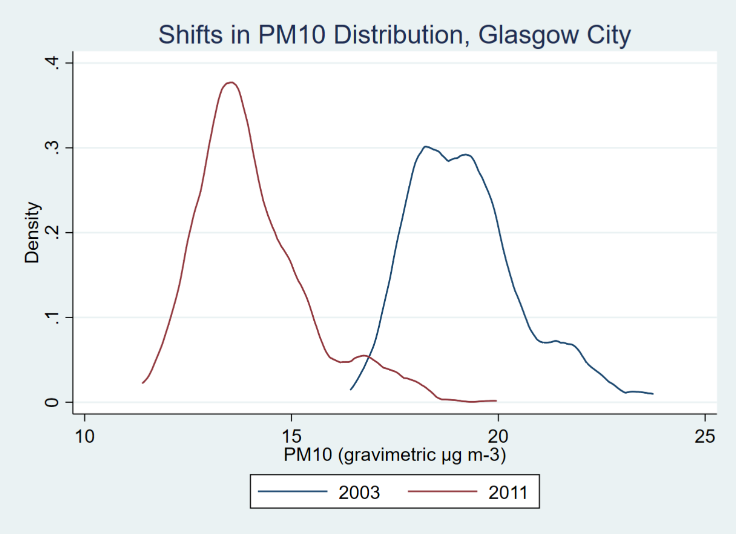
Using Panel Data to Recover Willingness to Pay (WTP) Function for Neighborhood Amenities

Chuhang Yin

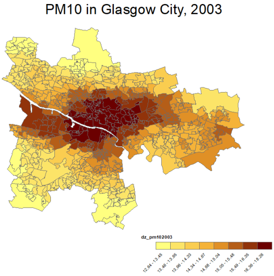
Duke University

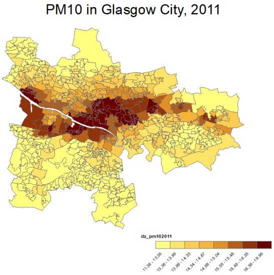
August 13, 2018

Motivation


Shifts in PM10 Distribution, Glasgow City
Density
0
.1
.2
.3
.4
10
15
20
25
PM10 (gravimetric µg m-3)
2003
2011

Motivation


PM10 in Glasgow City, 2003
dz_pm102003
12.64 - 13.45
13.46 - 13.95
13.96 - 14.33
14.34 - 14.67
14.68 - 15.04
15.05 - 15.48
15.49 - 16.35
16.36 - 18.26


PM10 in Glasgow City, 2011
dz_pm102011
11.38 - 13.35
13.36 - 13.95
13.96 - 14.33
14.34 - 14.67
14.68 - 15.04
15.05 - 15.48
15.49 - 16.35
16.36 - 19.95

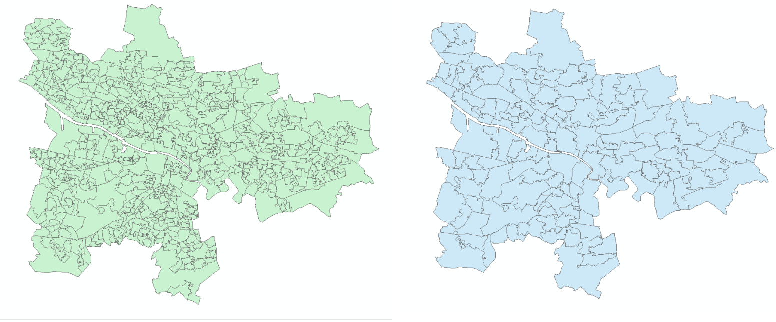
Motivation

- Lack of flexible approach to allow heterogeneity in WTP $\Rightarrow$ cannot predict heterogeneous welfare effects.
  - Rosen (1974) two stage: Omitted Variable Bias; identification; endogeneity.
  - Approaches to solve endogeneity: Ekeland, Heckman & Nesheim (2004), Heckman, Matzkin & Nesheim (2010).
  - Most related to our approach: Bajari & Benkard (2005), Bishop & Timmins (2018).
- Our contribution: propose an approach to estimate flexible, individual-level MWTP function with panel data.

