Spatial Health: How Smoking Shapes our Neighborhoods

Jesús Bueren[◇] Carla Revilla[*] Siyu Shi[*]

[◇] Católica-Lisbon

[*] EUI

PROSPER, October 2025

Introduction

- In the United states, not only income but also health are highly spatially segregated
  - Within New York City, the life expectancy difference between someone living in Tribeca (Manhattan) versus Brownsville (Brooklyn) is 11 years
- Local spillovers (school funding and peer effects) have been identified as the main drivers of income segregation through human capital accumulation
- On the other hand, we don't know much about the mechanisms driving health segregation
- Health outcomes are to a large extent determined by the health behaviors
- This paper proposes a theory of health segregation that is based on the idea that:
  - Richer individuals are more willing to adopt healthier behaviors
  - Health behaviors pose externalities on the health behaviors of others

This Paper

- A recent empirical literature has identified peer effects play a key role in smoking
  Argys and Rees (2008), Lundborg (2006), Card and Giuliano (2013), Arduini et al (2025)
- We extend previous equilibrium models on neighborhood segregation to incorporate endogenous health outcomes through smoking
  Benabou (1996), Durlauf (1996), Fernandez and Rogerson (1996)
- We calibrate the model using Add Health data and use it to:
  1. Understand why health is segregated space: income vs peer effects
  2. Evaluate the welfare effects of place-based policies: consumption versus health

Overview of the Model

- OLG model
  - Endogenous choices:
    1. Neighborhood
    2. Education
    3. Smoking
  - Individuals derive utility from consumption and being alive:
    - Higher educated individuals have a stronger incentive to avoid smoking as they enjoy a higher consumption flow
  - Local externalities:
    - Health: living in an area with a larger fraction of smokers increases the utility flow from smoking.
    - Education: individuals living in a more affluent neighborhoods (better schools) have larger probabilities of attending college.
  - House prices in each neighborhood adjust in equilibrium.

Preview of the Results

Literature

- Understanding differences in health inequalities across SES through health behavior choices

  Cole et al. (2019); Mahler, Yum (2023); Margaris and Wallenius (2023); Bueren et al. (2025)

  → We focus on the spatial dimension to understand health inequalities

- Recent spatial quantitative literature analyzing the welfare consequences of income segregation

  Chyn and Daruich (2025); Fogli et al. (2025); Eckert and Kleineberg (2025)

  → We analyze how smoking can amplify economic segregation

- Health in space

  Margaris and Wallenius (2025)

  → Write an equilibrium model where spillovers are endogenous to analyze how placed based policies affect welfare

Model

- We build an OLG model with 3 phases:
  1. Adolescence: individuals are born in a given neighborhood and decide education and smoking decision.
  2. Prime: enter the labor market, then have a kid and decide where to live taking as given education and smoking decisions.
  3. Retirement: individuals face survival uncertainty.
- Individuals discount the future at a rate β
- The utility function is given by:

$$u(c) = \frac{c^{1-\gamma} - 1}{1-\gamma} + b,$$

where b is a constant capturing the joy of living.

Adolescence

- A teenager is born in a given neighborhood n.
- Receives an parental transfer a_0
- Decides whether to smoke or not and whether to go to college or not.
- Are heterogeneous in terms of the utility of smoking and going to college
- Utility depends on:
  - Parental background: college ($\mathcal{C}^p$) and smoking ($\mathcal{S}^p$)
  - Neighborhood characteristics: fraction of smokers ($\bar{S}_n$) and average income ($\bar{y}_n$)
  - Idiosyncratic shocks related to college (ϵ_e) and smoking (ϵ_s)

