Sea Level Rise and Urban Inequality

By Allan Hsiao*

Sea level rise threatens coastal cities around the world. Will it exacerbate inequality in these already unequal places? The rich can adapt by moving to higher ground, leaving the poor behind with flooding and inundation. I study this spatial sorting with a simple empirical model and granular data from Jakarta, a flood-prone megacity of 32 million that faces clear dangers from sea level rise.

I. Model

Individuals i of wage groups j choose locations k to maximize residential utility, solving

$$\max_j \{v_{jk} + \varepsilon_{ijk}\}.$$

Residential utility includes representative utility v_{jk} and logit taste shocks ε_{ijk}, where representative utility for a given location is

$$v_{jk} = \alpha_j p_k + \beta f_k + x_k \gamma + \delta_{jk}. \tag{1}$$

It depends on housing prices p_k, flooding f_k, and observed amenities x_k, and unobserved amenities δ_{jk}. Price elasticities α_j can differ by wage group. Logit shocks imply location choice probabilities

$$\pi_{ijk} = \pi_{jk} = \frac{e^{v_{jk}}}{\sum_\ell e^{v_{j\ell}}}. \tag{2}$$

Individuals within wage groups have common wages and thus common choice probabilities.

Equilibrium is as follows. Housing demand by location is

$$n_k = \sum_i \pi_{ijk},$$

and I consider fixed housing supply $\bar{n}_k$. Prices $p = \{p_k\}$ clear housing markets in equilibrium.

$$n_k(p^*) = \bar{n}_k \quad \forall k \tag{3}$$

The denominator of equation 2 captures spatial interdependence: prices in each location affect choice probabilities in every location.

Wage-specific price elasticities and endogenous prices generate sorting on flood exposure. High-wage individuals pay for flood safety, which is an amenity that commands high prices. But if low-wage individuals are more price sensitive, then they may prefer the low prices of flood-prone locations. This sorting creates inequality. Sea level rise exacerbates sorting if coastal flooding leads high-wage individuals inland, bidding up prices and pushing low-wage individuals elsewhere. The rich crowd out the poor in pursuit of higher ground.

II. Data

I compile fine-grained spatial data for the city of Jakarta. I obtain populations, housing prices, flooding, and geographic variables by 300m cell from Hsiao (2023). Populations for 2015 are from the Global Human Settlement Layer, and housing prices for 2015 are constructed from transaction records and online listings. Flooding for 2013 to 2020 is from city government data. I compute flood frequency as the average number of flood days per year. Geographic variables include coordinates, administrative regions, elevation, distance to the coast, and distance to the nearest river.

To study inequality, I construct wage groups by 300m cell. I proceed in three steps. First, I begin with full count population census data from 2010, and I geocode the data based on household addresses. I find that 86% of households can be assigned to an administrative block (*rukun tetangga*). The average such unit contains 350 people, and I map households to cells using block centroids.

Second, I define high- and low-wage groups $j \in \{H, L\}$. The census data do not record wages, and so I proxy with education. I define the high-wage group as individuals with post-secondary education, and the low-wage group as individuals without. For each group, I drop cells

* Princeton University, Department of Economics, 20 Washington Road, Princeton, NJ 08540, ajhsiao@princeton.edu.
Wenqing Yu provided exceptional research assistance.

TABLE 1—WAGES BY EDUCATION

	None	Primary	Middle	High	College
Mean monthly wages (2015 USD)	131	149	182	262	624
Proportion of SUSENAS sample (%)	5	15	17	42	21
Proportion of census sample (%)	16	18	19	35	12

Note: Each column corresponds to a given level of educational attainment. Middle is lower secondary schooling, and high is upper secondary schooling, inclusive of vocational training. College includes all forms of post-secondary schooling. Mean monthly wages are from the 2011, 2012, 2013, and 2014 waves of the SUSENAS socioeconomic survey. These wages measure monthly net income, both money and goods, from an individual's main job. The second row reports educational composition in this sample. For comparison, the third row reports educational composition in the geocoded sample of the 2010 population census.

without at least 10 individuals in the census data. I do so to reduce noise in the population measurements, at the cost of losing 2.8% of group-cell observations. I then compute high- and low-wage population shares, and I multiply by 2015 populations from the Global Human Settlement Layer. I thus obtain 2015 populations by wage group and 300m cell.