Geographical Units of Analysis

- Datazones (neighborhoods): 500-1000 people
- Intermediate Zones: 4000 people on average

Figure 1: 2001 Datazones and Intermediate Zones in Glasgow City

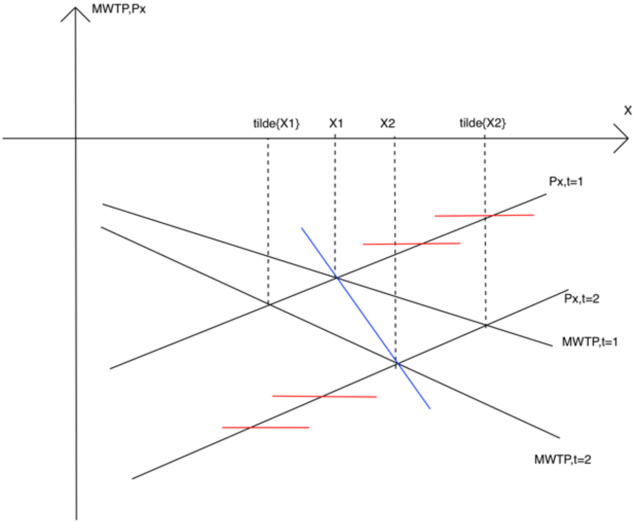


(a) Datazones (b) Intermediate Zones

Data

- Joint distribution of housing features and prices from Iterative Proportional Fitting details
  - Housing sales in Glasgow City Council (2003, 2011)
  - Breakdown of housing attributes by datazone (2001, 2011)
  - Housing sales & attributes from a leading mortgage lender (2003, 2011)
- Modelled air quality data
  - Raw measures at 1km$*$1km grid level, aggregated to datazone
  - Focus on number of days with 8-hr ozone exceeding 120 μgm^{-3}, PM10 (gravimetric μgm^{-3}).
- Individual-level data: Scottish Longitudinal Study (SLS)
  - 5.3% sample of Scottish population (2001, 2011)
  - Datazone of enumeration, demographics, education, employment, housing tenure & attributes
- Additional neighborhood amenities: proximity to landfill sites, land use, school performance (S4)

Empirical Framework: Intuition



Estimation Algorithm

- Estimate flexible hedonic price functions for each year $\Rightarrow$ hedonic gradient functions of O_3.
- Estimate relationship between consumed O_3 and household characteristics flexibly including FE.
- Predict O_3 consumption using alternative household characteristics in another t.
- Calculate equilibrium hedonic prices using hedonic gradient functions.
- Solve for constant and slope of MWTP function by year.

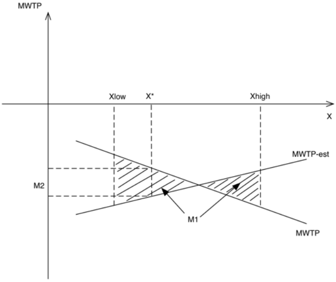
Additional details

Monte Carlo Exercise

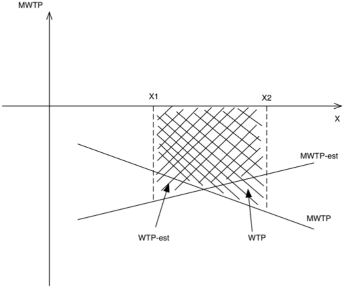
- Evaluate the performance of our approach relative to Rosen (1974) and Bajari & Benkard (2005).
- Simulate T=1000 samples for t=1,2 where
  - House prices in 500 neighborhoods based on size and number of days with high O_3.
  - MWTP functions for 12,000 households depends on their time-varying characteristics $Z_{i,t}$.
  - Equilibrium O_3 consumption where hedonic price equals MWTP.

Monte Carlo parameters

Monte Carlo Performance Evaluation



(a) M1, M2



(b) M3

Monte Carlo Results

Medians with 95% confidence intervals in brackets, $t = 1$.