Adolescence

- The teenager solves:

$$V_0(a_0, \mathcal{C}_p, \mathcal{S}_p, \epsilon_c, \epsilon_s, n) = \max_{c,\, a',\, \mathcal{C}\in\{1,0\},\, \mathcal{S}\in\{1,0\}} u(c) + u^{\mathcal{S}}(\mathcal{S}_p, \bar{\mathcal{S}}_n, \epsilon_s) + u^{\mathcal{C}}(\mathcal{C}_p, \bar{y}_n, \epsilon_c) + \beta \mathbb{E} V_1(a', \mathcal{C}, \mathcal{S}, z')$$

$$\text{s.t.} \quad c + a' = a_0 - \mathcal{C}.\tau_{\mathcal{C}} + (1-\mathcal{C})y_0$$

$$\underline{a} \le a' \le 0$$

$$u^{\mathcal{S}}(\mathcal{S}_p, \bar{\mathcal{S}}_n, \epsilon_c) = \mathcal{S}\Big[\alpha_s + \gamma_s \bar{\mathcal{S}}_n^{\rho_s} + \delta_s \mathcal{S}_p + \epsilon_s\Big], \;\; \epsilon_s \sim \mathcal{U}[-\sigma_s, \sigma_s]$$

$$u^{\mathcal{C}}(\mathcal{C}_p, \bar{y}_n, \epsilon_c) = \mathcal{C}\Big[\alpha_c + \gamma_c \bar{y}_n^{\rho_c} + \delta_c \mathcal{C}_p + \epsilon_c\Big], \;\; \epsilon_c \sim \mathcal{U}[-\sigma_c, \sigma_c]$$

where $\epsilon_{\mathcal{C}}$ and $\epsilon_{\mathcal{S}}$ are assumed to be independent.

Prime Age

- The prime age phase is divided into two periods:

  Period 1: individuals enter the labor market.

  - receive an exogenous income depending on education and a persistent shock (z)
  - make a consumption savings decision

  Period 2: on top individuals have a kid, decide where to live, and how much wealth to transfer to the kid.

Prime Age

Period 1

- In period 1 they solve:

$$V_1(a, \mathcal{C}, \mathcal{S}, z) = \max_{c, a'} u(c) + \beta \mathbb{E} V_2(a', \mathcal{C}, \mathcal{S}, z')$$
$$\text{s.t.} \quad c + a' = T(y_1(\mathcal{C}, \mathcal{S}, z)) + (1 + r)a$$

Prime Age

Period 2

- In period 2, the value function is given by:

$$V_2(a, \mathcal{C}, \mathcal{S}, z) = \mathbb{E}\Big[\max_n \Big\{ \tilde{V}_2^n(a, \mathcal{C}, \mathcal{S}, z) + \epsilon_n \Big\} \Big],$$

  where $\tilde{V}_2^n(a, \mathcal{C}, \mathcal{S}, z)$ the neighborhoods-specific value functions.

- We assume ϵ_n to be i.i.d across neighborhoods and type-I extreme value distributed with scale parameter σ_n.

Prime Age

Period 2

- Choice specific value functions are given by:

$$\tilde{V}_2^n(a,\mathcal{C},\mathcal{S},z) = \max_{c,a',a_0^k} u(c) - \mu_n + \beta\delta_2(\mathcal{C},\mathcal{S})\mathbb{E}_z\Big[V_3(a',\mathcal{C},\mathcal{S},z)\Big] + \alpha\mathbb{E}_{\epsilon_\mathcal{C},\epsilon_\mathcal{S}}\Big[V_0(a_0^k,\mathcal{C},\mathcal{S},\epsilon_\mathcal{C},\epsilon_\mathcal{S},n)\Big]$$

$$\text{s.t.} \quad c + a' + a_0^k = T(y_2(\mathcal{C},\mathcal{S},z)) + (1+r)a - P_n$$

μ_n: utility flow from living in neighborhood n

P_n: price of living in neighborhood n

α: altruism parameter

a_0^k: initial assets of the kid

$\delta_2(\mathcal{C},\mathcal{S})$: survival probability at period 2

Retirement

- During retirement (two subperiods), individuals receive a pension which is a function of their education and the last productivity shock
- They make a consumption/saving decision

$$V_3(a,\mathcal{C},\mathcal{S},z) = \max_{c,a'} u(c) + \beta\delta_3(\mathcal{C},\mathcal{S})V_4(a',\mathcal{C},\mathcal{S},z)$$
$$\text{s.t.} \quad c + a' = T(\omega(\mathcal{C},\mathcal{S},z)) + (1+r)a$$
$$V_4(a,\mathcal{C},\mathcal{S},z) = u\Big((1+r)a + T(\omega(\mathcal{C},\mathcal{S},z))\Big)$$

