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* $(\mathbf{p}^* = (p^*_{jt}), \mathbf{v}^* = (\mathbf{v}^*_{jt}))_{j=1,2,\ldots,J,\, t=1,2,\ldots,T}$ *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 $\eta^s = 0$. □

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. □

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. □

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* $0, 1s$ *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:

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

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 $\boldsymbol{\beta} - \breve{\boldsymbol{\beta}} = 0$. □

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 $\boldsymbol{\delta}_0 \in \mathbb{R}^{J-1}$. □

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 $\left(\widehat{\boldsymbol{\psi}}, \widehat{\boldsymbol{\gamma}}, \widehat{\Psi}, \widehat{\alpha}, \widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{\xi}}, \widehat{\Omega}, \widehat{\Sigma}\right)$ 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:

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

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 $\boldsymbol{\Delta}\mathbf{v}$. 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:

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

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.

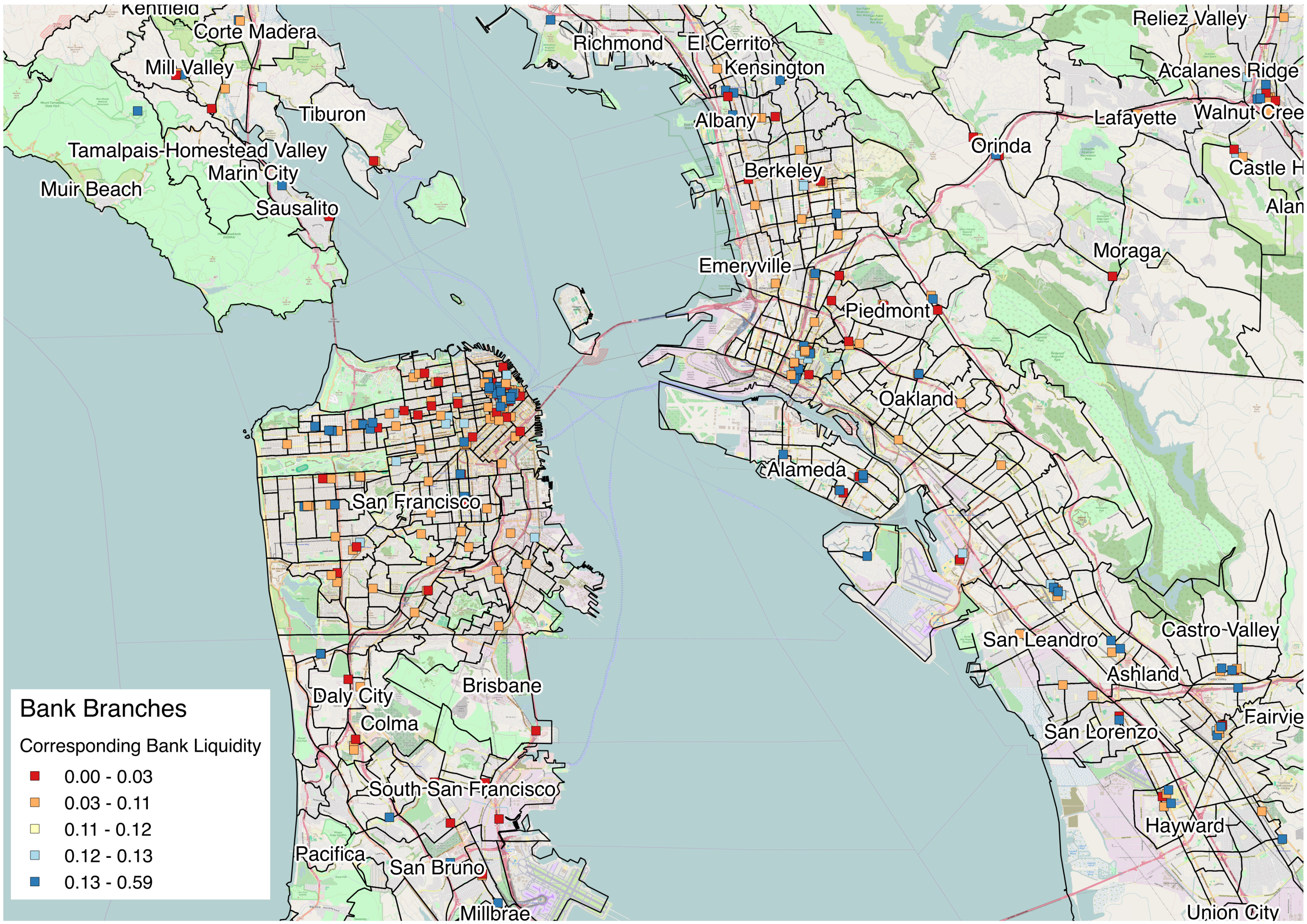
Swamy, P. A. (1970), 'Efficient inference in a random coefficient regression model', *Econometrica: Journal of the Econometric Society* pp. 311–323.

