Econometrica Supplementary Material

SUPPLEMENT TO "ESTIMATION OF AN EQUILIBRIUM MODEL WITH EXTERNALITIES: POST-DISASTER NEIGHBORHOOD REBUILDING"

(*Econometrica*, Vol. 87, No. 2, March 2019, 387–421)

CHAO FU
Department of Economics, University of Wisconsin and NBER

JESSE GREGORY
Department of Economics, University of Wisconsin and NBER

APPENDIX B: DATA IMPUTATIONS

TO SOLVE OUR MODEL NUMERICALLY, we must impute values for several of the model's exogenous variables we do not observe in our estimation data set, which covers the full universe of home-owning households in New Orleans when Katrina occurred. This appendix describes our imputation procedures.

B.1. *Wages*

We impute a New Orleans annual household earnings offer (i.e., the wage offer w_i^1) and an "outside option" annual household earnings offer (i.e., the wage offer w_i^0) for each household using geocoded microdata on households' pre-Katrina labor earnings from the Displaced New Orleans Residents Survey (DNORS)[34] and information about occupation-specific differences in prevailing wages across labor markets and across time from the 2005–2010 American Community Survey. The procedure involves two steps. In the first step, we match each household in our data set to a "donor" DNORS record using nearest Mahalanobis distance matching on a set of variables that are available for all households,[35] and impute to each record the labor market variables (household head and spouse's occupations and pre-Katrina annual earnings) of its DNORS donor record. We then compute w_i^1 and w_i^0 using the expressions

$$w_i^0 = w_{i,t<0}^{\text{head}}\left(\frac{\exp\left(\theta^0_{\text{occ}(i,\text{head}),t>0}\right)}{\exp\left(\theta^1_{\text{occ}(i,\text{head}),t<0}\right)}\right) + w_{i,t<0}^{\text{spouse}}\left(\frac{\exp\left(\theta^0_{\text{occ}(i,\text{spouse}),t>0}\right)}{\exp\left(\theta^1_{\text{occ}(i,\text{spouse}),t<0}\right)}\right),$$

$$w_i^1 = w_{i,t<0}^{\text{head}}\left(\frac{\exp\left(\theta^1_{\text{occ}(i,\text{head}),t>0}\right)}{\exp\left(\theta^1_{\text{occ}(i,\text{head}),t<0}\right)}\right) + w_{i,t<0}^{\text{spouse}}\left(\frac{\exp\left(\theta^1_{\text{occ}(i,\text{spouse}),t>0}\right)}{\exp\left(\theta^1_{\text{occ}(i,\text{spouse}),t<0}\right)}\right),$$

Chao Fu: cfu@ssc.wisc.edu
Jesse Gregory: jmgregory@ssc.wisc.edu

[34]Fielded by RAND in 2009 and 2010, the Displaced New Orleans Residents Survey located and interviewed a population-representative 1% sample of the population who had been living in New Orleans just prior to Hurricane Katrina.

[35]To allow us to match based on the OPAO property record variables that we observe for all households, we first merged the DNORS data with respondents' OPAO property records. We then performed the matching procedure, matching on the following variables; appraised pre-Katrina home values, pre-Katrina neighborhood demographic variables, block-level flood exposure, the extent of Katrina-related home damages measured by the decline in appraised home values from 2005 (prior to Katrina) to 2006, indicators for whether and when post-Katrina home repairs occurred, and indicators for whether and when a home was sold after Katrina.

 https://doi.org/10.3982/ECTA14246

where $w_{i,t<0}^{\text{head}}$ is the household head's pre-Katrina annual earnings, $w_{i,t<0}^{\text{spouse}}$ is his or her spouse's pre-Katrina annual earnings (zero if the household head is single), and the terms $\theta_{\text{occ},\tau}^{m}$ are log-wage indices estimated with data from the 2005–2010 ACS specific to labor markets $m \in \{0, 1\}$ (with $m = 0$ refering to the "outside" option, defined as the pooled group of all metro areas in the Census-defined South region—the typical destination of households displaced from New Orleans—and $m = 1$ referring to New Orleans) and time periods τ (with $\tau < 0$ referring to pre-Katrina wages and $\tau > 0$ referring to post-Katrina wages).[36]

