides elements of the distribution of household characteristics $\mathbf{x}_{it}$ in each neighborhood $j \in \mathbb{J}$. For instance, the Census provides log median family income per neighborhood (blockgroup in the empirical section). Such information provides a third set of moments that matches the predicted demand by demographic subgroup for each neighborhood to the actual demand by this same subgroup. If $\mathcal{X}$ is a subset of households such as, for example, the set of households with income above \$50,000, then the following moment states that the number of households with characteristics $\mathcal{X}$ predicted by the model must be equal to the number of such households in each neighborhood.

$$E\left(\sum_{x_{it}\in\mathcal{X}} D\left(j, t|\hat{\boldsymbol{\delta}}_{\cdot t}, \mathbf{z}_{\cdot t}, \mathbf{x}_{it}, C_{it}\right)\right) - D_{jt}^{\mathcal{X}} = 0 \tag{19}$$

with $D_{jt}^{\mathcal{X}}$ the observed number of such households in neighborhood j in year t.

The sum $\sum_{\mathbf{x}_{it}\in\mathcal{X}} D\left(j, t|\hat{\boldsymbol{\delta}}_{\cdot t}, \mathbf{z}_{\cdot t}, \mathbf{x}_{it}, C_{it}\right)$ is the demand for neighborhood j from individuals with characteristics in $\mathcal{X}$ as predicted by the model. A number L' of subsets $\mathcal{X}_1, \ldots, \mathcal{X}_{L'}$ of households provides a set of L' moment conditions, $E\left[G_2\left(\boldsymbol{\theta}\right)\right] = 0$.

(iv) Structural Estimation

Lending standards $(\boldsymbol{\psi}, \boldsymbol{\gamma}, \Psi)$ are estimated jointly with preferences $(\alpha, \boldsymbol{\beta}, \boldsymbol{\xi}, \Omega, \Sigma)$. We stack the L instrumental variable moment conditions of the approval specification, the L' moment conditions of base utility analysis, and the L'' moment conditions for subpopulations. Writing G as (G_0, G_1, G_2), parameters satisfy:

$$E\left(G\left(\boldsymbol{\psi}, \boldsymbol{\gamma}, \Psi, \alpha, \boldsymbol{\beta}, \boldsymbol{\xi}, \Omega, \Sigma\right)\right) = 0 \tag{20}$$

A consistent and asymptotically normal estimator of household preferences and banks' lending standards follows from Hansen (1982) as the minimand $\hat{\boldsymbol{\theta}} = \text{argmin}_{\theta \in \Theta} G^*(\boldsymbol{\theta})' G^*(\boldsymbol{\theta})$, where $\boldsymbol{\theta} = (\boldsymbol{\psi}, \boldsymbol{\gamma}, \Psi, \alpha, \boldsymbol{\beta}, \boldsymbol{\xi}, \Omega, \Sigma)$ and $G^*(\boldsymbol{\theta}) = A(\tilde{\boldsymbol{\theta}})\hat{G}(\boldsymbol{\theta})$, and A is a consistent estimate of the square root of the inverse of the asymptotic variance-covariance matrix of the moments,[12] obtained using $\tilde{\boldsymbol{\theta}}$, a preliminary consistent estimate of $\boldsymbol{\theta}$. $\hat{G}(\boldsymbol{\theta})$ is the simulated sample analogue of $G(\boldsymbol{\theta})$.

3.2 Estimation Steps

The model is estimated using three sources of data: the mortgage application and approval data, the neighborhood data, and the household micro data. As the first set of moments is taken from a different sampling process than the two other moments, estimating the lending standards $\hat{\boldsymbol{\psi}}$, $\hat{\boldsymbol{\gamma}}$,$\hat{\Psi}$ can be performed separately in a first step. We then proceed in the following way for the estimation of household preferences $\hat{\alpha}, \hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\xi}}, \hat{\Omega}, \hat{\Sigma}$:

1. For each observation i, t of the household sample:
   (a) Predict the probability of approval $\hat{\phi}_{ijt}$ for each neighborhood j using the estimate $\Lambda\left(\mathbf{x}_{it}'\hat{\boldsymbol{\psi}} + \mathbf{z}_{jt}'\hat{\boldsymbol{\gamma}} + \mathbf{x}_{it}'\hat{\Psi}\mathbf{z}_{jt}\right)$.
   (b) Draw $\mathcal{S}$ choice sets for individual i by drawing a dummy variable for each neighborhood j equal to 1 with probability $\hat{\phi}_{ijt}$. This defines a matrix of 0s and 1s, noted C_{ijt}.

[12] Section 2.4 of the Online Appendix discusses the implication of correlations between the lending standards' moments and moments identifying households' preferences for amenities.

(c) Draw a vector of i.i.d. multivariate standard normal random coefficients $\tilde{\mathbf{b}}_{it}$.[13]

2. Then we minimize the GMM objective function $G^*(\boldsymbol{\theta})'G^*(\boldsymbol{\theta})$. To obtain $\hat{G}(\boldsymbol{\theta})$ for a given vector of parameters $\boldsymbol{\theta}$:

   (a) Estimate the vector $\hat{\boldsymbol{\delta}}_t$ of base utilities using the contraction mapping (16) in each period $t = 1, 2, \ldots, T$.

   (b) At each iteration of the contraction mapping, simulated total demand is $\hat{D}(j,t) = \frac{1}{N_t}\sum_{i=1}^{N_t}\hat{D}(j,t|\mathbf{x}_{it},\mathbf{z}_{jt},\hat{\boldsymbol{\delta}}_{\cdot t},\tilde{\boldsymbol{\beta}}_{it})$, the average of individual demands. And simulated individual demand is $\hat{D}(j,t|\mathbf{x}_{it},\mathbf{z}_{jt},\hat{\boldsymbol{\delta}}_{\cdot t},\tilde{\boldsymbol{\beta}}_{it}) = \mathbf{1}(C_{ijt}=1)\cdot\frac{\exp(\delta_{jt}+\boldsymbol{x}'_{it}\Omega\mathbf{z}_{jt}+\tilde{\boldsymbol{\beta}}'_{it}\mathbf{z}_{jt})}{\sum_{k,s.t.C_{ikt}=1}\exp(\delta_{kt}+\boldsymbol{x}'_{it}\Omega\mathbf{z}_{kt}+\tilde{\boldsymbol{\beta}}'_{it}\mathbf{z}_{kt})}$. The $\tilde{\boldsymbol{\beta}}_{it}$ are obtained by multiplying $\tilde{\mathbf{b}}_{it}$ by $\Sigma^{1/2}$.

   (c) Perform the panel instrumental variable regression of base utilities $\hat{\boldsymbol{\delta}} = (\hat{\boldsymbol{\delta}}_{\cdot t}; t = 1, 2, \ldots, T)$ on neighborhood covariates instrumented by the vector Ξ, to obtain the panel residuals $\Delta\hat{\zeta}(\boldsymbol{\theta})$ and compute the second empirical moments, i.e. the orthogonality of instruments and residuals $\Delta\Xi\Delta\hat{\zeta}(\boldsymbol{\theta})$.

   (d) For each demographic subgroup $\mathcal{X}_1, \ldots, \mathcal{X}_L$, estimate the demand of that subgroup $\mathcal{X}_l$ for each neighborhood j using step (b)'s individual demands. This gives the second set of moment conditions.

   (e) Steps (c) and (d) together give the last two moments G_1 and G_2 of the objective function $\hat{G}(\boldsymbol{\theta})$. The weighting matrix is diagonal in the first step estimation, and, in the second step, equal to an estimate of the inverse of the variance covariance matrix of the moments based on the first step estimate $\hat{\boldsymbol{\theta}}$.

Further details of the optimization algorithm are presented in Appendix section 2.2. Standard errors for GMM estimates are computed as described in Section 2.3 of the Online Appendix.

[13] We use antithetic acceleration as in Geweke (1988) and Goeree (2008). Antithetic acceleration enables the use of simulated GMM standard errors.

4 Empirical Findings

4.1 Data

The model is estimated for the 9 counties that are contiguous to the San Francisco Bay: Alameda, Contra Costa, Marin, Napa, San Francisco, San Mateo, Santa Clara, Solano, and Sonoma. Data is gathered for three decades: in 1990, 2000, and 2010.[14]

Mortgage Application and Property Transaction Data

Mortgage application and approval data derives from data collected in accordance with the Home Mortgage Disclosure Act (HMDA). The act mandates reporting of mortgage application data by most depository and non-depository lending institutions.[15] Thus, HMDA disclosure requirements apply to more than 90% of all mortgage applications and originations (Dell'Arriccia, Igan & Laeven 2009). We focus on credit standards for first lien mortgages on homeowner occupied houses.[16] Each mortgage lender reports the loan amount, the applicant's income, the applicant's race and gender, the census tract of the house, and the approval decision.[17] Such geographical information allows this paper to present lending standards that vary across locations and across the distribution of household income and across races.

HMDA data are geographically matched to data on individual property transactions. Data on individual property transactions was obtained from the mortgage company FNC Inc.,[18] which compiles transaction data based using public records and real estate sales. FNC data reports the transaction price and the street address for the complete universe of transactions. Street addresses are geocoded using the transaction's street address and linked to each of our census tracts.[19]

[14] As the borders of geographies such as blockgroups and tracts change over time, we use constant 2000 borders throughout the analysis, i.e. for the three decennial waves of 1990 to 2010.

[15] Specifically, HUD regulates for-profit lenders that have combined assets exceeding $10 million and/or originated 100 or more home purchase loans (including refinancing loans) in the preceding calendar year.

[16] HMDA data contains information on the seniority of mortgage liens only starting in 2004. To eliminate second lien mortgages, also known as piggyback loans, throughout the sample, the dataset includes applications with a minimum LTI of 1.8; such threshold, according to the 2004 HMDA data, eliminates about 96 percent of second lien mortgages, while discarding only 5 percent of first lien mortgages.

[17] HMDA filings also report some information on whether or not loans have been pre-approved but only for some years (post 2004 inclusive) and for a subset of originators covering only 18 percent of applications (2004). The approval rate in that (potentially biased) sample was 70 percent for pre-approved loans vs. 68 percent for loans that have not been pre-approved.

[18] FNC Inc., headquartered in Oxford Mississippi, collects data that provides collateral information to the mortgage industry.

[19] Geocoding was performed using Texas A&M's geocoding services, with a higher than 90% success rate.

Housing and Amenities Data

Households choose dwelling location based on the quality of housing structures, local amenities, and neighborhood demographics.

Neighborhood-level data is derived from Census block-group data of the 1990 Summary Files 1 and 3, the 2000 Census Summary File 1, and the 2010 Summary File 1.[20] There are 4,418 block groups covering the San Francisco Bay area, with a median of 486 housing units per block group, and a median land area of 0.57 square mile. The block group is the smallest geographical unit for which the Census provides housing stock and demographic characteristics across three decades. As the borders of block groups may change across the three decennial census waves, we keep constant 2000 block group borders by building new block group-level census relationship files. Neighborhood characteristics are thus made comparable across the three waves.

The housing and amenities data comprises information on three dimensions that affect households' choice: characteristics of the housing structures (age of structure, median number of rooms, price), school test score data, and neighborhood demographics (race, ethnicity, household median income, fraction with college or associate degree, fraction with high school or more).

California's Department of Education provides Academic Performance Index (API) test score data for each school of the Bay area, for the last two waves, 2000 and 2010 of our neighborhood data. We use school district borders from the Census Bureau to match a block group to each of the 141 elementary or unified school districts, and assign the elementary school API to each block group. APIs are standardized by year to a mean of 0 and a standard deviation of 1.

For the distribution of house values, we choose to use FNC transaction price data rather than Census house values because, unlike the former, the latter are upper-censored and self-reported, with the upper-censoring threshold depending on the census year.[21] We verify that the distribution of FNC transaction prices matches well the distribution of Census house values up to the censoring points. With 126,104 transactions in 2000, a typical blockgroup of the San Francisco Bay Area has about 27.6 transactions in 2000. Thus, each block group has a house value measure that is FNC's

[20] We thank the National Historical Geographic Information System at the University of Minnesota for providing formatted Census files.

[21] Census values comprise the entire universe of houses, while FNC transaction data only measures prices for the subset of houses that transact. However, census values are self-declared and may thus be subject to non-classical measurement error.

median transaction price[22] for each year of observation.

Block group data also includes a measure of the distance to either the San Francisco Central Business District or the San Jose C.B.D. as a proxy for access to retail and employment,[23] and land area in square miles to account for varying land surfaces across block groups.

Household data

This paper's third data source is the 1% Census sample in 1990 and 2000, and the American Community Survey in 2010, which are representative samples of households of the San Francisco Bay area.[24] Household-level data provides an empirical estimate of the San Francisco Bay area distribution $f(\mathbf{x}_{it})$ of household characteristics. It thus allows for the estimation of household-level observable preference heterogeneity $\mathbf{x}_{it}'\Omega\mathbf{z}_{jt}$, as the third set of identifying moments (20) matches the predicted and the observed neighborhood demands according to $\mathbf{x}_{it}$. Individual characteristics include race (White, Black, Hispanic, Asian),[25] and income as measured by total money income of all household members aged 15 or above during the previous year.

The representative sample of households that the 1% Census provides also allows us to predict approval probabilities for all households and all neighborhoods, using the estimated lending standards, regardless of whether they applied for a mortgage in a given year or not (first moment of the model's identifying conditions (20)). In such prediction exercise, household-level data comes from the 1% Census and neighborhood-level data comes from FNC (for transaction prices) and from the Census and the California department of Education (for neighborhood amenities). This typically yields mortgage approval probabilities that are substantially below the observed approval probabilities, as the subset of applicants is a self-selected subset of households.

[22]We also estimated the 25th percentile and the 75th percentile of transaction prices. Although the heterogeneity of housing within a neighborhood could be a potential identification issue, estimation including such moments in base utility regression did not reveal that such heterogeneity plays a substantial role in determining utility.

[23]Census tracts of the two CBDs are set according to the 1982 Census of Retail Trade as in Glaeser, Gottlieb & Tobio (2012). Results can also use the population- or employment-weighted center of each metro area, with no significant difference in results.

[24]We acknowledge support from the Integrated Public Use Microdata Series center at the University of Missouri for providing comparable micro series across three decades.

[25] We use consistent definitions for races, from the 2000 Census and as defined by the Office of Management and Budget's 1997 Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity. Thus a household is either "White, non-Hispanic," "Black, non-Hispanic," "Hispanic, of any race," "Asian, non-Hispanic," or of "Any other race, non-Hispanic."

4.2 Estimation of Approval Probabilities

This section's goal is to estimate the approval regression model (6) using actual mortgage approval decisions. These estimates are then used to predict approval probabilities for each household of the micro household data, across all 4,418 neighborhoods. HMDA (described in the previous section) is, to our knowledge, the only data source that informs approval decisions for the universe of mortgage applications in the metropolitan area. HMDA data report, for each application, the approval decision, the income and the race of the applicant, the size of the loan,[26] the tract-level location of the application, and the financial institution where the application was filed. Other mortgage data sources, such as Dealogic or Black Box, report more information such as the loan-to-value ratio, the credit score, or the transaction price, but only for the universe of originated mortgages loans.

Identification Strategy

A regression of lenders' approval or denial decisions on mortgage loan, household, and housing characteristics will typically not yield causal estimates of the impact of such characteristics on the mortgage lender's decision. A key reason for the lack of a causal interpretation is that *observed* mortgage approval decisions are driven by both households' application behavior and by mortgage lenders' standards. Two endogenous selection mechanisms bias the estimates based on observational data: first, unobservable characteristics of loan and of households drive both lenders' approval decision and households' application decision; such characteristics include, the applicant's creditworthiness, and the house's unobservable quality dimensions. Second, households tend not to choose randomly the financial institutions to which they apply. They could, for instance, select banks endogenously so as to maximize the likelihood of obtaining the type of mortgage loan they need.[27] Therefore unobservable drivers of lenders' approval decisions are correlated with the observable loan, household, and housing characteristics.

We propose here an IV approach in order to isolate the component of mortgage approval rates' changes that is driven by shifts in banks' lending standards rather than by households' application

[26] Post 2004 HMDA data also report the interest rate of the loan.

[27] Households could also file multiple applications for the same house. In practice however the observed number of applications per household is small. According to the estimates of Ouazad & Rancière (2016), the average U.S. household that changes house files between 2 and 3 applications.

decisions. Appropriate instruments provide sources of variation in household and mortgage characteristics that are independent of households' decision to apply for a mortgage. Such identification strategy is formalized in Section 2.5 of the online appendix (equation 12).

We build on on previous work by the authors (Ouazad & Rancière 2016) as well as on Loutskina & Strahan (2009) and Loutskina (2011) to construct an instrument to predict the loan-to-income ratio of mortgage applicants. In these papers, a more favorable liquidity position of the national bank leads to an increase in their capacity to extend credit at the local level. The key innovation in this paper, compared to the papers cited above, is to use, for each loan application, the geographical information of the loan's corresponding neighborhood to match it to the five geographically closest bank branches.[28] Each of those five closest bank branches' mortgage characteristics is strongly correlated with the liquidity ratio of the corresponding national banks. The idea underlying this approach is that liquidity, measured as the ratio of liquid assets over total assets, on the national balance sheet of a bank, is exogenous to households' local decision to apply for a mortgage, but is a significant predictor of the mortgage's characteristics. [29]

Given that households' purchasing in a given neighborhood may select the corresponding branch endogenously, we use the five closest bank branches rather than the bank branch of the loan itself. The bank closest to the property is arguably more likely to be independent of approval unobservables than the bank chosen by the household itself. A key channel through which the liquidity conditions of neighboring bank branches affect the lending standard of a given bank branch is *competition*. Neighboring branches of different banks are competing between themselves to attract borrowers. They do so by balancing the benefits of lowering lending standards, thus increasing market share, and its costs, in terms of increased default risk. When a neighboring branch become more aggressive in cutting lending standards (because the national bank it belongs to is improving its liquidity position), then other branches are induced to also reduce their lending standards to a certain extent to maintain market share. In other terms, through the competition channel, the lending standard of

[28] In order to do so we pair each bank branch with the five geographically closest bank branches. The location of bank branches and the date of their opening is from the Summary of Deposits, which includes the branches of banks regulated by the Federal Reserve System, the Office of the Comptroller of the Currency (OCC) and the Federal Deposit Insurance Corporation (FDIC). Appendix Figure A shows, in 1994, the distribution of bank branches for a specific subarea of the San Francisco Bay Area, the distance between a census tract and bank branches uses the centroids of census tracts and the latitude and longitude of bank branches.

[29] Note that such liquidity mechanism also deals with the potential lending capacity constraints of banks: banks which are facing tighter liquidity constraint are likely to be more sensitive to the need to diversify their lending portfolio, and thus impose lending limits on their local branches.

a given bank is exogenously shifted by a change in liquidity position of neighboring bank branches. This causal chain leads to the first stage of the IV. The liquidity of the five closest bank branches predicts the loan-to-income ratio of each mortgage application, which is this paper's measure of lending standards.[30]

We deal with the endogenous unobserved drivers of the match between applicants and banks in two ways.[31] First, averaging information over the five closest bank branches helps dealing with some unobserved reason for which some applicants are systemically paired with some banks at which they might receive a preferential treatment (e.g. imagine a credit union that has been formed to service the specific banking needs of a sub-segment of the population). Second, in constructing a variant of our instrument, excluding the branches that opened after 1985, i.e. five years before our first observation. In doing so, we wish to control for demand-driven endogenous branching development through which new branches are opened as a response to changes in the pool of potential applicants.

A final endogeneity concern is due to the heterogeneity of mortgage lending standards across households. As households self-select into the pool of applicants for observable and unobservable reasons, the estimated coefficients of the lending standards equation might not be a meaningful average of lending standards coefficients for individual households.[32] In order to compare lending standards between 2000 and 2006, we weigh the likelihood by the distribution of observable applicants to obtain comparable pools. Such balancing is achieved by running a pre-estimation logit regression that regresses the probability that an observation is in the 2006 pool on the set of applicant characteristics. Then the IV probit estimation of the probability of approval on applicant and mortgage characteristics is weighted by the prediction of the pre-estimation logit.

[30]Appendix Figure A pictures bank branches according to their banks' liquidity. Liquidity is derived following Loutskina & Strahan (2009) and Loutskina (2011) as the ratio of securities over total assets obtained from the Federal Reserve of Chicago's Reports of Income andCondition (the Call Reports). Red dots indicate branches whose national bank has low levels of liquidity while blue squares indicate branches whose national bank have high levels of liquidity. The map suggests that bank branches with high or low liquidity levels are not specifically located in high- or low-income neighborhoods, or white, African-American, Hispanic, or Asian neighborhoods.

[31]We use here the matching terminology rather that the endogenous selection of banks by applicants because the selection can also go the other way with some banks specifically targeting some potential clients.

[32]That is, the estimate is not a weighted average of treatment effects with positive weights that sum to 1. A related discussion of policy relevant parameters is presented in Heckman, Urzua & Vytlacil (2006).

Results

In the baseline logit specification,[33] we predict the decision of mortgage approval, for the panel of mortgage applications filed in 1990, 2000, and 2010, based on a number of borrower and loan characteristics, including the loan-to-income ratio (LTI), the race of the applicant, the type of loan (Conventional, FHA insured, VA insured, FSA-RHS insured[34]), time effects, and census-tract fixed effects. Results are presented in Table 1. The first column of the upper part of the table shows results with census-tract fixed effects in the logit specification. We use probit for the instrument variable estimation (bottom part of Table 1). Probit instrumental variable moment conditions follow equation (15) and IV probit estimation is more convenient than IV logit.[35] For each estimation, the marginal probabilities, computed at the means, and standard errors computed using a delta method, are presented next to the logit coefficients. The coefficients of the race dummies (resp., the type of loan) should be interpreted as relative to the dummy for white households (resp., relative to the dummy for conventional loans). Standard errors are clustered at the census tract level.

In the uninstrumented logit specification, the coefficients for the LTI Ratio and for race dummies are significant at one percent. As expected, the LTI ratio coefficient is negative: applying for a larger loan relative to the applicant's income reduces the chances of obtaining approval. A one-standard deviation increase in LTI (+1.28) is associated in a reduction in the probability of approval ranging from 2.7 to 3.7 percentage points, depending on the specification. Conditional on the observable characteristics included, black and hispanic applications face a lower probability of approval.

The second column of the upper panel of Table 1 presents the results of the probit estimation when instrumenting the loan-to-income ratio of the application. The column presents the results where the loan-to-income ratio is instrumented by the average liquidity of the corresponding national banks, for the five geographically-closest branches. The lower panel, first column, presents the results obtained by constructing the same instrument but restricting bank branches that were set up prior to 1985 (first two columns), thus alleviating the potential endogeneity of branching decisions, produces

[33] Using linear instrumental variable regression (with or without fixed effect) yields coefficients that are similar to the marginal probability effects of the probit and logit specification. However linear methods' predictions typically fall outside the $[0, 1]$ interval with non zero probability and cannot be used to estimate choice probabilities.

[34] FHA: Federal Housing Administration; VA: Veteran Administration; FSA-RHS: Farm Service Agency or Rural Housing Service.

[35] The instrumental variable probit model is presented in Newey (1987). The IV probit likelihood has a closed form explicit expression unlike IV logit.

comparable results. The sample is made of observations for which we can obtain liquidity measures for banks and for which the geographic identifier (census tract) can be matched to corresponding census data. The second column of the lower panel presents a robustness test of the IV strategy in which we also instrument the race of the applicant (indicator variables for asian, black, Hispanic applicant) by the racial composition of neighboring tracts (last two columns).

In all IV regressions (i.e for the regressions of the second column of the upper panel, and both columns of the bottom panel of Table 1), the F statistics for each first stage, for both the loan-to-income and race indicator variables, are well above 10. In the corresponding IV estimators in the linear probability model, Cragg Donald finite sample bias F statistics are substantially above the threshold for a maximum 5% finite sample bias; this suggests that finite sample bias and weak instrument issues are unlikely to affect the IV estimates.[36]

The IV results of Table 1 indicate that an increase of the loan-to-income ratio by 1 standard deviation (+1.28) lowers the probability of approval by 15 to 28 percentage points. IV results thus suggest that the magnitude of non-instrumented estimates is substantially downward biased. This is consistent with a case where more creditworthy households tend to choose higher loan-to-income ratios. In such a scenario, unobservable characteristics of the applicant are positively correlated with the loan-to-income ratio and have a positive impact on the approval rate, leading to an upward bias on the coefficient.

Such approval regression estimates (coefficient vector and estimated neighborhood fixed effects) are then used to predict the probability of approval *for all households* in each census block group. In this out-of-sample estimation, the race of the household comes from the Census, and the loan-to-income ratio is computed using FNC housing prices and the total household income of the Census. Predictions are made under an assumption of a loan-to-value of 0.8, that is, the norm for conforming loans used for home purchase (Adelino, Gerardi & Willen 2013).[37] The mean (median) approval rates for all households is 15 percentage points lower than when considering the population of applicants observed in HMDA only. Such difference suggests that many households do not apply for loans because they either expect not to be approved or because they have not been pre-approved. In other terms, the intensity of borrowing constraints appears much more severe when we consider

[36]Statistics available from the authors.

[37]Results are robust to the inclusion or exclusion of loans above the conforming loan limit.

the overall population rather than the sample of loan applicants.

4.3 Estimated Preferences and Willingness to Pay for Amenities

This section presents estimated preferences with and without approval origination constraints. Estimated household preferences are the outcome of the optimization of the moment conditions (Equation 20). Preferences are allowed to vary across neighborhoods (this utility component defines the *base utility* δ), across individuals, and to vary across neighborhoods according to individual characteristics (Equation 1). Base utility depends on the neighborhood characteristics (Equation 2) which include: log median transaction price $log(p_j)$, the neighborhood's demographic composition (race, median income, and residents' education), school test scores (California's Academic Performance Index), and the quality of housing construction (age structure and number of rooms). As discussed in Section 3, the estimation of the relationship between base utility, log price, and neighborhood characteristics requires an IV approach as observable variables are likely to be correlated with unobserved determinants of neighborhood utility.