Part III

Tables and Figures

Appendix Figure A: Bank Branches and Liquidity

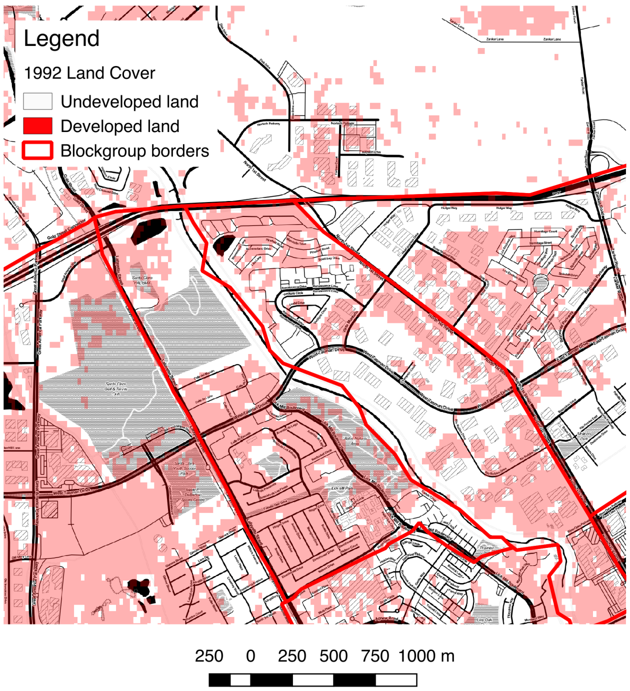
This map presents the set of bank branches from the Summary of Deposits in 1994. The map only includes bank branches whose corresponding bank is regulated by the Federal Reserve, the Office of the Comptroller of the Currency, or the Federal Deposit Insurance Corporation. The bank's liquidity level is 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.



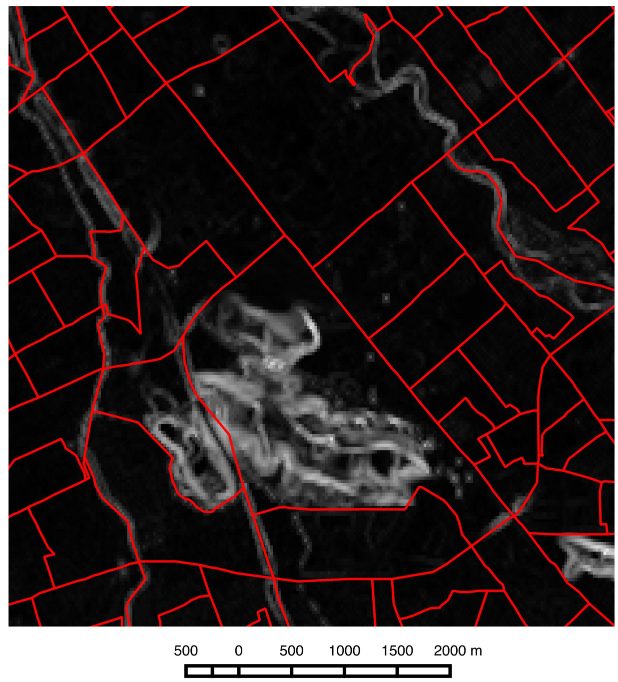
Appendix Figure B: Undeveloped Land - San Francisco Bay Area

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 $\eta_j = \hat{b} + \hat{c} \cdot Undeveloped\,Share_j + \hat{d} \cdot Ruggedness_j$.