B.2. *Non-housing Assets*

We impute an initial asset holding ($A_{it=0}$) for each household using asset data from Displaced New Orleans Residents Survey and the 2005 Panel Study of Income Dynamics.[37] First, using data from the PSID, we estimate a flexible statistical model of the *distribution* of non-housing assets conditional on a household's observable characteristics. We use a logistic regression to estimate the probability that a household has zero liquid assets conditional on the household's observable traits,[38] and we estimate a sequence of 99 quantile regressions (one for each quantile 1 to 99) to recover the distribution of assets conditional on the asset holding being positive. Then, using this estimated asset model, we draw 500 simulated asset holdings for each DNORS household from the *conditional* distribution of assets given the household's observable characteristics. Last, we match each household in our analysis data set to a "donor" DNORS record using nearest Mahalanobis distance matching on a set of variables that are available for all households,[39] and impute to each record a random draw from the DNORS donor record's simulated asset distribution.

[36]The composition adjusted log-wage indices $\theta_{\text{occ},\tau}^{m}$ are the estimated two-digit occupation by time period (either pre-Katrina or post-Katrina) by labor market (New Orleans or the pooled "other metro South") fixed effects from the regression

$$\ln(\text{earn}_{i,\tau}) = X_{i,\tau}' a + \theta_{\text{occ}(i,\tau),\tau}^{m} + e_{i,\tau},$$

where $\text{earn}_{i,\tau}$ is a worker's annual labor earnings, measured in the 2005–2010 ACS, and X is a vector of flexibly interacted demographic and human capital variables.

[37]Liquid assets are defined to be the sum of a household's non-IRA stock holdings, bond holdings, and holdings in checking accounts, savings accounts, money market accounts, and CDs.

[38]The explanatory variables include: indicators for solo-female headed household, solo-male headed household, the more educated household head being a high school dropout, the more educated household head having attended college but not received a bachelor's degree, the more educated household head having a bachelor's degree, a household head being black, the household residing in an urban area, the household residing in the south, an interaction of southern and urban, indicators for each of the four highest housing value quintiles, the age of the male head if present and the female head's age otherwise, and the square of the age of the male head if present and the square of the female head's age otherwise. When linking these estimates back to DNORS households, all DNORS households are classified as Southern and urban. The other inputs depend on the household's survey responses.

[39]To allow us to match based on the OPAO property record variables that we observe for all households, we first merged the DNORS data with respondents' OPAO property records. We then performed the matching procedure, matching on the following variables; appraised pre-Katrina home values, pre-Katrina neighborhood demographic variables, block-level flood exposure, the extent of Katrina-related home damages measured by the decline in appraised home values from 2005 (prior to Katrina) to 2006, indicators for whether and when post-Katrina home repairs occurred, and indicators for whether and when a home was sold after Katrina.

B.3. *Home Damages for Non-Road Home Households*

Last, we impute home replacement cost estimates and home repair cost estimates for households who did not apply to RH (and thus did not undergo RH damage appraisals). We first impute estimated replacement costs using the predicted values from a regression estimated among RH applicants of the log RH replacement cost estimate on log pre-Katrina appraised home value, pre-Katrina neighborhood demographic traits, and flood exposure. We then impute a damage fraction using the predicted estimate from nonlinear least squares estimates ($r^2 \approx 0.9$) of the statistical model

$$\text{DamageFraction}_i = \frac{\exp(\widetilde{X}_i' a)}{1 + \exp(\widetilde{X}_i' a)},$$

where $\widetilde{X}_i$ includes a polynomial in flood exposure, a polynomial in the percentage drop in the OPAO appraised value, and interactions of the two. Note that this imputation model is a smooth function of continuously distributed exogenous variables, and thus imputed records for nonapplicants do not contribute to any observed "jumps" in outcomes at the 51% grant formula threshold.

APPENDIX C: BASELINE MODEL—GOODNESS OF FIT

See Figures A2 and A3 and Table A.III.