The instrumental variable strategy is as follows. First, the regression includes a location-specific blockgroup fixed effect, so that the impact of housing characteristics, local amenities, and neighbors' characteristics are estimated in changes across time rather than in cross section. Second, we use time-variation in the observable characteristics of two-step geographically adjacent block groups as instruments for the decennial change in log price.[38] Such instrumental variable strategy, in the spirit of Berry et al. (1995), is based on the assumption that neighboring blocks' changes in observable characteristics are less confounded by changes in unobservables than the block's own characteristics. Neighbors also sort endogenously in unobservable and observable dimensions across neighborhoods. Finally, we also instrument the demographic composition (% black, % Hispanic, % Asian, log(median household income), % college educated), by the demographic composition of two-step geographically adjacent blockgroups as in Bayer & Timmins (2007).

Selection of neighborhood characteristics of interest for inclusion in the utility regressions (1) and (2) builds on existing literature. The racial composition of neighborhoods and its interaction with household race allows to test for heterogenous racial preferences (Bayer, McMillan & Rueben 2004,

[38] Two-step adjacent block groups are adjacent to neighboring blockgroups. We thus exclude the characteristics of the immediately adjacent blockgroups in the calculation of the instrument.

Bayer et al. 2007, Becker & Murphy 2009). The emphasis on the age of dwellings and its interaction with income is stressed in filtering theory (Rosenthal 2014). The inclusion of the number of rooms, as a proxy for the size of the dwellings, and its interaction with the distance to the metropolitan area's Central Business District, relates to the monocentric model, in which households sort at a particular distance from the CBD according to the relative importance of their taste for either larger houses, or for closer proximity with the city center and shorter commutes (Fujita 1989). The preference for more educated neighborhoods, in addition to more affluent neighbors, is linked to the literature on parents' involvement in school and education neighborhood peer effects (Durlauf 2004, Goux & Maurin 2007). Several of the neighborhood characteristics are interacted with individual characteristics such as race and income, thus allowing for a multidimensional sorting.

Tables 2 and 3 present estimation results for social preferences (first panel, Table 2), neighborhood income, education level and school quality (top panel, Table 3), and price-sensitivity (bottom panel, Table 3). For the sake of clarity, the variables capturing housing characteristics, e.g. the structure's median age, the number of rooms or the distance to the Central Business District are included in the estimation but the coefficients are not reported.[39]

In each panel, the first two columns present the estimated coefficient and the willingness to pay according to the model with borrowing constraints. The next two columns present the corresponding estimates for the model without borrowing constraints. In that latter model, the probability that any neighborhood j belongs to the choice set of any household i in year t is set equal to one. In both models, the willingness to pay is the change in price that offsets a given change in the quality of amenities. Formally for a given amenity k change Δz_{jtk}, the willingness to pay is the log price change $\Delta \log(p_{jt})$ that maintains utility constant. Such log price change satisfies $-\alpha_i \Delta \log(p_{jt}) = \beta_{ik} \cdot \Delta z_{jtk}$, where β_i is the individual-specific preference for amenity z_{jtk}, and α_i is the individual-specific log price coefficient. This yields an individual-specific willingness-to-pay $\Delta \log(p_{jt})_i = -\left(\beta_{ik}/\alpha_i\right) \cdot \Delta z_{jtk}$. For each dimension, we report the median willingness to pay $Median_i \left[\frac{\beta_{ik}}{\alpha_i} \cdot IQR(z_k)\right]$ corresponding to an inter-quartile increase $IQR(z_k)$ in the amenity z_{jtk}.

Estimates of the preferences for neighborhoods' racial composition reveal, in both models, a similarly strong preference for same-race neighbors. Such same-race preference can also partly

[39] Unsurprisingly these coefficients indicate that households have a preference for larger and newer homes, and trade-off their preference larger home against the proximity to the central business district .

reveal preferences for unobservable individual characteristics strongly correlated with neighbors' race (Bayer et al. 2004). Regarding other social preferences, the two models differ significantly in only one dimension, that is the preference of non-Hispanics to live in a neighborhood with a higher fraction of Hispanics. The willingness to pay of Non-Hispanics to move into a neighborhood with an interquartile range increase in the fraction of Hispanics is 25 percent lower in the model with than in the model without borrowing constraints.

The two models do not significantly differ with respect to the preference for neighborhoods with higher median income and a higher fraction of college educated individuals. However, the model estimated with borrowing constraints exhibits a significantly larger willingness to pay for neighborhoods with better performing public schools. The willingness to pay for an interquartile increase in the Academic Performance Index is $22,877 for the average household in the model with borrowing constraints, and $8,552 in the model without borrowing constraints. The existence of approval constraints mitigates how preferences gets reflected into demand differences. Accounting for approval constraints can thus reveal sharper underlying differences in school quality preferences.

The bottom panel of Table 3 presents the price-sensitivity of the base utility in both models. We allow for the price-sensitivity of base utility to be heterogeneous in two dimensions. First, we allow for a price-sensitivity coefficient to vary according to a random coefficient specification, as in specification 2. Second, we allow for the price-sensitivity to vary according to the level of household's income. The estimation results reject the first type of heterogeneity by estimating a non-statistically significant variance of the coefficient distribution, but does not reject the second type of heterogeneity: there is negative and strongly significant interaction term between log(price) and log(income) of similar magnitude in both models. Since log(price) is interacted with standardized log income, the coefficient on log prices captures the price-sensitivity of a median income household. This coefficient is about three times lower in the model with than in the model without borrowing constraints (-0.089 vs -0.299) and the difference is highly significant. Such large difference means that accounting for sensitivity of the approval rate to a price change leads to a dramatic downward revision of the sensitivity of the base utility to price changes. This result suggests that borrowing constraints are a key driver of price elasticity as we shall see next.

4.4 Price Elasticity

The model with borrowing constraint allows separate estimation of two additive components of the price elasticity of demand. By price elasticity of demand, we refer here to the elasticity of the demand for a neighborhood with respect to its *own price* but we shall emphasize that the model provides estimates for the complete set of cross-price and income elasticities.[40] The first component, the conditional elasticity of demand, is computed at given approval-based choice sets, i.e. assuming that approval probabilities are not affected by a change in house prices. Absent heterogeneity, the conditional elasticity is simply computed as one minus the market share of a given neighborhood times the price-sensitivity of the base utility, and given the large number of neighborhoods (4,416), can be well approximated by the log price coefficient. As before, conditional price elasticity has a non-degenerate distribution made possible by the impact of the interaction of household income with the log price. The second component, the borrowing elasticity, captures the fall in demand associated with a fall in approval rates following an increase in housing prices and is heterogeneous in all the variables relevant for mortgage approval decision. The sum of the two components gives the total (unconditional) price elasticity. Figure 1 illustrates the elasticity decomposition with the distribution of elasticities across neighborhood (top panels of the figure) and their sensitivity to the initial log price (bottom panels of the same figure). The top panel suggests that borrowing elasticity is an important driver of total elasticity. Indeed the ratio of borrowing elasticity to total elasticity has a mean of 0.69, and a median of 0.55. The bottom panel show that both conditional elasticity and borrowing elasticity are negatively correlated with log price. The negative correlation between borrowing elasticity and log price is however much lower than the correlation between conditional demand elasticity and log price (-0.54 vs. -0.24). This difference shall play a role in the general equilibrium analysis that comes next.

4.5 Robustness

The Online Appendix's Table E (Section 6) presents estimation results when neighborhood utility depends on log non-housing consumption rather than log price, that is log of income minus mortgage payments. Table G in the online appendix presents estimates of the relationship between the base

[40]The full matrix of total demand derivatives is used in the general equilibrium analysis of Section 5.

utility of rental and amenities. In this specification the log rent replaces the log price, and households have an option to rent in each neighborhood (Online Appendix Section 6.1). In both case signs of the impact of amenities on base utilities are the same as the signs of the base utility regressions of the baseline model but with some variations in the magnitude of the coefficients. As expected, non-housing consumption has a positive impact on utility, the log rent a negative one, and the demand for rental is substantially more elastic than the demand for homeownership.

5 General Equilibrium Impact of Lending Standards

In this section we use the model, estimated with borrowing constraints, to run a general equilibrium analysis of the impact of a change in lending standards on the distribution of prices across neighborhoods, and on households' spatial segregation.

The model predicts that relaxing lending standards in the credit market typically leads to an increase in housing demand, at initial house prices, for most neighborhoods. This generally leads to excess demand that outstrips housing supply. The general equilibrium price response in each neighborhood reduces excess demand to the point where each neighborhood's demand equals each neighborhood's housing supply. Even when total neighborhood housing demand remains unchanged, which must be the case at equilibrium when housing supply is fixed, the composition of neighborhood demand changes by race, ethnicity, and income. Thus the model predicts substantial reallocation of households across space and changes in segregation.

We use here the estimated model here to run a general equilibrium analysis of a change in lending standards. In Section 7 of the online appendix, we show and give examples on how the model could be used, in partial equilibrium, to study the effects of policy programs targeting a subpopulation small enough to avoid the price impacts of lending supply policies.

5.1 Comparative Statics: Price Impacts of Lending Standards

This section derives the impact of lending standards $(\boldsymbol{\psi}, \boldsymbol{\gamma}, \Psi)$ on neighborhood log prices $\log(\mathbf{p})$. We start by focusing on one specific coefficient, say, the coefficient of the log loan-to-income ratio in the approval specification, noted as a scalar ψ. A change in the coefficient ψ of the approval specification causes three different demand effects: first, a change in neighborhood demand at given prices. Second, the change in lending standards causes a change in each neighborhood's price (*the general equilibrium change in prices)* that in turn causes a change in demand. Third, the change in lending standards causes a change in neighborhoods' racial composition (*the general equilibrium change in neighborhood demographics)*, which affects households' preferences for these specific neighborhoods. Such three effects formally translate into:

$$\frac{d\mathbf{D}}{d\psi} = \underbrace{\frac{\partial \mathbf{D}}{\partial \psi}}_{Direct\, Impact} + \underbrace{\frac{\partial \mathbf{D}}{\partial \log(\mathbf{p})} \cdot \frac{d \log(\mathbf{p})}{d\psi}}_{Impact\, due\, to\, price\, changes} + \underbrace{\frac{\partial \mathbf{D}}{\partial \mathbf{v}} \cdot \frac{\partial \mathbf{v}}{\partial \psi}}_{Impact\, due\, to\, social\, preferences} \tag{21}$$

For the sake of clarity, we consider here the case of perfectly inelastic housing supply ($\frac{\partial \mathbf{S}}{\partial \log(\mathbf{p})} = 0$), and extend the analysis to the non-perfectly inelastic housing supply case in Section 6.2.[41] The price change that maintains demand equal to supply in each neighborhood is such that:

$$\frac{d \log(\mathbf{p})}{d\psi} = \left[\frac{\partial}{\partial \log(\mathbf{p})}\mathbf{D}(\log(\mathbf{p}); \psi) - \frac{\partial \mathbf{S}}{\partial \log(\mathbf{p})}\right]^{-1} \left[-\frac{\partial}{\partial \psi}\mathbf{D}(\log(\mathbf{p}); \psi) - \frac{\partial \mathbf{D}}{\partial \mathbf{v}} \cdot \frac{\partial \mathbf{v}}{\partial \psi}\right] \tag{22}$$

a vector of size $\sum_t J_t$, i.e. one vector of log price changes per year of analysis. The vector of prices $\log(\mathbf{p})$ affects demand through its impact on the base utility vector $\boldsymbol{\delta}$ and the interaction terms Ω and $\tilde{\boldsymbol{\beta}}_{it}$.

(i) Impact of Lending Standards on Demand, at Given Prices

We start with the impact of lending standards on demand at given prices. Lending standards affect both the probability of each choice set, and the choice of the household within the choice set. We start by focusing on the impact of lending standards on the probability of the choice set.

$$\frac{\partial}{\partial \psi} D(j, t|\boldsymbol{\delta}, \mathbf{z}_{\cdot t}; \psi) = \sum_{C \in \mathbb{C}} \int_X \left[\frac{\partial}{\partial \psi} P(C|\mathbf{z}_{\cdot t}, \mathbf{x}_{it}; \psi)\right] \cdot P(j, t|\boldsymbol{\delta}, \mathbf{z}_{\cdot t}, \mathbf{x}_{it}, C) \cdot f(\mathbf{x}_{it}) d\mathbf{x}_{it} \tag{23}$$

Empirically the sum $\sum_{C \in \mathbb{C}}$ is taken over the set of simulated choice sets. The derivative of the probability of the choice set $P(C|\mathbf{z}_{\cdot t}, \mathbf{x}_{it}; \psi)$ is the derivative of a product of J terms $\Pi_{j \in C} \phi_{jt} \Pi_{k \notin C}(1-\phi_{kt})$ w.r.t. lending standards. Each individual probability ϕ_{jt} has a simple binomial derivative, and thus

[41] The evidence presented in Section 6.2 suggests indeed that housing supply is fairly inelastic in most neighborhoods of San Francisco. Data from Saiz (2010) suggests that the metropolitan-area wide housing supply elasticity is 0.66 for the entire San Francisco metropolitan area and 0.76 for the entire San Jose metropolitan area.

the impact of lending standards on choice set probabilities:

$$\frac{\partial P(C|\mathbf{z}_{\cdot t}, \mathbf{x}_{it}, \psi)}{\partial \psi} = P(C|\mathbf{z}_{\cdot t}, \mathbf{x}_{it}, \psi) \cdot \underbrace{\sum_{j} w_{ijt} \cdot \{\mathbf{1}(j \in C) - \phi_{jt}(\mathbf{z}_{jt}, \mathbf{x}_{it}; \psi)\}}_{\Upsilon_{ijt}(C)} \tag{24}$$

where w_{ijt} is the value corresponding to the coefficient ψ. The impact of lending standards on the probability of approval in any one neighborhood (a single term in the sum Σ_j) is of the sign of w_{ijt}. For instance, if w_{ijt} corresponds to the loan-to-income ratio, w_{ijt} is positive, the coefficient ψ of the loan-to-income ratio in the approval specification is negative. An increase in ψ will lead to an increase in the probability ϕ_{ijt} for all households i and all neighborhoods j of the city. And the probability $P(C)$ of all choice sets C increases. The Monte Carlo simulated estimate of (23) is obtained by averaging (23) over simulated choice sets.

(ii) Impact of Prices on Demand, at Given Lending Standards

We then turn to the impact of log prices on demand. The matrix of demand derivatives $\frac{\partial \mathbf{D}}{\partial \log(\mathbf{p})}$ is considered for year t. $\log(\mathbf{p})$ is a line vector of size J_t while $\mathbf{D}(\mathbf{p}; \psi)$ is a column vector of size J_t. We distinguish diagonal elements, which are derivatives of demand w.r.t.[42] its own price (proportional to own-price demand elasticity), and off-diagonal elements, which are derivatives of demand w.r.t. another neighborhood's price (proportional to cross-price demand elasticity).

The derivation of the cross-price demand elasticity yields a decomposition between borrowing elasticity and conditional price elasticity, which is the analog to the one presented in 2 for the own-price demand elasticity:

$$\begin{aligned} \frac{\partial D(j,t|\boldsymbol{\delta}_{\cdot t}, \mathbf{z}_{\cdot t}; \psi)}{\partial \log(p_k)} &= \underbrace{\sum_{C \in \mathbb{C}} \int_X \frac{\partial}{\partial \log(p_k)} P(C|\mathbf{z}_{\cdot t}, \mathbf{x}_{it}; \psi) \cdot P(j,t|\boldsymbol{\delta}_{\cdot t}, \mathbf{z}_{\cdot t}, \mathbf{x}_{it}, C) f(\mathbf{x}_{it}) d\mathbf{x}_{it}}_{\text{Shift in choice set}} \\ &+ \underbrace{\sum_{C \in \mathbb{C}} \int_X P(C|\mathbf{z}_{\cdot t}, \mathbf{x}_{it}; \psi) \cdot \frac{\partial}{\partial \log(p_k)} P(j,t|\boldsymbol{\delta}_{\cdot t}, \mathbf{z}_{\cdot t}, \mathbf{x}_{it}, C) f(\mathbf{x}_{it}) d\mathbf{x}_{it}}_{\text{Shift in utility}} \end{aligned} \tag{25}$$

Appendix Section 3 provides closed-form derivations for both demand derivative terms.

[42] We note "with respect to" as w.r.t. in the remaining parts of this paper.

(iii) Impact of Lending Standards on Social Preferences

We derive the impact of lending standards on neighborhood composition. Neighborhood composition affects neighborhood demand as long as households have preferences for particular neighbor demographics (e.g. education, race, ethnicity). Getting such social interaction effects is the last step in completing the calculation of general equilibrium effects on prices. Formally, such social interaction effect is a first-order impact on households' demand $-\frac{\partial \mathbf{D}}{\partial \mathbf{v}} \cdot \frac{\partial \mathbf{v}}{\partial \psi}$. The second factor, the impact of lending standards on neighborhood demographics, $\frac{\partial \mathbf{v}}{\partial \psi}$ is simply the vector of demands scaled by the total demand for the neighborhood.

$$\frac{\partial \mathbf{v}}{\partial \psi} = \left(\frac{n^k}{n} \cdot \left(\frac{\partial \mathbf{D}^k}{\partial \psi} / \mathbf{D} \right) - \frac{\mathbf{D}^k}{\mathbf{D}} \cdot \left(\frac{\partial \mathbf{D}}{\partial \psi} / \mathbf{D} \right) \right)_{k=1,2,\ldots,M} \tag{26}$$

where n^k/n is the fraction of population k in the overall population of the metropolitan area. The first factor is the impact of neighborhood demographics on demand. $\frac{\partial \mathbf{D}}{\partial \mathbf{v}}$ is also straightforwardly derived. For a group k, say Hispanics, $\gamma^k(\mathbf{x}_{it})$ is the coefficient of social interactions for an individual with characteristics $\mathbf{x}_{it}$. $\gamma^k(\mathbf{x}_{it})$ depends on $\mathbf{x}_{it}$ as Asians, Blacks, Hispanics, Whites do not necessarily exhibit similar preferences. Then:

$$\begin{aligned} \frac{\partial}{\partial \mathbf{v}^k} D(j,t|\boldsymbol{\delta}_{\cdot t}, \mathbf{z}_{\cdot t}) \quad &= \quad \sum_{C \in \mathbb{C}} \int_X P(C|\mathbf{z}_{\cdot t}, \mathbf{x}_{it}) \\ &\cdot \gamma^k(\mathbf{x}_{it}) P(j,t|\boldsymbol{\delta}_{\cdot t}, \mathbf{z}_{\cdot t}, \mathbf{x}_{it}, C) \cdot (1 - P(j,t|\boldsymbol{\delta}_{\cdot t}, \mathbf{z}_{\cdot t}, \mathbf{x}_{it}, C)) f(\mathbf{x}_{it}) d\mathbf{x}_{it} \end{aligned} \tag{27}$$

$\gamma^k(\mathbf{x}_{it})$ varies across individuals: for instance, results suggest households' preferences for same-race neighbors.

(iv) Empirical Results:

General Equilibrium Impacts of Lending Standards on Prices

The general equilibrium impact of lending standards on prices is the shift in prices that keeps demand for each neighborhood constant. We combine the analytic formulas of (i)–(iii) above to derive the following: first, the impact of lending standards on demand at constant prices; second, the impact of prices on demand; and third, the impact of lending standards on neighborhood racial and income composition. All three effects lead to the general equilibrium impact of lending standards on prices, following equation (22) above.

The derivations presented in the previous subsections (i)–(iii) were done by looking at the marginal change of one arbitrary coefficient of the approval specification. Table C estimates the change in lending standards between 2000 and 2006. The regression keeps the distribution of applicant characteristics constant in 2000 and 2006, by weighting 2006 observations by the share of 2000 observations with identical characteristics. The results suggests that there were significant changes in underwriters' sensitivity to the log(price) of the house, the log(income) of the applicant, and to the race of the applicant. The regression is a panel logit regression combining data from the Home Mortgage Disclosure Act of 2000 and 2006. The specification is similar to the main baseline analysis of approval rates (Table 1), with added interactions between the 2006 year dummy and respectively: the log(price), the log(income) of the applicant, and the applicant's race. Results suggest a relaxation of lending standards in each of those dimensions. We combine the comparative statics in each of these dimensions to simulate a marginal change in lending standards comparable to the 2000–2006 relaxation.[43]

Figure 2 illustrates the general equilibrium impact of the simulated change in lending standards on house prices changes as resulting from the three component of housing demand changes according to equation (21): partial equilibrium effects, price effects, and social preferences effects.

Figure 2 (c) plots the implied partial equilibrium effect, where the increase in demand (vertical axis) at constant prices is plotted against the initial log(price) in 2000. While a relaxation of lending standard leads to an increase in housing demand across all neighborhoods, such demand shifts displays almost no correlation with initial log prices. Partial equilibrium demand effects cannot explain the *negative correlation between price changes and initial log prices* that arise in our general equilibrium simulations (Figure 2 (a)) as well in the data (Figure 3). The housing demand shifts resulting from social preferences (Figure 2 (d)) display a negative correlation with initial log prices, and thus contributes to explain the greater increase in prices in neighborhoods

[43]This section estimates the impact of a change in lending standards $d\boldsymbol{\psi}$ on prices. The theoretical derivation is done for a *small* change in lending standards. The impact of a *large* shift in lending standards such as the one observed between 2000 and 2006 is then estimated by extrapolating the impact of the small shift around the initial equilibrium. This method does not account for the shift in lending standards *driven by the equilibrium change*, and in particular driven by equilibrium price and demographic changes. Section 5 in the online appendix shows the robustness of the results to a methodology estimating of the impact of a large change in lending standards $\Delta\boldsymbol{\psi} = \boldsymbol{\psi}_{2006} - \boldsymbol{\psi}_{2000}$ and the equilibrium shifts for each small shift along the path from the 2006 lending standards coefficients $\boldsymbol{\psi}_{2000}$ to $\boldsymbol{\psi}_{2006}$. Section 4 in the online appendix deals with the robustness of our comparative statics results to a possible shift in unobservable fixed effects between 2000 and 2006.

with lower prices. The remaining change in the distribution of price changes result from the impact of prices on neighborhood demand associated with shifts in choice sets and shifts in utility. In that dimension, the negative correlation between price changes and initial log price is consistent with negative correlation between (own-) price elasticity and initial log prices (Figure 1f), which is mostly driven by the negative correlation between borrowing elasticity and initial log prices.[44]

House Price Changes: Model vs. Data

Next, we compare the prediction of the model on log(price) changes, computed according to equation 22, to actual price changes during the same the period computed from property transaction data. While the *change* in lending standards is taken as realized in the data, the estimation of households' preferences for neighborhoods, and estimation of initial lending standards did not use annual information for the boom years of 2001-2006. The last observation in our decennial panel is 2010; in 2010 the Case-Shiller price index for the San Francisco MSA was almost back to its 2000 level (136.99 vs. 130.06).

Figure 3 plots the actual log price change in the model (red dots) and in the data (black dot) as a function of the initial log price in 2000. The bottom panel of Figure 3 summarizes the key moments of the distribution of price changes in the model and in the data. In Figure 2 we use the average annual log(price) change over 2000-2006. Figure 3 shows that the negative correlation between price changes and initial log prices predicted by the model is also a feature of the data. The model predicts a correlation of price changes with initial log(price) of -0.286, which is slightly less than 40% of the correlation observed in the data-0.748. The model's predictions matches rather closely other moments of the actual distribution of price changes: the mean (0.101 in the model vs 0.111 in the data), the standard deviation (0.045 vs 0.055), the median (0.087 vs. 0.110), the lower quantile (0.076 vs. 0.072), and the upper quantile (0.131 vs 0.149)

The predicted compression of the price distribution has also been observed in San Diego by Landvoigt, Piazzesi & Schneider (2015). Appendix Figure C uses additional transaction-level data from Los Angeles and Washington DC to document such compression of the price distribution in two other major metro areas. Each figure depicts the difference between the upper and the lower quartile of the log(price) distribution and suggests that such difference has decreased over 30% in

[44]Recall however that price effects depend not only on own price elasticities but also on cross-price elasticities.

Los Angeles and San Francisco, 17% in San Jose, and 18% in Washington DC between 2000 and 2006.

5.2 Impacts of Lending Standards on Spatial Segregation

The general equilibrium price change maintains demand in each neighborhood constant. However, the demand of particular racial or other demographic subgroups for each neighborhood is typically affected by changes in lending standards.

Consider, for example, the demand D_{rjt} of racial group $r \in \{asian, black, hispanic, white, other\}$ for a specific neighborhood j in year t. The total change in demand caused by a change in lending standards is decomposed into a partial equilibrium effect, an effect due to social preferences, and a change due to shifts in the distribution of prices:

$$\frac{dD_{rjt}}{d\psi} = \underbrace{\frac{\partial D_{rjt}}{\partial \psi}}_{partial\,equilibrium} + \underbrace{\frac{\partial D_{rjt}}{\partial \log(\mathbf{p})} \cdot \frac{d\log(\mathbf{p})}{d\psi}}_{general\,equilibrium} + \underbrace{\frac{\partial D_{rjt}}{\partial \mathbf{v}} \cdot \frac{\partial \mathbf{v}}{\partial \psi}}_{preferences\,for\,same\,race\,neighbor} \tag{28}$$

where the price change $d\log(\mathbf{p})/d\psi$ was obtained in the earlier general equilibrium analysis. In general $dD_{rjt}/d\psi \neq 0$, while $dD_{jt}/d\psi = 0$. Given racial groups' changing demand, the model gives a structural estimate of how changes in lending standards affect racial segregation across neighborhoods within the metro area.

As measure of Bay Area-wide racial segregation we use the exposure indices, as in Massey & Denton (1988) and Cutler, Glaeser & Vigdor (1999). The exposure index of a given racial subgroup (say, Blacks) to another (say, Whites) measures the average fraction of white neighbors for an average black resident. Thus:

$$Exposure(Whites|Blacks)_t = \sum_{j=1}^{J} \frac{Black_{jt}}{Black} \cdot \frac{White_{jt}}{Population_{jt}} \tag{29}$$

where the first factor $Black_{jt}/Black$ is the share of black population living in neighborhood j in year t; and the second factor $White_{jt}/Population_{jt}$ is the fraction white in neighborhood j in year t. The isolation of Blacks is the exposure of Blacks to their own racial group. Therefore an increase in black isolation is due to a decline in black exposure to other races.