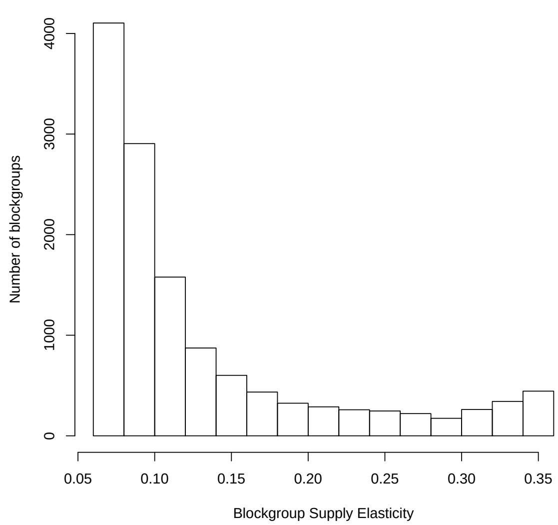
(a) Undeveloped Land – Lafayette and North 1st St.



(b) Ruggedness – Oak Hill Memorial Park

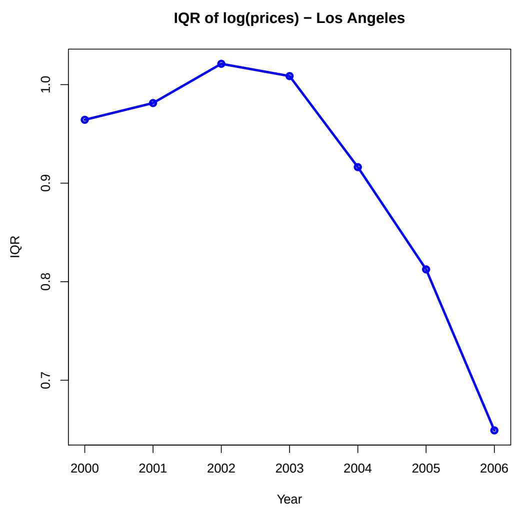


(c) Distribution of Housing Supply Elasticities

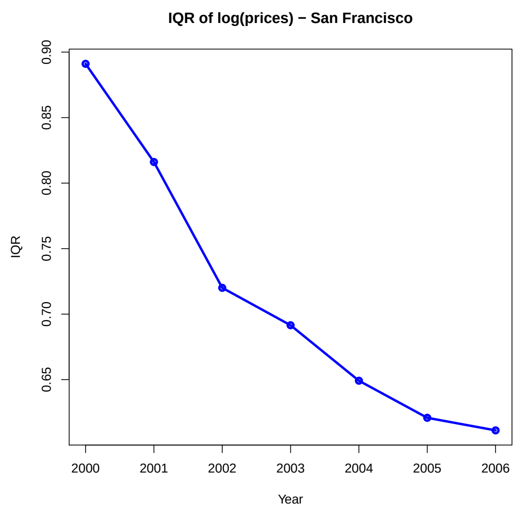


Appendix Figure C: The Compression of the Price Distribution in 4 Metropolitan Areas

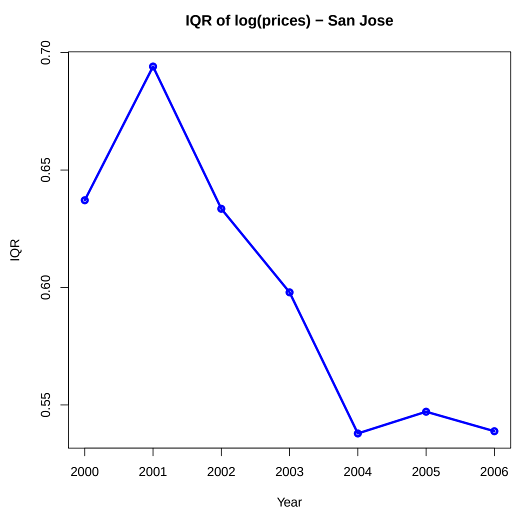
Each figure presents the Interquartile Range (IQR) of the distribution of log(price) for Los Angeles, San Francisco, San Jose, and Washington DC metropolitan areas. Individual transaction data from FNC Inc. and calculations from the authors.