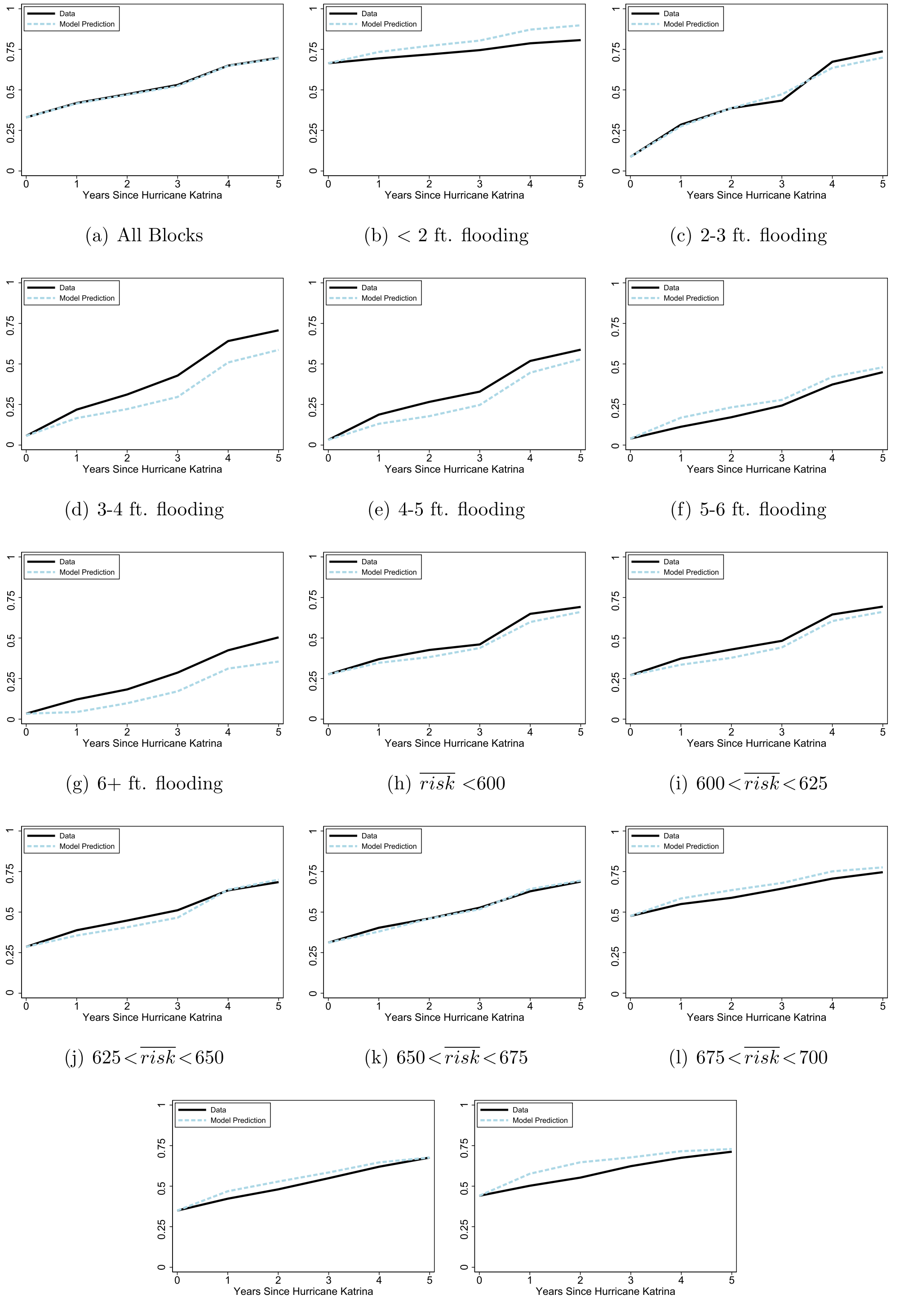
APPENDIX D: ADDITIONAL MODEL SPECIFICATIONS AND SIMULATIONS

D.1. *Model With More Flexible Specification of Amenities*

To assess the model's robustness, we re-estimated the model allowing amenity utility to follow separate linear time trends within each of the six flood-depth categories. Figure A4 shows the fit of this version of the model to rebuilding time trends, which is slightly improved relative to the more parsimonious specification of the model. Table A.IV compares the impact of RH on rebuilding rates in this model relative to the baseline model. The alternative specification predicts an equilibrium impact of 8.4 percentage points, compared to an impact of 8.0 percentage points in the baseline model. Table A.V compares the welfare implications of RH versus an unconditional grant policy as assessed by the two model specifications. The two models yield similar predictions about the fraction of households that are marginal (9.1% and 8.7% in the baseline and more-flexible models) and similar predictions about the per-household welfare impact of RH's distortionary structure (+\$2177 and +\$2754 in the baseline and more-flexible models).