Definition (29) can be simply tied to the structural model: black population $Black_{jt}$ is simply the total black population count in the metro area multiplied by black neighborhood demand for that area, i.e. $Black_{jt} = Black \cdot D_{black,jt}$, as $D_{black,jt} \in (0,1)$. Similarly, $White_{jt} = White \cdot D_{white,jt}$.

Thus the model provides a structural prediction of the impact of lending standards on racial isolation and exposure. For instance, deriving black exposure to Whites with respect to the lending standard parameter ψ:

$$\frac{d}{d\psi} Exposure(Whites|Blacks)_t = \frac{White}{Population} \cdot \frac{d}{d\psi} \left[\sum_{j=1}^{J} \frac{D_{black,jt} \cdot D_{white,jt}}{D_{jt}} \right] \tag{30}$$

The change in exposure is driven by both white and black demand changes.

The results of the analysis of the impact of lending standards on neighborhoods' racial composition and on racial segregation are presented in Figure 4, Figure 5, and in Table 4a. The starting point of such comparative statics is the spatial distribution of racial groups in the Bay area in 2000. In both Figure 4 and Figure 5, black dots are demand changes *at given prices*, as a percentage of initial demand. The red dots are the neighborhoods' predicted change in white demand *accounting for the general equilibrium* change in prices. In Table 4a the diagonal of each table is the impact on the isolation of Whites, Blacks, Hispanics, Asians (the exposure of a racial group to itself), while off-diagonal elements present the impact on the exposure of a racial group to another. All estimates are in percentage points, as isolation and exposure measures are expressed in percentages.

Figure 4 presents, for each racial group, partial and general equilibrium housing demand changes (top panels) and price changes (bottom panels) as a function of the initial fraction of the same racial group in each neighborhood.

Since we have estimated for each group a strong preference for having more same-race neighbors, we are expecting for their partial equilibrium demand to be increasing in the fraction of same-race population. This is actually the case for all the four racial groups. They are however substantial differences across racial groups in how partial equilibrium demand changes translate into general equilibrium demand changes.

For Whites, general equilibrium demand changes for mostly whites neighborhoods tend to be *smaller* than the corresponding partial equilibrium demand changes. The reason is that mostly white neighborhoods are generally more expensive neighborhoods which, as we have seen in section 4.4 (Figure 1f), display a larger price elasticity of demand. Therefore a fraction of white households shift their demand towards more mixed neighborhoods, which are initially less expensive, thus contributing to the price increase in such neighborhoods. Indeed the correlation between neighborhood price changes and the initial fraction of Whites is negative, and similar in magnitude to the negative correlation between initial prices and price changes (Figure 4). A consequence of such general equilibrium price effect pushing a fraction of Whites towards more mixed neighborhoods is a (limited) *reduction in the isolation of Whites* as shown in Table 4a (−0.63 percentage points).

By contrast, general equilibrium demand changes of black, African-American households are generally *larger* than their partial equilibrium counterparts in neighborhoods with 40 percent of more of black households in 2000. This reflects the fact that neighborhoods with a large fraction of Blacks tend to be initially less expensive neighborhoods and display a lower price elasticity. As a consequence the general equilibrium demand by Blacks for such neighborhoods has increased far beyond their partial equilibrium demand, pushing prices up, and contributing to an increase in the isolation of blacks (+2.84 percentage points) as shown in Table 4a. General equilibrium demand changes for Asian and Hispanics are more or less in line with partial equilibrium demand changes. For those two groups, the increase in isolation reflect therefore their preference to live in neighborhoods with a larger fraction of same-race neighbors, Table 4a indicates an increase in the isolation of Hispanics (+1.04 percentage points) and in the isolation of Asians (+1.15 percentage points).

Figure 5 enables a better understanding of the change white exposure to other groups. The top panels shows the demand of Whites for neighborhoods ranked according to their fraction asian, Hispanic and black. A reduction of Whites' demand for neighborhoods with a low fraction of Asians and Hispanics coexists with an increase in Whites' demand for neighborhoods with a moderate fraction of Hispanics and Asians. There is also a increase in the demand of Whites for neighborhoods with a moderate share of Blacks coexisting with a reduction in the demand of Whites for neighborhoods with a large fraction of blacks. As a consequence the exposure of Whites to Hispanics (+0.29 percentage points), to Asians (+0.2 percentage points) and to Blacks (+0.15 percentage points) exhibit a moderate increase. These changes in Whites demand appear related to the gentrification observed in some parts of the San Francisco Bay during the credit boom.

Segregation: Model vs. Data

How do these model results compare with actual changes in segregation (Table 4b)? Note that such comparison is not straightforward as changes in the racial demographics of the metropolitan area also affect measures of exposure and isolation.[45] The model reproduces however qualitatively well the reduction in the isolation of Whites, a finding in line with evidence of gentrification in the San Francisco Bay, as well as the increase in the isolation of Hispanics and Asians.

Overall the ability of the estimated model to simultaneously predict, out-of-sample, observed changes in the house price distribution and observed changes in the spatial distribution of households provides considerable support to the view that general equilibrium effects of changes in credit standards play an important role in the medium-run dynamics of cities.

6 Extensions

We introduce here extensions to take into account the option to rent, the role of local housing supply elasticity, and of demographic changes (population growth and changes in metro-level racial composition).

6.1 The Role of Rental

The relaxation of lending standards may affect both neighborhood choice and tenure choice. Changes in lending standards, by affecting the probabilities to obtain credit, are also likely to change the relative demand for ownership vs. rental housing both within and across neighborhoods.

Tenure choice features in the model as follows. In each neighborhood j, household i can choose either homeownership $s = ownership$ or rental, $s = rental$. The utility of neighborhood j with tenure s is noted U_{ijst}, as in equation (1), and the corresponding base utility is noted δ_{jst}. The difference $U_{ij,ownership,t} - U_{ij,rental,t}$ is partly driven by unobservable quality differences between the

[45] A way to see the importance of city-wide demographic changes is to look at the relationship between the change in the exposures of Whites to Blacks to that of the exposure of Blacks to Whites:

$$\Delta \log E(Whites|Blacks) = \Delta \log(\frac{w}{b}) + \Delta \log E(Blacks|Whites) \quad (31)$$

Where w/b is ratio of the Whites-to-Blacks population in the metropolitan area. In absence of any demographic change in the racial composition of the metropolitan area ($\Delta \log(\frac{w}{b}) = 0$), the log change in both measures are identical. When they are demographic changes however, the evolution of the two exposure measures can diverge as it is the case for several exposure measures on Table 4, which reports change in actual exposure measures between 2000 and 2006.

rental housing stock and the owner-occupied housing stock; and driven by households' preference for homeownership vs. rental. Such model with neighborhood and tenure choice reduces to the paper's baseline model of Section 2 whenever (i) a uniform arbitrage relationship ties the price of owner-occupied units and rental values, such as the the price of owner-occupied units equal to the discounted value of rental payments with a constant discount factor; and (ii) households are indifferent between homeownership and rental.

The model with neighborhood choice, tenure choice, and borrowing constraints relaxes those two assumptions. Rents and the number of renter-occupied housing units are observed in neighborhood-level Census files. Also, by estimating a different base utility for rental and homeownership in the same neighborhood (backed out by contraction mapping on the population of households that rent and own), the model allows for a neighborhood- and time-specific value of homeownership. Estimates of the relationship between the base utility of rental and amenities have been discussed in Section 4.5. General equilibrium results, presented in Online Appendix Figure F, suggest no economically significant difference between the two models in the predicted magnitude of the compression of the price distribution. Both models predict a similar (and not statistically different) correlation between log price changes and the initial log price in 2000.

6.2 The Role of Housing Supply Elasticity and of Demographic Shifts

Housing Supply

The general equilibrium analysis of Section 5 assumed that the supply of housing units was constant during the period. We relax here this assumption and check whether the compression of the price distribution, observed between 2000 and 2006, could be due to the heterogeneity in housing supply elasticities across neighborhoods rather than to the differential impact of lending standards changes.

We build neighborhood- (i.e. blockgroup-) level estimates of housing supply elasticities by combining satellite data on land cover and elevation in order to measure the share of each blockgroup that is not developed in 1992 and can likely be developed given the geographic features of the land. Then, in a similar fashion as in Saiz (2010), but in our case at the blockgroup level, we perform the following regression of log housing units on log price:

$$\log(housing\,units_{j,t}) = a + (b + c \cdot Undeveloped\,Share_j + d \cdot Ruggedness_j) \cdot \log(price_{j,t}) + Blockgroup_j + Residual_{j,t} \quad (32)$$

where j is a neighborhood index, t is either 2000 or 2010; $Undeveloped\,Share_j$ is the undeveloped share of the blockgroup's surface in 1992, 8 years prior to the first cross section used in the general equilibrium analysis. The measure of undeveloped share of land corresponds to the share of the surface of the blockgroup (in squared meters) that is not developed and that is not covered by water. The measure is derived from satellite data of the United States Geological Survey (U.S.G.S.) database. The landcover data set of 1992 measures, in each 30m×30m cell, whether the cell is developed (low, medium, or high intensity) and the nature of the land (forest, water, barren, rock, grass, wetland, crops and pasture). Appendix Figure Ba illustrates the construction of our underdeveloped share measure by mapping the underlying satellite landcover data in the city of San Jose.

Then we use a set of set of elevation measures ($Ruggedness_j$) based on the slope and the ruggedness of the blockgroup to proxy for the cost of building on undeveloped cells.[46] Such elevation measures come from a second satellite data set, the U.S.G.S.'s digital elevation model (D.E.M.), which measures elevation for each 30m×30m cell. Such blockgroup-level ruggedness is the average of ruggedness across cells, as illustrated in Appendix Figure Bb.

$Blockgroup_j$ is a blockgroup fixed effect. Neighborhood-level housing supply elasticity is proxied by the coefficient $\eta_j = b + c \cdot Undeveloped\,Share_j + d \cdot Ruggedness_j$ in specification 32 above. Appendix Figure Bc presents the distribution of estimated supply elasticity across neighborhoods. About half of the neighborhoods exhibit a supply elasticity inferior to 0.1, and about 82% percent an elasticity below 0.2. As expected neighborhood housing supply elasticity is positively correlated with the distance to the Central Business District.[47] These elasticities are typically lower than those estimated at the metro-area level (Saiz 2010). Section 8 of the online appendix explores this issue

[46] The ruggedness index measures terrain irregularity and was initially developed by Riley, DeGloria & Elliot (1999) and used in economics by Nunn & Puga (2012). The ruggedness at a given cell (30mx30m) is the square root of the sum of the squared differences of the elevation of the cell with its eight adjacent cells.

[47] Results available from the authors.

further, building up on Imbs & Mejean (2015) who shows that, when micro-level elasticities have both an observed and an unobserved component, the average of micro-elasticities does not typically match macro-estimated elasticities. The estimation of the variance of the unobserved component proceeds as in Swamy (1970), which shows that micro-level unobservable coefficient heterogeneity can lead to upward biases in aggregate regressions.

We then replace each neighborhood housing supply change $dS_{jt}/d\log(p_{jt}) = S_{jt}\eta_j$ by its empirical counterpart in the general equilibrium price change equation 22. This allows a robustness check for the estimates of Section 5, i.e. the robustness of estimated general equilibrium price effects to the introduction of neighborhood-level housing supply elasticity (η_j).

Demographic Shifts

Between 2000 and 2006 the San Francisco Bay area experienced both an increase in population and a change in its racial demographics. Comparison of data from the 2000 Census with estimates from the 2005–2009 American Community Survey suggests the the number of households increased 3.6%, and that, within that population, the fraction of Asian households increased 2.6 percentage points, the fraction of Black households increased 0.6 percentage points, and the fraction of Hispanic households increased 2.3 percentage points, with a corresponding decline in the fraction of white households (−5.5 percentage points). Both population growth and changes in the racial make up of the Bay area are likely impacting the distribution of house prices and segregation levels.

We simulate the impact of population growth first. At the initial 2000 equilibrium, the model assumes that demand and supply are equal, where $N \cdot \mathbf{D} = N_{hu} \cdot \mathbf{s}$. The right-hand side factor N_{hu} is the total number of housing units in the Bay area in 2000. An increase in population dN causes an increase in demand $\mathbf{D} + N \cdot \frac{\partial \mathbf{D}}{\partial \log(\mathbf{p})} \cdot \frac{\partial \log(\mathbf{p})}{\partial N}$, where $\frac{d\mathbf{D}}{dN}$ includes the impact of prices on demand $\frac{\partial \mathbf{D}}{\partial \log(\mathbf{p})} \cdot \frac{\partial \log(\mathbf{p})}{\partial N}$. Such population growth also causes an increase in supply driven by price changes, $N_{hu} \cdot \frac{\partial \mathbf{s}}{\partial \log(\mathbf{p})} \cdot \frac{d\log(\mathbf{p})}{dN}$. Hence the general equilibrium impact of population growth on prices:

$$\frac{d\log(\mathbf{p})}{dN} = -\left[\frac{\partial \mathbf{D}}{\partial \log(\mathbf{p})} - \frac{\partial \mathbf{s}}{\partial \log(\mathbf{p})}\right]^{-1} \cdot \frac{1}{N} \cdot \mathbf{D} \tag{33}$$

where $\partial \mathbf{D}/\partial \log(\mathbf{p})$ is the Jacobian of demand, $\frac{\partial \mathbf{s}}{\partial \log(\mathbf{p})}$ is the supply response to price changes,

Thus population growth typically leads to non-linear increases in prices, which can lead to shifts in racial exposure. Similarly, a change in the Bay area's racial composition, where the fraction π_k of each racial group $k = 1, 2, \ldots, K$ changes by $d\pi_k$ such that $\sum_k d\pi_k = 0$, leads to the following price change:

$$\frac{d\log(\mathbf{p})}{d\boldsymbol{\pi}} = -\left[\frac{\partial\mathbf{D}}{\partial\log(\mathbf{p})} - \frac{\partial\mathbf{s}}{\partial\log(\mathbf{p})}\right]^{-1} \cdot \left[\frac{\partial\mathbf{D}}{\partial\boldsymbol{\pi}} + \frac{\partial\mathbf{D}}{\partial\boldsymbol{v}} \cdot \frac{\partial\mathbf{v}}{\partial\boldsymbol{\pi}}\right] \tag{34}$$

Appendix Table A displays the estimated moments of log price increases and the magnitude of the compression of the price distribution $Corr(\Delta\log(\mathbf{p}), \mathbf{p})$ as predicted by the data (first line), the baseline model (last line), and by all possible combination of extensions incorporating local housing elasticity, population change, and racial composition change. Extensions are either incorporated one at a time or in combination. The extended model yield results that are rather similar to those obtained with the baseline model. The correlation capturing the compressions of the house price distribution ranges between -0.2 and -0.38. The most comprehensive extension which combines local elasticity, population growth, and change in racial demographics of the SF Bay, yields the same price compression as the baseline model, and produces a similarly good match with the data on the key moments of the distribution of house price changes.

7 Conclusion

This paper's goal was to improve our understanding of the equilibrium relationships between lending standards, location choices, house prices and segregation. We do so by designing and estimating a model of neighborhood choice that explicitly takes into account the role of borrowing constraints. A key estimation result is that borrowing elasticity, the component of price elasticity associated with price-induced changes in mortgage approval probabilities, explains a larger share of demand elasticity than conditional demand elasticity, the component of price elasticity that reflects price-induced changes in utility. Such result validates the idea that differences in probability of mortgage approval across neighborhoods strongly shape the choice set of prospective home buyers.

The comparative static analysis of the model provides measures of the general equilibrium effects of a change in lending standards on the distribution of prices and households. The model's predictions are well in line with the data regarding the key moments of the distribution of house price changes and the compression of the price distribution between 2000 and 2006. The model also predicts key changes in the observed patterns of segregation such as the reduction in the isolation of

Whites. The ability of the estimated model to jointly predict observed changes in the distribution of prices and in the distribution of households strengthens the view that general equilibrium effects associated with changes in credit conditions play an important role in the dynamics of cities.

The estimated model could be used to understand other important dimensions of city dynamics. For example, the model could be informative about the effects of a change in the level or distribution of income, or the effects of a change in the spatial distribution of amenities such as schools or access to transportation. This paper looks at the causal effect of changes in lending standards on city outcomes. The model could also be used "in reverse" in order to trace down the fundamental causes of salient city transformations such as the gentrification of some neighborhoods.

References

Abaluck, J. & Adams, A. (2017), What do consumers consider before they choose? identification from asymmetric demand responses, Working Paper 23566, National Bureau of Economic Research.

Adelino, M., Gerardi, K. & Willen, P. S. (2013), 'Why don't lenders renegotiate more home mortgages? Redefaults, self-cures and securitization', *Journal of Monetary Economics* **60**(7), 835–853.

Adelino, M., Schoar, A. & Severino, F. (2012), Credit supply and house prices: evidence from mortgage market segmentation, Technical report, National Bureau of Economic Research.

Bayer, P., Ferreira, F. & McMillan, R. (2007), 'A unified framework for measuring preferences for schools and neighborhoods', *Journal of Political Economy* .

Bayer, P., McMillan, R., Murphy, A. & Timmins, C. (2016), 'A dynamic model of demand for houses and neighborhoods', *Econometrica* **84**(3), 893–942.

Bayer, P., McMillan, R. & Rueben, K. S. (2004), 'What drives racial segregation? new evidence using census microdata', *Journal of Urban Economics* **56**, 514–535.

Bayer, P. & Timmins, C. (2007), 'Estimating equilibrium models of sorting across locations', *The Economic Journal* **117**(518), 353–374.

Becker, G. S. & Murphy, K. M. (2009), *Social economics: Market behavior in a social environment*, Harvard University Press.

Berry, S., Levinsohn, J. & Pakes, A. (1995), 'Automobile prices in market equilibrium', *Econometrica: Journal of the Econometric Society* pp. 841–890.

Brock, W. A. & Durlauf, S. N. (2001), 'Discrete choice with social interactions', *The Review of Economic Studies* **68**(2), 235–260.

Conlon, C. T. & Mortimer, J. H. (2013), 'Demand estimation under incomplete product availability', *American Economic Journal: Microeconomics* **5**(4), 1–30.

Cutler, D. M., Glaeser, E. L. & Vigdor, J. L. (1999), 'The rise and decline of the american ghetto', *Journal of Political Economy* **107**(3), 1–52.

Debreu, G. (1970), 'Economies with a finite set of equilibria', *Econometrica: Journal of the Econometric Society* pp. 387–392.

Dell'Arriccia, G., Igan, D. & Laeven, L. (2009), 'Credit booms and lending standards: Evidence from the subprime mortgage market', *European Banking Center Working Paper* pp. 1–44.

Durlauf, S. N. (2004), 'Neighborhood effects', *Handbook of regional and urban economics* **4**, 2173–2242.

Favara, G. & Imbs, J. (2015), 'Credit supply and the price of housing', *The American Economic Review* **105**(3), 958–992.

Fujita, M. (1989), *Urban economic theory: land use and city size*, Cambridge university press.

Geweke, J. (1988), 'Antithetic acceleration of monte carlo integration in bayesian inference', *Journal of Econometrics* **38**(1), 73–89.

Geweke, J. (1996), 'Monte carlo simulation and numerical integration', *Handbook of computational economics* **1**, 731–800.

Glaeser, E. L., Gottlieb, J. D. & Gyourko, J. (2012), Can cheap credit explain the housing boom?, *in* 'Housing and the Financial Crisis', University of Chicago Press, pp. 301–359.

Glaeser, E. L., Gottlieb, J. D. & Tobio, K. (2012), 'Housing booms and city centers', *The American Economic Review* **102**(3), 127.

Goeree, M. S. (2008), 'Limited information and advertising in the us personal computer industry', *Econometrica* **76**(5), 1017–1074.

Goux, D. & Maurin, E. (2007), 'Close neighbours matter: Neighbourhood effects on early performance at school', *The Economic Journal* **117**, 1–23.

Hansen, L. P. (1982), 'Large sample properties of generalized method of moments estimators', *Econometrica: Journal of the Econometric Society* pp. 1029–1054.

Heckman, J. J., Urzua, S. & Vytlacil, E. (2006), 'Understanding instrumental variables in models with essential heterogeneity', *The Review of Economics and Statistics* **88**(3), 389–432.

Holmes, T. & Sieg, H. (2015), 'Structural estimation in urban economics', *Handbook of Urban and Regional Economics* **5**.

Imbens, G. W. & Lancaster, T. (1994), 'Combining micro and macro data in microeconometric models', *The Review of Economic Studies* **61**(4), 655–680.

Imbs, J. & Mejean, I. (2015), 'Elasticity optimism', *American Economic Journal: Macroeconomics* **7**(3), 43–83.

Landvoigt, T., Piazzesi, M. & Schneider, M. (2015), 'The housing market(s) of San Diego', *American Economic Review* **105**(4), 1371–1407.

Loutskina, E. (2011), 'The role of securitization in bank liquidity and funding management', *Journal of Financial Economics* **100**(3), 663–684.

Loutskina, E. & Strahan, P. E. (2009), 'Securitization and the declining impact of bank finance on loan supply: Evidence from mortgage originations', *The Journal of Finance* **64**(2), 861–889.

Massey, D. S. & Denton, N. A. (1988), 'The dimensions of residential segregation', *Social Forces* **67**(2), 281–315.

McFadden, D. (1973), 'Conditional logit analysis of qualitative choice behavior', *Frontiers in Econometrics* pp. 105–142.

Newey, W. K. (1987), 'Efficient estimation of limited dependent variable models with endogenous explanatory variables', *Journal of Econometrics* **36**(3), 231–250.

Nunn, N. & Puga, D. (2012), 'Ruggedness: The blessing of bad geography in Africa', *Review of Economics and Statistics* **94**(1), 20–36.

Ouazad, A. & Rancière, R. (2016), 'Credit standards and segregation', *Review of Economics and Statistics* **98**(5), 880–896.

Petrin, A. (2002), 'Quantifying the benefits of new products: The case of the minivan', *Journal of political Economy* **110**(4), 705–729.

Riley, S. J., DeGloria, S. D. & Elliot, R. (1999), 'A terrain ruggedness index that quantifies topographic heterogeneity', *Intermountain Journal of Sciences* **5**(1-4), 23–27.

Rosenthal, S. S. (2014), 'Are private markets and filtering a viable source of low-income housing? estimates from a "repeat income" model', *The American Economic Review* **104**(2), 687–706.

Saiz, A. (2010), 'The geographic determinants of housing supply', *The Quarterly Journal of Economics* pp. 1253–1296.

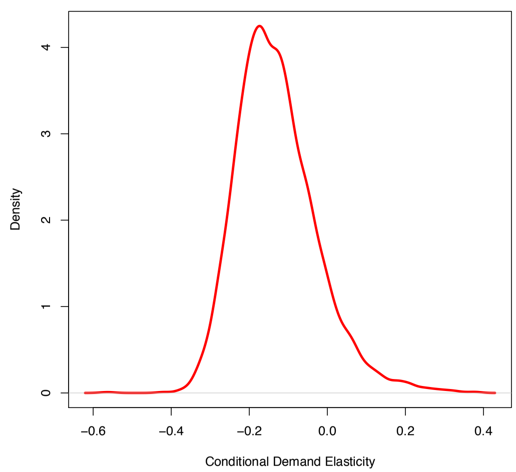
Swamy, P. A. (1970), 'Efficient inference in a random coefficient regression model', *Econometrica: Journal of the Econometric Society* pp. 311–323.

Train, K. (2009), 'Discrete choice models', pp. 1–42.