	M_1	M_2	$M_{3,25\%}$	$M_{3,50\%}$	$M_{3,75\%}$
A1: our approach	0.0932	0.000158	0.0962	0.000387	-0.00141
	(0.0581, 0.2313)	(0.000025, 0.000385)	(-0.0137,0.0315)	(-0.014183,0.015358)	(-0.00946,0.00141)
A2: Rosen	0.1603	0.000639	0.102	0.0615	0.0347
	(0.1480, 0.1872)	(0.000153, 0.001197)	(0.0628,0.1216)	(0.0301,0.0963)	(0.00853,0.0871)
A3: Bajari & Benkard	0.396	0.00418	-0.08	0.0418	0.139
	(0.2661, 0.6198)	(0.00257, 0.00677)	(-0.204,0.0114)	(-0.00623,0.0749)	(0.0611,0.265)

Hedonic Regression Results

	2003	2011
Housing type: semi-detached	-0.278***	-0.272***
	[0.005]	[0.006]
Housing type: terraced	-0.253***	-0.372***
	[0.006]	[0.006]
Housing type: flats	-0.254***	-0.317***
	[0.006]	[0.007]
Housing type: converted/mobile/shared	0.085***	0.153***
	[0.008]	[0.010]
Housing size: 3-4 rooms	0.293***	0.268***
	[0.004]	[0.005]
Housing size: 5+ rooms	0.650***	0.593***
	[0.005]	[0.007]
% S4 pupils w/ 5 awards at SCQF level 5+	0.000**	-0.000
	[0.000]	[0.000]
% land as urban green space	-0.001***	-0.001***
	[0.000]	[0.000]
% land for agricultural use	-0.000	0.001***
	[0.000]	[0.000]
Hazardous landfill sites in 1 mile	0.023*	0.030*
	[0.012]	[0.016]
Non-hazardous landfill sites in 1 mile	0.001	-0.021
	[0.014]	[0.064]
Log(PM10)	-0.298*	-0.050
	[0.166]	[0.158]
Log(days with high O_3)	-0.156***	-0.256**
	[0.040]	[0.120]
N	137,280	66,680
R^2	0.248	0.301

More specifications

Empirical Implementation: Iterative Proportional Fitting

- Goal: estimate the joint distribution of housing price and features by datazone for Glasgow City in 2001 and in 2011.
- Iteratively adjust the joint distribution until the marginal distributions of the selected variables match observed population distributions.
- Divide Glasgow into 6 sub areas to match aggregate to micro distributions.
- Restrictions for each sub area-year
  1. For each datazone with observed housing sales, the distributions of housing types & housing sizes match Census.
  2. Distribution of price quartiles matches that observed in the housing sales data.
  3. The joint distribution of price quartiles, housing types and housing sizes matches that observed in Nationwide.

Back

Empirical Implementation: Hedonic Regression

- Restrict the individuals to be in Glasgow City in 2001 and 2011.
- Drop students/living rent free/living in mobile housing.
- Simulate 1000 samples and get hedonic estimates through OLS.
- Include Intermediate Zone fixed effects in hedonic regressions.
- Specifications with raw pollution levels/standardized z-scores/logs.
- Calculate standard errors from bootstrapping.

Back

Additional Regression Results

	2003I	2003II	2003III	2011I	2011II	2011III
Housing type: semi-detached	-0.278***	-0.278***	-0.278***	-0.272***	-0.272***	-0.272***
	[0.005]	[0.005]	[0.005]	[0.006]	[0.006]	[0.006]
Housing type: terraced	-0.254***	-0.254***	-0.253***	-0.372***	-0.372***	-0.372***
	[0.006]	[0.006]	[0.006]	[0.007]	[0.007]	[0.006]
Housing type: flats	-0.254***	-0.254***	-0.254***	-0.317***	-0.317***	-0.317***
	[0.007]	[0.007]	[0.007]	[0.008]	[0.008]	[0.007]
Housing type: converted/mobile/shared	0.084***	0.084***	0.085***	0.154***	0.154***	0.153***
	[0.008]	[0.008]	[0.008]	[0.010]	[0.010]	[0.010]
Housing size: 3-4 rooms	0.293***	0.293***	0.293***	0.268***	0.268***	0.268***
	[0.004]	[0.004]	[0.004]	[0.005]	[0.005]	[0.005]
Housing size: 5+ rooms	0.650***	0.650***	0.650***	0.593***	0.593***	0.593***
	[0.006]	[0.006]	[0.005]	[0.007]	[0.007]	[0.007]
PM10	-0.085			-0.103		
	[0.195]			[0.087]		
$PM10^2$	0.002			0.003		
	[0.005]			[0.003]		
Days with high ozone	-0.072			-0.052		
	[0.081]			[0.516]		
Days with high $ozone^2$	0.006			-0.032		
	[0.009]			[0.175]		
Standardized PM10		-0.010			-0.009	
		[0.018]			[0.018]	
Standardized $PM10^2$		0.004			0.006	
		[0.009]			[0.005]	
Standardized days with high ozone		-0.014			-0.028	
		[0.019]			[0.018]	
Standardized days with high $ozone^2$		0.006			-0.001	
		[0.009]			[0.006]	
Log(PM10)			-0.298*			-0.050
			[0.166]			[0.158]
Log(days with high ozone)			-0.156***			-0.256**
			[0.040]			[0.120]
N	137,280	137,280	137,280	66,680	66,680	66,680
R^2	0.248	0.248	0.248	0.301	0.301	0.301