D.2. *The Impact of Removing the RH Grant Formula Discontinuity*

To assess whether the discontinuity in RH's grant formula itself was an important factor in determining the program's impact, we simulated rebuilding choices under a version of RH where all grants are based on *damage* estimates (as opposed to *replacement cost* estimates). This is equivalent to calculating all grants based on the first of the two grant formulas on page 7. Table A.VI compares the rebuilding rate impacts of these two versions of RH. Overall, the impacts are similar, with slightly smaller impacts occurring under the "smooth" RH policy. This is because the "smooth" formula offers somewhat smaller grants for households with >51% home damage.



(a) All Blocks

(b) < 2 ft. flooding

(c) 2-3 ft. flooding

(d) 3-4 ft. flooding

(e) 4-5 ft. flooding

(f) 5-6 ft. flooding

(g) 6+ ft. flooding

(h) $\overline{risk}$ <600

(i) $600 < \overline{risk} < 625$

(j) $625 < \overline{risk} < 650$

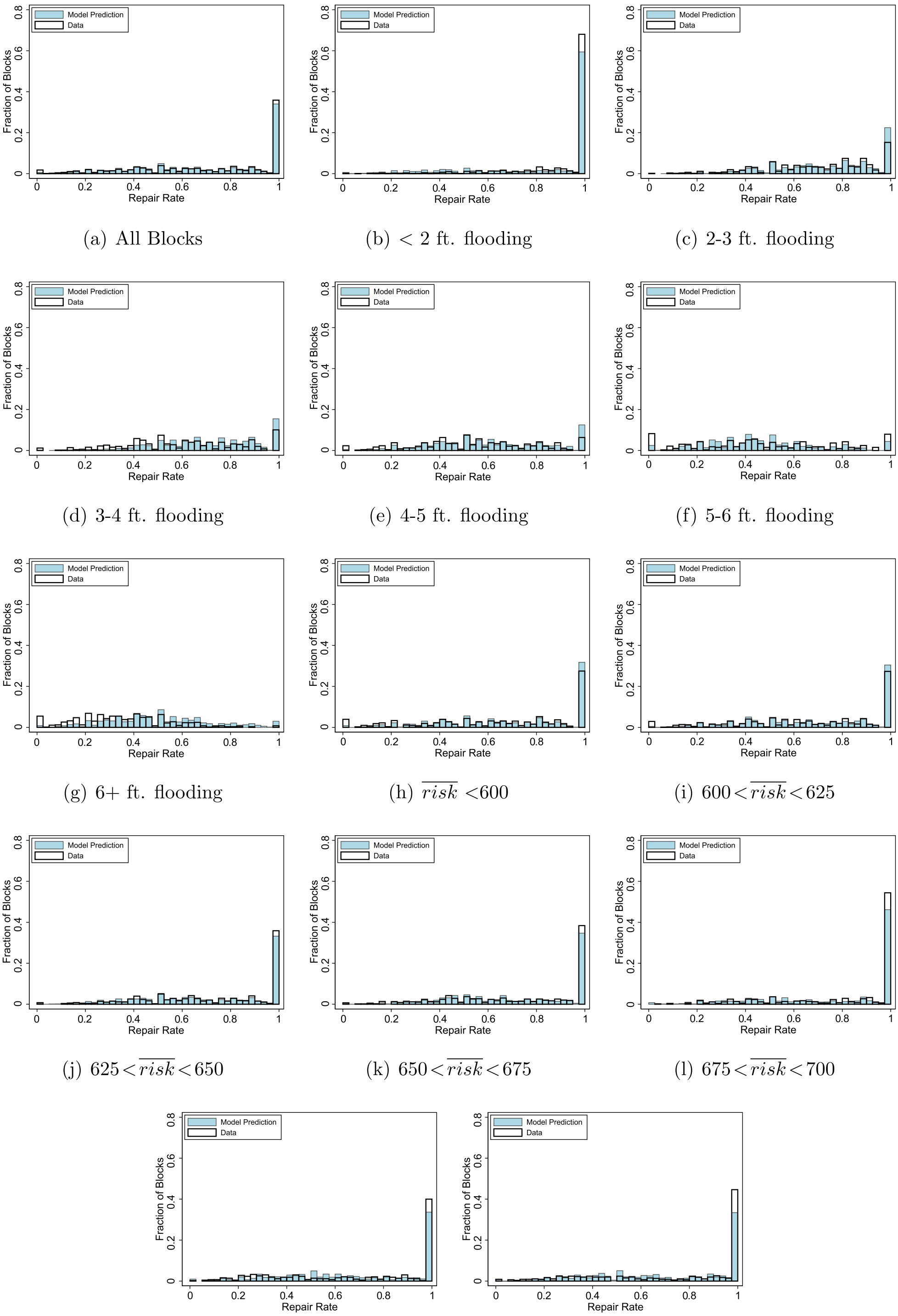
(k) $650 < \overline{risk} < 675$

(l) $675 < \overline{risk} < 700$

(m) $700 < \overline{risk} < 725$

(n) $\overline{risk}$ >725

FIGURE A2.—Goodness of fit: trends in fraction of homes livable by neighborhood characteristics.



(a) All Blocks