Third, I evaluate education as a proxy measure of wages. I merge the 2011, 2012, 2013, and 2014 waves of the SUSENAS socioeconomic survey to obtain a sample of 26,401 individuals for Jakarta. This sample is less comprehensive than the full count census, and it can only be geocoded to the district level. But it records both education and wages.

Table 1 presents average monthly wages by educational attainment. Wages are defined as net monthly income from an individual's main job, and college includes all forms of post-secondary schooling. Wages increase monotonically with education, and I take my high-wage cutoff from the large wage increase at post-secondary schooling. The sample of SUSENAS wage-earners is more highly educated than the broader population, but estimation and counterfactuals avoid this sample selection by relying solely on census data.

III. Estimation

I estimate the model on observed data to recover parameters α_j, β, γ, and δ_{jk}. With the estimated model, I can characterize sorting and evaluate welfare for any given pattern of flooding. Inverting equation 2,

$$\ln \pi_{jk} - \ln \pi_{j0} = v_{jk} - v_{j0}$$

for reference location $k = 0$. Substituting equation 1, I obtain a linear estimating equation.

$$\Delta \ln \pi_{jk} = \alpha_j \Delta p_k + \beta \Delta f_k + \Delta x_k \gamma + \Delta \delta_{jk} \tag{4}$$

for $\Delta y_{jk} = y_{jk} - y_{j0}$ and $\Delta y_k = y_k - y_0$.

I estimate equation 4 with data. I can compute choice probabilities $\pi_{jk} = n_{jk} / \sum_\ell n_{j\ell}$ from populations n_{jk}, which I observe by wage group and location. Wages w_j, housing prices p_k, and flooding f_k are data. Observed amenities x_{jk} include distance to the coast, distance to the nearest river, elevation, and district fixed effects. These variables act as controls. Unobserved amenities δ_{jk} represent structural errors.

The identification problem is that prices are correlated with unobserved amenities are correlated. The reason is sorting: high-amenity locations attract high-wage individuals that bid up prices. I thus require a price instrument. Typical candidates for demand estimation include cost shifters, prices in other markets, characteristics of competing products, and demographics in other markets (Berry and Haile, 2021). This context calls for shifters of housing construction costs in a given location, as well as housing prices and resident demographics in nearby locations. I choose ruggedness as a cost shifter, as construction must flatten rugged terrain. The exclusion restriction argument is that modest ruggedness in Jakarta does not impede transportation, and thus is less salient to residents.

I take flooding as uncorrelated with amenities. In practice, coastal areas may enjoy pleasant coastal views despite elevated flood risk. Conversely, flood-prone areas may suffer from disinvestment in public amenities. Controls help to mitigate this concern, and the baseline specification includes coastal distance, river proximity,

TABLE 2—DEMAND ESTIMATION

	IV		OLS	
	Estimate	SE	Estimate	SE
Log price, low wages	-2.734	(0.450)	-0.250	(0.088)
Log price, high wages	-1.951	(0.612)	0.074	(0.146)
Flooding	-0.091	(0.039)	-0.055	(0.022)
Observations	10,421		10,421	
p-value, low = high	0.046		0.002	
F-statistic	15.29			

Note: Each pair of columns is one regression. Each observation is a cell-group with 300m cells and low- and high-wage groups. Log prices comes from 2015 property prices per square meter, measured in units of 1M IDR (roughly 75 USD). The IV specification instruments for log prices with ruggedness. I proxy for wages with education. The high-wage group is those with post-secondary education, and the low-wage group is those without. Flooding is the average number of flood days per year, as observed from 2013 to 2020. Controls include distance to the coast, distance to the nearest river, elevation, and district fixed effects. I report *p*-values for the null hypothesis that low- and high-wage price elasticities are equal.

elevation, and district fixed effects. At the same time, omitting these controls does not greatly affect the estimated flooding coefficient.

Estimation proceeds by stacked linear regression. First, I choose a reference location and compute the differenced regressors relative to this location. I do so for each wage group. Second, I stack the data and construct a group indicator. I also construct price-group interaction terms, and I do the same for ruggedness as a price instrument. Third, I regress choice probabilities on the price-group interactions, flooding, observed amenities, and the group indicator. I instrument for the price-group interactions with the ruggedness-group interactions. This regression yields $\hat{\alpha}_j$, $\hat{\beta}$, and $\hat{\gamma}$ as coefficients and $\Delta\hat{\delta}_{jk}$ as residuals. I thus obtain δ_{jk} relative to δ_{j0}, but not in levels, noting that the group indicator allows δ_{j0} to vary freely across wage groups.

