Do Public Goods Actually Reduce Inequality?

Micael Castanheira & Giovanni Paolo Mariani

March 18, 2024

Carlos III P.E. & Micro Seminar Series

Introduction

- Growing literature on income and wealth inequality
- Main objective is to determine who-gets-what from each element of GDP
- So far, main effort was to allocate *post-tax disposable income*: $y_i^d = y_i - \tau(y_i)$
  (DINA $\leftarrow$ Atkinson, Piketty, Saez 2011; Piketty, Saez, Zucman 2018, Bozio et al 2022...)
- Alternatively: allocate *final consumption expenditure* (private consumption)

Introduction

- Growing literature on income and wealth inequality
- Main objective is to determine who-gets-what from each element of GDP
- So far, main effort was to allocate *post-tax disposable income*: $y_i^d = y_i - \tau(y_i)$
  (DINA $\leftarrow$ Atkinson, Piketty, Saez 2011; Piketty, Saez, Zucman 2018, Bozio et al 2022...)
- Alternatively: allocate *final consumption expenditure* (private consumption)
- Big issue (30% of GDP): **how to allocate public consumption expenditure?**
  In-Kind Transfers and *Public Spending* (aka STiKs and "public goods and services")
- "Comprehensive income": $cy_i = y_i - \tau(y_i) + g_i$

Introduction

- To be clear, DINA and other research do take public goods into account
- They mainly contrast two *imputation methods*:
  1. Equal distribution (like a *pure public good*: highly redistributive)
  2. Proportional to disposable income (does not affect income distribution)

Introduction

- To be clear, DINA and other research do take public goods into account
- They mainly contrast two *imputation methods*:
    1. Equal distribution (like a *pure public good*: highly redistributive)
    2. Proportional to disposable income (does not affect income distribution)
- DINA opts for method 2

Introduction

- To be clear, DINA and other research do take public goods into account
- They mainly contrast two *imputation methods*:
  1. Equal distribution (like a *pure public good*: highly redistributive)
  2. Proportional to disposable income (does not affect income distribution)

- DINA opts for method 2
- Other research opts for refined and improved method 1
  - Usage surveys
    Smeeding 1984: medical+housing subsidies,
    Gethin 2024: global effects of public goods, building on 1300 national surveys
  - Refined imputations, e.g. #children to proxy benefits from education
    Aaberge et al. 2010; Verbist 2011; Verbist & Forster 2012, 2019 (SILC data)...

Introduction

- To be clear, DINA and other research do take public goods into account
- They mainly contrast two *imputation methods*:
  1. Equal distribution (like a *pure public good*: highly redistributive)
  2. Proportional to disposable income (does not affect income distribution)

- DINA opts for method 2
- Other research opts for refined and improved method 1
  * Usage surveys
    Smeeding 1984: medical+housing subsidies,
    Gethin 2024: global effects of public goods, building on 1300 national surveys
  * Refined imputations, e.g. #children to proxy benefits from education
    Aaberge et al. 2010; Verbist 2011; Verbist & Forster 2012, 2019 (SILC data)...
  * **Method 1** concludes that public goods are highly redistributive

Introduction

"Highly redistributive" ?

- Convincing from a macro point of view
- Conflicts with a lot of micro/meso evidence
  - Educational performance varies strongly across socio-economic groups
  - Chetty's "moving to opportunity" $\rightarrow$ very strong neighborhood effects
  - Life expectancy varies substantially across education & income groups (Case and Deaton 2021, Eggerickx et al. 2018, Renard et al. 2019)
  - Substantial geographic variability (e.g. well-served neighborhoods v medical deserts)

Research question

- Does Samuelsonian definition of public goods distort our vision ?
  - Samuelson (1954, 1955): $g_i = G, \forall i$

Research question

- Does Samuelsonian definition of public goods distort our vision ?
  - Samuelson (1954, 1955): $g_i = G, \forall i$
  - Intense debate in the 1950s and 60s: Tiebout (1956), Williams (1966) ...
    $\rightarrow$ "vote with your feet" $\Leftrightarrow$ public goods have a significant local component

Research question

- Does Samuelsonian definition of public goods distort our vision ?
  - Samuelson (1954, 1955): $g_i = G, \forall i$
  - Intense debate in the 1950s and 60s: Tiebout (1956), Williams (1966) ...
    $\rightarrow$ "vote with your feet" $\Leftrightarrow$ public goods have a significant local component
  - Musgrave 1959: *The principle, that goods are consumed in equal amounts by all, now encounters a spacial limitation*
    - Public goods v. pork barrel politics, (fiscal) federalism, ...

Research question

- Does Samuelsonian definition of public goods distort our vision ?
  - Samuelson (1954, 1955): $g_i = G, \forall i$
  - Intense debate in the 1950s and 60s: Tiebout (1956), Williams (1966) ...
    $\rightarrow$ "vote with your feet" $\Leftrightarrow$ public goods have a significant local component
  - Musgrave 1959: *The principle, that goods are consumed in equal amounts by all, now encounters a spacial limitation*
    - Public goods v. pork barrel politics, (fiscal) federalism, ...
- From conceptualization to measurement of pork barrel v. public good ?
- We propose to use *geography*:
  - Model: how *should* you allocate public goods across population groups?
  - Assessment: *measure* inequality in *supply* of public goods across space (Belgium)

Research angle

- Pork or Public good?
  - What *symptoms* to observe the difference?
  - Geography: public goods located close to target groups

Research angle

- Pork or Public good?
  - What *symptoms* to observe the difference?
  - Geography: public goods located close to target groups

- **Observation 1**: richer people locate close to "good schools" to gain access (Tiebout → Banzhaf & Walsh 2008, Burgess et al. 2019, 2023)

Research angle

- Pork or Public good?
  - What *symptoms* to observe the difference?
  - Geography: public goods located close to target groups

- **Observation 1**: richer people locate close to “good schools” to gain access (Tiebout $\rightarrow$ Banzhaf & Walsh 2008, Burgess et al. 2019, 2023)
- **Observation 2**: most public spending is on similar G&S: require *physical access* (Healthcare: hospitals. Public transport: stops)

Case of Spain according to AI (Google's Gemini)

Public Good	% Budget
Social Security (pensions, unemployment benefits)	1/3
Healthcare	1/5
Education	1/7
Public Debt Service (interest payments)	1/10
National Defense	1/15
Infrastructure (roads, bridges, public transport)	1/20
Law Enforcement & Justice System	1/20
Environmental Protection	1/30
Culture & Research	1/40
Foreign Aid & Development	1/50
Public Administration	1/60
Housing & Urban Development	1/75
Disaster Relief & Emergency Services	1/100
Consumer Protection & Regulation	1/125
Rural Development & Agriculture	1/150

Total: GREEN / (GREEN + RED)	**88%**

Sweden Satisfaction Survey:
Growing differences between rural and urban

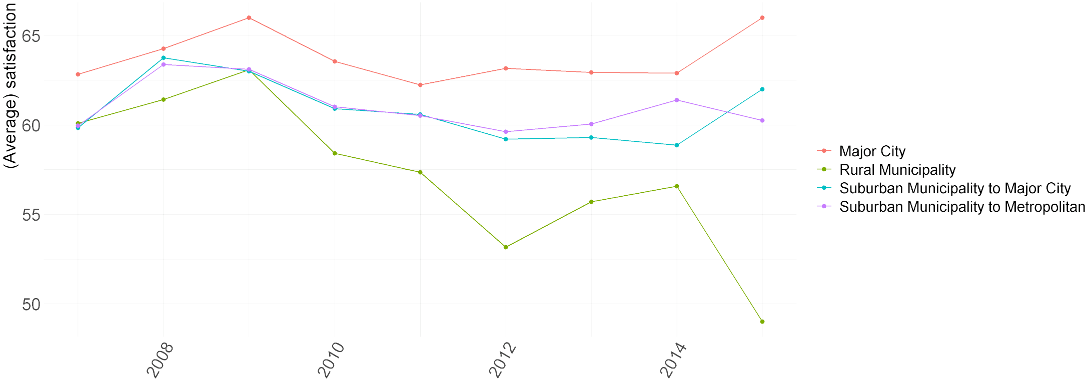


Figure : Average Satisfaction on Public Goods Delivery (Authors' calculations based on Sweden Satisfaction Survey)