(b) < 2 ft. flooding

(c) 2-3 ft. flooding

(d) 3-4 ft. flooding

(e) 4-5 ft. flooding

(f) 5-6 ft. flooding

(g) 6+ ft. flooding

(h) $\overline{risk}$ <600

(i) 600<$\overline{risk}$<625

(j) 625<$\overline{risk}$<650

(k) 650<$\overline{risk}$<675

(l) 675<$\overline{risk}$<700

(m) 700<$\overline{risk}$<725

(n) $\overline{risk}$ >725

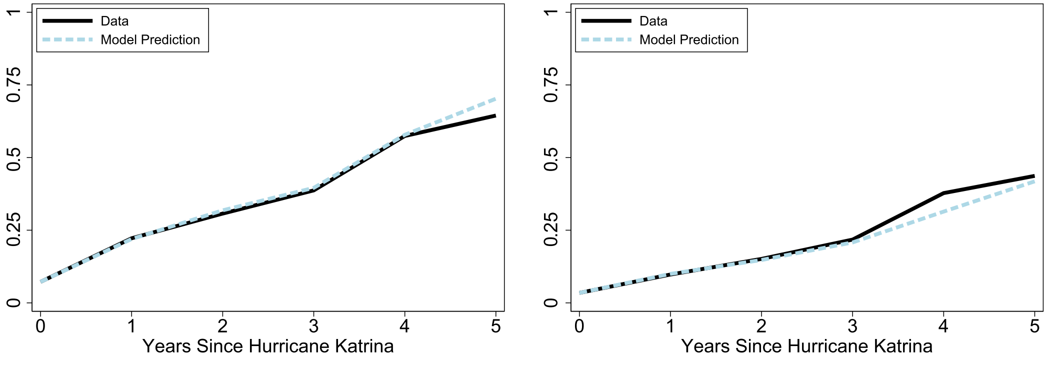
FIGURE A3.—Goodness of fit: histogram of fifth-anniversary block repair rates by neighborhood characteristics.

TABLE A.III

GOODNESS OF FIT TO NON-TARGETED MOMENTS: FIFTH-ANNIVERSARY REBUILDING RATE BY SUBGROUPS[a]

Subgroup	Data	Model
Home damages: < median	84.7 (0.21)	83.5
Home damages: > median	54.2 (0.29)	55.9
Insurance payout: < median	78.8 (0.24)	78.9
Insurance payout: > median	60.2 (0.28)	60.5
Tract poverty: < median	65.8 (0.27)	67.8
Tract poverty: > median	73.2 (0.26)	71.7
Tract majority noncollege	66.0 (0.27)	68.4
Tract majority college	72.8 (0.25)	71.0
Tract majority nonblack	79.4 (0.26)	75.6
Tract majority black	63.0 (0.25)	65.9
Not low/moderate income household	70.2 (0.24)	70.0
Low/moderate income household	68.3 (0.31)	69.4
Uninsured damages	54.8 (0.4)	57.9
No uninsured damages	74.4 (0.21)	73.7

[a]Source: Authors' calculations using the estimated equilibrium model.