Housing Market

- Housing prices are determined in equilibrium
- The supply of housing is given by:

$$S_n = A_n P_n^{\tau},$$

where τ drives the price elasticity of housing supply, A_n are housing supply shifters

Equilibrium

- A stationary equilibrium in this economy is:
  - a set of neighborhood, education, smoking, parental transfer decisions
  - house prices
  - average smoking rates and income in each neighborhood

  such that:
  - Given, prices and shares, hh decision are optimal
  - Given house prices, housing supply equals housing demand
  - Average smoking rates and income in each neighborhood are consistent with the households decision (endogenous neighborhood quality)

Data

- To estimate the model we are going to use AddHealth data.
- Tracks a cohort of adolescents from grades 7-12 in the US during the 1994-95 school year through several follow-up waves: 1996, 2001-02, 2008-09, and 2016-18.
- Wave 1 includes comprehensive data on the parents of these adolescents (smoking and college).
- Tract-level statistics using Census on income and house prices
- Sample: 6,034 parent-child pairs

Data

- We set the number of neighborhoods equal to three: $n \in \{A, B, C\}$
- We assign kids in Addhealth to neighborhoods depending on the income distribution of the tract with respect to the MSA in which they lived:
  - A: income in the bottom 30%
  - B: income in 30%-90%
  - C: income in the top 10%

Estimation

- We follow a standard 2 step procedure:

  External:

  - Wages $y_t(\mathcal{C}, z)$ (PSID)
  - Survival probabilities $\delta_t(\mathcal{C}, \mathcal{S})$ (HRS)
  - College costs $\tau_{\mathcal{S}} = 60k$
  - House Prices: $P_A = 174k$; $P_B = 253k$; $P_C = 374k$
  - Elasticity of housing supply: 1.75 (Saiz, 2010)
  - Replacement rate: 0.494 (OECD, 2019)

Estimation

Internal: The remaining parameters are estimated internally.

We use Add Health data to match:

- College graduation rates conditional on neighborhood and parental education
- Smoking rates conditional on neighborhood and parental smoking
- Parental transfers
- Value of Statistical Life
- Asset to income ratio
- Average marginal tax rate
- Intergenerational mobility

Estimation

Estimated parameters: smoking and college decision

Parameter	Description	Value	
		$\mathcal{S}$	$\mathcal{C}$
$\alpha_{\mathcal{S}\|\mathcal{C}}$	constant	-1.72	-4.15
$\delta_{\mathcal{S}\|\mathcal{C}}$	parental influence	0.27	0.44
$\gamma_{\mathcal{S}\|\mathcal{C}}$	neighborhood effect	1.55	2.75
$\rho_{\mathcal{S}\|\mathcal{C}}$	curvature	0.37	0.08
$\sigma_{\mathcal{S}\|\mathcal{C}}$	support of $\epsilon_{\mathcal{S}\|\mathcal{C}}$	0.96	0.83