Research angle

- Pork or Public good?
  - What *symptoms* to observe the difference?
  - Geography: public goods located close to target groups
- **Observation 3**: econ phenomena (e.g. GDP) measured through geo/satellite data

This paper:

The geographic distribution of public amenities helps reveal actual allocation of public goods to different population groups

This Paper

Conceptualize + actually measure allocation of public good "accessibility"

- Focus is on **supply side**
  - Directly assesses **to whom** government supplies G $\rightarrow$ measure of access
  - Weakness: we cannot measure the demand side
  - Actual inequality combines access & indiv. preferences

This Paper

Conceptualize + actually measure allocation of public good "accessibility"

- Focus is on **supply side**
  - Directly assesses **to whom** government supplies G $\rightarrow$ measure of access
  - Weakness: we cannot measure the demand side
  - Actual inequality combines access & indiv. preferences

Contributions:

- Novel geographical methodology to *measure* allocation of government expenditure
  - Fine-grained analysis of $\approx$20'000 statistical sectors ($\in$ 581 municipalities)
- We find:
  - Surprisingly high inequalities in this allocation
  - Supply $\uparrow$ with density (expected) and with income (a prior unexpected)

Literature 1: On Public Goods

- Concept of “pure public goods”: Samuelson (1954, 1955)
- Intense discussion around purity, universality, and gov’t organization:
  - Tiebout (1956), Banzhaf & Walsh (2008), Chapelle et al. (2023)
  - Williams (1966), Alesina & Spolaore (1997), Bolton & Roland (1997)...
  - Persson, Roland, Tabellini (1997, 2000), Persson & Tabellini (2000, 2003), Lizzeri & Persico (2001, 2004)...
- Productivity of public goods into utility: Jones & Klenow (2016), Gethin (2024)
- Link btw culture and local dvp (nightlights): Michalopoulos&Papaioannou (2013, 2014)

Literature 2: On Inequality

- Inequality measure: National Accounts
  - DINA: Piketty and Saez (2003); Piketty, Saez and Zucman (2018); Bruil (2018); Saez and Zucman (2020); Christl et al. (2020); **Chatterjee, Czajka and Gethin (2021)**; **Chanchel et al. (2021)**; Bozio et al. (2022), **Gethin (2024)**
  - EGDNA: Fesseau and Mattonetti (2013), Zwijnenburg et al. (2017)

- Extended income from Survey data
  - Smeeding et al. (1993); Evandrou et al. (1993); Garfinkel et al. (2006); Aaberge and Langoren (2006); Maricaal et al. (2008); Paulus et al. (2010); Aaberge et al. (2010, 2019); Verbist and Forster (2012, 2019); Andreou (2014); Figari and Paulus (2015); Lustig (2018), Aaberge at al. (2017)

Literature 3: Access to Public services

- Methodologies for Floating Catchment Areas (Geography literature)
  - Radke and Mu (2000); Luo and Wang (2003); Delameter (2013); **Paez et al. (2019)**
- Specialized Literature:
  - Education: Kirby (1979); Parsons et al. (2020); Tigre et al. (2016); Dickerson and McIntosh (2010); Flach et al. (2012); Burgess et al. (2019); Burgess et al. (2023), Pennerstorfer and Pennerstorfer (2019)
  - Healthcare: Ingram et al. (1978), Oliveira et al. (2015), Higgs (2019), Paez et al. (2019), Lindo et al. (2019), Myers (2021), Irlacher et al. (2023)
  - Transportation: Martin and Van Ommeren (2016), Anderson (2014), Chen and Whalley (2012), Gu et al. (2021)

Basic Model
Simple Setup

A social planner has a given budget G on hand:

- She must allocate this budget G across L locations
- Each location $l \in \{1, \ldots, L\}$ has a population n_l with average income y_l
- Indiv. utility: CES w/ public good g_l and private consumption(=income) y_l

$$U(g_l, y_l) = \left(\gamma \left(\frac{g_l}{n_l^{\alpha}}\right)^{\eta} + (1-\gamma) y_l^{\eta}\right)^{1/\eta},$$

  where α allows for congestion ($\alpha = 0$ is public good); γ is weight on g_l

- Elasticity of substitution is $1/(1-\eta)$

Basic Model

Objective function and Optimal allocation

- Efficiency-Maximizing Social Planner:

$$\max_{\mathbf{g}} EF(\mathbf{g}; n_l, y_l) = \sum_l n_l U(g_l, y_l) \text{ s.t. } \sum_l g_l = G$$

- Solution:

$$\left(\frac{\left(\frac{g_l}{n_l^{\alpha}}\right)^{\eta}}{\gamma \left(\frac{g_l}{n_l^{\alpha}}\right)^{\eta} + (1-\gamma) y_l^{\eta}} \right) \Bigg/ \left(\frac{\left(\frac{g_j}{n_j^{\alpha}}\right)^{\eta}}{\gamma \left(\frac{g_j}{n_j^{\alpha}}\right)^{\eta} + (1-\gamma) y_j^{\eta}} \right) = \left(\frac{n_l}{n_j}\right)^{\frac{\eta}{1-\eta}(1-\alpha)}$$

- Implications: $\frac{g_l}{n_l^{\alpha}}$ $\uparrow$ in n_l fixing y_l and $\uparrow$ in y_l fixing n_l $(g_l/g_j = y_l/y_j, \ \forall \eta)$

Basic Model
Objective function and Optimal allocation

Proposition

For homothetic utility functions, ceteris paribus (hence, fixing n_l): $g_l/y_l = constant$.
Implication: supply of g_l is inequality neutral.

Direct application of consumption theory: expansion path of c and g is linear in income.

Basic Model

Objective function and Optimal allocation

Proposition

For homothetic utility functions, ceteris paribus (hence, fixing n_l): $g_l/y_l = constant$.
Implication: supply of g_l is inequality neutral.

Direct application of consumption theory: expansion path of c and g is linear in income.

Reason here: relative price of g_l versus c_l $(= y_l)$ does not vary with y_l.

$\Rightarrow$ **Model with "Tiebout goods" in line with DINA's hypothesis that G leaves inequality unchanged...**

Basic Model: alternative objective function

Objective function and Optimal allocation

- Equity-Maximizing Social Planner:

$$\max_{\mathbf{g}} EQ(\mathbf{g}; n_l, y_l) = \sum_l -n_l(U(g_l, y_l) - \bar{U})^2 \text{ s.t. } \sum_l g_l = G$$

- Unconstrained solution is to equalize $U(g_l, y_l)$ across locations:

$$\gamma \left(\frac{g_l}{n_l^\alpha}\right)^\eta + (1-\gamma)y_l^\eta = \gamma \left(\frac{g_j}{n_j^\alpha}\right)^\eta + (1-\gamma)y_j^\eta$$

- Implications: $\frac{g_l}{n_l^\alpha}$ $\rightarrow$ in n_l fixing y_l and $\downarrow$ in y_l fixing n_l

How can we test such a model?

Long-term objective is:

$$cy_i = y_i - \tau(y_i) + g_i$$

Today's focus is on inequality of g_l / n_l:

1. Collect data on geographic allocation of public amenities and of population
2. E2SFCA method (adapted): determine *catchment area* of each amenity
3. E2SFCA method: how does access decrease with distance? ("impedance" of geography)
4. E2SFCA method: measure who benefits from which public amenity
5. Back to econ: Lorenz curves for g_l / n_l & correlations with income...