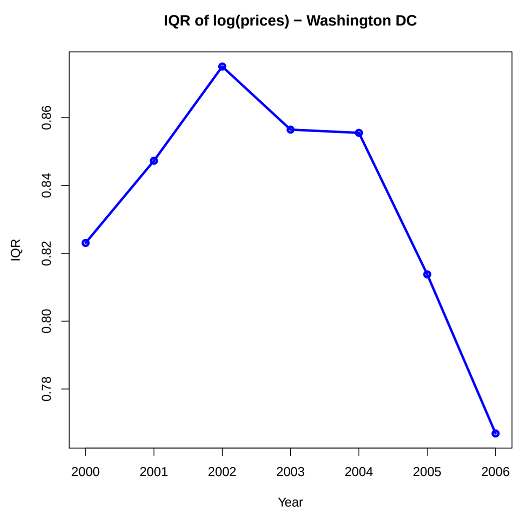
(a) Los Angeles



(b) San Francisco



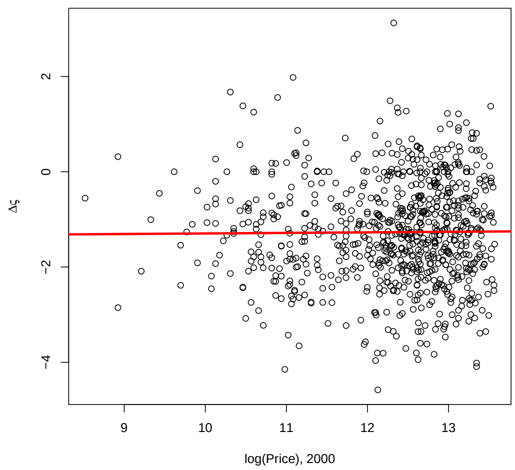
(c) San Jose



(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* $\log(p_{jt})$



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

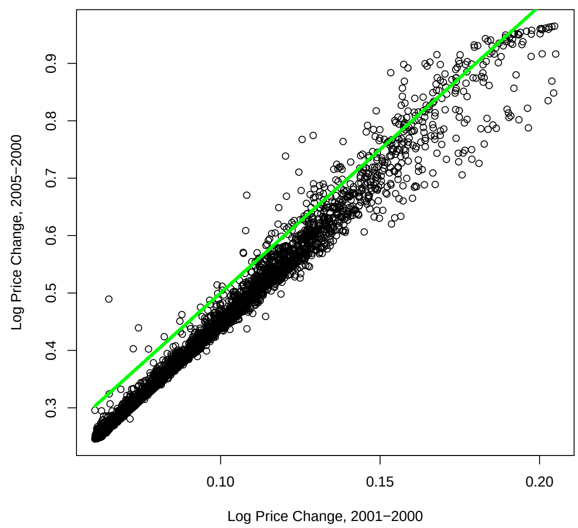
This accompanying table presents summary statistics on the unobservable fixed effects in the mortgage approval analysis.

	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

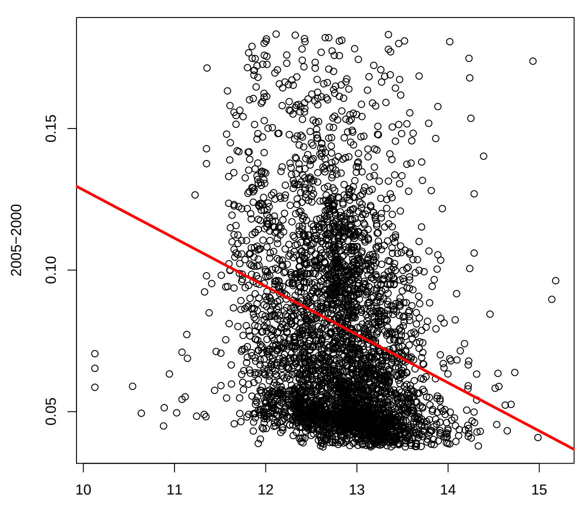
Appendix Figure E: Equilibrium Invariance vs. Equilibrium Shifts