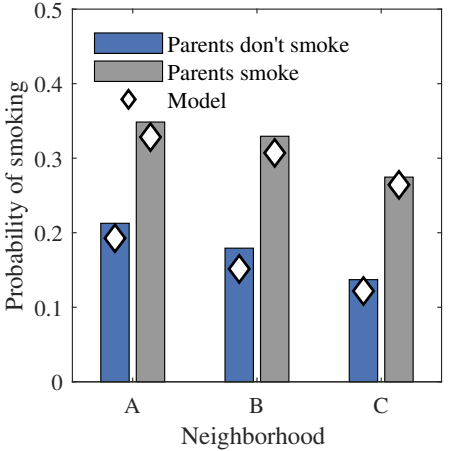
Estimation

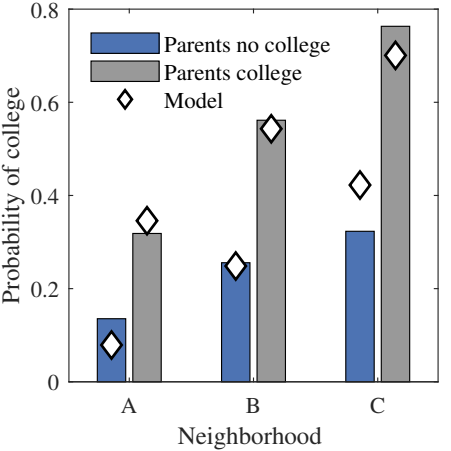
Remaining estimated parameters

Parameter	Description	Value
b	value of life	10.30
β	annual discount factor	0.98
α	altruism parameter	0.68
μ_n	amenity in NBH B, C	0.14, 0.15
σ_n	Gumbel taste shock for NBH	0.02
λ	tax rate	0.77

Estimation

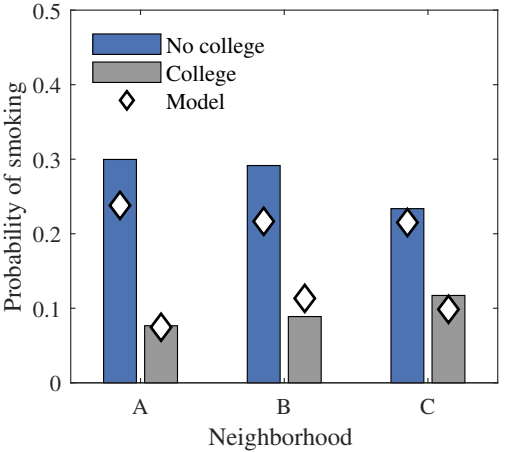
Model Fit: Smoking and College by Neighborhood and Parental characteristics





Estimation

Model Fit: Smoking by Neighborhood and College Decision



Estimation

Model Fit

Moments		Model			Data		
VSL		$ 6.6 million			$ 6 millions		
A/Y at age 40		4.33			4.32		
Cost of raising a kid		462k			301k		
% living in good NBH A,B,C		0.29, 0.62, 0.09			0.3, 0.6, 0.1		
Average tax rate		0.34			0.35		
		$n' = A$	$n' = B$	$n' = C$	$n' = A$	$n' = B$	$n' = C$
Transition probabilities	$Pr(n'\|A)$	0.36	0.60	0.04	0.49	0.48	0.03
	$Pr(n'\|B)$	0.28	0.62	0.10	0.25	0.69	0.06
	$Pr(n'\|C)$	0.17	0.65	0.18	0.18	0.67	0.16

Decompositions

The Role of Smoking Externalities

- To which extent space is segregated in health because the returns of non-smoking are larger for the rich vs local externalities?
  - Set $\gamma_S = 0$ and see how the spatial equilibrium is affected

Benchmark

	NBH A	NBH B	NBH C
Smoking	0.22	0.18	0.14
College	0.10	0.35	0.65
Housing Price	237k	345k	511k
Fraction	0.29	0.62	0.09

No spillover in smoking

	NBH A	NBH B	NBH C
Smoking	0.22	0.20	0.18
College	0.10	0.35	0.66
Housing Price	243k	343k	493k
Fraction	0.31	0.61	0.08

Counterfactuals

Small scale experiments

- What are the welfare consequence of placed-based policies like MTO?
- To which extent increases in welfare are driven by changes in economic vs changes in health outcomes.

COUNTERFACTUALS

Small scale experiments

- We simulate an experiment where we give a voucher so that reduces the prices of living in neighborhood B
- We target individuals in the bottom 10% and living in neighborhood A

Outcomes of the kids of eligible individuals

	Without voucher	With vouchers
Smoking rate	22.6%	20.2%
College rate	17.9%	21.4%

- Compute welfare gains from income and health

Counterfactuals: Large-Scale Experiments

- Small experiments affect a few individuals.
- Large-scale policies reshape entire neighborhoods.
- Composition changes. Neighborhood quality shifts. Prices adjust.
- We ask: How do these equilibrium effects modify the results?

Conclusions

- A lot of work to be done, we will get there.
- To be continued in another Prosper meeting...