Dataset

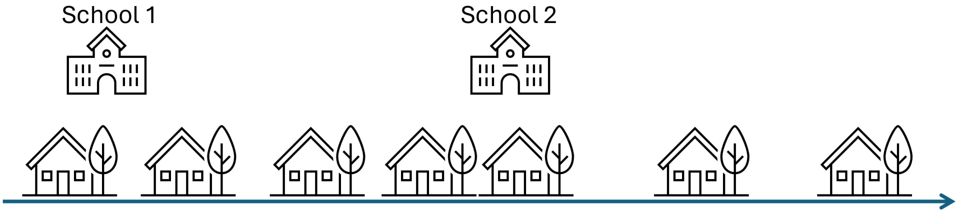
1. Geolocation of public goods (from official statistics and OpenStreetMaps)

   ▸ Schools; ▸ hospitals; ▸ public transport; ▸ stations;
   ▸ highways; ▸ elderly care services; ▸ police stations

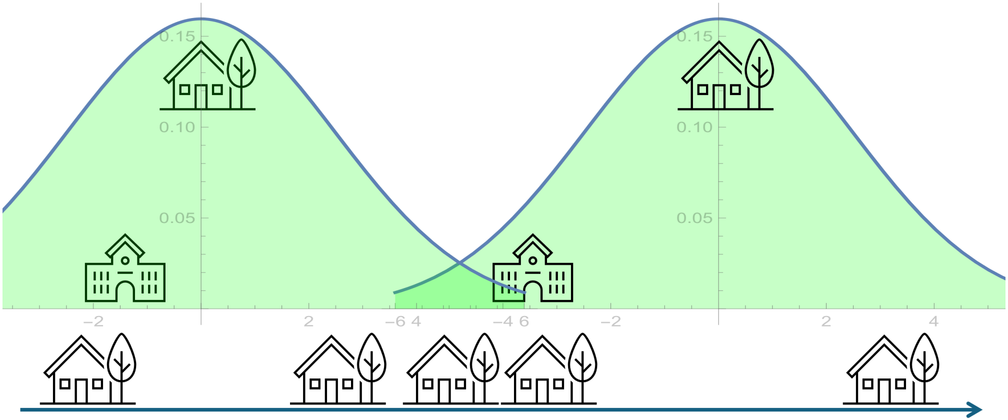
2. Socio-economic data: statistical sector level ($\approx$ 20'000, for 581 municipalities), from STATBEL ▸ Summary Statistics

   * Median and Average Net Taxable Income ▸ Maps
   * Population and Age structure

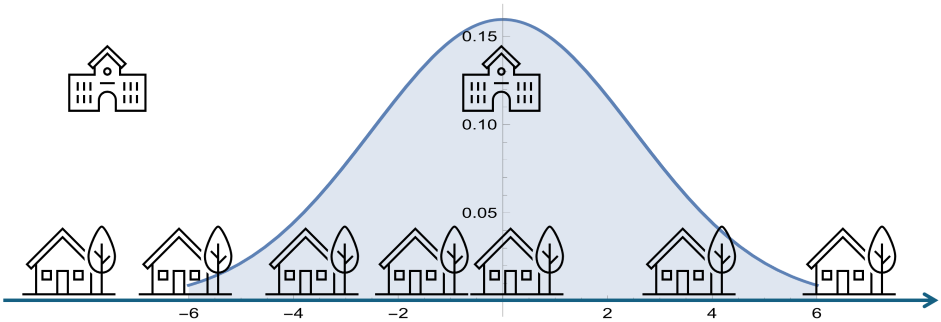
E2SFCA Methodology in Pictures



E2SFCA Methodology in Pictures

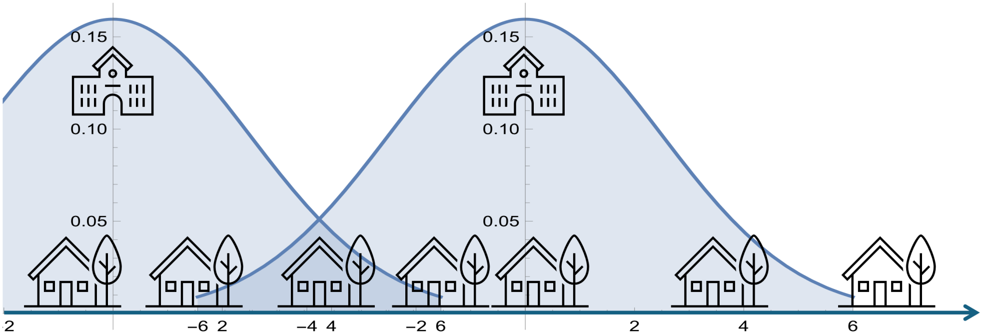


E2SFCA Methodology in Pictures



Burgess et al (2019, 2023): British parents list more schools in dense environments.
When less dense: 1st school is closest in >70% of cases

E2SFCA Methodology in Pictures



Enhanced Two-Step Floating Catchment Areas (E2SFCA)

We compute an **accessibility index** at the statistical sector level in **two steps** (2SFCA)

1. Compute level of service (L_j) for each amenity $\rightarrow$ Supply (e.g., beds in a hospital) over the demand by the **weighted** surrounding population in its catchment area
   ▸ 1st and 2nd step
2. Compute accessibility (▸ A) by summing the level of services in each stat. sectors catchment area

The accessibility measure has nice and intuitive features:

- It **increases** if the number of amenities $\uparrow$ in the stat. sector catchment area
- It **decreases** if the weighted number of people $\uparrow$ within amenity's catchment area

Construction of the steps for the stepwise function

We compute **travel time** ($\approx$ cost) from each stat. sectors to each public amenities

- Travel-time matrix NxM with N stat. sectors and M public amenities
- We consider the J closest amenities from each stat. sectors (Matrix NxJ)
- Thresholds of travel time distances (d_{ij}) computed by quantiles analysis of this matrix
- We use the quantiles to calibrate the Gaussian function (Paez et al. 2019)
- Gaussian densities return the *Impedance Matrix* for each distance

Accessibility index: Schools

Preliminaries: #schools, unweighted
(impedance = 1)

- School count within C.A. (99% have 5+)
  (max travel time = 35min by car)
- Brussels and Flanders display the largest values
- Wallonia has several deprived areas

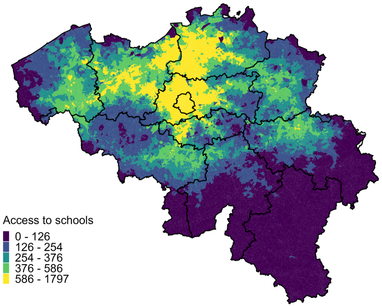


Figure : #Schools within C.A.

Accessibility index: Schools

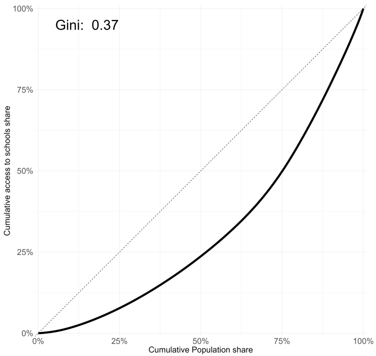


Figure : Lorenz Curve #Schools in Catchment Area

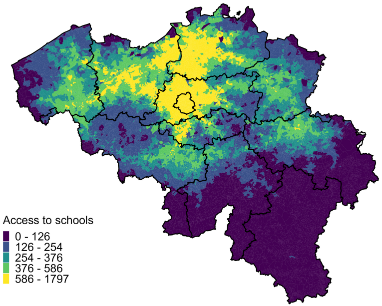


Figure : #Schools within C.A.

Accessibility index: Schools

Step 1: #schools, weighted by travel time (impedance ≤ 1)

- Map remains very similar (lower values)
- Brussels and Flanders still display the largest values
- Wallonia still shows deprived areas

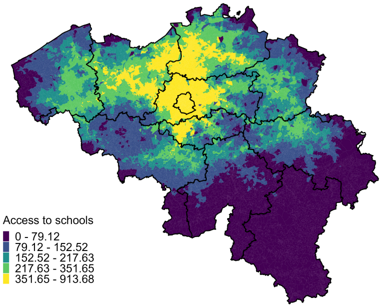


Figure : Travel-Time-Weighted #Schools within C.A.