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

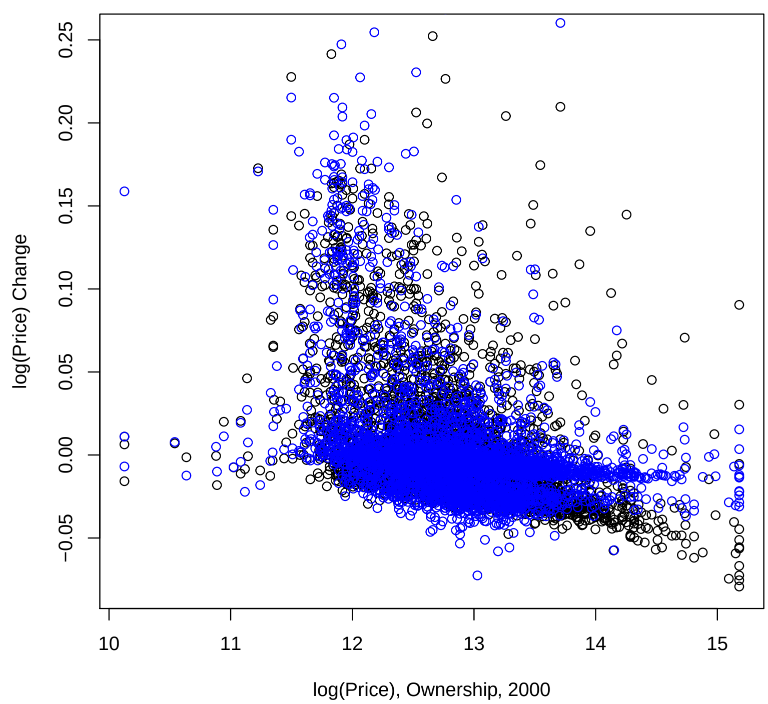


(b) Price Distribution Compression



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.



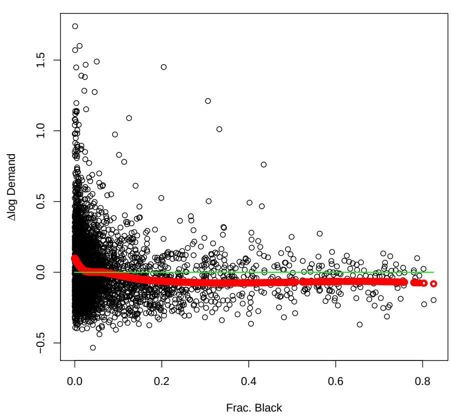
Appendix Figure G: Partial Equilibrium Impact of an Unconstrained Choice Set on Demand – Black Households

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.

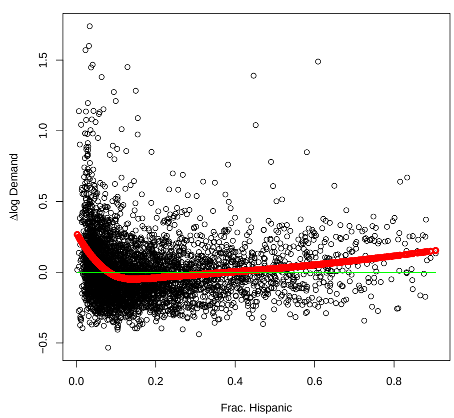
$$\Delta D = \log D(j,t|i, C = \{1,2,\ldots,J\}) - \log D(j,t|i) \quad (40)$$

This figure presents results for black households.