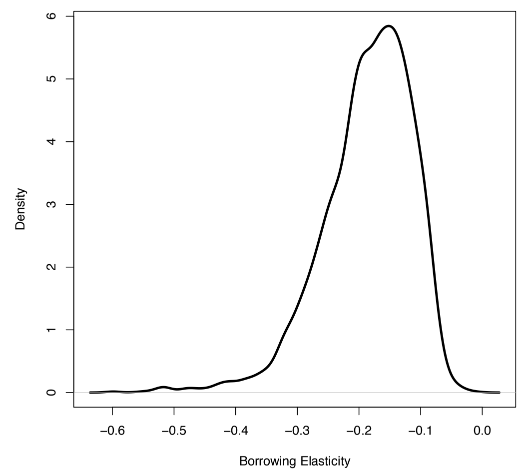
Figure 1: Demand Elasticity Distributions – Model with Borrowing Constraints

These figures present the total own-price demand elasticity η_j *(panel (c)) for each of the 4,416 neighborhoods, which is the sum of the conditional own-price demand elasticity* $\eta|_{C,j}$ *(panel (a)) and the borrowing elasticity* $\eta_j - \eta|_{C,j}$ *(panel (b)). The conditional demand elasticity is the elasticity at given borrowing constraints, i.e. at given choice set.*

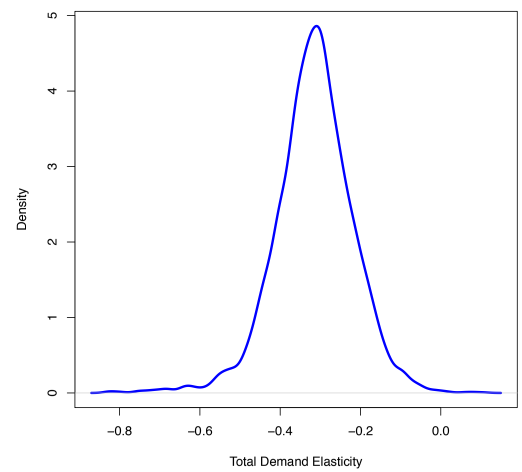
(a) Conditional Demand Elasticity



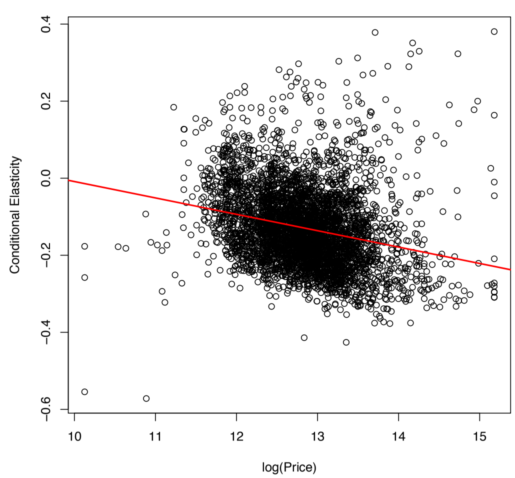
(b) Borrowing Elasticity



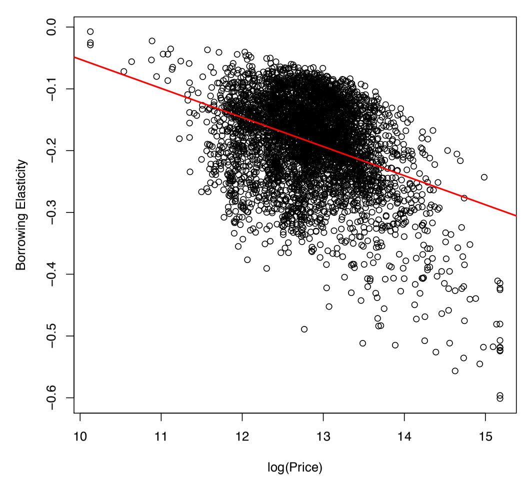
(c) Total Elasticity



(d) Conditional Elasticity vs. Price



(e) Borrowing Elasticity vs. Price



(f) Total Elasticity vs. Price

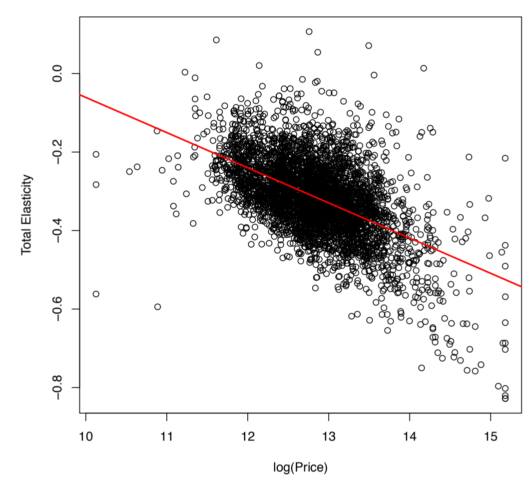
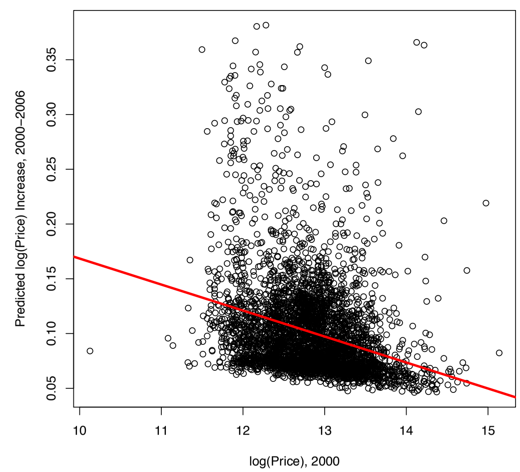


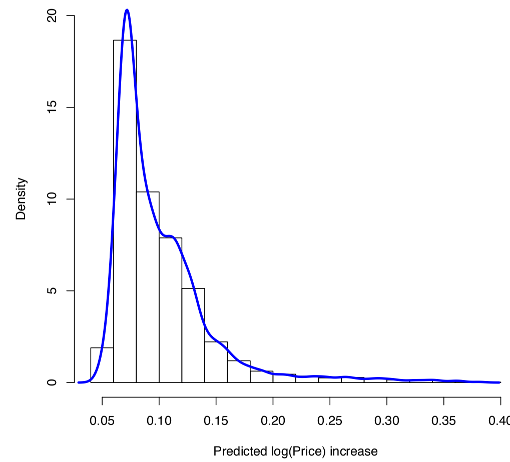
Figure 2: General Equilibrium Effects of Lending Standards on the Price Distribution

These figures illustrate the general equilibrium compression of the price distribution caused by the relaxation of lending standards. The relaxation of lending standards $d\psi$ is estimated in Appendix Table C. Figure (a) shows the relation between initial log(price) in 2000 and general equilibrium log(price) growth. Figure (b) shows the distribution of general equilibrium log price changes $d\log(\boldsymbol{p})/d\boldsymbol{\psi}$. Figure (c) shows the partial equilibrium effects $\partial\mathbf{D}/\partial\boldsymbol{\psi}$ of lending standards on demand at constant prices, in percentage of initial demand $\mathbf{D}$. Figure (d) shows the social interaction impact of lending standards on demand $\partial\mathbf{D}/\partial\boldsymbol{v}\cdot\partial\boldsymbol{v}/\partial\boldsymbol{\psi}$, in percentage of initial demand $\mathbf{D}$. Notations as in Section 5.

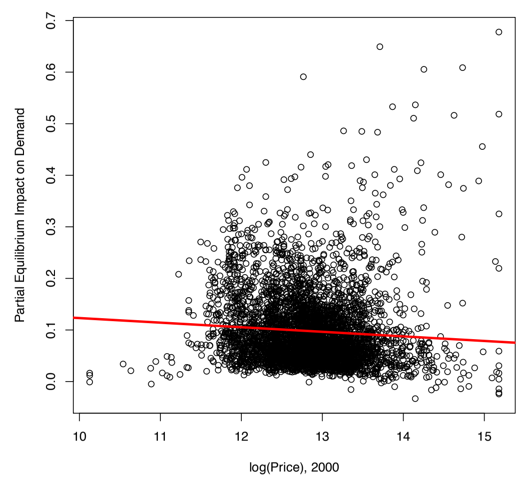
(a) G.E. Effect and Initial log(Price)



(b) Distribution of log(Price) Changes



(c) Partial Equilibrium Demand Effects $\frac{\partial \mathbf{D}}{\partial \boldsymbol{\psi}}/\mathbf{D}$



(d) Social Interaction Impacts $(\partial\mathbf{D}/\partial\mathbf{v}\cdot\partial\mathbf{v}/\partial\psi)/\mathbf{D}$

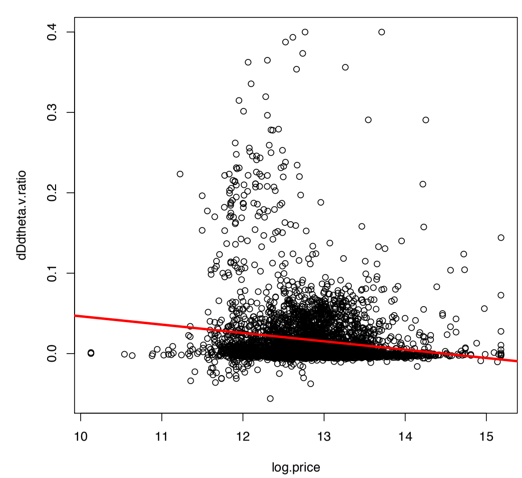
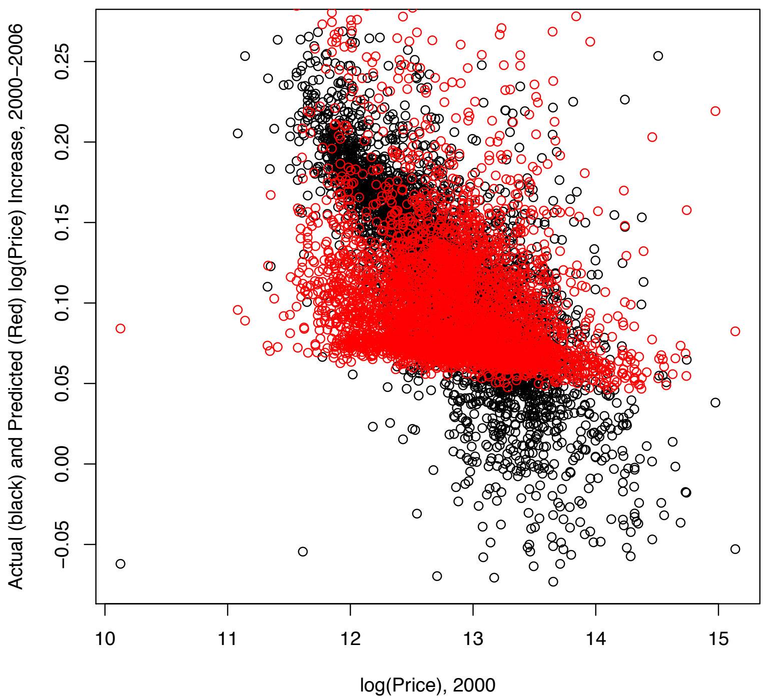


Figure 3: Compression of the Price Distribution – Comparison of General Equilibrium Predictions and the Actual 2000–2006 log(Price) Change

Black points below plot the actual *average annual change in the log(price) from our property-level transaction data against the initial log(price) in 2000. Red points plot the log(price) change as predicted by the model's general equilibrium comparative statics against the initial log(price) in 2000. Each point is a blockgroup.*

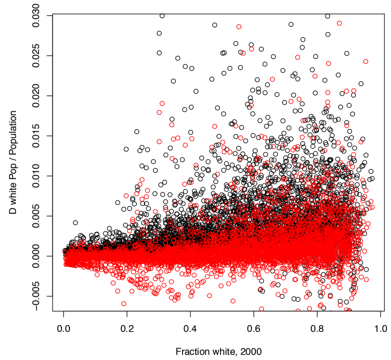


Actual log price change					
1st Qu.	Median	Mean	S.D.	3rd Qu.	Corr($\Delta log(p_{jt}), \log(p_{jt})$)
0.076	0.110	0.111	0.055	0.148	-0.748

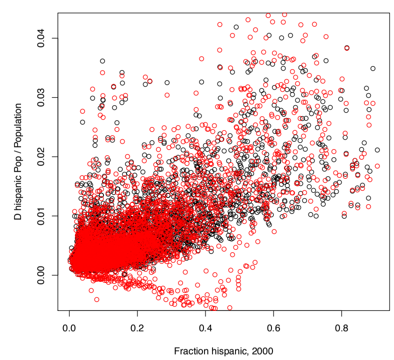
Predicted log price change					
1st Qu.	Median	Mean	S.D.	3rd Qu.	Corr($\Delta log(p_{jt}), \log(p_{jt})$)
0.072	0.087	0.101	0.045	0.116	-0.286

Figure 4: General Equilibrium Effects of Lending Standards on Segregation (1/2)

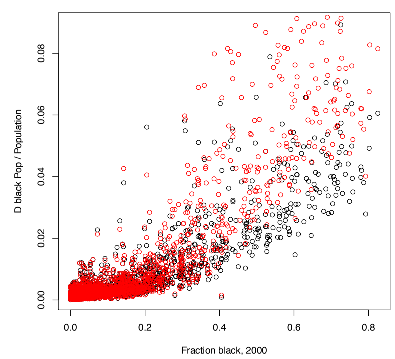
The four top figures (a)–(d) illustrate the impact of lending standards on segregation as predicted by the model's comparative statics. The red dots are the demand changes accounting for the general equilibrium change in prices and demographics. The black dots are the partial equilibrium demand changes, i.e. at given prices and demographics. The bottom four panels (e)–(h) display the correlation between general equilibrium price changes and initial neighborhood-level racial demographics.



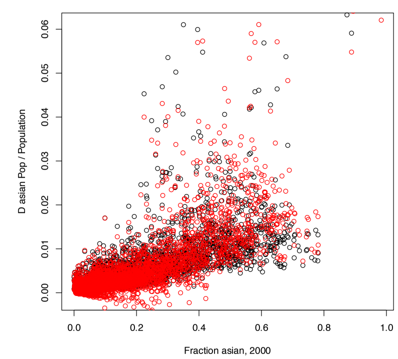
(a) White Demand Change and Initial Fraction White



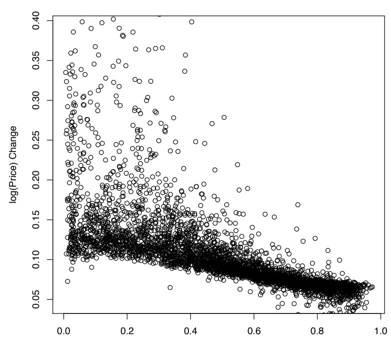
(b) Hispanic Demand Change and Initial Fraction Hispanic



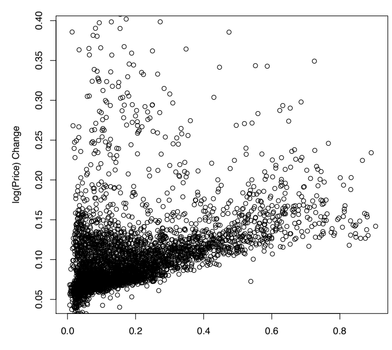
(c) Black Demand Change and Initial Fraction Black



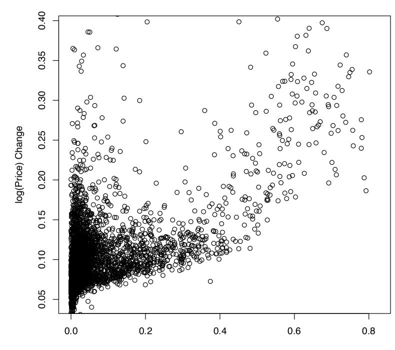
(d) Asian Demand Change and Initial Fraction Asian



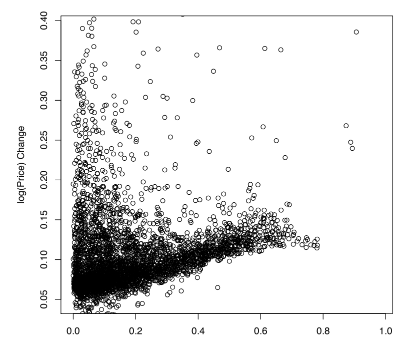
(e) Price Change and Initial Fraction White



(f) Price Change and Initial Fraction Hispanic



(g) Price Change and Initial Fraction Black

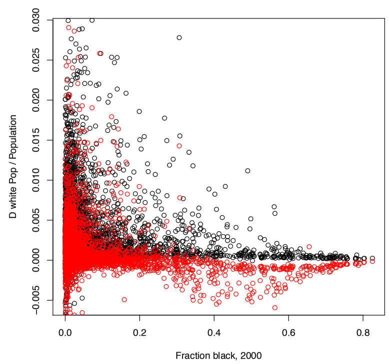


(h) Price Change and Initial Fraction Asian

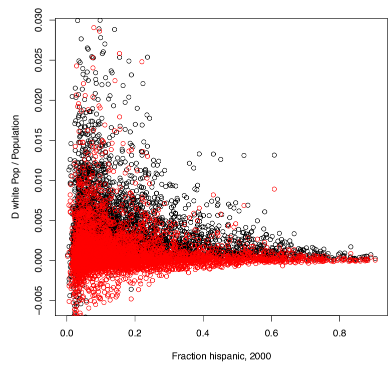
Figure 5: General Equilibrium Effects of Lending Standards on Segregation (2/2)

Figures (a)–(e) illustrate the impact of lending standards on segregation. The red dots are the demand changes accounting for the general equilibrium change in prices and demographics. The black dots are the partial equilibrium demand changes, i.e. at given prices and demographics.

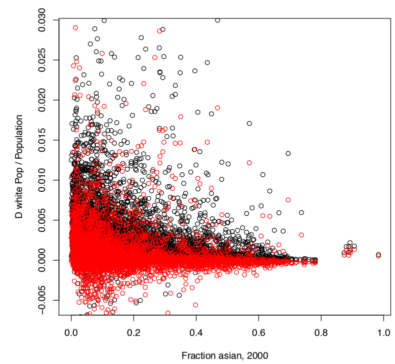
(a) White Demand Change and Initial Fraction Black



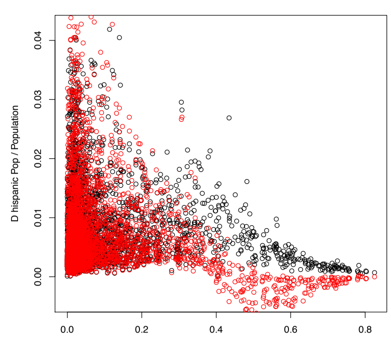
(b) White Demand Change and Initial Fraction Hispanic



(c) White Demand Change and Initial Fraction Asian



(d) Hispanic Demand Change and Initial Fraction Black



(e) Asian Demand Change and Initial Fraction Black

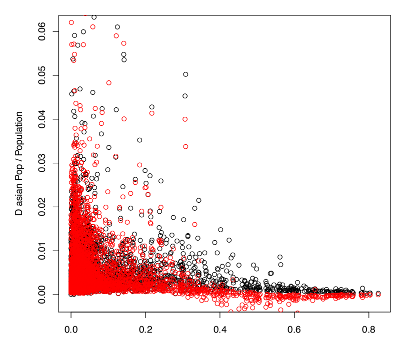


Table 1: Mortgage Approval Equation – Logit and IV Estimation – 1990–2010

The table presents the estimation of the approval model (Specification 6). In the top panel, column (1) presents the regression with tract fixed effect and no instrument. Column (2) presents the IV logit regression that uses the the liquidity of the nearest branches as IV. The top panel also includes VA and FSA-RHS indicators. The bottom panel presents two additional IV approaches.

	Specification:			
	Tract Fixed Effect		IV Liquidity of Nearby Branches	
	Logit Coefficients	Marginal Probabilities	Logit Coefficients	Marginal Probabilities
Loan to Income Ratio	−0.252***	−0.028***	−0.791***	−0.132***
	(0.008)	(0.001)	(0.183)	(0.025)
Black	−0.718***	−0.080***	−0.461***	−0.077***
	(0.031)	(0.003)	(0.054)	(0.009)
Asian	−0.066***	−0.007***	−0.012	−0.002
	(0.021)	(0.002)	(0.044)	(0.007)
Hispanic	−0.331***	−0.037***	−0.199***	−0.033***
	(0.029)	(0.003)	(0.046)	(0.007)
Other Race	−0.410***	−0.046***	−0.159***	−0.026***
	(0.024)	(0.003)	(0.058)	(0.009)
Year 2000	0.026	0.003	0.022	0.003
	(0.024)	(0.003)	(0.070)	(0.011)
Year 2010	0.031	0.003	0.388**	0.065**
	(0.031)	(0.003)	(0.173)	(0.026)
FHA Insured Loan	0.150***	0.017***	0.062	0.010
	(0.029)	(0.003)	(0.061)	(0.010)
Observations	208,206	208,206	41,153	41,153
Census Tracts	1,606	1,606	1,606	1,606
Pseudo R Squared	0.0436	0.0436	159.84	159.84

	IV Specification:			
	Liquidity of Nearby Branches Set Up before 1985		Branches IV and Racial Composition of Adjacent Neighborhoods	
	Probit Coefficients	Marginal Probabilities	Probit Coefficients	Marginal Probabilities
Loan to Income Ratio	−1.049***	−0.225***	−1.037***	−0.228***
	(0.209)	(0.042)	(0.205)	(0.044)
Black	−0.364***	−0.078***	−1.247***	−0.274***
	(0.056)	(0.011)	(0.266)	(0.056)
Asian	0.038	0.008	0.320**	0.070**
	(0.044)	(0.009)	(0.153)	(0.035)
Hispanic	−0.121**	−0.026**	−0.713***	−0.156***
	(0.049)	(0.010)	(0.222)	(0.046)
Other Race	−0.064	−0.013	−0.080	−0.017
	(0.063)	(0.013)	(0.086)	(0.018)
FHA Insured Loan	0.004	0.001	0.170***	0.037***
	(0.057)	(0.012)	(0.060)	(0.013)
VA-guaranteed	−0.025	−0.005	0.097	0.021
	(0.081)	(0.017)	(0.098)	(0.022)
FSA-RHS	0.446	0.095	0.466	0.102
	(0.318)	(0.067)	(0.297)	(0.064)
Year 2000	0.075	0.016	0.118***	0.026***
	(0.051)	(0.011)	(0.038)	(0.008)
Year 2010	0.686	0.147***	0.545***	0.119***
	(0.203)	(0.042)	(0.193)	(0.042)
Observations	41,153	41,153	41,153	41,153
Census Tracts	1,606	1,606	1,606	1,606
Wald χ^2	150.25	150.25	152.37	152.37

***: Significant at 1%, **: Significant at 5%, *: Significant at 10%.
Robust standard errors clustered at the census tract level in parenthesis.

Table 2: Household Preferences and Willingness to Pay for Neighborhood Amenities: Neighborhoods' Racial Demographics (1/2)

The table reports household preferences for neighborhoods' racial demographics. We consider the model with borrowing constraints (columns (1)–(2)), and the model without borrowing constraints (columns (3)–(4)). Columns (1) and (3) report the utility coefficient β_{ik} and the interaction terms. Columns (2) and (4) report the median willingness to pay (WTP) $Median_i\left[\frac{\beta_{ik}}{\alpha_i}\cdot IQR(z_k)\right]$ *for an interquartile range (IQR) change in the value of the amenity z_k, where α_i is the (conditional) price elasticity coefficient.*

	Model			
	with borrowing constraints		without borrowing constraints	
	Utility Coefficient	Median WTP for one IQR	Utility Coefficient	Median WTP for one IQR
Frac. Asian	−0.278	−9.090	−0.189	−6.96
	(0.063)	−\$39,984	(0.062)	−\$30,615
Standard deviation	0.072	–	0.006	–
	(0.931)		(1.112)	
x Asian	4.510	–	4.476	–
	(0.187)		(0.121)	
Frac. Black	−1.698	−21.930	−1.732	−20.83
	(0.079)	−\$96,463	(0.078)	−\$91,624
Standard deviation	0.105	–	0.043	–
	(1.281)		(0.972)	
x Black	6.217	–	6.298	–
	(0.107)		(0.082)	
Frac. Hispanic	0.350	12.880	0.413	17.34
	(0.077)	\$56,655	(0.077)	\$76,273
Standard deviation	0.127	–	0.085	–
	(1.150)		(1.404)	
x Hispanic	3.470	–	3.414	–
	(0.145)		(0.131)	
Frac. White	Ref.	Ref.	Ref.	Ref.
x White	2.933	–	2.967	–
	(0.090)		(0.085)	

Standard errors clustered by block group in parenthesis.

Table 3: Household Preferences and Willingness to Pay for Local Amenities (2/2)

The table reports household preferences for neighborhoods' education, income, and school test scores. We consider the model with borrowing constraints (columns (1)–(2)), and the model without borrowing constraints (columns (3)–(4)). Columns (1) and (3) report the utility coefficient β_i *and the interaction terms. Columns (2) and (4) report the median willingness to pay* (WTP) $Median_i\left[\frac{\beta_{ik}}{\alpha_i}\cdot IQR(z_k)\right]$ *for an interquartile range (IQR) change in the value of the amenity* z_k*, where* α_i *is the (conditional) price elasticity coefficient.*

(a) Neighborhood Education, Income, and School Performance Index

	Model			
	with borrowing constraints		without borrowing constraints	
	Utility Coefficient	Median WTP for one IQR	Utility Coefficient	Median WTP for one IQR
log(median household Income)	0.093 (0.018)	11.510 $50,629	0.094 (0.018)	11.48 $50,497
x log(household income)	0.465 (0.007)	–	0.436 (0.006)	–
Frac. College Educated	0.322 (0.048)	24.070 $105,876	0.322 (0.048)	23.73 $104,380
Standardized API	0.023 (0.005)	5.201 $22,877	0.008 (0.005)	1.944 $8,552
Standard deviation	0.009 (0.070)	–	0.001 (0.097)	–

The table reports conditional price elasticity estimates, interacted with household log income. This conditional price elasticity measure is the coefficient of log price in the utility specification equation 1. Including a random coefficient for log(price) yields similar results.

(b) Price Elasticity Measures

	Model	
	with borrowing constraints	w.o. borrowing constraints
log(Price)	−0.086 (0.032)	−0.299 (0.046)
x Standardized log(Income)	−0.363 (0.008)	−0.324 (0.011)

Standard errors clustered by block group in parenthesis.

Table 4: General Equilibrium Effects of Lending Standards on Racial Segregation

The table presents the comparative statics impact of the marginal change in lending standards on racial segregation. The numbers presented are the general equilibrium impacts, accounting for price and demographic changes, on racial exposure (Section 5.2). The tables read as follows: the cell in the second row and the first column is the impact of lending standards on the exposure of Blacks to Asians, in percentage points. The diagonal of each table is the impact on isolation, i.e. the exposure of a racial group to same-race neighbors.