Accessibility index: Schools

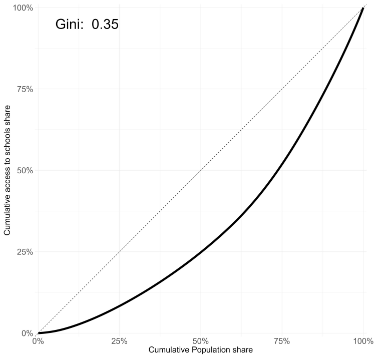


Figure : Lorenz Curve weighted #schools in C.A.

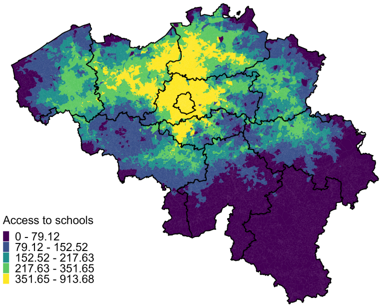


Figure : Travel-Time-Weighted #Schools within C.A.

Accessibility index: Schools

Step 2: pop'n-&-time weighted #schools
= **Accessibility Index** using the E2SFCA

- School access per pop'n aged 0-15:
  - boosts lower-density areas
  - depresses high-density areas
  - $A.I. \in [0.49;\ 4.79]$ schools/1'000 pupils
- Flanders $>$ Brussels $>>$ Wallonia
- Correlation with income = 0.05

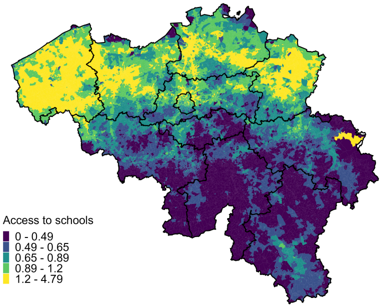


Figure : Accessibility to Education

Accessibility index: Schools

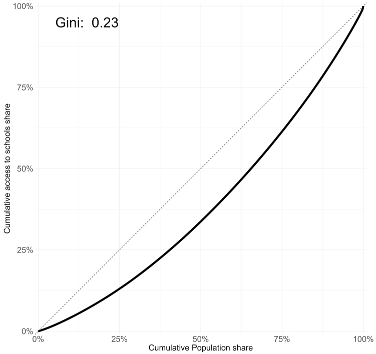


Figure : Lorenz Curve Accessibility to Education

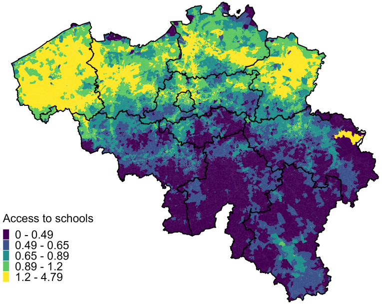


Figure : Accessibility to Education

Accessibility index: Hospital beds

Takeaway: $\neq$'l distribution of hospital beds
(3+ hosp. Max = 60min by car)

- $A.I.(hosp) \in [0;\ 8.06]$ beds/50K people
- Brussels & Flanders $>>$ Wallonia
- Correlation with income $= 0.13$
- Correlation with $A.I.(schools) = 0.60$

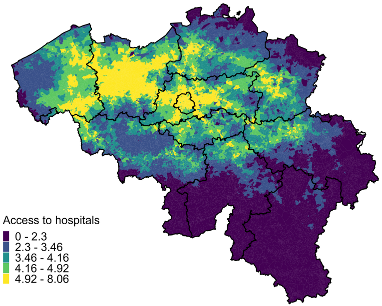


Figure : Accessibility to Healthcare

Accessibility index: Hospital beds

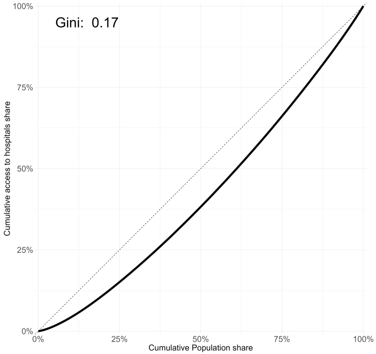


Figure : Lorenz Curve Accessibility to Healthcare

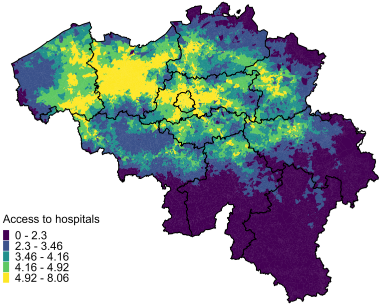


Figure : Accessibility to Healthcare

Accessibility index: Public Transport

Very unequal public transport distribution

- Thick right tail: high concentration
- Cities & West Flanders: large A.I.
- 4th quintile: both in Wallonia and in Flanders
- Correlation with income is -0.12

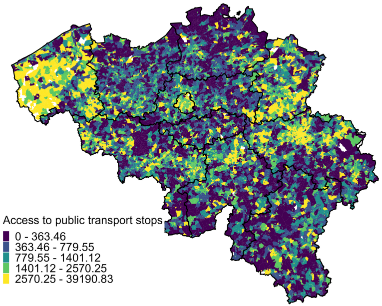


Figure : Accessibility to Public Transport

Accessibility index: Public Transport

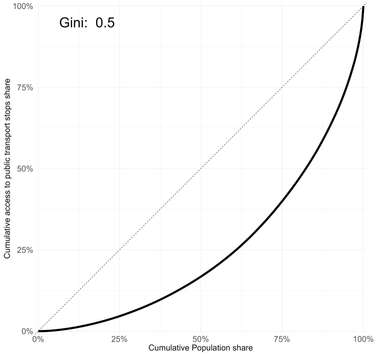


Figure : Lorenz Curve A.I.(Public Transport)

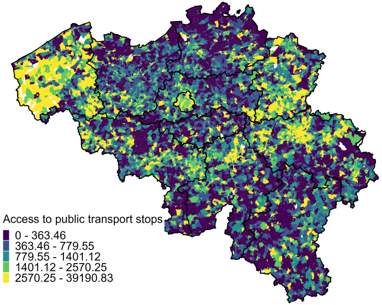


Figure : Accessibility to Public Transport

Accessibility index: Security (Police stations)

Quite a different picture for police stations

- 5th quintile prevalent in South-eastern Wallonia
- Brussels and much of Flanders in 4th quintile
- Correlation with income is again -0.12

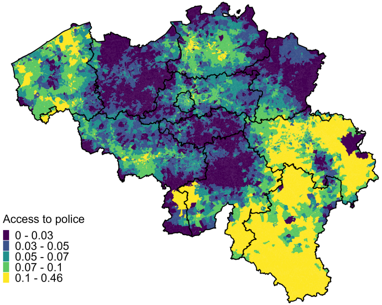


Figure : Accessibility to Security/Police

Accessibility index: Security (Police stations)

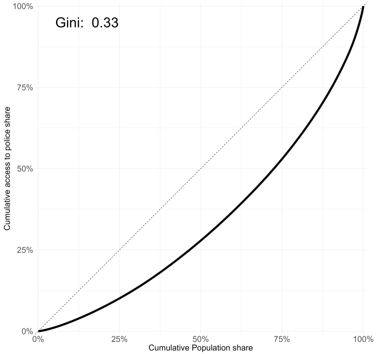


Figure : Lorenz Curve A.I.(Police)

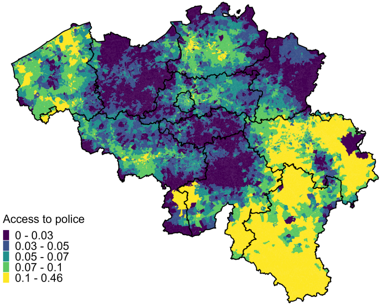


Figure : Accessibility to Security/Police

Regressions: Log Number of Schools (#) & Access to Schools (Ai)