Table 2 presents parameter estimates. IV estimates show that high prices and severe flooding each reduce residential demand. Regressing on log prices allows me to interpret the coefficients as elasticities, and indeed both low- and high-wage groups have elastic demand. But the low-wage group is 40% more price sensitive than the high-wage group, and this difference is statistically significant. Ruggedness serves as a strong instrument, increasing prices in the first stage with an F-statistic of 15.29.

OLS estimates ignore the price endogeneity that arises from sorting, which generates strong upward bias in the estimated price coefficients. Locations with high unobserved amenities also have high prices, and so individuals may choose these locations despite high prices. Not accounting for this possibility leads to the false conclusion that individuals are not price sensitive. For the high-wage group, estimation bias even results in a positive demand elasticity.

IV. Sea Level Rise

Will sea level rise exacerbate inequality? I project flooding in Jakarta under sea level rise with a simple, elevation-based hydrological model. For a given degree of sea level rise, I identify 300m cells that fall below sea level. To these inundated cells, I assign the maximum flooding observed in the data: 24.5 flood days per year. Other cells retain their observed flooding values. These projections are likely underestimates, as inundation will be worse than 24.5 flood days per year. And I make no adjustment to flooding for cells that fall near sea level, but not below. At the same time, I assume no adaptation via government intervention.

I consider 1, 3, and 5m of relative sea level rise, which push 1.3%, 5.9%, and 19.7% of cells below sea level. Government plans anticipate 3 to 5m by 2050, citing a combination of 8mm of global mean sea level rise and 7 to 14cm of local land subsidence annually (NCICD 2014). The global rate is consistent with scientific estimates, as surveyed by Depsky et al. (2023). The local rates are consistent with older estimates of land subsidence from 1982 to 2010 (Abidin

TABLE 3—INEQUALITY WITH SEA LEVEL RISE

	Flooding			Prices
	Low wages	High wages	L/H ratio	Low-high difference
Current	0.86	0.61	1.42	-0.09
Projected				
1m sea level rise	0.99	0.66	1.49	-0.10
3m sea level rise	1.87	0.94	1.98	-0.12
5m sea level rise	5.50	2.37	2.31	-0.19
Projected, no sorting				
1m sea level rise	0.98	0.70	1.40	-0.09
3m sea level rise	1.82	1.18	1.55	-0.11
5m sea level rise	5.29	3.47	1.52	-0.15

Note: The first row computes flooding exposure and price incidence from observed data on the low- and high-wage groups. Each other row is one counterfactual. The second panel solves the model for equilibrium prices and choice probabilities under projected flooding from sea level rise. Prices are normalized against reference location $k = 0$ and thus can only be interpreted in changes. Flooding can be interpreted in levels. The third panel suppresses the impact of sorting. It computes flood exposure with counterfactual flooding but current choice probabilities, and similarly it computes price incidence with counterfactual prices but current choice probabilities.

et al., 2011), although newer estimates are more modest for 2014 to 2020 (Tay et al., 2022).

I calculate flood exposure by wage group j as the average faced by individuals in each group.

$$F_j = \sum_k f_k \pi_{jk} \tag{5}$$

Current flood exposure is $F_j(f, \pi_j)$ for $f = \{f_k\}$ and $\pi_j = \{\pi_{jk}\}$. I observe flooding f and choice probabilities π as data, and so I can calculate this measure directly. Projected flood exposure is $F_j(f', \pi'_j)$ for $f' \in \{f^{1\mathrm{m}}, f^{3\mathrm{m}}, f^{5\mathrm{m}}\}$. The hydrological model gives flooding f', and the sorting model gives choice probabilities π'.

I obtain choice probabilities π' by solving conditions 3, which pin down equilibrium prices and thus choice probabilities. These conditions represent a system of nonlinear equations, which can be large and difficult to solve when there are many locations. I compute the prices needed to compensate for projected flooding in each location as $\log p'_k = \log p_k - \beta(f'_k - f_k)/\bar{\alpha}$, for average price elasticity $\bar{\alpha}$, and I use these non-equilibrium prices as a starting point in solving the system. I normalize the price of reference location $k = 0$ to zero, as uniform price increases do not affect choice probabilities (absent an outside option).