(a) Predicted Segregation Changes

	Exposure to			
Exposure Of	Asian	Black	Hispanic	White
Asian	+1.15 ppt	−0.05 ppt	−0.02 ppt	−1.08 ppt
Black	−0.51 ppt	+2.84 ppt	−0.25 ppt	−2.07 ppt
Hispanic	−0.11 ppt	+0.07 ppt	+1.04 ppt	−1.00 ppt
White	+0.15 ppt	+0.20 ppt	+0.29 ppt	−0.63 ppt

(b) Actual Segregation Changes

	Exposure to			
Exposure Of	Asian	Black	Hispanic	White
Asian	+5.34 ppt	−0.95 ppt	+0.42 ppt	−4.19 ppt
Black	+1.29 ppt	−2.59 ppt	+2.54 ppt	−0.56 ppt
Hispanic	+0.87 ppt	−0.83 ppt	+4.48 ppt	−3.76 ppt
White	+1.91 ppt	−0.09 ppt	+1.54 ppt	−3.04 ppt

Online Appendix

October 19, 2017

Contents

Part I

Proofs and Estimation Details

1 Equilibrium Properties

1.1 Existence

Lemma 1. ***(Equilibrium Existence)*** *There exists an equilibrium vector of prices and neighborhood demographics* $(\mathbf{p}^* = (p^*_{jt}), \mathbf{v}^* = (\mathbf{v}^*_{jt}))_{j=1,2,...,J,\, t=1,2,...,T}$ *that satisfies conditions (13) and (14).*

Proof. We start by proving existence with an elastic housing supply, and then turn to equilibrium existence with a perfectly inelastic housing supply. First, note that the equilibrium of the city can be equivalently rewritten as:

$$\begin{aligned} s_1^{-1}(D_1(p_1, \ldots, p_J, \mathbf{v}_1, \ldots, \mathbf{v}_J) &= p_1 \\ &\vdots \\ s_J^{-1}(D_J(p_1, \ldots, p_J, \mathbf{v}_1, \ldots, \mathbf{v}_J)) &= p_J \\ D_1^x(p_1, \ldots, p_J, \mathbf{v}_1, \ldots, \mathbf{v}_J) &= v_1^x s_1 \\ &\vdots \\ D_J^x(p_1, \ldots, p_J, \mathbf{v}_1, \ldots, \mathbf{v}_J) &= v_J^x s_J \end{aligned}$$

for all neighbor characteristics $x \in X^+$ that affect individual demand. We note $|X^+| = K$. Define the mapping: $\phi : \mathbb{R}^J \times [0,1]^{KJ} \to \mathbb{R}^J \times [0,1]^{KJ}$. The mapping is continuous given the functional forms of the supply curves, the demand curves, and the probability of origination. Notice that the upper bound of D_j is 1 for each function, hence the mapping ϕ takes its values in $[s^{-1}(0); s^{-1}(1)]^J \times [0,1]^{KJ}$, which is a closed compact subset of $\mathbb{R}^J \times [0,1]^{KJ}$. We can thus consider the mapping $\tilde{\phi}$ of $[s^{-1}(0); s^{-1}(1)]^J \times [0,1]^{KJ}$ to itself, equal to ϕ on $[s^{-1}(0); s^{-1}(1)]^J \times [0,1]^{KJ}$. Such a mapping is continuous, and $[s^{-1}(0); s^{-1}(1)]^J \times [0,1]^{KJ}$ is a convex set of a Banach space. Hence, by the Brouwer fixed point theorem, $\tilde{\phi}$ admits a fixed point, i.e. a vector $(p_1^*, \ldots, p_J^*, \mathbf{v}_1^*, \ldots, \mathbf{v}_J^*)$ that satisfies the $J + KJ$ equations that define the equilibrium.

We then consider the case of a perfectly inelastic housing supply. Set $s_j(p) = s_j \cdot p^{\eta^s}$ with η^s the elasticity of housing supply. For a given η^s consider the set of equilibrium vectors $E(\eta^s) = \{(p_1^*, \ldots, p_J^*, \mathbf{v}_1^*, \ldots, \mathbf{v}_J^*)\}$. We just showed that $E(\eta^s) \neq \emptyset$ for any $\eta^s > 0$. Consider a sequence of equilibrium vectors for a sequence of $\eta^s \to 0$. Such sequence of equilibrium vectors converges to a vector $(\mathbf{p}^* = (p_{jt}^*), \mathbf{v}^* = (\mathbf{v}_{jt}^*))_{j=1,2,\ldots,J,\, t=1,2,\ldots,T}$ that is an equilibrium price vector when $\eta^s = 0$. $\square$

1.2 Local and Global Uniqueness

Lemma 2. ***(Local Equilibrium Uniqueness)*** *An equilibrium vector* $(\mathbf{p}^*, \mathbf{v}^*)$ *is locally unique almost surely. Formally, write*

$$\mathcal{D}_0 = \Bigg\{ \mathbf{D} = (D_{jt}) \quad s.t. \quad \exists\, (\mathbf{p}^*, \mathbf{v}^*)\, \forall j, t \quad D(j, t; \mathbf{p}^*, \mathbf{v}^*) = D_{jt},$$
$$\forall x \in X^+, \quad \mathbf{v}_{jt}(x) = \frac{D(j, t|x; \mathbf{p}^*, \mathbf{v}^*)}{D(j, t; \mathbf{p}^*, \mathbf{v}^*)} \;\; and \;\; rank\left(\frac{\partial \mathbf{D}}{\partial(\mathbf{p}, \mathbf{v})}\right) < J - 1 \Bigg\}$$

the set of neighborhood demands such that there is a corresponding price and demographics vector for which the Jacobian of demand is not of full rank. Then $Prob(\mathbf{D} \in \mathcal{D}_0) = 0$.

Proof. The result follows from Sard's (1942) theorem. Given that all derivatives of the mapping $D : \mathbb{R}^{J-1} \times \mathbb{R}^{(J-1)K} \to \mathbb{R}^J \times \mathbb{R}^{JK}$ exist and are continuous, the Lebesgue measure of the set of critical values of D is zero. This implies that for any probability measure over $\mathbb{R}^J \times \mathbb{R}^{JK}$, the probability that $\mathbf{D} \in \mathcal{D}_0$ is zero. $\square$

Lemma 3. ***(Global Equilibrium Uniqueness with no Social Preferences)*** *Where there are no preferences for same-race neighbors* ($\gamma_i = 0$)*, the city equilibrium is unique, up to one price.*

Proof. To prove such equilibrium uniqueness, we extend the model by treating consumer income as a household endowment ω_i for each household i. Consumer demand for neighborhood j is $D_j(p_1, \ldots, p_J, p)$, where p is the price of the numeraire consumer good. Thus the value of household i's endowment is $p\omega_i$ in terms of the numeraire good. Now notice that demand for neighborhood j is homogeneous, as the approval specification depends on the ratio of price and income and neighborhood demands are left unchanged when all prices are multiplied by a constant. Thus p can be normalized to 1. Note also that as the price of neighborhood j increases, demand for neighborhood j strictly decreases (the probability of acceptance goes down strictly and the utility value of neighborhood j strictly decreases). When the price of neighborhood $-j$ strictly increases, demand for neighborhood j strictly increases (the probability of acceptance in neighborhood $-j$ goes down strictly, and the utility value of neighborhood $-j$ strictly decreases). Thus housing in neighborhood 1 and housing in neighborhood 2 are gross substitutes. By Proposition 17.F.3 of Mas-Colell, Whinston, Green et al. (1995), the equilibrium of the city is unique up to one neighborhood price. $\square$

2 Identification and Estimation

2.1 Proof of Identification

In the following proposition, K is the dimension of $\mathbf{x}_{it}$, L the dimension of $\mathbf{z}_{jt}$, and W the dimension of $\mathbf{x}_{it}\Omega\mathbf{z}_{jt}$.

Proposition 2. *Assume that:*

1. $\mathbf{x}_{it}$ *does not lie in a strict subspace of* $\mathbb{R}^K$.
2. $\mathbf{z}_{jt}$ *does not lie in a strict subspace of* $\mathbb{R}^L$.
3. *Noting* W *the number of non-zero elements in* Ω, *the* W *interaction terms between* $\mathbf{x}_{it}$ *and* $\mathbf{z}_{jt}$ *do not lie in a strict subspace of* $\mathbb{R}^W$.
4. *There is no linear combination between the columns of* $\mathbf{x}_{it}$, $\mathbf{z}_{jt}$, *and the* W *interaction terms of* $\mathbf{x}_{it}$ *and* $\mathbf{z}_{jt}$. *Note* Ξ *the matrix of* $0, 1$*s such that* $\Xi_{kl} = 1$ *if and only if* Ω_{kl} *is not constrained to* 0. *Then,* $(\mathbf{x}_{it}, \mathbf{z}_{jt}, \mathbf{x}_{it}\Xi\mathbf{z}_{jt})$ *does not lie in a strict subspace of* $\mathbb{R}^K \times \mathbb{R}^L \times \mathbb{R}^W$.

Then, the model's structural parameters are identified, i.e., there does not exist two observationally-equivalent vectors of structural parameters. Formally, note $\boldsymbol{\theta} = (\boldsymbol{\psi}, \boldsymbol{\gamma}, \Psi, \boldsymbol{\delta}, \Omega, \Sigma)$ *following the paper's notations. If there are two vectors of structural parameters* $\boldsymbol{\theta}$ *and* $\breve{\boldsymbol{\theta}}$ *such that:*

1. *Probabilities of approval for all households* i, *neighborhoods* j, *and years* t *are equal under* $\boldsymbol{\theta}$ *and* $\breve{\boldsymbol{\theta}}$.
2. *Utilities* U_{ijt} *for all* i, j, t *are equal under* $\boldsymbol{\theta}$ *and* $\breve{\boldsymbol{\theta}}$.

Then $\boldsymbol{\theta} = \breve{\boldsymbol{\theta}}$. Further, the base utility preference coefficients $(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\xi})$ are identified under the assumption that $(\log p_{jt}, \mathbf{z}_{jt}, \xi_j)$ do not lie in a strict subspace of $\mathbb{R}^{2+L}$.

Proof. The proposition is proven by contradiction. Assume there are two parameter vectors $\boldsymbol{\theta}$ and $\breve{\boldsymbol{\theta}}$ such that $\boldsymbol{\theta} \neq \breve{\boldsymbol{\theta}}$, and such that probabilities of approval and utility values are equal. The parameter vectors include (i) lending standards coefficients $(\boldsymbol{\psi}, \boldsymbol{\gamma}, \boldsymbol{\Psi})$, (ii) base utilities $\boldsymbol{\delta}$, and (iii) heterogeneous preferences for amenities (interaction terms Ω and the variance-covariance Σ of random coefficients $\tilde{\boldsymbol{\beta}}_i$, with $E(\tilde{\boldsymbol{\beta}}_i) = 0$).

Identification of Lending Standards

If probabilities of approval are equal under $\boldsymbol{\theta}$ and $\breve{\boldsymbol{\theta}}$, then:

$$\Lambda\left(\mathbf{x}'_{it}\boldsymbol{\psi} + \mathbf{z}'_{jt}\boldsymbol{\gamma} + \mathbf{x}'_{it}\boldsymbol{\Psi}\mathbf{z}_{jt}\right) = \Lambda\left(\mathbf{x}'_{it}\breve{\boldsymbol{\psi}} + \mathbf{z}'_{jt}\breve{\boldsymbol{\gamma}} + \mathbf{x}'_{it}\breve{\boldsymbol{\Psi}}\mathbf{z}_{jt}\right) \tag{1}$$

on the support of $\mathbf{x}_{it}$ and $\mathbf{z}_{jt}$. As the logit function Λ is invertible this implies:

$$\mathbf{x}'_{it}(\boldsymbol{\psi} - \breve{\boldsymbol{\psi}}) + \mathbf{z}'_{jt}(\boldsymbol{\gamma} - \breve{\boldsymbol{\gamma}}) + \mathbf{x}'_{it}(\boldsymbol{\Psi} - \breve{\boldsymbol{\Psi}})\mathbf{z}_{jt} = 0 \tag{2}$$

on the support of $\mathbf{x}_{it}$, $\mathbf{z}_{jt}$, and all products of columns of $\mathbf{x}_{it}$ and $\mathbf{z}_{jt}$. As $\boldsymbol{\theta} \neq \breve{\boldsymbol{\theta}}$, this implies that there exists a linear combination between the columns of $\mathbf{x}_{it}$, $\mathbf{z}_{jt}$, and the interaction terms of $\mathbf{x}_{it}$ and $\mathbf{z}_{jt}$, which violates assumption 1. Thus $\boldsymbol{\psi} = \breve{\boldsymbol{\psi}}$, $\boldsymbol{\gamma} = \breve{\boldsymbol{\gamma}}$, and $\boldsymbol{\Psi} = \breve{\boldsymbol{\Psi}}$.

Base Utilities Identification

If utilities U_{ijt} are equal under $\boldsymbol{\theta}$ and $\breve{\boldsymbol{\theta}}$, then:

$$\delta_{jt} - \breve{\delta}_{jt} = -\left(\boldsymbol{x}'_{it}(\Omega - \breve{\Omega})\mathbf{z}_{jt} + (\tilde{\boldsymbol{\beta}}_{it} - \breve{\boldsymbol{\beta}}_{it})'\mathbf{z}_{jt}\right) \tag{3}$$

where $\mathbf{z}_{jt}$ is the vector of all neighborhood amenities. The left-hand side varies at the neighborhood-level, while the right-hand side can take any individual-level set of values. Taking the expectation of the right-hand side, this implies that the interaction terms evolve in a strict subspace of $\mathbb{R}^W$, which violates the proposition's assumption, unless $\delta_{jt} - \breve{\delta}_{jt} = 0$.

Identification of Random Coefficients

Then, interaction terms are equal to the difference of random coefficients.

$$\mathbf{x}_{it}'(\Omega - \breve{\Omega})\mathbf{z}_{jt} = -(\tilde{\boldsymbol{\beta}}_{it} - \breve{\boldsymbol{\beta}}_{it})'\mathbf{z}_{jt} \tag{4}$$

The variance of the left-hand side is zero, hence $\text{Var}(\tilde{\boldsymbol{\beta}}_{it} - \breve{\boldsymbol{\beta}}_{it}) = 0$, and $\Sigma = \breve{\Sigma}$.

Social Interactions – Non Linear Terms

In equation (3), split $\mathbf{z}_{jt}$ into exogenous amenities and into the vector $\mathbf{v}_{jt}$ of neighborhood demographics, with support $[0,1]^S$, so that $\mathbf{z}_{jt} = (\overline{\mathbf{z}}_{jt}, \mathbf{v}_{jt})$. Following the paper, the s-th element of $\mathbf{v}_{jt}$ is:

$$v_{jt}^s = \frac{N_{x^s}}{N} \cdot \frac{D(j,t|x^s;\mathbf{p}^*,\mathbf{v}^*)}{D(j,t;\mathbf{p}^*,\mathbf{v}^*)} \in [0,1] \tag{5}$$

with N_{x^s} the mass of households of type s, and x^s their characteristics. Write ζ and $\breve{\zeta}$ the coefficient vectors for $\mathbf{v}_{jt}$. Then, isolate the second and the third terms of equation (3).

$$\mathbf{x}_{it}(\Omega - \breve{\Omega})\overline{\mathbf{z}}_{jt} = -(\boldsymbol{\zeta} - \breve{\boldsymbol{\zeta}})\mathbf{v}_{jt} \tag{6}$$

The right-hand side has support in $[0,1]$, while the left-hand side has support in $\mathbb{R}$, unless $\boldsymbol{\zeta} - \breve{\boldsymbol{\zeta}} = 0$.

Base Utility Preference Coefficients

We finally turn to the identification of base utility preference coefficients. Assume there are two vectors $(\alpha, \boldsymbol{\beta}, \boldsymbol{\xi})$ and $(\breve{\alpha}, \breve{\boldsymbol{\beta}}, \breve{\boldsymbol{\xi}})$ such that, for all j:

$$-\alpha \log p_{jt} + \mathbf{z}_{jt}\boldsymbol{\beta} + \xi_j = -\breve{\alpha} \log p_{jt} + \mathbf{z}_{jt}\breve{\boldsymbol{\beta}} + \breve{\xi}_j \tag{7}$$

This implies

$$-(\alpha - \breve{\alpha}) \log p_{jt} + \mathbf{z}_{jt}(\boldsymbol{\beta} - \breve{\boldsymbol{\beta}}) = -\left(\xi_j - \breve{\xi}_j\right) \tag{8}$$

As before, the right hand side varies at the neighborhood level, while the left-hand side varies at both the year and neighborhood level. In other words, (8) implies that $(\log p_{jt}, \mathbf{z}_{jt})$ lies in a strict subspace of $\mathbb{R}^{L+1}$, which violates the proposition's assumption. This implies that $\xi_j - \breve{\xi}_j = 0$, $\alpha - \breve{\alpha} = 0$ and $\boldsymbol{\beta} - \breve{\boldsymbol{\beta}} = 0$. $\square$

2.2 Estimation Technique

We estimate the model using the nested fixed point (NFP) method of Berry, Levinsohn & Pakes (1995). Su & Judd (2012) has shown that the NFP and the MPEC method of Dubé, Fox & Su (2012) are equivalent, but the NFP method requires a tight tolerance level to converge. The inner loop of our contraction mapping uses a 10^{-12} tolerance, and the outer loop a 10^{-6} tolerance level.

Estimating the model in double precision Fortran has considerable impacts on speed and precision. The algorithm is available through the corresponding author. The contraction mapping is also parallelized by acknowledging that the demand for housing is independent across the three decades of the data, 1990, 2000, and 2010. For each computation of $\boldsymbol{\delta} = (\boldsymbol{\delta}_{1990}, \boldsymbol{\delta}_{2000}, \boldsymbol{\delta}_{2010})$ therefore, the R package `snowfall` parallelizes the three contraction mappings.

Minimization of the objective function $G(\boldsymbol{\theta})$ proceeds using the Nelder-Mead algorithm, and the nature of the optimum $\hat{\boldsymbol{\theta}}$ (global vs. local) is assessed using profile GMM objective functions.

Proposition 3. ***(Contraction mapping)*** *Consider the mapping from $\hat{\boldsymbol{\delta}}$ to $f(\hat{\boldsymbol{\delta}})$, defined as in equation (16):*

$$\begin{aligned} f: \quad & \mathbb{R}^{J-1} \to \mathbb{R}^{J-1} \\ & \hat{\boldsymbol{\delta}} \mapsto \hat{\boldsymbol{\delta}} + \log(\hat{\mathbf{D}}_{-j^0}) - \log\left(\mathbf{D}_{-j^0}(\hat{\boldsymbol{\delta}})\right) \end{aligned}$$

where $\hat{\mathbf{D}}_{-j^0}$ is the $J-1$-vector of observed demands excluding neighborhood j^0, and $\mathbf{D}_{-j^0}(\hat{\boldsymbol{\delta}})$ is the vector of predicted demand given the $J-1$-vector of base utilities. Then: (i) f admits a unique fixed point $\boldsymbol{\delta}_0 \in \mathbb{R}^{J-1}$, and (ii) there is a value $\overline{\delta} \in \mathbb{R}$ such that $\hat{f}_j = \min\left\{f_j(x), \overline{\delta}\right\}$ is a contraction mapping with modulus strictly less than one and the sequence $\boldsymbol{\delta}^{k+1} = f(\boldsymbol{\delta}^k)$ with $\boldsymbol{\delta}^0 \in \mathbb{R}^{J-1}$ converges to the fixed point $\boldsymbol{\delta}_0$.

Proof. We prove the proposition by showing that the function $f(\cdot)$ satisfies the three conditions of

Berry et al.'s (1995) Theorem (page 887):

1. $\forall x \in \mathbb{R}^{J-1}$, f is continuously differentiable, with $\forall j$ and k, $\partial f_j(x)/\partial x_k \geq 0$ and $\sum_{k=1}^{J-1} \partial f_j/\partial x_k < 1$.

2. $\min_j \inf_x f(x) > -\infty$

3. There is a scalar $\overline{x} \in \mathbb{R}$ with the property that if for any j, $x_j \geq \overline{x}$, then for some k, $f_k(x) < x_k$.

Berry et al. (1995) proves that if conditions 1–3 are satisfied, f admits a unique fixed point $x_0 \in \mathbb{R}^{J-1}$; the truncated function $\hat{f} : \mathbb{R}^{J-1} \to \mathbb{R}^{J-1}$, $x \mapsto \min\{f_j(x), \overline{x}\}$ is Lipschitz with modulus less than one.

In this paper, demand for neighborhood j in year t is the probability-weighted sum of the demands $D(j,t|C)$ conditional on each choice set C, for all choice sets $C \subset \mathbb{R}^J$, i.e. $D(j,t) = \sum_{C \subset \mathbb{R}^J} P(C) \cdot D(j,t|C)$, with $P(C) \in (0,1)$ for all C. Then:

$$\frac{\partial f_{jt}}{\partial \delta_{jt}} = 1 - \frac{\sum_{C \subset \mathbb{R}^J} P(C)}{\sum_{C \subset \mathbb{R}^J} P(C) \cdot D_{jt|C}} \frac{\partial D_{jt|C}}{\partial \delta_{jt}}$$
$$\frac{\partial f_{jt}}{\partial \delta_{kt}} = -\frac{1}{D_{jt}} \frac{\partial D_{jt}}{\partial \delta_{kt}}$$

Berry et al. (1995) established the properties conditional on the choice set C, specifically that $\partial D_{jt|C}/\partial \delta_{jt} < D_{jt|C}$ for all j and all $C \subset \mathbb{R}^J$. Thus $\sum_{C \subset \mathbb{R}^J} P(C) \cdot \partial D_{jt|C}/\partial \delta_{jt} < \sum_{C \subset \mathbb{R}^J} P(C) \cdot D_{jt|C}$, and $\frac{\partial f_{jt}}{\partial \delta_{jt}} > 0$. Berry et al. (1995) also established that $\partial D_{jt|C}/\partial \delta_{kt} < 0$, which implies that $\sum_C P(C) \partial D_{jt|C}/\partial \delta_{kt} < 0$, and $\partial f_{jt}/\partial \delta_{kt} > 0$. Similarly, $\sum_{k=1}^{J} \partial D_{jt|C}/\partial \delta_{kt} < D_{jt|C}$ for all $C \subset \mathbb{R}^J$ implies that $\sum_{k=1}^{J} \partial D_{jt}/\partial \delta_{kt} < D_{jt}$. Thus the function $f(\cdot)$ satisfies the conditions of Assumption 1 of Berry et al.'s (1995) Theorem. Conditions 2 and 3 similarly hold when considering $D(j,t) = \sum_{C \subset \mathbb{R}^J} P(C) \cdot D(j,t|C)$: writing $D_j(\boldsymbol{\delta}) = \exp(\delta_j) \cdot F(\boldsymbol{\delta})$ shows that $f(\cdot)$ admits a lower bound. The proof of Berry (1994) applies to the function$f(\cdot)$ of this proposition and implies condition 3.

With conditions 1–3 satisfied, the function $\hat{f} : \mathbb{R}^{J-1} \to \mathbb{R}^{J-1}$ whose j-th component is $\hat{f}_j(x) = \min\{f_j(x), \overline{x}\}$ is a contraction mapping with modulus less than one that converges to the fixed point $\boldsymbol{\delta}_0 \in \mathbb{R}^{J-1}$. $\square$

2.3 Standard Errors for Structural Parameters

The model's structural parameter vector $\boldsymbol{\theta}$ is estimated with three sets of moment conditions, as described in Section 3.1 (iv) of the paper:

$$E\left[G\left(\boldsymbol{\psi}, \boldsymbol{\gamma}, \Psi, \alpha, \boldsymbol{\beta}, \boldsymbol{\xi}, \Omega, \Sigma\right)\right] = 0 \tag{9}$$

where (i) $\boldsymbol{\psi}, \boldsymbol{\gamma}, \Psi$ are the coefficients of lending standards, (ii) $\alpha, \boldsymbol{\beta}, \boldsymbol{\xi}$ are the base preference parameters and (iii) Ω, Σ are the non-linear preference parameters (interaction terms and variance-covariance of the random coefficients respectively).

There are three sources of sampling variation and errors in the data generating process that lead to a difference between the estimator $\left(\widehat{\boldsymbol{\psi}}, \widehat{\gamma}, \widehat{\Psi}, \widehat{\alpha}, \widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{\xi}}, \widehat{\Omega}, \widehat{\Sigma}\right)$ and the true value $(\boldsymbol{\psi}, \boldsymbol{\gamma}, \Psi, \alpha, \boldsymbol{\beta}, \boldsymbol{\xi}, \Omega, \Sigma)$.

1. Lending standards parameters $\boldsymbol{\psi}, \boldsymbol{\gamma}, \Psi$ are estimated on the HMDA dataset of mortgage applications by maximum likelihood.

2. Base utility parameters are estimated by a 2-stage least squares regression on the neighborhood data set.

3. Non-linear preference parameters Ω, Σ are estimated by the simulated method of moments, where the true moments are simulated given the large number of possible choice sets.

The paper draws on three data sets: (i) the HMDA data set, (ii) the 1 % micro Census, (iii) the blockgroup-level Census files.

We follow the bootstrap procedure of Shao & Tu (2012) to estimate standard errors. Shao & Tu's (2012) method samples by replacement from each of the data sets used in the GMM estimation. Such sampling yielding a new vector of structural parameters $\widehat{\boldsymbol{\theta}}_b$.

1. We sample by replacement a new mortgage application data set from the original HMDA mortgage application data set, of the same size. The logit regression by maximum likelihood yields new estimates $\widehat{\boldsymbol{\psi}}_b, \widehat{\gamma}_b, \widehat{\Psi}_b$.

2. We sample by replacement a new 1 % micro Census data set from the original 1% sample, of the same size.

3. We generate choice sets for each of the households in the newly sample 1 % micro Census data set obtained in the previous step, using the lending standards coefficients obtained in the first step.

4. We sample by replacement a new set of neighborhoods from the original set of neighborhoods, of the same size. Demands $\mathbf{D}_t$ for each neighborhood are computed so that $\sum_j D_{jt} = 1$.

5. The model's base and non-linear preference parameters are estimated using the choice sets obtained in step 3., and the samples obtained in steps 2. and 4.

Applying steps 1–5 yields a estimated vector of structural parameters $\widehat{\boldsymbol{\theta}}_b$. Repeating this procedure B times leads to an estimate F_b of the cumulative distribution function of the estimator $\widehat{\boldsymbol{\theta}}$. The standard error of each element of $\widehat{\boldsymbol{\theta}}$ is then the standard deviation of each elements of the set $\widehat{\boldsymbol{\theta}}_b$, $b = 1, 2, \ldots, B$.

2.4 Unobservables of Lending Standards and Demand Unobservables

The paper's identification strategies use two sets of instrumental variables to (i) estimate the impact of observable amenities on utilities, and (ii) to estimate the impact of observable household and mortgage characteristics on the probability of approval for mortgage applications. The key issue is that unobservables may be correlated with observables. While the IV approach may credibly orthogonalize the observables of *each* specification with the unobservable amenities, *common* unobservable amenities may *both* affect a neighborhood's utility level and the probability that a lender approves a mortgage.

For instance, unobservable predictors of future neighborhood price trends could affect both approval probabilities and the utility value of a neighborhood. They could affect approval probabilities as lenders consider the future value of the collateral in case of a foreclosure. Such unobservables could also affect the utility value of the neighborhood as capital appreciation lead to wealth and consumption gains. In such a case the correlation of unobservables would be positive.