(a) 2–3 ft. flooding (b) > 4 ft. flooding

FIGURE A4.—Goodness of fit: trends in fraction of homes livable predicted by the model allowing for flood-category-specific amenity time trends. Source: Authors' calculations using a re-estimated version of the model that allows for households' amenity valuations to follow separate linear time trends within each of the six flood categories.

TABLE A.IV

REBUILDING RATE IMPACTS IMPLIED BY THE BASELINE MODEL AND MODEL THAT ALLOWS NEIGHBORHOOD AMENITIES TO FOLLOW DIFFERENT TIME TRENDS BY FLOOD CATEGORY[a]

	(1)	(2)	(3)
		Rebuilding Rate Impacts	
Universe	No Grants Rebuilding Rate	Baseline Model	Beseline + Amenity Time Trends by Flooding
All households	61.7	+8.0	+8.4

[a]This table compares the predicted impact of RH on equilibrium rebuilding rates in the baseline model and a re-estimated version of the model that allows for households' amenity valuations to follow separate linear time trends within each of the six flood categories.

D.3. *Model With More Flexible Specification of Amenities*

With social spillover effects, multiple equilibria may exist (from the researcher's point of view), all of which can be computed given the structure of our model. One commonly assumed equilibrium selection rule for empirical applications is that agents agree on the equilibrium that maximizes their joint welfare (see, e.g., Jia (2008)). We use this equilibrium selection rule because we deem it reasonable in the context of a game among neighbors. As a robustness check, we have re-estimated our model selecting the equilibrium that minimizes joint welfare. Our counterfactual experiment results remain robust, as shown in Table A.VII.

APPENDIX E: IDENTIFICATION

We show identification of a simplified, one-period version of our model. Given our model assumption that neighborhood and household unobservables are permanent, having multiple-period data will only help identification.

E.1. *Simplified Model*

Households face a discrete choice of whether to rebuild and receive u_{i1} or relocate and receive u_{i0}:

$$u_{i1} = \ln c_1(z_i) + g(\mu_{j(i)}) + x_j'\beta + b_{j(i)} + \varepsilon_{ij}, \tag{13}$$

$$u_{i0} = \ln c_0(z_i), \tag{14}$$

TABLE A.V

WELFARE IMPACTS IN THE BASELINE MODEL AND THE MODEL THAT ALLOWS NEIGHBORHOOD AMENITIES TO FOLLOW DIFFERENT TIME TRENDS BY FLOOD CATEGORY[a]

	(1)	(2)	(3)	(4)
	% Marginal		Welfare Impacts ($ per Capita)	
Universe	Baseline Model	Beseline + Amenity Time Trends by Flooding	Baseline Model	Beseline + Amenity Time Trends by Flooding
All households	9.1	8.7	2177	2754

[a]This table compares the predicted impact of RH on household welfare in the baseline model and a re-estimated version of the model that allows for households' amenity valuations to follow separate linear time trends within each of the six flood categories.

TABLE A.VI

REBUILDING RATE IMPACTS OF A ROAD HOME PROGRAM THAT USES A "SMOOTH" FORMULA, PAYING ALL HOUSEHOLDS BASED ON DAMAGES ESTIMATES[a]

Subgroup	(1) No Grants Rebuilding Rate	(2) Equilibrium Rebuilding Impacts: Actual Road Home	(3) Equilibrium Rebuilding Impacts: "Smooth" Road Home
All	61.7	+8.0	+7.0
Flood depth:			
<2 feet	76.2	+4.5	+3.9
2–4 feet	59.6	12.7	11.1
>4 feet	42.3	9.8	8.4
Rebuilding Rate w/o RH:			
80–100%	88.8	4.0	2.5
60–80%	70.3	10.0	7.9
40–60%	50.4	11.5	10.6
20–40%	33.0	11.8	10.6
0–20%	6.2	14.9	14.4

[a]This table compares the predicted impacts of RH and a RH style with a "smooth" grant formula using the estimated model.

where z_i is household characteristics or household-level incentive shifters. In our context, the exogenous incentive shifter z is an indicator that a household's rebuilding cost assessment falls above the policy formula discontinuity. $j(i)$ is the neighborhood that i

TABLE A.VII

RH'S EQUILIBRIUM EFFECTS ON REBUILDING BY EQUILIBRIUM-SELECTION RULE[a]