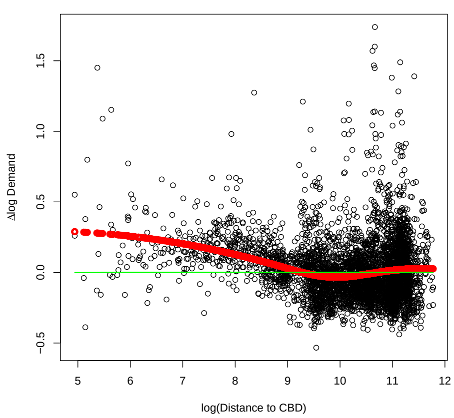
(a) By Fraction Black



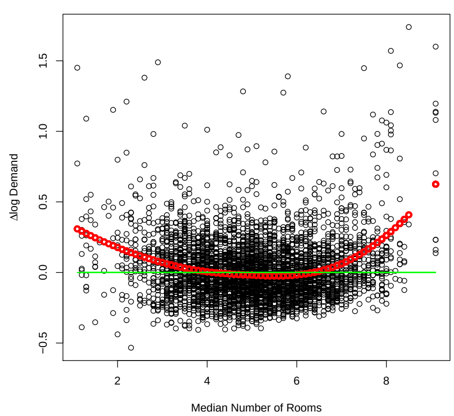
(b) By Fraction Hispanic



(c) By log(Distance) to CBD



(d) By Median Number of Rooms



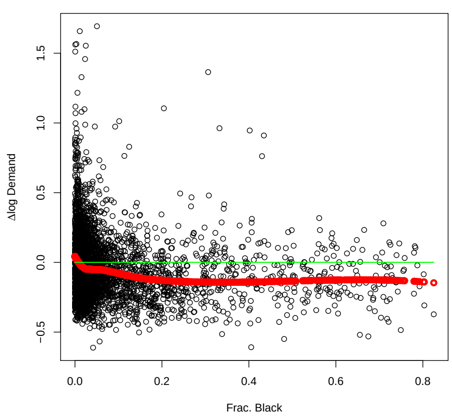
Appendix Figure H: Partial Equilibrium Impact of an Unconstrained Choice Set on Demand – Poor Households

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.

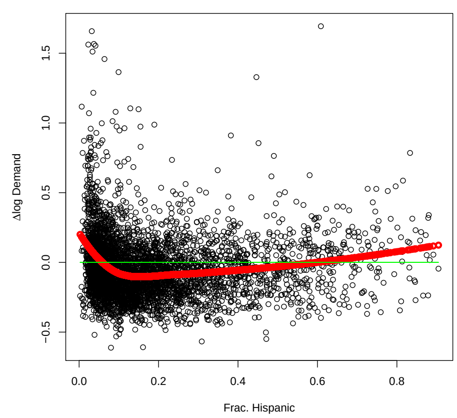
$$\Delta D = \log D(j,t|i, C = \{1, 2, \ldots, J\}) - \log D(j,t|i) \tag{41}$$

This figure presents results for poor households.

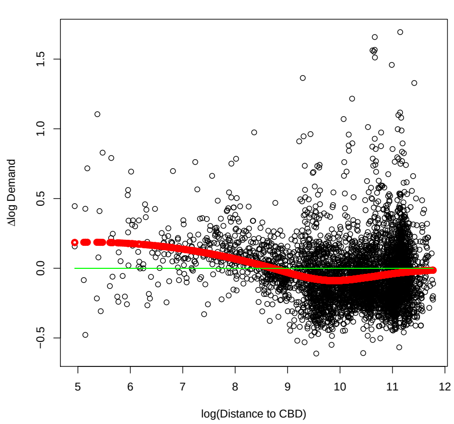
(a) By Fraction Black



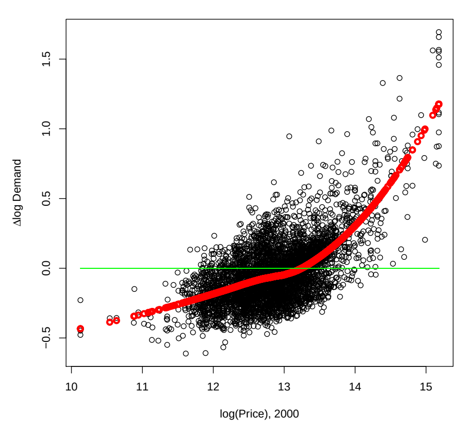
(b) By Fraction Hispanic



(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)
× 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	