The correlation of unobservables could be negative as well. That will be the case if, for instance, banks are more likely to extend loans into neighborhoods with a high expected increase in price, and those neighborhoods also have worse unobservable amenities.

The paper's point estimates suggests a statistically significant correlation between the non-time-varying unobservables that determine the probability of approval and the unobservables that determine neighborhood utility. Table F shows that the correlation between the fixed effects of the approval regression $\xi_j^{approval}$ and the fixed effects of the base utility term ξ_j is negative.

While such correlations affect the efficiency of the estimator, they typically do not affect the consistency of the estimator of $\boldsymbol{\theta}$ (Hansen 1982, Section 3). The bootstrap approach to the computation of standard errors described in the online appendix captures the potential correlation of lending standards' unobservables and utility unobservables.

2.5 Identification of Lending Standards

The data set provides information on lenders' decisions within the pool of applicants. Households' decision to apply can be modeled in a latent variable framework. In such a framework, household i applies for a mortgage in neighborhood j if and only if:

$$applies^*_{ijt} = \mathbf{x}_{it} b + \mathbf{z}_{jt} c + \mathbf{x}'_{it} \mathbf{h} \mathbf{z}_{jt} - \eta_{ijt} > 0, \tag{10}$$

with $G(\cdot)$ the cdf of η_{ijt}, and $\mathbf{h}$ a vector of interaction coefficients. Similarly, as in the paper, the lender approves the mortgage application if and only if:

$$approves^*_{ijt} = \mathbf{x}_{it} \boldsymbol{\psi} + \mathbf{z}_{jt} \boldsymbol{\gamma} + \mathbf{x}'_{it} \boldsymbol{\Psi} \mathbf{z}_{jt} - e_{ijt} > 0, \tag{11}$$

and the econometrician observe the approval decision conditional on application.

$$P(approves = 1 | applies = 1) = P(e_{ijt} < \mathbf{x}_{it} \boldsymbol{\psi} + \mathbf{z}_{jt} \boldsymbol{\gamma} + \mathbf{x}'_{it} \boldsymbol{\Psi} \mathbf{z}_{jt} | \eta_{ijt} < \mathbf{x}_{it} b + \mathbf{z}_{jt} c + \mathbf{x}'_{it} \mathbf{h} \mathbf{z}_{jt}) \tag{12}$$

which could boil down to $P(e_{ijt} < \mathbf{x}_{it} \boldsymbol{\psi} + \mathbf{z}_{jt} \boldsymbol{\gamma} + \mathbf{x}'_{it} \boldsymbol{\Psi} \mathbf{z}_{jt})$ provided households' decision to apply were independent of lenders' decision to approve the mortgage. In general however, unobservables of the application decision are correlated with unobservables of the approval decision.

The econometric approach introduces a vector of instruments $\boldsymbol{\xi}_{ijt}$ for $(\mathbf{x}_{it}, \mathbf{z}_{jt}) \equiv \mathbf{w}_{ijt}$. This leads to the two-step specification:

$$\begin{cases} \mathbf{w}_{ijt} & = & \boldsymbol{\xi}_{ijt}\boldsymbol{\varphi} + \nu_{ijt} \\ approves^*_{ijt} & = & \mathbf{w}_{ijt}\boldsymbol{\lambda} - \varepsilon_{ijt} \end{cases} \tag{13}$$

with the assumption that $\boldsymbol{\xi}_{ijt} \perp\!\!\!\perp \varepsilon_{ijt}$.

The liquidity of the national banks of branches closest to the house, noted $\boldsymbol{\xi}_{ijt}$ here, is used as an instrument that is arguably independent of the unobservable consumer-specific drivers of the approval decision ε_{ijt}. We use the liquidity of the neighboring banks and not the liquidity of the bank of the application itself, to avoid a correlation between $\boldsymbol{\xi}_{ijt}$ and household unobservables ε_{ijt} driven by households' self-selection into banks.

This paper's parameter of interest is $\boldsymbol{\lambda}$, the impact of mortgage and household characteristics on the approval decision. The vector of parameters $(\boldsymbol{\lambda}, \boldsymbol{\varphi}, \sigma^2_\nu, \sigma^2_\varepsilon, \rho)$ is estimated by maximum likelihood. With a bivariate normal distribution for (ν, ε), where ρ is the correlation of the observables of the two steps, the joint likelihood of the observations depends on the density of the bivariate distribution. The estimator is consistent under the assumption of independence of neighboring banks' liquidity $\boldsymbol{\xi}_{ijt}$ and unobservables ε_{ijt}.

A final point for the estimation of specification is that the impact of observables on the application may be heterogeneous, e.g. if $\boldsymbol{\lambda}$ depends on households' characteristics. If that is the case, the entry and exit of households into the applicant pool may lead to an estimate of $\boldsymbol{\lambda}$ that is not policy relevant.[1] We address this issue by weighing the likelihood to ensure that the distribution of household characteristics $f(\mathbf{x}_{it})$ stays constant across years, equal to the distribution of household characteristics at the beginning of the period of observation.

3 General Equilibrium Derivations

This Appendix's section provides the derivations for Section 5.1(ii): the impact of a log(price) change on demand, at given lending standards. Equation (25) has two parts: the shift in the choice set probabilities implied by the shift in price (without a change in lending standards), and the shift

[1] That is, the estimate is not a weighted average of treatment effects $\boldsymbol{\lambda}$s with positive weights that sum to 1. A related discussion of policy relevant parameters is presented in Heckman, Urzua & Vytlacil (2006).

in demand caused by a shift in utility at given choice set probabilities. We detail both below.

Shift in choice set induced by changing prices.

We first focus on the first term, which is the impact of the log price $log(p_k)$ on choice set probabilities. For every choice set that includes neighborhood k the probability of that choice set goes down, and for every choice set that does *not* include neighborhood k, the probability of that choice set goes up.

$$\frac{\partial}{\partial \log(p_{kt})} P(C|\mathbf{z}_{\cdot t}, \mathbf{x}_{it}; \psi) \quad = \quad -\alpha_{approval}(\mathbf{1}(k \in C) - \phi_{kt}) P(C|\mathbf{z}_{\cdot t}, \mathbf{x}_{it}; \psi)$$

Hence the impact of a price change on demand at given utility levels. When estimating the impact of prices on demand at given lending standards, this is the first term (*"shift in choice set"*) in equation 25:

$$\begin{aligned} & \frac{\partial}{\partial \log(p_{kt})} \ D(j,t|\boldsymbol{\delta}_{\cdot t}, \mathbf{z}_{\cdot t}; \psi)|_{fixed\, utility} \\ = \quad & -\sum_{C \in \mathbb{C}} \int_X \alpha_{approval} \left(\mathbf{1}(k \in C) - \phi_{kt} \right) P(C|\mathbf{z}_{\cdot t}, \mathbf{x}_{it}; \psi) \cdot D(j,t|\boldsymbol{\delta}_{\cdot t}, \mathbf{z}_{\cdot t}, \mathbf{x}_{it}, C_{it}) f(\mathbf{x}_{it}) d\mathbf{x}_{it} \end{aligned}$$

Effect of prices on neighborhood choice conditional on a choice set.

We then turn to the second term, i.e. the impact of prices on utility, thus to the impact of prices on households' choice *conditional on their initial choice set.* The derivative of conditional demand $D(j,t)$ w.r.t. its own price p_j is $\alpha_{utility} \cdot D(j,t) \cdot (1 - D(j,t))$. The coefficient α is the coefficient of log price in utility. Thus the second term (*"shift in utility"*) in equation 25:

$$\begin{aligned} \frac{\partial}{\partial \log(p_j)} D(j,t|\boldsymbol{\delta}_{\cdot t}, \mathbf{z}_{\cdot t}; \psi) \quad = \quad & -\sum_{C \in \mathbf{C}} \int_X \alpha(\mathbf{x}_{it}, \tilde{\boldsymbol{\beta}}_{it}) \cdot P(C|\mathbf{z}_{\cdot t}, \mathbf{x}_{it}; \psi) \\ & \cdot D(j,t|\boldsymbol{\delta}_{\cdot t}, \mathbf{z}_{\cdot t}, \mathbf{x}_{it}, C) \cdot (1 - D(j,t|\boldsymbol{\delta}_{\cdot t}, \mathbf{z}_{\cdot t}, \mathbf{x}_{it}, C)) f(\mathbf{x}_{it}) d\mathbf{x}_{it} \end{aligned}$$

We apply this same logic to calculate the impact of the price of another neighborhood on its own demand, and thus obtain cross-price elasticities.

Part II

Supplementary Material

4 Comparative Statics and the Shift in Unobservable Fixed Effects

The paper's Section 5.1 (iv) estimates the impact of a shift in lending standards on house prices. While the comparative statics exercise focuses on changes in the coefficients on observable neighborhood characteristics, unobservable determinants of lending standards, i.e. of mortgage approval probabilities, also shifted during the period of analysis, i.e. 2000-2006.

The first subsection below shows that the price impacts of the change in coefficient and the price impact of a shift in unobservables sum up to provide the total price shift. The next subsection then proceeds to estimate the shift in unobservables and check whether such shift can lower the estimated compression of the price distribution.

4.1 Formal Analysis

Formally, the probability of approval for a mortgage application is modeled as a function of household characteristics $\mathbf{x}_{it}$, neighborhood characteristics $\mathbf{z}_{jt}$, and an unobservable component ξ_{jt}:

$$approves^*_{ijt} = \mathbf{x}_{it}\boldsymbol{\psi} + \mathbf{z}'_{jt}\boldsymbol{\gamma} + \mathbf{x}'_{it}\Psi\mathbf{z}_{jt} + \varsigma_{jt} + e_{ijt} \tag{14}$$

where the latent variable $approves^*_{ijt}$ is lenders' net benefit of originating a mortgage in neighborhood j for household i in year t. The mortgage application is approved whenever $approves^*_{ijt} > 0$. With e_{ijt} extreme-value distributed, the probability of approval is thus:

$$\phi_{ijt} = \Lambda(\mathbf{x}_{it}\boldsymbol{\psi} + \mathbf{z}'_{jt}\boldsymbol{\gamma} + \mathbf{x}'_{it}\Psi\mathbf{z}_{jt} + \varsigma_{jt}) \tag{15}$$

In the paper the unobservable fixed effect ς_{jt} is included as part of a set of coefficients on neighborhood $\times$ year indicator variables included in $\mathbf{z}_{jt}$. This section makes it explicit as the term ς_{jt}.

The impact of a shift in the unobservable term ς_{jt} on the choice set probability is:

$$\frac{\partial \phi_{ijt}}{\partial \varsigma_{jt}} = \phi_{ijt}(1-\phi_{ijt}), \quad \text{and} \quad \frac{\partial P(C)}{\partial \varsigma_{jt}} = \left(\mathbf{1}(j \in C) - \phi_{ijt}\right) P(C) \tag{16}$$

Thus the impact of the shift on the demand for neighborhood j, as

$$D_{ijt} = \sum_{C \subset \mathbb{J}, j \in C} P(C) \cdot P(j|i,t,C), \quad \text{then} \quad \frac{\partial D_{ijt}}{\partial \varsigma_{jt}} = (1-\phi_{ijt}) D_{ijt}, \tag{17}$$

which is the impact of a shift in unobservable ς_{jt} on the demand for neighborhood j. Finally, the impact of a shift in unobservable ς_{kt} of another neighborhood, $k \neq j$ on the demand for neighborhood j is:

$$\begin{aligned} \frac{\partial D_{ijt}}{\partial \varsigma_{kt}} = & \sum_{C \subset \mathbb{J}, k \in C} (1-\phi_{ikt}) P(C) P(j|i,t,C) \\ & - \sum_{C \subset \mathbb{J}, k \notin C} \phi_{ikt} P(C) P(j|i,t,C) \end{aligned} \tag{18}$$

Together (17) and (18) determine the impact of a shift in all unobservable effects $d\boldsymbol{\varsigma_t}$ on the shift in the J-demand vector $d\mathbf{D}_{it}$ for all households i in year t.

The general-equilibrium price shift as a consequence of the shift in unobservable fixed effects follows:

$$\frac{d \log \mathbf{p}}{d\boldsymbol{\varsigma}} = \left(\frac{\partial \mathbf{D}}{\partial \log \mathbf{p}}\right)^{-1} \left(-\frac{\partial \mathbf{D}}{\partial \boldsymbol{\varsigma}} - \frac{\partial \mathbf{D}}{\partial \boldsymbol{v}} \frac{\partial \mathbf{v}}{\partial \boldsymbol{\varsigma}}\right) \tag{19}$$

That change in log price can be added to the change in log price due to the shift in coefficients, derived in the main body of the paper (Section (5)). The sum of both impacts on log price is simply $d \log \mathbf{p} = \frac{d \log \mathbf{p}}{d\boldsymbol{\psi}} \cdot d\boldsymbol{\psi} + \frac{d \log \mathbf{p}}{d\boldsymbol{\varsigma}} \cdot d\boldsymbol{\varsigma}$. The paper's results suggest a compression of the price distribution due to the shift in coefficients, i.e. $Corr(\frac{d \log \mathbf{p}_j}{d\boldsymbol{\psi}}, \log p_{j2000}) < 0$. The total compression is thus:

$$Corr(d \log \mathbf{p}, \log p_{j2000}) = \underbrace{Corr(\frac{d \log \mathbf{p}}{d\boldsymbol{\psi}} \cdot d\boldsymbol{\psi}, \log p_{j2000})}_{<0} + Corr(\frac{d \log \mathbf{p}}{d\boldsymbol{\varsigma}} \cdot d\boldsymbol{\varsigma}, \log p_{j2000}),$$

and the next subsection analyzes the sign of $Corr(\frac{d \log \mathbf{p}_j}{d\boldsymbol{\varsigma}} \cdot d\boldsymbol{\varsigma}, \log p_{j2000})$.

4.2 Empirical Analysis

The lower panel of Figure D presents statistics on the level and change of unobservable fixed effects ς_{jt} of the lending standards specification. The bottom table shows that, as the coefficient of log income went up and the coefficient of log price went down, the average and median fixed effect ς_{jt} went down from 2000 to 2006. The upper panel displays a scatter plot of the change in fixed effect $\Delta\varsigma$ from 2000 to 2006 and the initial level of log prices in 2000. Despite substantial variation in fixed effects, the fixed effects display no statistically significant correlation with the initial level of log prices. The coefficient of the regression is 0.012, with a standard error of 0.042.

Given the lack of correlation between log prices in 2000 and the change in fixed effects, it is unlikely that the general equilibrium shift in prices $d\mathbf{p}/d\boldsymbol{\varsigma}$ is correlated with the initial level of log prices. Therefore it is unlikely that the paper's Figure 3, which displays the compression of the price distribution, is affected by the addition of a shift in unobservable fixed effects $d\boldsymbol{\varsigma}$ between 2000 and 2006.

5 Comparative Statics with Lending Policy Shifts

Section 5.1 of the paper presents a derivation of comparative statics of lending standards. In particular, the section estimates the impact of a change in lending standards $d\boldsymbol{\psi}$ on (i) prices and (ii) neighborhood demographics. The theoretical derivation is done for a *small* change in lending standards. The impact of a *large* shift in lending standards such as the one observed between 2000 and 2006 is then estimated by extrapolating the impact of the small shift around the initial equilibrium.

While convenient, such a method does not account for the shift in lending standards *driven by the equilibrium change*, and in particular driven by equilibrium price and demographic changes. In other words, the method of small shifts $d\boldsymbol{\psi}$ considers lending policies invariant.

This section addresses this issue by considering the estimation of the impact of a large change in lending standards $\Delta\boldsymbol{\psi} = \boldsymbol{\psi}_{2006} - \boldsymbol{\psi}_{2000}$ and the equilibrium shifts for each small shift along the path from the 2006 lending standards coefficients $\boldsymbol{\psi}_{2000}$ to $\boldsymbol{\psi}_{2006}$.

5.1 Policy Shifts

A *large* change in lending standards $\Delta\boldsymbol{\psi} = \int d\boldsymbol{\psi}$ will cause a change in equilibrium prices and neighborhood demographics that will in turn cause changes *banks' lending policies* and thus households' demand over and above the direct impact of such lending standards changes.

Total demand is a function of approval probabilities, the log price vector, and social demographics. Each of those endogenous drivers of location change smoothly along the path to the peak of the lending boom in 2006. For instance, compare the probability of approval at constant prices and the shift in the probability of approval with shifting prices. The probability of approval for household i in neighborhood j at time t is:

$$\boldsymbol{\varphi}_{ijt} = \Lambda(-\alpha \log(p_{jt}) + \mathbf{x}'_{it}\boldsymbol{\xi} + \mathbf{z}'_{jt}\gamma + \mathbf{x}'_{it}\boldsymbol{\Psi}\mathbf{z}_{jt}) \tag{20}$$

The shift in the probability of approval from $\boldsymbol{\psi}_{2000}$ to $\boldsymbol{\psi}_{2006}$ is:

$$\begin{aligned}\Delta\boldsymbol{\varphi} = \int & \Lambda'(-\alpha \log(\mathbf{p}_{jt}) + \mathbf{x}'_{it}\boldsymbol{\xi} + \mathbf{z}'_{jt}\gamma + \mathbf{x}'_{it}\boldsymbol{\Psi}\mathbf{z}_{jt}) \\ & \cdot \left[-\frac{d\alpha}{d\boldsymbol{\psi}} \log(\mathbf{p}_{jt}) + \mathbf{x}'_{it}\frac{d\boldsymbol{\xi}}{d\boldsymbol{\psi}} + \mathbf{z}'_{jt}\frac{d\gamma}{d\boldsymbol{\psi}} + \mathbf{x}'_{it}\frac{d\boldsymbol{\Psi}}{d\boldsymbol{\psi}}\mathbf{z}_{jt} \right. \\ & \left. \qquad -\alpha\frac{d\log(\mathbf{p}_{jt})}{d\boldsymbol{\psi}} + \frac{d\mathbf{z}'_{jt}}{d\boldsymbol{\psi}}\gamma + \mathbf{x}'_{it}\boldsymbol{\Psi}\frac{d\mathbf{z}_{jt}}{d\boldsymbol{\psi}} \right] d\boldsymbol{\psi}\end{aligned} \tag{21}$$

which incorporates the shift in lending policies due to changes in lending standards (first term of the factor, on the second line) and the shift in lending policies due to the shift in market equilibrium (second term of the factor, third line of the expression). As described in the paper, amenities $\mathbf{z}_{jt}$ include the demographic composition $\mathbf{v}_{jt}$ of the neighborhood.

5.2 Comparative Statics with Policy Shifts

To express the impact of this large change on the price of housing and on demographics, we start by noticing that the large changes are the continuous sums of small changes in prices and demographics:

$$\Delta \log \mathbf{p} = \int \frac{d \log \mathbf{p}}{d\boldsymbol{\psi}} d\boldsymbol{\psi}, \qquad \Delta\mathbf{v} = \int \frac{d\mathbf{v}}{d\boldsymbol{\psi}} d\boldsymbol{\psi} \tag{22}$$

where each infinitesimal shift in lending standards $d\boldsymbol{\psi}$ shifts the equilibrium vector of prices $\mathbf{p}$ and social composition $\mathbf{v}$. At each step of the relaxation of lending standards, for each infinitesimal change in lending standards $d\boldsymbol{\psi}$ the *price change* $d\log\mathbf{p}$ causes a shift in lending standards through a shift in approval probabilities $d\boldsymbol{\varphi}$.

To each lending standard coefficient $\boldsymbol{\psi} \in [\boldsymbol{\psi}_{2000}, \boldsymbol{\psi}_{2006}]$ corresponds an equilibrium.[2] Write $(\mathbf{p}^*(\boldsymbol{\psi}), \mathbf{v}^*(\boldsymbol{\psi}))$ such an equilibrium. The paper provides $d\log p/d\boldsymbol{\psi}$ *starting from* a given equilibrium.

We can expand the integral of infinitesimal shifts to find the total change in prices $\Delta\log\mathbf{p}$ from 2000 to 2006:

$$\Delta\log\mathbf{p} = -\int_{\boldsymbol{\psi}_{2000}}^{\boldsymbol{\psi}_{2006}} \left[\frac{\partial\mathbf{D}}{\partial\log\mathbf{p}}\right]^{-1} \left[\frac{\partial\mathbf{D}}{\partial\boldsymbol{\psi}} + \frac{\partial\mathbf{D}}{\partial\mathbf{v}}\cdot\frac{\partial\mathbf{v}}{\partial\boldsymbol{\psi}}\right] d\boldsymbol{\psi} \tag{23}$$

And similarly for the other equilibrium variables, the change in neighborhood demographics $\Delta\mathbf{v}$. In contrast, Section 5.1 of the paper estimated $\Delta\log\mathbf{p}$ by extrapolating:

$$\Delta\log\mathbf{p} \simeq -\left[\frac{\partial\mathbf{D}}{\partial\log\mathbf{p}}\right]^{-1} \left[\frac{\partial\mathbf{D}}{\partial\boldsymbol{\psi}} + \frac{\partial\mathbf{D}}{\partial\mathbf{v}}\cdot\frac{\partial\mathbf{v}}{\partial\boldsymbol{\psi}}\right]\cdot\Delta\boldsymbol{\psi} \tag{24}$$

There are significant differences between (23) and (24). In (23), both the demand sensitivity matrix $\frac{\partial\mathbf{D}}{\partial\log\mathbf{p}}$ and the partial equilibrium impacts $\frac{\partial\mathbf{D}}{\partial\boldsymbol{\psi}}$ change along the path from $\boldsymbol{\psi}_{2000}$ to $\boldsymbol{\psi}_{2006}$. In (24) in contrast, both of those elements are constant.

Another difference between (23) and (24) lies in the fact that there are multiple ways for lending standards to evolve from $\boldsymbol{\psi}(t = 2000)$ to $\boldsymbol{\psi}(t = 2006)$. For instance, a different timing of the relaxation of the loan-to-income constraint and of the increase in the volume of loans matters in the price effects.

A potentially difficulty in the estimation of (23) is that the price sensitivity $\frac{\partial\mathbf{D}}{\partial\log\mathbf{p}}$ may not be invertible at some point on the path from lending standards $\boldsymbol{\psi}_{2000}$ to lending standards $\boldsymbol{\psi}_{2006}$.

Indeed, on the path $\boldsymbol{\psi} : \boldsymbol{\psi}_{2000} \to \boldsymbol{\psi}_{2006}$ the economy could reach an equilibrium $(\mathbf{p}^*, \mathbf{v}^*)$ which is not locally unique, i.e. a saddle point where the demand derivative matrix $\frac{\partial\mathbf{D}}{\partial\log\mathbf{p}}$ is not invertible. Proposition 2 of this Appendix has shown that the vector of structural parameters (including $\boldsymbol{\psi}$) such that there are non-locally unique equilibria is of measure zero. Nevertheless at a given point $\boldsymbol{\psi}$ along the path, the demand derivative matrix might not be statistically different from a singular

[2] Potential singularity and issues related to equilibrium multiplicity are discussed below.

matrix. In such a case the path $\boldsymbol{\psi}_{2000} \to \boldsymbol{\psi}_{2006}$ could move the city along two possible paths.

Another possibility is that a small shift $\boldsymbol{\psi} \to \boldsymbol{\psi} + d\boldsymbol{\psi}$ can cause a large change in equilibrium prices $\mathbf{p}$ and neighborhood demographics $\mathbf{v}$ ('tipping'). This would happen for instance if, at some $\boldsymbol{\psi}$, two equilibria coincide and then separate, as in a 3-equilibrium Schelling model.[3] At such a point $\boldsymbol{\psi}$, the derivative of demand w.r.t. prices $\partial \mathbf{D}/\partial \log \mathbf{p}$ would not be invertible. We test empirically that the rank of $\frac{\partial \mathbf{D}}{\partial \log \mathbf{p}}$ is full on the points $\boldsymbol{\psi} \in [\boldsymbol{\psi}_{2000}, \boldsymbol{\psi}_{2006}]$ that we consider in the approximation below.

5.3 Empirical Analysis

It is typically difficult to obtain the change in price $\Delta \log \mathbf{p}$ in (23), as (i) there is no closed form expression, and (ii) the change in lending standards is continuous from $\boldsymbol{\psi}_{2000}$ to $\boldsymbol{\psi}_{2006}$.

We address these two challenges in two ways. First we estimate lending standards changes each year from 2000 to 2006 to obtain the lending standards coefficients $\boldsymbol{\psi}_{2000}, \boldsymbol{\psi}_{2001}, \ldots, \boldsymbol{\psi}_{2006}$. We then interpolate coefficients linearly for monthly changes. This yields a sequence of lending standards coefficients:

$$(\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, \ldots, \boldsymbol{\psi}_K) \text{ with } \boldsymbol{\psi}_1 = \boldsymbol{\psi}_{2000} \text{ and } \boldsymbol{\psi}_K = \boldsymbol{\psi}_{2006} \tag{25}$$

We use the finite-difference approximation to get an estimate $\widehat{\Delta \log \mathbf{p}}$ of $\Delta \log \mathbf{p}$, such that $\widehat{\Delta \log \mathbf{p}} = \sum_{k=0}^{K} (\Delta \log \mathbf{p})_k$, and similarly for $\widehat{\Delta \mathbf{v}}$ as the sum of $(\Delta \mathbf{v})_k$; each $(\Delta \log \mathbf{p})_k$ and $(\Delta \mathbf{v})_k$ is obtained under the assumption of policy invariance from $\boldsymbol{\psi}_k$ to $\boldsymbol{\psi}_{k+1}$, i.e. following (24).

The equilibrium shifts (prices and demographics) at each step k are estimated as follows.

Estimation Procedure

The estimation procedure for neighborhood log price changes $(\Delta \log \mathbf{p})_k$ and demographic changes $(\Delta \mathbf{v})_k$ is as follows. For $k = 1, 2, \ldots, K$:

1. We find the partial equilibrium shift in demand in response to the lending standards change $\partial \mathbf{D}/\partial \boldsymbol{\psi} \cdot \Delta \boldsymbol{\psi}_k$.