Subgroup	(1) Baseline Model	(2) Alternative Eqm.- Selection Rule
All	+8.0	+7.6
Flood depth:		
<2 feet	+4.5	+4.4
2–3 feet	+14.1	+13.1
3–4 feet	+11.2	+10.5
4–5 feet	+12.6	+11.4
5–6 feet	+9.3	+8.8
>6 feet	+8.0	+7.6
Rebuilding Rate w/o RH:		
90–100%	+0.2	+0.2
80–90%	+5.3	+5.0
70–80%	+8.8	+7.7
60–70%	+11.0	+9.9
50–60%	+11.2	+10.3
40–50%	+11.8	+11.4
30–40%	+11.7	+11.5
20–30%	+11.9	+11.1
10–20%	+14.7	+13.4
0–10%	+14.9	+12.7

[a]This table compares the simulated equilibrium impacts of the Road Home grant program on rebuilding rates using the baseline model (column (1)), which assumes that the total-welfare-maximizing equilibrium is selected on blocks with multiple self-consistent equilibria, to the simulated impact of RH using a (re-estimated) version of the model that assumes the total-welfare-*minimizing* equilibrium occurs in such cases (column (2)). Source: Authors' calculations using the estimated equilibrium models.

belongs to, x is neighborhood observable characteristics, such as flood exposure. $\mu_{j(i)}$ is rebuilding rate in the neighborhood. b is unobservable neighborhood characteristics, ε_{ij} is household's idiosyncratic taste for moving back.

Household i will move back $d_i = 1$ if $u_{i1} > u_{i0}$; therefore, the following holds:

$$\Pr(d_i = 1|x, z, \mu) = F_{b+\varepsilon}\big(\underbrace{\ln c_1(z_i) - \ln c_0(z_i)}_{\Delta c(z_i)} + g(\mu) + x_j'\beta\big).$$

This implies the expected rebuilding rate in the neighborhood is determined by x and the distribution of z in neighborhood $j(i)$. In our context, given that z is an indicator function, the average of z would serve as a sufficient summary statistic, which we denote by $Z_{j(i)}$. Therefore, the expected rebuilding rate is given by $\mu(x_{j(i)}, Z_{j(i)})$.

E.2. *Identification*

Assumptions:

1. ε_{ij} is independent of (z, b, x).
2. z_i is independent of $b_{j(i)}$ for all i, and thus $Z_{j(i)}$ is independent of $b_{j(i)}$.
3. x is independent of b.

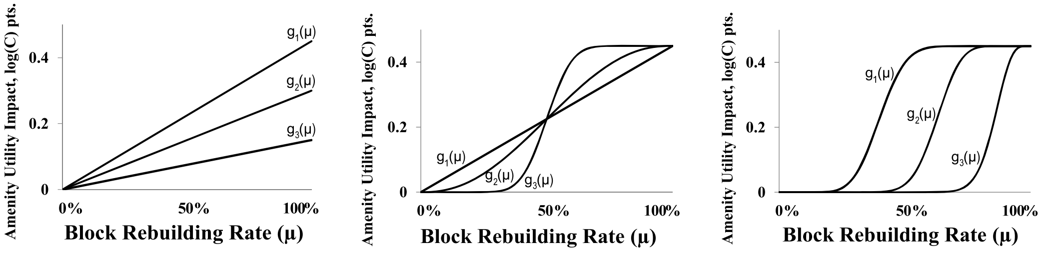
CLAIM: *Given Assumptions* 1 *to* 3, *marginal rate of substitution between neighbors' rebuilding* $\mu(x_{j(i)}, Z_{j(i)})$ *and private consumption* c; $\mathrm{MRS} = \frac{\Delta u}{\Delta\mu(x_{j(i)}, Z_{j(i)})} / \frac{\Delta\mu(x_{j(i)}, Z_{j(i)})}{\Delta c}$ *is identified*.