I focus on the impacts of flooding via sorting and housing prices. In solving the model, I fix wage groups j, amenities x_k and δ_k, and housing supply $\bar{n}_k$ at current levels. It is equivalent to assume that amenities change uniformly across space, that population grows proportionally across wage groups, and that housing supply grows proportionally across locations.

Table 3 presents current and projected flood exposure. At current levels, low-wage individuals are 42% more exposed to flooding than high-wage individuals. But both groups are highly exposed, experiencing an average of one flood day every one to two years. At projected levels, sea level rise widens the gap between wage groups. For sea level rise of 3 and 5m, low-wage flood exposure is twice as high as high-wage flood exposure. Moreover, the absolute increase in flood exposure is substantial. For sea level rise of 5m, the low- and high-wage groups experience 5.50 and 2.37 flood days per year, relative to 0.86 and 0.61 at current levels.

I also calculate price incidence by wage group j as an average analogous to equation 5.

$$P_j = \sum_k p_k \pi_{jk} \tag{6}$$

Table 3 presents the difference between low- and high-wage measures, which I compute with log prices. I can compare across scenarios, as price normalizations difference out. I find that sea level rise widens the gap between wage groups, but now to the benefit of the low-wage group. Lower prices compensate for higher flood exposure, potentially narrowing the welfare gap.

Lastly, I isolate the role of sorting by conditioning on current choice probabilities.

$$F_j(f', \pi_j), \quad P_j(p', \pi_j)$$

Inequality in flood exposure is greatly attenuated. For sea level rise of 5m, low-wage flood exposure is 1.52 times as high as high-wage flood exposure, relative to 2.31 times with equilibrium sorting. Ignoring mobility leads to understated inequality, as sorting drives inequality. Sorting reduces high-wage flood exposure to 2.37 flood days per year, relative to 3.47 without sorting, as high-wage individuals seek out flood-safe areas. Sorting also raises low-wage flood exposure to 5.50 flood days, relative to 5.29 without sorting, as higher prices push low-wage individuals toward flood-prone areas. Inequality in price incidence is similarly attenuated. Suppressing sorting keeps high-wage individuals from high-priced higher ground.

V. Conclusion

This paper studies the distributional consequences of sea level rise. I use a simple empirical model to show that sea level rise will exacerbate inequality in flood exposure in Jakarta. Sorting drives this inequality: high-wage individuals seek out flood-safe areas, bidding up prices and pushing low-wage individuals out. Policymakers must weigh these distributional considerations as sea level rise reshapes urban landscapes worldwide.

REFERENCES

Abidin, Hasanuddin, Heri Andreas, Irwan Gumilar, Yoichi Fukuda, Yusuf Pohan, and T. Deguchi. 2011. "Land Subsidence of Jakarta (Indonesia) and Its Relation with Urban Development." *Natural Hazards*, 59(3): 1753–1771.

Berry, Steven, and Philip Haile. 2021. "Foundations of Demand Estimation." In *Handbook of Industrial Organization*. Vol. 4, 1–62. Elsevier. Ed. Kate Ho, Ali Hortaçsu, and Alessandro Lizzeri.

Depsky, Nicholas, Ian Bolliger, Daniel Allen, Jun Ho Choi, Michael Delgado, Michael Greenstone, Ali Hamidi, Trevor Houser, Robert Kopp, and Solomon Hsiang. 2023. "DSCIM-Coastal V1.1: An Open-Source Modeling Platform for Global Impacts of Sea Level Rise." *Geoscientific Model Development*, 16(14): 4331–4366.

Hsiao, Allan. 2023. "Sea Level Rise and Urban Adaptation in Jakarta."

National Capital Integrated Coastal Development Project. 2014. "Master Plan."

Tay, Cheryl, Eric Lindsey, Shi Tong Chin, Jamie McCaughey, David Bekaert, Michele Nguyen, Hook Hua, Gerald Manipon, Mohammed Karim, Benjamin Horton, Tanghua Li, and Emma Hill. 2022. "Sea-Level Rise from Land Subsidence in Major Coastal Cities." *Nature Sustainability*, 5(12): 1049–1057.