[3] Consider for instance the simple Schelling model whose equilibrium is a solution to $\theta = F(\frac{\theta-\mu}{\sigma})$ where μ is the mean of the distribution of same-race preferences, σ its standard deviation and θ the equilibrium point. There are up to three equilibria. A shift in μ leads to a shift $d\theta$ such that $(\frac{1}{\sigma} f(\frac{\theta-\mu}{\sigma}) - 1)\frac{d\theta}{d\mu} = \frac{1}{\sigma} f(\frac{\theta-\mu}{\sigma})$. At a point where there are exactly 2 equilibria (2 out of the 3 equilibria coincide at the same θ), then $\frac{1}{\sigma} f(\frac{\theta-\mu}{\sigma}) - 1 = 0$ and such implicit function theorem cannot be applied. Similarly that could happen in the paper if $\partial \mathbf{D}/\partial \log \mathbf{p}$ is not invertible.

2. The Slutsky matrix ($\partial \mathbf{D}/\partial \mathbf{p}$) and the sensitivity to social composition ($\partial \mathbf{D}/\partial \mathbf{v}$) are computed at step k.

3. We solve for the shift in prices $d\mathbf{p}/d\boldsymbol{\psi} \cdot \Delta\boldsymbol{\psi}_k$ from equilibrium k to equilibrium $k+1$. This affects base utilities δ_{jt}, which are updated to reflect such log price changes.

4. We solve for the shift in demographics $d\mathbf{v}/d\boldsymbol{\psi} \cdot \Delta\boldsymbol{\psi}_k$ from equilibrium k to equilibrium $k+1$. This affects base utilities δ_{jt}, which are updated to reflect such demographic changes.

5. Finally the shift in log prices affects choice set probabilities through the shift in approval probabilities $d\boldsymbol{\varphi}/d\boldsymbol{\psi} \cdot \Delta\boldsymbol{\psi}_k$.

Then, the total log price change and the total demographic change from 2000 to 2006 is estimated as:

$$\widehat{\Delta \log \mathbf{p}} = \sum_{k=1}^{K} (\Delta \log \mathbf{p})_k \quad \text{and} \quad \widehat{\Delta \mathbf{v}} = \sum_{k=1}^{K} (\Delta \mathbf{v})_k \tag{26}$$

This approximation to the estimation of the continuous integral (22) is akin to the finite-difference method (Smith 1985).

Empirical Results

The results of such an approach are described in Figure E, panels (a) and (b). Panel (a)'s vertical axis is the log price change from 2000 to 2006, while the horizontal axis is the log price change from 2000 to 2001 extrapolated in a linear fashion to 2000–2006. The latter log price change is the one adopted in the main body of the paper, while the former log price change is the one estimated in equation (26). Thus panel (a) compares the log price changes obtained with the linear extrapolation method to the equilibrium shifts method. The green line is the 45 degree line. The panel suggests that the equilibrium shifts method yield slightly lower log price increases than the linear extrapolation method. This is likely intuitive: indeed, the relaxation of lending standards gradually leads to price increases that more quickly offset the increase in demand. The second panel, panel (b), suggests that the compression of the price distribution is a fact that is robust to using the equilibrium shift method described in this section.

6 Consumption, Mortgage Payments, and Credit Constraints

6.1 Location Choice in a Mortgage Payment-Amenity Trade-Off, with Borrowing Constraints

We adopt here an approach in the spirit of Berry, Levinsohn, and Pakes (1995), applied to the specific case of housing. Household i derives utility $U(\mathbf{h}_j, c_{ij}, \varepsilon_{ij})$ from a vector of amenities $\mathbf{h}_j$, consumption c_{ij}, and unobservables ε_{ij}. The vector of amenities $\mathbf{h}_j$ includes the characteristics of the housing stock in j, e.g. housing size. Utility is Cobb Douglas, with α the preference for non-housing consumption:

$$U(\mathbf{h}_j, c_{ij}, \varepsilon_{ij}) = c_{ij}^{\alpha} \cdot G(\mathbf{h}_j) \cdot e^{\varepsilon_{ij}} \tag{27}$$

Unobservables ε_{ij} are introduced as in BLP. Here the budget constraint splits household income Y_i into neighborhood- and household-specific mortgage payments M_{ij} and consumption:

$$c_{ij} + m_{ij} \leq y_i \tag{28}$$

Households choose a trade-off between neighborhood amenities and consumption, i.e. maximize $U(\mathbf{h}_j, y_i - m_{ij}, \varepsilon_{ij})$ or, equivalently, maximize:

$$\log U_{ij} = \alpha \log(y_i - m_{ij}) + \log G(\mathbf{h}_j) + \varepsilon_{ij} \tag{29}$$

when ε_{ij} is extreme-value distributed, location choice is according to a McFadden (1974) discrete-choice model.

Data described in the next section provides us with the relationship between mortgage payments, price, and household characteristics. In the simple case where the mortgage is a fixed-rate perpetuity with interest rate r, and the loan-to-value ratio is λ, then $m = r\lambda p$. Instead of assuming a specific structure of the mortgage (fixed or variable rate, teaser period, interest-only mortgages), we use the estimation of mortgage payments M from the 1pct Census and thus potentially allow for a variety of mortgage contracts.

The term $\alpha \log(y_i - m_{ij})$ is both household- and neighborhood-specific. In contrast, base utility δ_{jt} is the utility derived by an average household of the metropolitan area from living in the specific neighborhood j in year t. When the joint distribution of household income y and mortgage payments in neighborhood j has distribution $f(y, m_j; t)$ in year t, base utility is:

$$E(\log U_{ij}|j) = \alpha \int \log(y - m_j) f(y, m_j; t) dy dm_j + \log G(\mathbf{h}_j) \tag{30}$$

The next section estimates such joint distribution of mortgage payments and income in each neighborhood.

Utility is thus split into a base utility term and an individual-specific term as follows:

$$\begin{aligned} \log U_{ijt} &= \delta_{jt} + \alpha_h \left(\log(y_{it} - m_{ijt}) - \int \log(y - m_j) f(y, m_j; t) dy dm_j \right) & (31) \\ & \quad + \mathbf{x}_{it} \Omega \mathbf{z}_{jt} + \tilde{\boldsymbol{\beta}} \mathbf{z}_{jt} + \varepsilon_{ijt} & (32) \end{aligned}$$

with the base utility term as:

$$\delta_{jt} = \alpha \int \log(y - m_j) f(y, m_j; t) dy dm_j + \mathbf{z}_{jt} \boldsymbol{\gamma} + \xi_j + \zeta_{jt} \tag{33}$$

In the specific case where household income and mortgage payments in neighborhood j are independent, e.g. if the LTV and interest rates are independent of income, and with a log normal distribution for income of mean μ and standard deviation σ, then $\int \log(y - m_j) f(y) dy = \log(e^{\mu+\sigma} - m_j)$. [4]

6.2 Estimating Mortgage Payments and Consumption

The micro Census 1% sample provides a set of N observations indexed by i, $(\mathbf{x}_{it}, m_{ij(i)t}, \mathbf{z}_{j(i)t})$. In those observations, $\mathbf{x}_{it}$ are the household characteristics, $m_{ij(i)t}$ is the mortgage payment for household i in his neighborhood of residence $j(i)$, and $\mathbf{z}_{jt}$ are neighborhood characteristics. Mortgage payments m_{ijt} are observed for the neighborhood $j(i)$ of household i, but we are interested in m_{ijt} for all potential neighborhoods $j = 1, 2, \ldots, J$ that i can choose from.

When $N \to \infty$ such Census 1% data set can provide a consistent estimate of the conditional distribution of mortgage payments $f(m|\mathbf{x}, \mathbf{z})$ given household and neighborhood characteristics $(\mathbf{x}, \mathbf{z})$, as long as unobservable confounders do not bias the estimate of the distribution $f(\cdot|\cdot)$.

[4] This is the case considered in Berry, Levinsohn, and Pakes (1995) as the price of a car can be reasonably assumed to be independent of buyer income. In the case of housing however, mortgage payments typically depend on household income.

We impute mortgage payments in the dataset as follows. For each household i of the paper's micro sample, the expected log mortgage payment in each neighborhood $j = 1, 2, \ldots, J$ is assumed to be the linear combination of household and neighborhood characteristics:

$$\log m_{ijt} = \mathbf{x}_{it}\boldsymbol{\beta} + \mathbf{z}_{jt}\boldsymbol{\gamma} + \mathbf{x}_{it}\Psi\mathbf{z}_{jt} + \varepsilon_{ijt} \tag{34}$$

Estimates of $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$, Ψ, and σ_ε^2 are obtained using the Census 1 pct files.

The average log consumption in neighborhood j (where the average is taken over all potential households of the metropolitan area) is then estimated as:

$$\hat{E}(\log(y_{it} - m_{ijt})) = \frac{1}{N}\sum_{i=1}^{N}\log(y_{it} - \exp(\mathbf{x}_{it}\hat{\boldsymbol{\beta}} + \mathbf{z}_{jt}\hat{\boldsymbol{\gamma}} + \mathbf{x}_{it}\hat{\Psi}\mathbf{z}_{jt} + \widehat{\varepsilon_{ijt}})) \tag{35}$$

where $\boldsymbol{z}_{jt}$ is fixed and both x_{it} and y_{it} are draws from the metropolitan area wide distribution. ε_{ijt} is taken from a normal distribution with mean zero and standard deviation estimated in regression (34).

6.3 Empirical Analysis

Table D presents the relationship between mortgage payments on one hand, and household, housing, and neighborhood characteristics on the other hand. Such regression is performed on the observed Census 1pct. The regression is saturated and includes, in addition, higher polynomial powers of the covariates to capture non-linear relationships between covariates and log mortgage payments. This enables us to impute log consumption measures in the main dataset and perform base utility regressions where the log price is replaced by the average log non-housing consumption following specification 35.

The corresponding base utility regression is presented in Table E. Signs of the impact of amenities on base utility are similar to the paper's results, and higher log non-housing consumption leads to higher base utility.

7 Partial Equilibrium Relaxation of Borrowing Constraints for Targeted Households

The paper focuses on the impact of relaxing borrowing constraints in general equilibrium, i.e. for all households simultaneously, according to the marginal shifts estimated using mortgage application data from 2000 to 2006. Another noteworthy application of the model is to simulate the impact of relaxing borrowing constraints for *one* household, thus allowing all neighborhoods in the choice set, while avoiding price responses to such relaxation. This partial equilibrium analysis of the relaxation of borrowing constraints may be policy relevant if a policy's goal is to target a subpopulation sufficiently small to avoid price impacts of the policy.

This partial equilibrium approach takes preference parameters as they are estimated in the paper's Section 3 accounting for borrowing constraints. We then estimate the unconstrained demand for each household i, taking all other households' demands as fixed (constrained). This yields:

$$\Delta \log D_{ijt} = \log D_{ijt}(C = \{1, 2, \ldots, J\}) - \log D_{ijt} \tag{36}$$

This Appendix's Figure G (resp., Figure H) presents such average changes for black households (resp., poor households), i.e. $E(\Delta \log D_{ijt} | i \in black)$ (resp., $E(\Delta \log D_{ijt} | i \in poor)$).

Figure G suggest that black households' demand for neighborhoods would decline for most neighborhoods except for neighborhoods with less than 3% of black households. This, despite the strong estimated preference for black neighbors, is the outcome of the correlation between race and the quality of amenities. Hence, black households would move to better-amenity neighborhoods *despite* the negative impact of more white, Hispanic, or asian neighborhoods on utility. Panel (c) suggests that black households would move to more central locations, and panel (d) suggests that households would move either to smaller units (closer to the CBD) or to larger units (further away from the CBD), implying that the tradeoff between location and size plays out differently for different households. Such results also suggest that the more substantial segregation effects obtained in general equilibrium are due to the competition of all racial groups for similar amenities.

Figure H presents similar results for poor households. Given the significant correlation between poverty and race, panels (a), (b), and (c) depict similar dynamics as for black households. The pool of poor households is, however, racially diverse, implying that the impact of the relaxation of the choice set on demand pushes demand even more towards non-black neighborhoods. Panel (d) shows that the relaxation of the choice set would unambiguously lead to higher demand for higher priced neighborhoods, and to a decline in demand for lower-priced neighborhoods.

8 Comparing Neighborhood-Level and City-Level Housing Supply Elasticities

The paper estimates neighborhood-level supply elasticities ranging from 0.05 to 0.35, with a strong, positive, and statistically significant correlation between neighborhood-level supply elasticity and the log distance to the Central Business District. Saiz (2010) finds that, at the metro-level, the elasticity of housing supply is 0.66 in San Francisco MSA. Such metro-level supply elasticity is thus higher than the average of the neighborhood-level estimates. Aggregation of micro parameters to the macro level is a concern in other fields such as trade and macroeconomics. In particular, Imbs & Mejean's (2015) shows that, when micro-level elasticities have both an observed and an unobserved component, the average of micro-elasticities does not typically match macro-estimated elasticities. Similarly, Swamy (1970) shows that micro-level unobservable coefficient heterogeneity can lead to upward biases in aggregate regressions.

Write for instance, in this paper's notation, the relationship between log supplies and log prices:

$$\Delta \log s_{jt} = c_j + \eta_j^s \cdot \Delta \log p_{jt} + \varepsilon_{jt} \tag{37}$$

where j indexes neighborhoods, c_j is a neighborhood-specific constant, η_j^s is neighborhood j's supply elasticity, and ε_{jt} is a residual that is orthogonal to $\log p_{jt}$. Then split elasticity η_j^s into an observable and an unobservable component:

$$\eta_j^s = \eta^s + \omega_j \tag{38}$$

where η^s is here a constant, but can be made a function of observable neighborhood characteristics $\mathbf{z}_j$. ω_j is an unobservable residual orthogonal to $\mathbf{z}_j$. Then, the aggregation of 37 to the metro-area level yields:

$$\Delta \log S_t = c + \eta^s \cdot \Delta \log P_t + e_t \tag{39}$$

which is the metro-level regression that yields metro-level elasticity. In such a metro-level specification, $\Delta \log S_t = \sum_{j=1}^{S} \frac{s_j}{S} \cdot \Delta \log s_j$, $\Delta \log P_t = \sum_{j=1}^{J} \frac{s_j}{S} \cdot \log p_{jt}$, and $e_t = \sum_{j=1}^{J} \frac{s_j}{S} \cdot (\omega_j \Delta \log p_{jt} + \varepsilon_{jt})$.

The aggregation of the specification at the macro-level thus leads to a specification 39 where the sign of the bias depends on the sign of the covariance between the residual e_t and the change in log price $\Delta \log P_t$:

$$Cov(e_t, \Delta \log P_t) = \sum_{j=1}^{S} \frac{s_j}{S} E(\omega_j (\Delta \log p_{jt})^2),$$

whose value is 2.06 in the dataset: neighborhoods with the largest swings in log prices are also the most elastic neighborhoods. Such positive covariance typically leads to an upward bias in the estimation of the average supply elasticity η^s when using city-level data. This is consistent with the result whereby Saiz's (2010) elasticity is substantially higher than the average of micro-level elasticities.

References

Berry, S., Levinsohn, J. & Pakes, A. (1995), 'Automobile prices in market equilibrium', *Econometrica: Journal of the Econometric Society* pp. 841–890.

Berry, S. T. (1994), 'Estimating discrete-choice models of product differentiation', *The RAND Journal of Economics* pp. 242–262.

Dubé, J.-P., Fox, J. T. & Su, C.-L. (2012), 'Improving the numerical performance of static and dynamic aggregate discrete choice random coefficients demand estimation', *Econometrica* **80**(5), 2231–2267.

Hansen, L. P. (1982), 'Large sample properties of generalized method of moments estimators', *Econometrica: Journal of the Econometric Society* pp. 1029–1054.

Heckman, J. J., Urzua, S. & Vytlacil, E. (2006), 'Understanding instrumental variables in models with essential heterogeneity', *The Review of Economics and Statistics* **88**(3), 389–432.

Imbs, J. & Mejean, I. (2015), 'Elasticity optimism', *American Economic Journal: Macroeconomics* **7**(3), 43–83.

Mas-Colell, A., Whinston, M. D., Green, J. R. et al. (1995), *Microeconomic theory*, Vol. 1, Oxford university press New York.

Saiz, A. (2010), 'The geographic determinants of housing supply', *The Quarterly Journal of Economics* **125**(3), 1253–1296.

Sard, A. (1942), 'The measure of the critical values of differentiable maps', *Bull. Amer. Math. Soc* **48**(12), 883–890.

Shao, J. & Tu, D. (2012), *The jackknife and bootstrap*, Springer Science & Business Media.

Smith, G. D. (1985), *Numerical solution of partial differential equations: finite difference methods*, Oxford university press.

Su, C.-L. & Judd, K. L. (2012), 'Constrained optimization approaches to estimation of structural models', *Econometrica* **80**(5), 2213–2230.

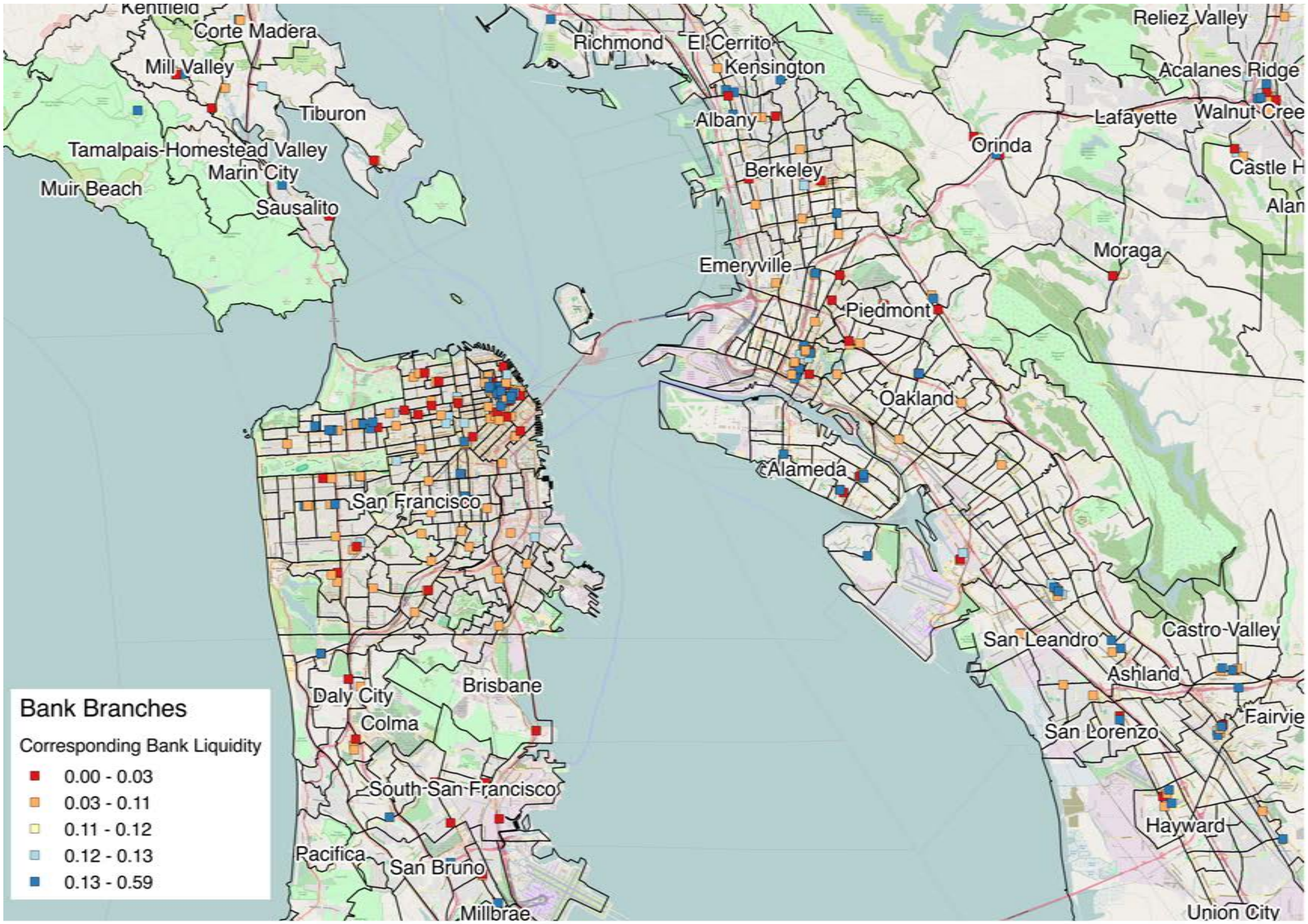
Swamy, P. A. (1970), 'Efficient inference in a random coefficient regression model', *Econometrica: Journal of the Econometric Society* pp. 311–323.

Part III

Tables and Figures

Appendix Figure A: Bank Branches and Liquidity

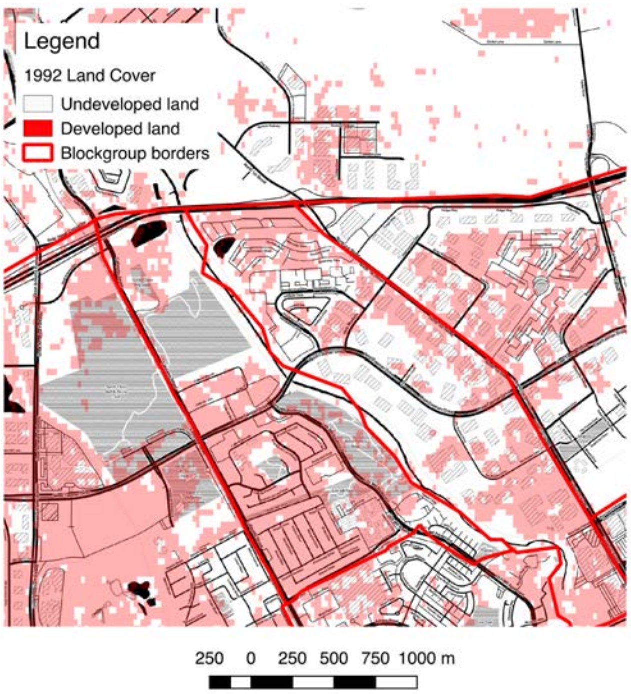
This map presents the set of bank branches from the Summary of Deposits in 1994. The map only includes bank branches whose corresponding bank is regulated by the Federal Reserve, the Office of the Comptroller of the Currency, or the Federal Deposit Insurance Corporation. The bank's liquidity level is derived from the Federal Reserve Bank of Chicago's Reports of Income and Condition (Call Reports). Color codes indicate the national bank's liquidity level.



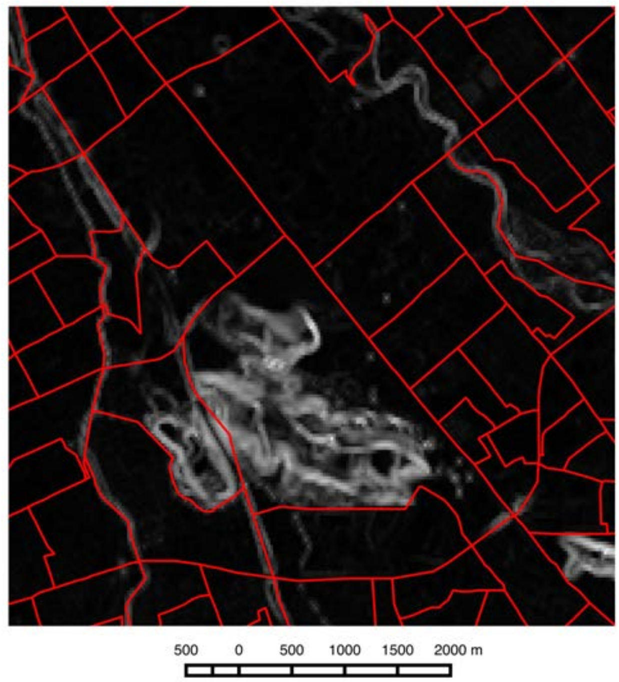
Appendix Figure B: Undeveloped Land - San Francisco Bay Area

Figures (a) and (b) illustrate the computation of our housing supply proxies in two sample areas. Figure (a) presents satellite data on developed land in 1992 for each 30m by 30m cell, for the neighborhood of Lafayette and North 1st St. in San Jose. The background street map is for 2016. Red cells are developed land. Figure (b) presents ruggedness values based on USGS satellite data on elevation, for the neighborhood of Oak Hill in San Jose. Lighter cells exhibit higher ruggedness. Figure (c) displays the distribution of housing supply elasticities $\eta_j = \hat{b} + \hat{c} \cdot Undeveloped\,Share_j + \hat{d} \cdot Ruggedness_j$.

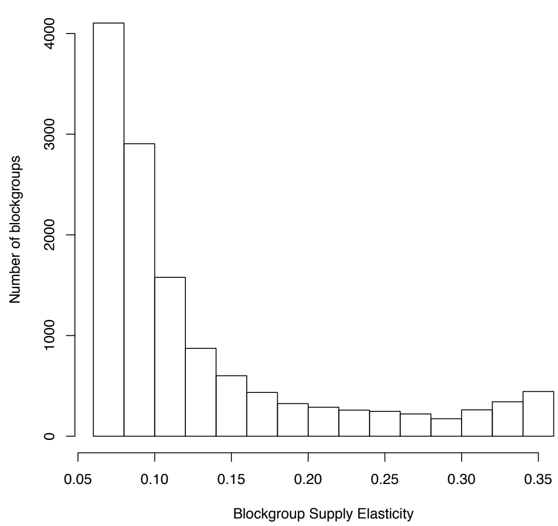
(a) Undeveloped Land – Lafayette and North 1st St.