	Schools (#)	Schools (ai)	Schools (#)	Schools (ai)	Schools (#)	Schools (ai)	Schools (#)	Schools (ai)
(Intercept)	0.259*	-2.190***	0.380***	-2.417***	-0.259**	-2.603***	0.778***	-2.175***
	(0.107)	(0.060)	(0.110)	(0.062)	(0.094)	(0.060)	(0.093)	(0.060)
Log Avg Income	1.055***	0.380***	0.981***	0.519***	1.026***	0.532***	0.882***	0.435***
	(0.028)	(0.016)	(0.033)	(0.018)	(0.028)	(0.018)	(0.028)	(0.018)
Log pop. density	0.260***	0.093***	0.261***	0.091***	0.141***	0.056***	0.144***	0.058***
	(0.004)	(0.002)	(0.004)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)
Inc. Ineq.			0.001***	-0.002***	0.0004	-0.003***	0.001***	-0.002***
			(0.0003)	(0.0002)	(0.0003)	(0.0002)	(0.0003)	(0.0002)
Degree of Urb					0.703***	0.205***		
					(0.009)	(0.006)		
URB. FE							Y	Y
Num.Obs.	16593	16593	16593	16593	16593	16593	16593	16593
R2	0.244	0.114	0.245	0.124	0.460	0.191	0.476	0.217

Regressions: Log Number of Schools (#) & Access to Schools (Ai)

Rosults robust when weighting by population per stat. sector

	Schools (#)	Schools (ai)	Schools (#)	Schools (ai)	Schools (#)	Schools (ai)	Schools (#)	Schools (ai)
(Intercept)	-0.081	-2.286***	0.397***	-2.771***	-0.043	-2.872***	0.637***	-2.639***
	(0.101)	(0.063)	(0.108)	(0.066)	(0.098)	(0.065)	(0.098)	(0.065)
Log Avg Income	1.003***	0.399***	0.744***	0.661***	0.897***	0.696***	0.788***	0.619***
	(0.024)	(0.015)	(0.031)	(0.019)	(0.029)	(0.019)	(0.029)	(0.019)
Log pop. density	0.349***	0.109***	0.347***	0.111***	0.214***	0.081***	0.217***	0.083***
	(0.004)	(0.002)	(0.004)	(0.002)	(0.004)	(0.003)	(0.004)	(0.003)
Inc. Ineq.			0.004***	-0.004***	0.002***	-0.005***	0.003***	-0.004***
			(0.0003)	(0.0002)	(0.0003)	(0.0002)	(0.0003)	(0.0002)
Degree of Urb					0.516***	0.118***		
					(0.009)	(0.006)		
URB. FE							Y	Y
Num.Obs.	16593	16593	16593	16593	16593	16593	16593	16593
R2	0.317	0.107	0.324	0.131	0.442	0.152	0.450	0.165

Regressions: Access to Hospitals (Ai)

	Hospitals (ai) 1	Hospitals (ai) 2	Hospitals (ai) 3	Hospitals (ai) 4	Hospitals (ai) 1 (w)	Hospitals (ai) 2 (w)	Hospitals (ai) 3 (w)	Hospitals (ai) 4 (w)
(Intercept)	-1.172***	-1.141***	-1.444***	-0.920***	-1.033***	-0.891***	-1.092***	-0.764***
	(0.057)	(0.059)	(0.053)	(0.052)	(0.051)	(0.054)	(0.050)	(0.050)
Log Avg Income	0.471***	0.452***	0.473***	0.390***	0.402***	0.325***	0.395***	0.333***
	(0.015)	(0.018)	(0.016)	(0.016)	(0.012)	(0.016)	(0.015)	(0.015)
Log pop. density	0.115***	0.116***	0.059***	0.060***	0.138***	0.137***	0.076***	0.078***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Inc. Ineq.		0.0003*	-0.00008	0.0002+		0.001***	0.0003+	0.0007***
		(0.0002)	(0.0001)	(0.0001)		(0.0002)	(0.0002)	(0.0002)
URB. FE				Y				Y
Num.Obs.	16593	16593	16593	16593	16593	16593	16593	16593
R2	0.180	0.181	0.361	0.381	0.223	0.226	0.337	0.348

Regressions: Access to Police stations (Ai)

	Police (ai) 1	Police (ai) 2	Police (ai) 3	Police (ai) 3	Police (ai) 1 (w)	Police (ai) 2 (w)	Police (ai) 3 (w)	Police (ai) 4 (w)
(Intercept)	-1.630***	-1.553***	-1.457***	-1.925***	-2.357***	-2.113***	-2.141***	-2.368***
	(0.091)	(0.094)	(0.094)	(0.093)	(0.092)	(0.099)	(0.099)	(0.097)
Log Avg Income	-0.387***	-0.434***	-0.441***	-0.283***	-0.272***	-0.404***	-0.394***	-0.221***
	(0.023)	(0.028)	(0.028)	(0.028)	(0.022)	(0.029)	(0.029)	(0.029)
Log pop. density	0.024***	0.025***	0.043***	0.040***	0.063***	0.062***	0.054***	0.049***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)
Inc. Ineq.		0.0008**	0.0009***	0.0003		0.002***	0.002***	0.0008**
		(0.0003)	(0.0003)	(0.0003)		(0.0003)	(0.0003)	(0.0003)
URB. FE				Y				Y
Num.Obs.	16583	16583	16583	16583	16583	16583	16583	16583
R2	0.025	0.026	0.034	0.068	0.049	0.052	0.053	0.086

Regressions: Access to Transport (Ai)

	Transport (ai)	Transport (ai)	Transport (ai)	Transport (ai)	Transport (w)	Transport (w)	Transport (w)	Transport (w)
(Intercept)	8.586***	8.570***	8.481***	8.270***	9.055***	9.354***	9.226***	9.080***
	(0.159)	(0.165)	(0.165)	(0.166)	(0.157)	(0.168)	(0.168)	(0.167)
Log Avg Income	-0.734***	-0.724***	-0.717***	-0.582***	-0.903***	-1.064***	-1.019***	-0.814***
	(0.041)	(0.049)	(0.049)	(0.050)	(0.037)	(0.049)	(0.049)	(0.050)
Log pop. density	0.145***	0.144***	0.128***	0.126***	0.145***	0.143***	0.104***	0.098***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)	(0.007)
Inc. Ineq.		-0.0002	-0.0003	-0.0008+		0.003***	0.002***	0.0006
		(0.0005)	(0.0005)	(0.0005)		(0.0005)	(0.0005)	(0.0005)
Degree of Urb			0.094***				0.153***	
			(0.015)				(0.015)	
URB. FE				Y				Y
Num.Obs.	16257	16257	16257	16257	16257	16257	16257	16257
R2	0.078	0.078	0.080	0.088	0.119	0.120	0.126	0.141

Gethin (2024):
Imputation of Public goods through estimated consumption

$$cy_i = y_i - \tau(y_i) + g_i$$

- DINA used to consider two imputation methods:
  - Equal distribution (pure, universal, public goods)
  - Proportional to income (neutral effect on income distribution)
- Gethin (2024) combines 1300 national surveys to estimate *G's consumption*

$$g(y_i) = \underbrace{\sum_j G^j}_{\text{Cost function j}} \times \overbrace{\gamma^j(y_i)}^{\text{progressivity}} \times \underbrace{\theta^j(y_i)}_{\text{productivity}}$$

- Our approach: progressivity of **provision** (closer to actual spending per hh)

Imputation of Public goods through estimated provision

- How (un)equal is the allocation of StiKs ($\perp \; y_I$)?
  - 1st: neutral distribution (blue)
  - 2d: geographical approach (green)

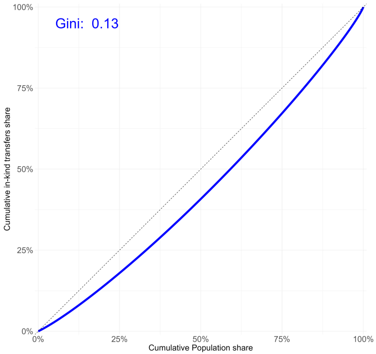


Figure : Lorenz Curve, STiKS distributed as pre-tax income

Imputation of Public goods through estimated provision

- How (un)equal is the allocation of StiKs ($\perp$ y_l)?
  - 1st: neutral distribution (blue)
  - 2d: geographical approach (green)