PROOF: The attractiveness of a block varies with x, which generates variation in expected rebuilding rates $\mu(x_{j(i)}, Z_{j(i)})$. We can trace out the spillover function $g(\mu)$ by performing the following calculation over a range of values of the exogenous vector x that yield different predicted rebuilding rates $\mu(x_{j(i)}, Z_{j(i)})$ and exploiting experimental variation in z_i and $Z_{j(i)}$ (the discontinuity):

$$\Pr(d_i = 1|X, Z) = F_{b+\varepsilon}\big(\underbrace{\ln c_1(z_i) - \ln c_0(z_i)}_{\Delta c(z_i)} + g\big(\mu(x_{j(i)}, Z_{j(i)})\big) + x_j'\beta\big),$$

$$\begin{aligned}\delta_1(x, z, \overline{z}) &= \frac{\Delta \Pr(d_i = 1|X, Z)}{\Delta z_i} \\ &\approx F'_{b+\varepsilon}\big(\underbrace{\ln c_1(z_i) - \ln c_0(z_i)}_{\Delta c(z_i)} + g\big(\mu(x_{j(i)}, Z_{j(i)})\big) + x_j'\beta\big) \\ &\quad \times \left(\frac{\Delta c(z_i)}{\Delta z_i} + \underbrace{\frac{\Delta E(b_{j(i)}|z_i)}{\Delta z_i} + \frac{\Delta E(\varepsilon_i|z_i)}{\Delta z_i}}_{\text{assumed}=0}\right),\end{aligned} \tag{15}$$

$$\begin{aligned}\delta_2(x, z, \overline{z}) &= \frac{\Delta \Pr(d_i = 1|X, Z)}{\Delta Z_{j(i)}} \\ &\approx F'_{b+\varepsilon}\big(\underbrace{\ln c_1(z_i) - \ln c_0(z_i)}_{\Delta c(z_i)} + g\big(\mu(x_{j(i)}, Z_{j(i)})\big) + x_j'\beta\big)\end{aligned} \tag{16}$$



(a) Amplitude: $S^{(3)} > S^{(2)} > S^{(1)}$ (b) Shape: $\lambda_2^{(3)} > \lambda_2^{(2)} > \lambda_2^{(1)}$ (c) Location $\lambda_1^{(3)} > \lambda_1^{(2)} > \lambda_1^{(1)}$

FIGURE A5.—Parameterization of the amenity spillover function.

$$\times \Bigg(\frac{\Delta g\big(\mu(x_{j(i)}, Z_{j(i)})\big)}{\Delta \mu} \times \underbrace{\frac{\Delta \mu(x_{j(i)}, Z_{j(i)})}{\Delta Z_{j(i)}}}_{\delta_3(x,z,\overline{z})}$$

$$+ \underbrace{\frac{\Delta E(b_{j(i)}|\overline{z})}{\Delta \overline{z}} + \frac{\Delta E(\varepsilon_i|\overline{z})}{\Delta \overline{z}}}_{\text{assumed}=0} \Bigg).$$

Note that many variables in the vector X vary continuously, letting us nonparametrically identify $E(g(\mu(x_{j(i)}, Z_{j(i)})))$.[40] Each of these three δ's are nonparametrically identified by flexibly measuring how these conditional probabilities depend on the right-hand-side variables (x, z, Z). The marginal rate of substitution is thus, as well, given by

$$\text{MRS}(x, z, Z) = \frac{\delta_2(x, z, Z)/\delta_3(x, z, Z)}{\delta_1(x, z, Z)}. \qquad \textit{Q.E.D.}$$

Note that this expression is only consistent for $\text{MRS} = \frac{\Delta u}{\Delta \mu(x_{j(i)}, Z_{j(i)})} / \frac{\Delta \mu(x_{j(i)}, Z_{j(i)})}{\Delta c}$ under the assumption that the incentive shifters z_i and $Z_{j(i)}$ are uncorrelated with ε_i and $b_{j(i)}$, as illustrated with the above Equations (15) and (16) with the terms labeled "assumed=0," which motivated us to exploit quasi-experimental variation to the incentives of households and neighboring households that, as shown in the body, appears consistent with these assumptions.

REFERENCES

JIA, P. (2008): "What Happens When Wal-Mart Comes to Town: An Empirical Analysis of the Discount Retailing Industry," *Econometrica*, 76, 1263–1316. [7]

Co-editor Liran Einav handled this manuscript.

Manuscript received 16 March, 2016; final version accepted 25 June, 2018; available online 8 August, 2018.

[40] We use Δ's instead of partial derivatives because the discontinuity we exploit as an instrument can generate large jumps in rebuilding incentives in some cases.