(b) Ruggedness – Oak Hill Memorial Park

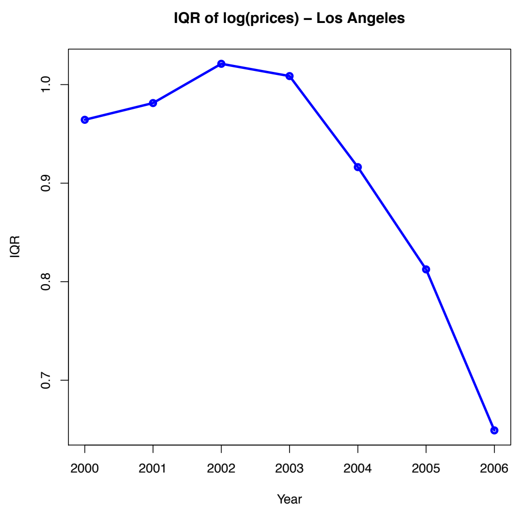


(c) Distribution of Housing Supply Elasticities

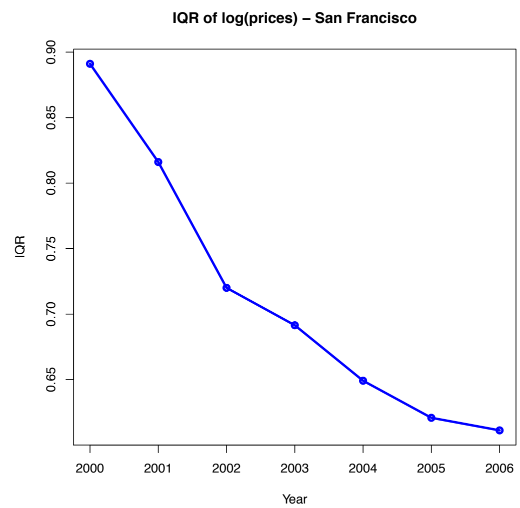


Appendix Figure C: The Compression of the Price Distribution in 4 Metropolitan Areas

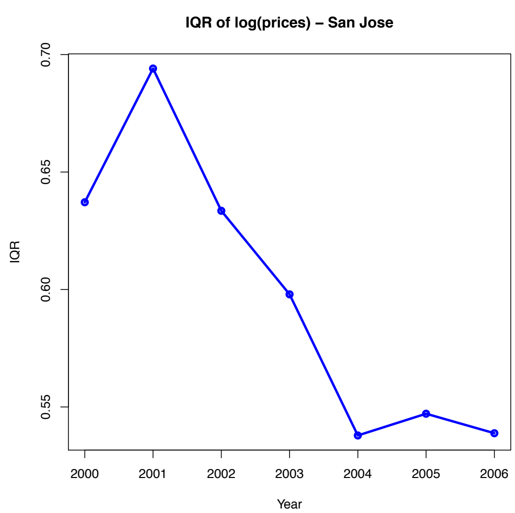
Each figure presents the Interquartile Range (IQR) of the distribution of log(price) for Los Angeles, San Francisco, San Jose, and Washington DC metropolitan areas. Individual transaction data from FNC Inc. and calculations from the authors.



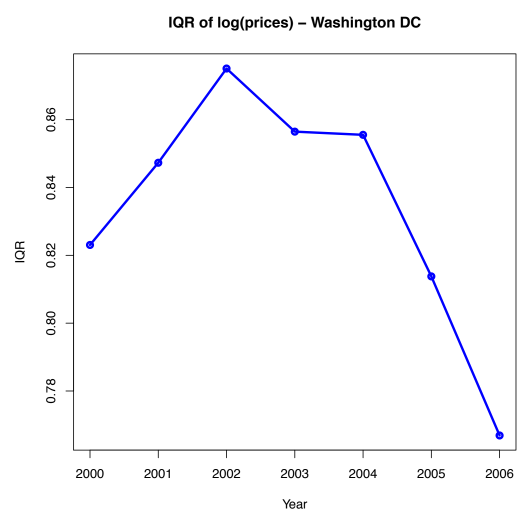
(a) Los Angeles



(b) San Francisco



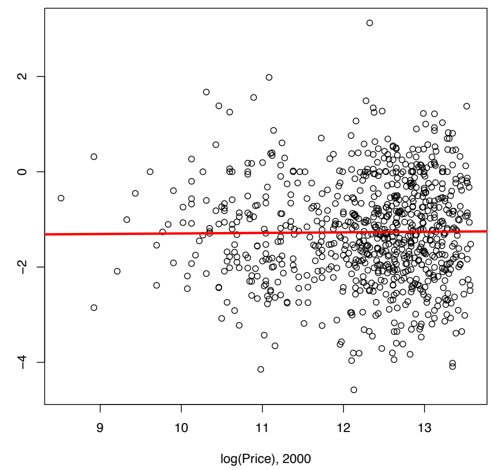
(c) San Jose



(d) Washington DC

Appendix Figure D: Shift in Unobservable Fixed Effects and the Initial Price Level

This figure presents a scatterplot, where one point is one census tract, relating the variation $\Delta\varsigma_{j,2000-2006}$*unexpected''inmath to the initial log price level* $\log(p_{jt})$



$$\Delta\varsigma_j = \underset{(0.516)}{-1.410} + \underset{(0.042)}{0.012}\Delta\log p_j + \varepsilon_j$$

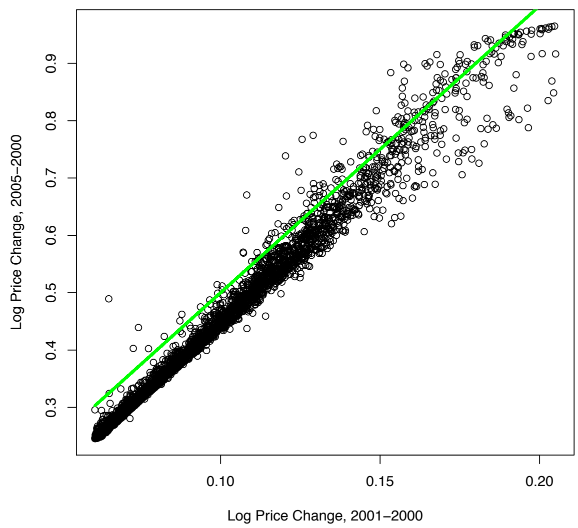
This accompanying table presents summary statistics on the unobservable fixed effects in the mortgage approval analysis.

	Mean	Median	S.D.
Fixed effect $\hat{\varsigma}$, 2000	0.000	−0.113	1.030
Fixed effect $\hat{\varsigma}$, 2006	−1.239	−1.261	0.653
Change, $\Delta\hat{\varsigma}$, 2000 to 2006	−1.239	−1.254	1.094

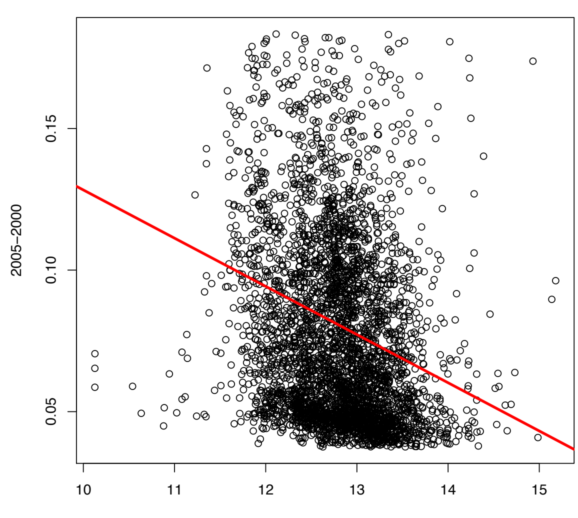
Appendix Figure E: Equilibrium Invariance vs. Equilibrium Shifts

These figures and table compare the log price change obtained using the simple extrapolation $\Delta \log p \simeq \frac{d \log \mathbf{p}}{d\psi} \Delta\psi$*, where* $\Delta\psi$ *is the 2000-2006 change in lending standards, to the log price change obtained using gradual equilibrium shifts* $\Delta \log \mathbf{p} = \int_{\psi_{2000}}^{\psi_{2006}} \frac{d \log \mathbf{p}}{d\psi} d\psi$*. In Figure (a), the bold green line is the log price change in 2000-2006 extrapolated to 2000-2006. Each black dot is a neighborhood. In Figure (b), the compression of the price distribution is obtained using the gradual equilibrium shifts. In contrast, the paper's Figure 3 presents the compression of the price distribution with extrapolation.*

(a) Log(price) Changes

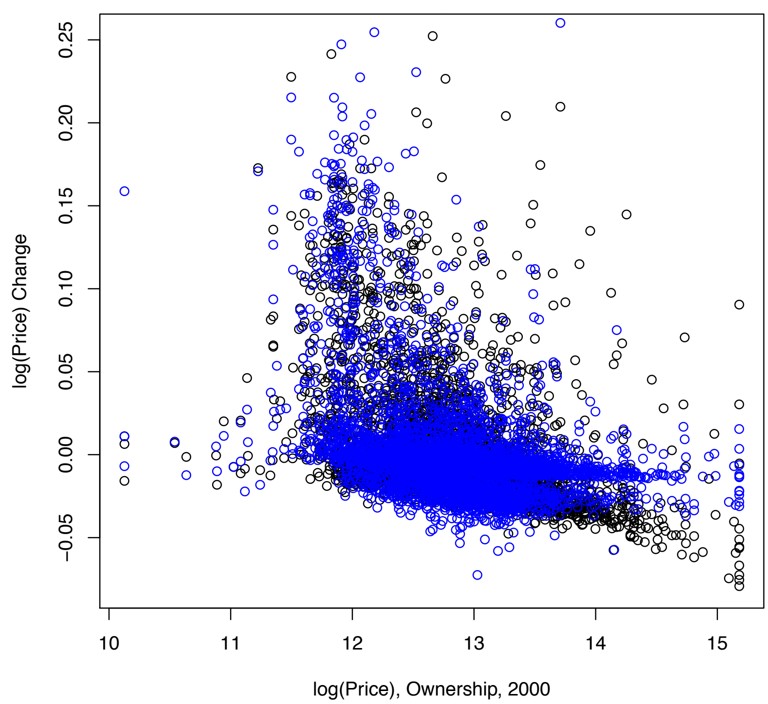


(b) Price Distribution Compression



Appendix Figure F: Tenure Choice and The Relaxation of Lending Standards

The Figure shows the general equilibrium impact of the relaxation of lending standards on prices, with the option of renting in each neighborhood. A relaxation of lending standards increases households' access to homeownership in partial equilibrium, but leads to price increases for owner-occupied units. Blue points are the general equilibrium log(price) changes in the model with both tenure and neighborhood choice, and credit constraints for access to homeownership. Black points as in the baseline model Figure 3.



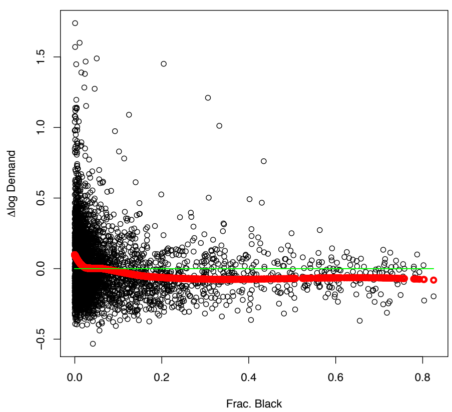
Appendix Figure G: Partial Equilibrium Impact of an Unconstrained Choice Set on Demand – Black Households

These figures depict the impact of an unconstrained choice set on household demands. Each black point is a neighborhood (blockgroup). The figures differ in the amenity or neighborhood characteristic on the horizontal axis.

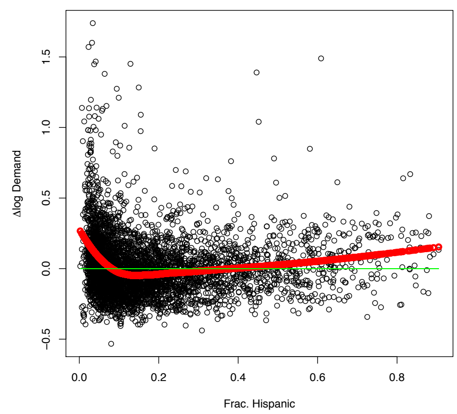
$$\Delta D = \log D(j,t|i, C = \{1,2,\ldots,J\}) - \log D(j,t|i) \tag{40}$$

This figure presents results for black households.

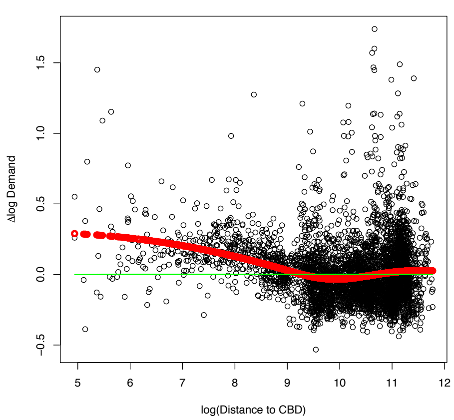
(a) By Fraction Black



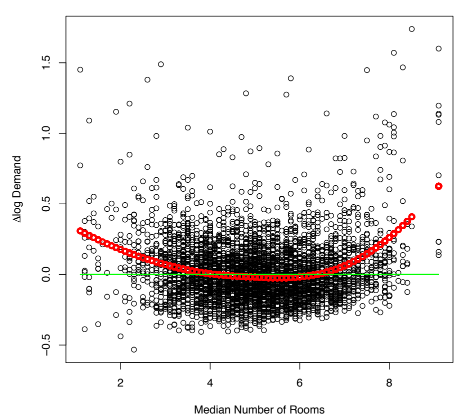
(b) By Fraction Hispanic



(c) By log(Distance) to CBD



(d) By Median Number of Rooms



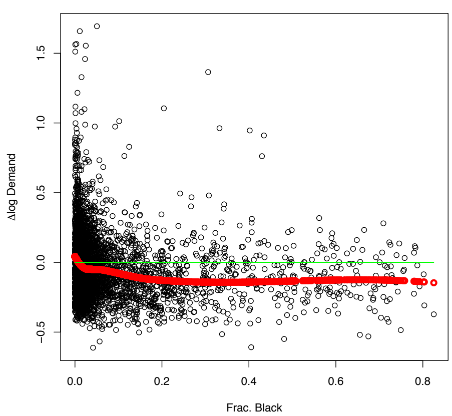
Appendix Figure H: Partial Equilibrium Impact of an Unconstrained Choice Set on Demand – Poor Households

These figures depict the impact of an unconstrained choice set on household demands. Each black point is a neighborhood (blockgroup). The figures differ in the amenity or neighborhood characteristic on the horizontal axis.

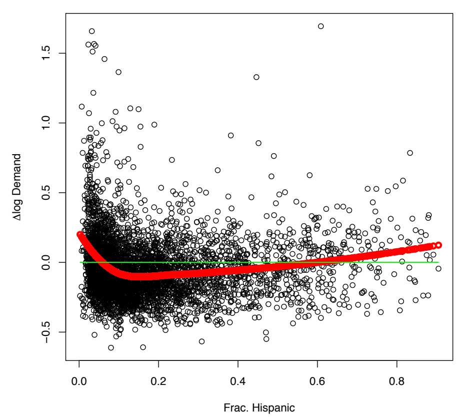
$$\Delta D = \log D(j,t|i, C = \{1, 2, \ldots, J\}) - \log D(j,t|i) \tag{41}$$

This figure presents results for poor households.

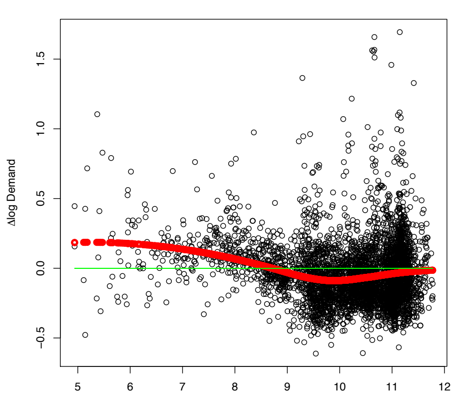
(a) By Fraction Black



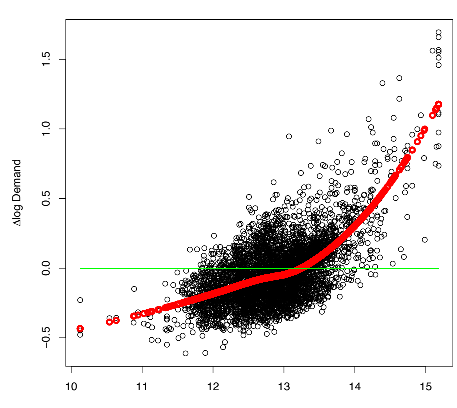
(b) By Fraction Hispanic



(c) By log(Distance) to CBD



(d) By log Price



Appendix Table A: General Equilibrium Price Changes in Model Extensions

This table presents the change in log prices between 2000-2006 as predicted by the baseline model when lending standards change as estimated in Appendix Table C, with either elastic or perfectly inelastic housing supply, with or without population growth, and with or without shifts in the racial composition of the Bay area. Population growth and shifts in the racial composition of the Bay area between 2000 and 2006 are obtained from the 2000 Census and the 2005-2009 American Community Survey.

Impact of credit	Population Growth	Racial Shifts	Supply Elasticity	1st Qu.	Median	Mean	3rd Qu.	$Cor(\log(\Delta p_{jt}), \log(p_{jt}))$
Yes	No	No	Inelastic	0.072	0.088	0.105	0.118	−0.267
Yes	No	No	Local	0.037	0.053	0.068	0.081	−0.200
Yes	No	Yes	Inelastic	0.121	0.137	0.149	0.161	−0.378
Yes	No	Yes	Local	0.063	0.080	0.092	0.106	−0.236
Yes	Yes	No	Inelastic	0.088	0.133	0.161	0.221	−0.301
Yes	Yes	No	Local	0.041	0.080	0.102	0.151	−0.250
Yes	Yes	Yes	Inelastic	0.137	0.179	0.205	0.262	−0.343
Yes	Yes	Yes	Local	0.067	0.105	0.127	0.175	−0.267
Actual Annualized Change				0.076	0.110	0.111	0.148	−0.748

1st Qu., Median, Mean, 3rd Qu: moments of the distribution of log price changes $\Delta \log(p_{jt})$.

Appendix Table B: Summary Statistics and Data Sources

The first part of this table ('Blockgroup data') presents neighborhood data from a variety of sources: Summary File 3 of the 1990 and 2000 Censuses and employment numbers from the ZIP-level County Business Patterns. The second part ('Micro Census Data') is household observations from the 1% Census microfiles in 1990, 2000, 2010 for the San Francisco and the San Jose MSAs. The third part ('Mortgage Application Data') presents mortgage application information from the Home Mortgage Disclosure Act of 1990 and 2000.

	Mean	Median	S.D.	Min	Max	Obs.
Blockgroup Data						
log(Median price)	12.757	12.737	0.688	7.601	18.258	13,154
Median number of rooms	5.156	5.062	1.254	0.074	9.100	13,154
Median age of structure	33.562	35.000	17.052	0.000	61.000	13,154
Frac. Black	0.078	0.025	0.140	0.000	0.948	13,154
Frac. Hispanic	0.278	0.170	0.289	0.002	1.818	13,154
Frac. Asian	0.175	0.116	0.163	0.000	0.983	13,154
log(Median household income)	11.050	11.053	0.542	6.787	12.429	13,154
Frac. college educated	0.439	0.426	0.211	0.000	0.972	13,154
Frac. more than high school	0.582	0.517	0.251	0.033	1.000	13,154
Frac. denied	0.132	0.141	0.052	0.000	0.615	13,154
Micro Census Data						
Household Income	168,300	68,440	864,759	10,000	100,000	120,029
White	0.555	1.000	0.497	0.000	1.000	120,029
Black	0.062	0.000	0.240	0.000	1.000	120,029
Hispanic	0.157	0.000	0.364	0.000	1.000	120,029
Asian	0.201	0.000	0.401	0.000	1.000	120,029
Mortgage Application Data						
Approved	0.769	1.000	0.421	0.000	1.000	163,630
Loan to Income Ratio	3.380	3.267	0.874	1.883	6.117	163,630
Loan Amount ('000)	424.503	436.000	132.960	78.000	638.000	163,630
Applicant Income ('000)	162.710	150.000	73.531	44.000	406.000	163,630
Black	0.063	0.000	0.243	0.000	1.000	163,630
Hispanic	0.021	0.000	0.143	0.000	1.000	163,630
Asian	0.209	0.000	0.406	0.000	1.000	163,630
FHA Loan	0.001	0.000	0.018	0.000	1.000	163,630

Appendix Table C: Change in Lending Standards between 2000 and 2006

The table presents estimation of the approval model (Specification 6), with the 2000 and 2006 mortgage application data. The second column ('change 2000–2006') presents the coefficients of the interaction between the right-hand side borrower characteristics and the year 2006 dummy. The 2006 observations are weighted to match the share of income×race borrowers in 2000. Thus the estimates of the change in lending standards are adjusted for demographic shifts in the applicants' pool.

	Approved	Change 2000 – 2006
log(Price)	–0.352*	+0.159
	(0.127)	(0.140)
log(Income)	0.759*	–0.503*
	(0.110)	(0.115)
Constant	–2.111	+2.967*
	(0.962)	(1.113)
Black or African American, Nonhispanic	-1.154*	+0.345*
	(0.086)	(0.093)
Hispanic, Any Race	–0.716*	+0.294**
	(0.103)	(0.117)
Asian, Nonhispanic	–0.286*	+0.404*
	(0.069)	(0.075)
Observations	103,176	
Census Tracts	1,315	
Pseudo R Squared	0.0407	

***: Significant at 1%, **: Significant at 5%, *: Significant at 10%.
Robust standard errors clustered at the census tract level in parenthesis.

Appendix Table D: Mortgage Payment, Household, Housing, and Neighborhood Characteristics

	Dependent variable: log(Mortgage Payment)		
Year:	1990	2000	2010
log(Value)	0.455***	0.460***	0.348***
	(0.011)	(0.006)	(0.005)
log(Household Income)	0.198***	0.155***	0.178***
	(0.008)	(0.005)	(0.005)
Median Number of Rooms	0.003	0.007***	−0.006***
	(0.003)	(0.002)	(0.002)
Built 2−5 years ago	−0.119***	−0.090***	0.320***
	(0.033)	(0.029)	(0.030)
Built 6 to 10 years ago	−0.188***	−0.159***	0.245***
	(0.033)	(0.028)	(0.016)
Built 11−20 years ago	−0.538***	−0.228***	0.030**
	(0.031)	(0.027)	(0.013)
Built 21−30 years ago	−0.791***	−0.330***	−0.025
	(0.031)	(0.026)	(0.000)
Built 31−40 years ago	−0.712***	−0.344***	0.023*
	(0.032)	(0.026)	(0.013)
Built 41−50 years ago	−0.677***	−0.330***	0.027**
	(0.034)	(0.026)	(0.012)
Built 51−60 years ago	−0.617***	−0.316***	−0.019
	(0.033)	(0.028)	(0.015)
American Indian/Alaska Native	0.115	0.008	−0.112
	(0.070)	(0.065)	(0.078)
Asian and/or Pacific Islander	0.211***	0.074***	0.081***
	(0.013)	(0.008)	(0.008)
Black	−0.075***	0.024*	0.178***
	(0.021)	(0.015)	(0.016)
Hispanic, Any Race	0.049***	0.089***	0.148***
	(0.015)	(0.009)	(0.009)
Other race, non−Hispanic	0.292**	0.107	−0.04
	(0.137)	(0.069)	(0.054)
County Fixed Effects	Yes	Yes	Yes
Observations	27,007	29,584	32,177
R^2	0.227	0.317	0.258
Adjusted R^2	0.227	0.316	0.257
F Statistic	361.100***	652.492***	398.799***

Appendix Table E: Household Preferences with log Non-Housing Consumption

This table presents the estimated preference parameters in the model with log non-housing consumption.

	Dependent variable:	
	Base Utility δ_{jt}	
Median Age of Structure	-0.011^{***}	(0.0005)
Median Number of Rooms	0.132^{**}	(0.055)
$\times$ log(Distance)	-0.017^{***}	(0.005)
Academic Performance Index	0.024^{***}	(0.008)
Frac. Black	-1.267^{***}	(0.057)
Frac. Hispanic	-0.162^{***}	(0.051)
Frac. Asian	0.607^{***}	(0.047)
log(Median Household Income)	0.008	(0.013)
Frac. College	0.203^{***}	(0.036)
Av. log(Non-Housing Consumption)	1.070^{***}	(0.096)
Observations	13,081	
R^2	0.897	
Adjusted R^2	0.846	
Residual Std. Error	0.229 (df = 8691)	
Note:	*p<0.1; **p<0.05; ***p<0.01	

Y_{it}: Household i's income in year t.
m_{ijt}: Mortgage Payment of household i in neighborhood j in year t.
$\overline{Y_{\cdot t} - m_{\cdot jt}}$: average log of non-housing consumption in neighborhood j across households.

Appendix Table F: Unobservables in Lending Decisions and Unobservables in Utility

	δ_{jt}	ξ_j	ϕ_{ijt}	$\xi_j^{approval}$
δ_{jt}	1.000			
ξ_j	0.479	1.000		
ϕ_{ijt}	-0.148	-0.146	1.000	
$\xi_j^{approval}$	-0.059	-0.073	0.520	1.000

δ_{jt}: base utility term. ξ_j: fixed effect of the regression of base utility on neighborhood amenities. ϕ_{ijt}: probability of approval for household i in neighborhood j in year t. $\xi_j^{approval}$: fixed effect of the approval regression.

Appendix Table G: Base Utility Analysis in the Model with Rental

This table presents the regression of the base utility of renting in neighborhood j in year t on neighborhood amenities and log(rent). The log(rent) is the log gross rent from the Census files. The procedure for the imputation for upper-censored rents is discussed in the body of the appendix.

	Dependent variable:	
	Base Utility of Rental $\delta_{jt,rental}$	
Median Age of Structure	-0.031^{***}	(0.003)
Median Number of Rooms	0.866^{***}	(0.298)
$\times$ log(Distance)	-0.088^{***}	(0.025)
Academic Performance Index	0.016	(0.052)
log(Median Household Income)	0.670^{***}	(0.165)
Frac. College	1.647^{***}	(0.262)
log(Rent)	-3.027^{***}	(1.017)
Frac. Black	-9.418^{***}	(0.597)
Frac. Hispanic	2.212^{***}	(0.831)
Frac. Asian	-7.405^{***}	(0.727)
Observations	13,080	
R^2	0.806	
Adjusted R^2	0.708	
Residual Std. Error	0.903 (df = 8690)	
Note:	*p<0.1; **p<0.05; ***p<0.01	