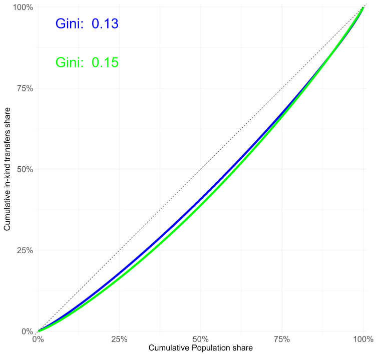
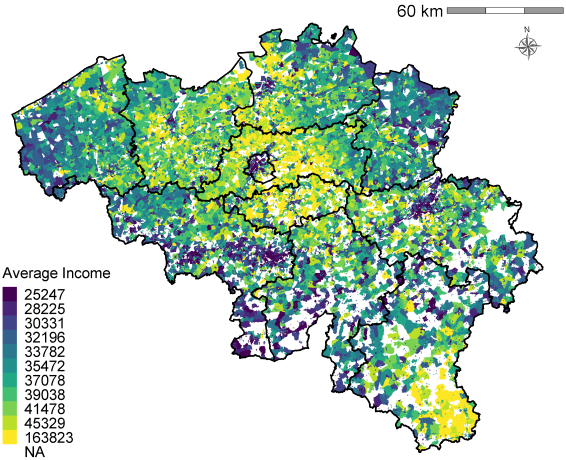


Figure : Lorenz Curve, STiKS distributed as pre-tax income

Geographical distribution of (taxable) income

We now move from estimating distribution of provision to estimating redistribution



Imputation of STiKs combined with disposable income

- **Combine data on y_l with g_l^t**
- GNP = C (≈50.7%) + I (≈24%) + G (≈24%) + (X-M) (≈-1.6%)
  - In GNP data, HH income ≈ consumption
- We rescale: $C = \sum_l n_l y_l = 0.507Y$ and $G = \sum_l g_l = 0.24Y$
- And compute the distribution of:

$$y_l^{ext'd} = y_l + g_l \times \frac{.24C}{.507G\ n_l}$$

  - Check: $\sum_l n_l y_l^{ext'd} = \sum_l n_l\ y_l + \frac{.24C}{.507G}\sum_l g_l = \frac{C}{.507}(.507 + .24) = Y \times (.507 + .24)$

Imputation of STiKs combined with disposable income

- **Distribution of $y_l^{ext'd}$**

- In **red** we have *extended income* following the equality assumption

- In **blue** we have *extended income* following the neutral assumption

- In **green** we have *extended income* following our geographical approach

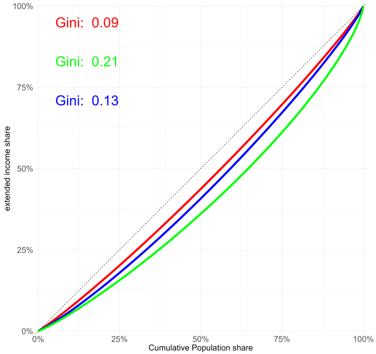


Figure Lorenz Curve, extended income

Conclusion

- Trad'l measures of income distribution do not account for allocation of public goods
- Given the local dimension of STiKs, geography can tell us something

Conclusion

- Trad'l measures of income distribution do not account for allocation of public goods
- Given the local dimension of STiKs, geography can tell us something
  - We find: provision of main two public goods $\uparrow$ in income and density
  - Unequal spatial distribution of public services ...
  - ... translates into an unequal spatial distribution of government expenditures
- 61.5% increase in Geo'l Gini if we scale STiKs to match C and G in GNP data
- Worse if we focus on hospitals and/or schools

Conclusion

- Trad'l measures of income distribution do not account for allocation of public goods
- Given the local dimension of STiKs, geography can tell us something
  * We find: provision of main two public goods $\uparrow$ in income and density
  * Unequal spatial distribution of public services ...
  * ... translates into an unequal spatial distribution of government expenditures
- 61.5% increase in Geo'l Gini if we scale STiKs to match C and G in GNP data
- Worse if we focus on hospitals and/or schools
- Results are not driven by miscounting population

Limitations and next steps

- Main limitations:
  - Gini results sensitive to size of Catchment Areas
  - Lack detailed income data and consumption by inc. group at local level

Limitations and next steps

- Main limitations:
  - Gini results sensitive to size of Catchment Areas
  - Lack detailed income data and consumption by inc. group at local level
- Next steps:
  - Consumption very $\neq$ from Allocation: reconciliation?
  - Use cost of commuting to determine impedance?
  - Robustness: Outliers + cities vs rural areas?
  - Policy implications?

Thank you!

mcasta@ulb.be & gmariani@ulb.be

Appendix: Location of Schools

Education: 6.710 (French/Dutch) schools

- Addresses on French and Dutch schools from Federation Wallonie-Bruxelles (FWB) and the Flemish and German Ministry of Education

- Pre-elementary, Primary and Secondary schools

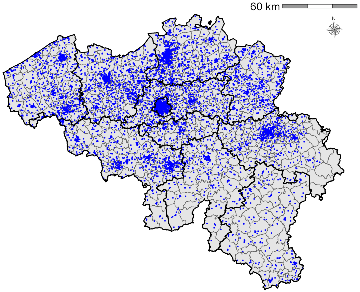


Figure : Location of schools in Belgium

Appendix: Location of Schools

Education: 6.710 (French/Dutch) schools

- Addresses on French and Dutch schools from Federation Wallonie-Bruxelles (FWB) and the Flemish and German Ministry of Education

- Pre-elementary, Primary and Secondary schools

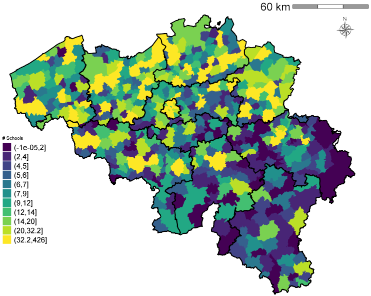


Figure : Number of schools per municipality in Belgium

Appendix: Location of Hospitals

Healthcare: 257 Hospitals (General, University and Psychiatric)

- General hospitals addresses from the Federal Public Service (FPS) Health, Food Chain Safety and Environment
- Majority of them have also emergency service and SMUR ($\approx$ ambulances)
- We include the number of beds and functions provided

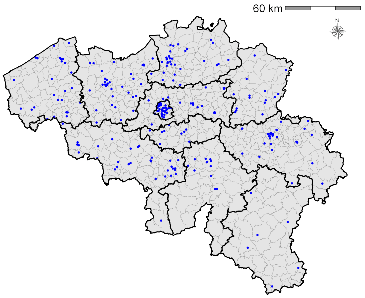


Figure : Location of hospitals

Appendix: Location of Hospitals

Healthcare: 257 Hospitals (General, University and Psychiatric)

- General hospitals addresses from the Federal Public Service (FPS) Health, Food Chain Safety and Environment
- Majority of them have also emergency service and SMUR ($\approx$ ambulances)
- We include the number of beds and functions provided

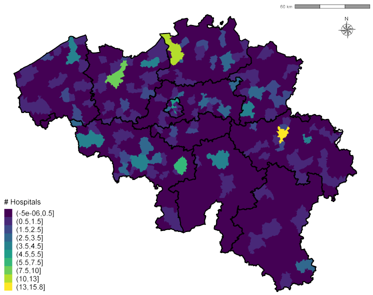


Figure : Number of hospitals per municipality in Belgium

Appendix: Location of Public Transport and Train Stations

- Public Transport: GTFS data (on Bus, Tram and Metro) from MVIB - STIB (Bruxelles), TEC (Wallonie) and De Ljin (Flanders) (74.369 stops)
- Railway Stations: Geolocation of train stations (568) from iRail project

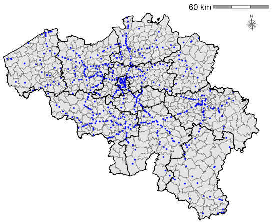


Figure : Location of train stations in Belgium

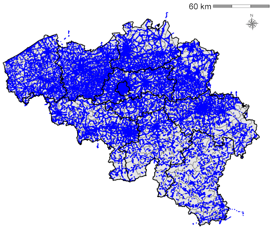


Figure : Location of public transport stops in Belgium

Appendix: Location of Public Transport and Train Stations

- Public Transport: GTFS data (on Bus, Tram and Metro) from MVIB - STIB (Bruxelles), TEC (Wallonie) and De Ljin (Flanders) (74.369 stops)
- Railway Stations: Geolocation of train stations (568) from iRail project

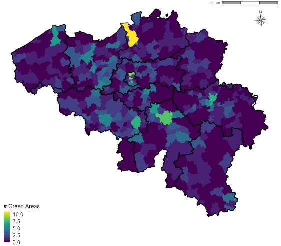


Figure : Number of train stations per municipality in Belgium

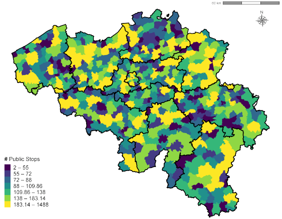


Figure : Number of public transport stops per municipality in Belgium

Appendix: Location of Highways Entrances

- Highways entrances: geolocation of highways networks from OSM
- We then select the point where OSM indicates the presence of an entrance (7.181 links)

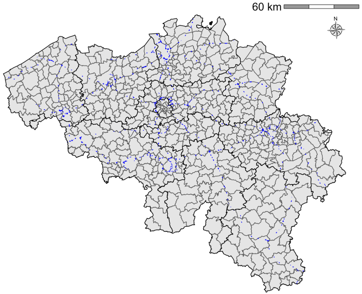


Figure : Location of highways entrances in Belgium

Appendix: Location of Highways Entrances

- Highways entrances: geolocation of highways networks from OSM
- We then select the point where OSM indicates the presence of an entrance (7.181 links)

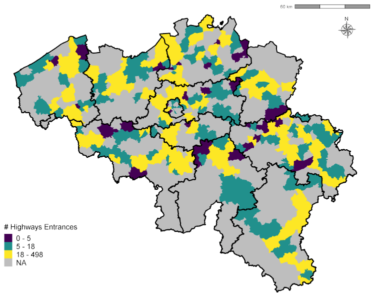


Figure : Number of highway entrances per municipality in Belgium

Appendix: Location of Police Stations

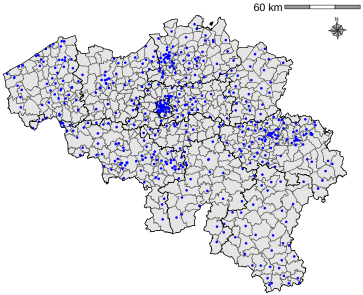


Figure : Location of police stations in Belgium

Appendix: Location of Police Stations

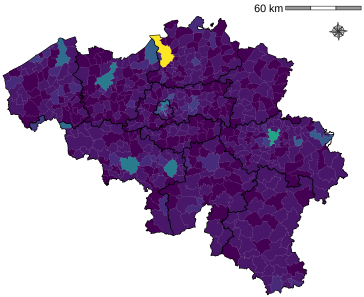


Figure : Number of police stations per municipality in Belgium

Spatial Distribution of Income and Income inequality in Belgium

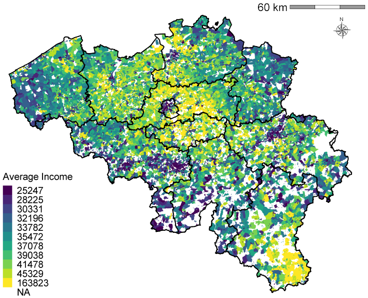


Figure : Average Net Taxable Income, statistical sector level (Belgium)

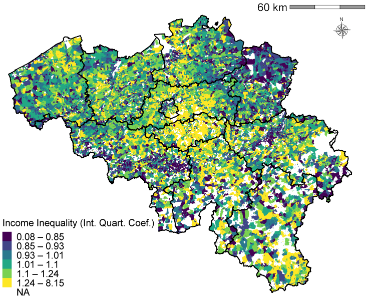


Figure : Income Inequality, statistical sector level (Belgium)

Bivariate Map

Average Taxable Income and Income Inequality (Int. Quart. Coefficient)

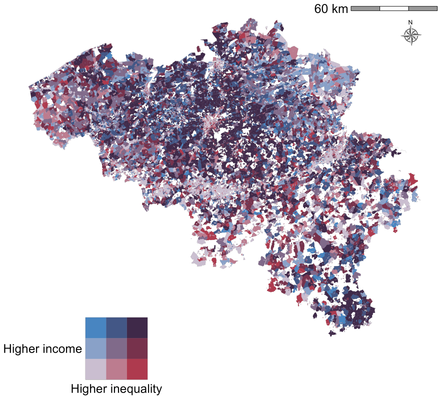


Figure : Bivariate map of Income and Income Inequality, Belgium

Summary Statistics

Table Summary Statistics

	N	NUnique	Mean	SD	Min	Max
Average Net Taxable Income	16595	12503	35202.32	8417.47	7884	163823
Median Net Taxable Income	17575	11491	27117.82	5470.48	1677	83197
Area (km^2)	19794	19792	1.55	2.70	0.01	58.51
Income Inequality (Int. Coefficient)	17575	17571	1.05	0.25	0.08	8.15
People with > 65 years	19794	771	108.85	138.49	0.00	1805.00
People between 15 and 64 years	19794	1950	374.76	469.22	0.00	5475.00
People with < 15 years	19794	767	98.50	137.41	0.00	2506.00
Total Population	19794	2640	582.55	721.75	0.00	8254.00
Male Population	19781	1593	286.84	355.03	0	4212
Female Population	19781	1629	296.01	368.01	0	4356
Population Density (km^2)	19781	19287	1699.15	3278.40	0.00	46681.75

All the variables are collected at the statistical sector level.

1st step 2SFCA, Level of service

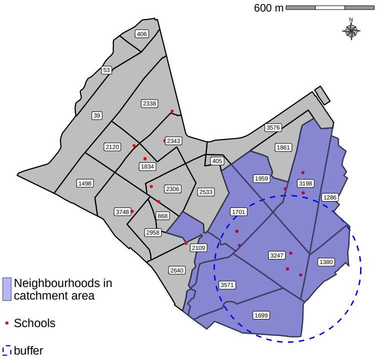


Figure : 1st step, amenity catchment area

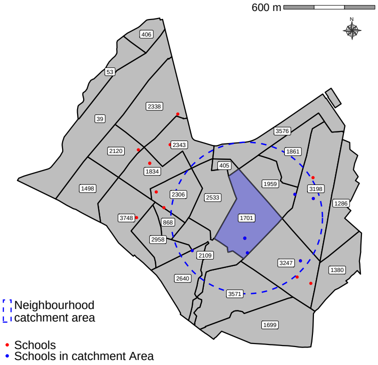


Figure : 2nd step, stat. sect. catchment area

Results from 2SFCA

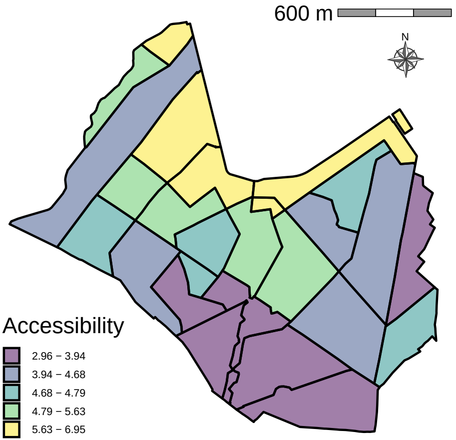


Figure Accessibility to schools, Saint-Gilles

Two-Step Floating Catchment Areas (2SFCA)

$$W(d_{ij}) \leq d_0 = \begin{cases} 1 & \text{if } d_{ij} \leq d_o \\ 0 & \text{if } d_{ij} > d_0 \end{cases} \quad \text{for stat. sector } i \text{ and amenity } j$$

1st step: compute **Level of Service** (L_j):

- Demand (D_j) is given by $D_j = \sum_i D_{ij} = \sum_i P_i W(d_{ij} \leq d_0)$
- Supply S of the service at location j is dived by D_j
- $L_j = \frac{S_j}{D_j} = \sum_i L_{ij}$

Two-Step Floating Catchment Areas (2SFCA)

$$W(d_{ij}) \leq d_0 = \begin{cases} 1 & \text{if } d_{ij} \leq d_o \\ 0 & \text{if } d_{ij} > d_0 \end{cases} \quad \text{for stat. sector } i \text{ and amenity } j$$

1st step: compute **Level of Service** (L_j):

- Demand (D_j) is given by $D_j = \sum_i D_{ij} = \sum_i P_i W(d_{ij} \leq d_0)$
- Supply S of the service at location j is dived by D_j
- $L_j = \frac{S_j}{D_j} = \sum_i L_{ij}$

2nd step: accessibility index (A_i) sums all L_j:

- $A_i = \sum_j L_j W(d_{ij} \leq d_0)$

Allowing for the decay effect of distance

Introducing decay effect of distance: ***Enhanced*** *Two-Step Floating Catchment Areas* (E2SFCA)

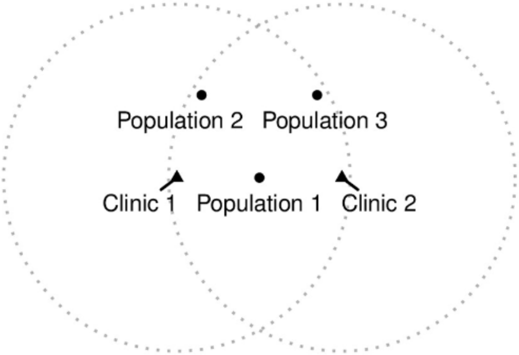
$$W(d_{ij}|d_1, d_2, ..., d_R) = \begin{cases} k_1 & \text{if } d_{ij} \leq d_1 \\ k_2 & \text{if } d_1 < d_{ij} \leq d_0 \\ ... & \\ k_{R-1} & \text{if } d_{R-1} < d_{ij} \leq d_R \\ 0 & \text{if } d_{ij} > d_R \end{cases}$$

- Demand becomes $\rightarrow D_j = \sum_i D_{ij} = \sum_i P_i W(d_{ij}|d_1, d_2, ..., d_R)$
- Accessibility is $\rightarrow A_i = L_j W(d_{ij}|d_1, d_2, ..., d_R)$

Addressing the "inflation" issue

Inflation happens when D_j or L_j are overestimated

- Some stat. sectors contribute to the level of demand for multiple amenities
- Amenities can be counted multiple times if they are "caught" by multiple stat. sectors



Addressing the inflation issue

To address inflation we adopt the modification to the E2SFCA by Paez et al. (2019)

- Main assumption: Population in a given stat. sector is indifferent between two amenities of the same type (if they are inside the same buffer)
- We can assume population sort proportionally (e.g., 50-50 if there are 2)
- This leads to the following weights for stat sector i and amenity j

$$W_{ij}^{i} = \frac{W_{ij}}{\sum_j W_{ij}} \quad \text{and} \quad W_{ij}^{j} = \frac{W_{ij}}{\sum_i W_{ij}} \quad \text{with:}$$

$$\sum_j W_{ij} = 1 \quad \text{and} \quad \sum_i W_{ij} = 1 \quad ; \quad \sum_j P_i W_{ij}^{i} = P_i \quad \text{and} \quad \sum_i L_j W_{ij}^{j} = L_j$$

DINA vs Geography

- DINA $\longrightarrow$ Govt. Consumption distributed:
  - Neutral: Gini remains the same
  - Max equality: Gini decreases to 0.09

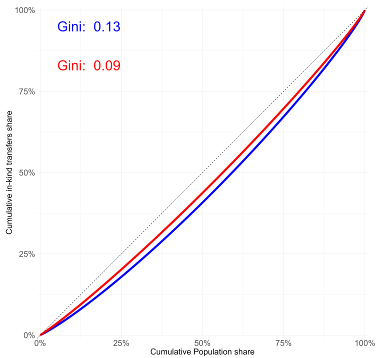


Figure : Lorenz Curve, Average Taxable Income and Extended Income (gov. spending distributed in an egalitarian way)

DINA vs Geography

- DINA $\rightarrow$ Govt. Consumption distributed:
  - Neutral: Gini remains the same
  - Max equality: Gini decreases to 0.09
  - Geography $\rightarrow$ Govt. Consumption distributed:
    - Schools accessibility $\rightarrow$ Gini rises to 0.209 when we impute G as school AI (60.7% increase)

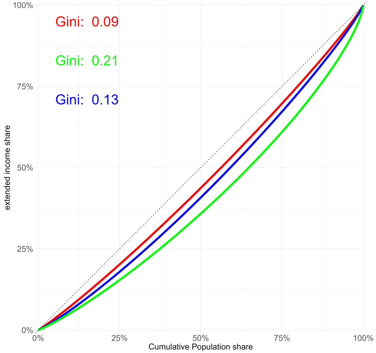


Figure : Lorenz curve Extended Income (gov. spending distributed like schools)

DINA vs Geography

- DINA → Govt. Consumption distributed:
  - Neutral: Gini remains the same
  - Max equality: Gini decreases to 0.09
  - Geography → Govt. Consumption distributed:
    - Hospitals accessibility → Gini rises to 0.207 when we impute G as hospital AI (59.2% increase)

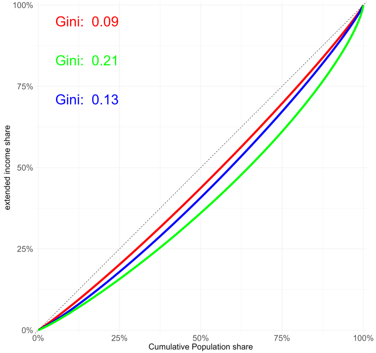


Figure : Lorenz curve Extended Income (gov. spending distributed like hospitals)

DINA vs Geography

- DINA $\rightarrow$ Govt. Consumption distributed:
  - Neutral: Gini remains the same
  - Max equality: Gini decreases to 0.09
    - Public transport accessibility $\rightarrow$ Gini rises to 0.26 (= doubles) when G = transport AI

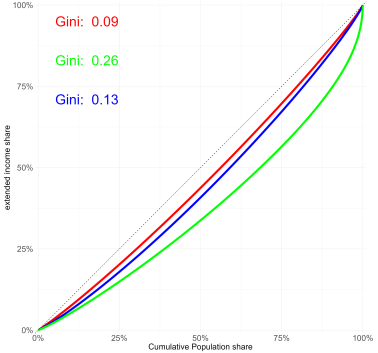


Figure : Lorenz curve Extended Income (gov. spending distributed like public transport stops)

DINA vs Geography

- DINA → Govt. Consumption distributed:
  - Neutral: Gini remains the same
  - Max equality: Gini decreases to 0.09
    - Police stations accessibility → Gini rises to 0.24 (85% increase) when G = police AI

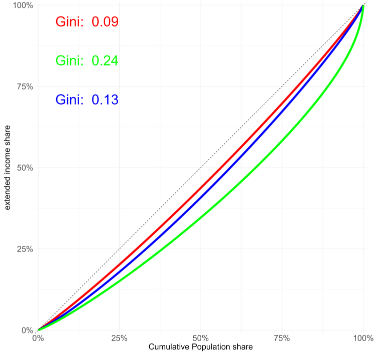


Figure : Lorenz curve Extended Income (gov. spending distributed like police stations)