Back

Empirical Framework

- Estimate a flexible hedonic price function

$$P_{i,t} = f_t(X_{it}; \bar{\alpha_t}), t = 1, 2 \tag{1}$$

- In a hedonic equilibrium, MWTP = hedonic price for neighborhood amenities $\rightarrow$ relationship between consumed amenity $X_{i,t}$ and individual attributes $Z_{i,t}$, e.g.

$$\beta_{0,i,t} + \beta_{1,i,t} X_{i,t} = \frac{dP_{i,t}}{dX_{i,t}} = P^X_{i,t} = MWTP^X_{i,t} = b_0 + b_1 X_{i,t} + b_2 Z_{i,t} \tag{2}$$

- Estimate the relationship between $X_{i,t}$ and $Z_{i,t}$ flexibly

$$X_{i,t} = g_t(Z_{i,t}, \mu_i; \gamma) \tag{3}$$

Empirical Framework (Cont'd)

- Use estimation results to predict the equilibrium $X_{i,t}$ in $t = 2$ if the individual attributes were $Z_{i,1}$

$$\tilde{X_{i,2}} = g_2(Z_{i,1}, \hat{\mu_i}; \hat{\gamma}) \tag{4}$$

- In hedonic equilibrium, $MWTP_i^X(\tilde{X_{i,2}}) = P_{i,1}^X(\tilde{X_{i,2}})$. Solve for $(\beta_{0,i,1}, \beta_{1,i,1})$ for each i

$$P_{i,1}^X(X_{i,1}) = \beta_{0,i,1} + \beta_{1,i,1}X_{i,1} \tag{5}$$
$$P_{i,2}^X(\tilde{X_{i,2}}) = \beta_{0,i,1} + \beta_{1,i,1}\tilde{X_{i,2}} \tag{6}$$

- Similarly, we can solve for $(\beta_{0,i,2}, \beta_{1,i,2})$ using the estimated $\tilde{X_{i,1}} = g_1(Z_{i,2}, \hat{\mu_i}; \hat{\gamma})$
  - No theoretical foundation regarding which one to prefer

Back

Monte Carlo Parameters

House prices determined by

$$
\begin{aligned}
LogP_{j,1} &= 10.36 + 0.012 \cdot size_{j,1} - 0.0000136 \cdot size_{j,1}^2 - 0.1 \cdot Oz_{j,1} + 0.01 \cdot Oz_{j,1}^2 + \nu_{j,1} \\
LogP_{j,2} &= 10.89 + 0.014 \cdot size_{j,2} - 0.0000257 \cdot size_{j,2}^2 - 0.22 \cdot Oz_{j,2} + 0.015 \cdot Oz_{j,2}^2 + \nu_{j,2}
\end{aligned}
$$

with the following distributional parameters

	t=1	t=2
size	logN(4.49, 0.34)	logN(4.62, 0.33)
Oz	N(2.5, 1) on [0, 4.5]	N(3,1) on [0, 5]
ν	N(0, 0.35)	N(0, 0.4)

Monte Carlo Parameters (Cont'd)

$$\begin{aligned} MWTP_{i,t} &= \alpha_{i,t} + \beta_{i,t} \cdot Oz_{i,t} + \eta_{i,t} \\ \alpha_{i,t} &= 0.3 - 0.01 \cdot Z_{i,t} + \epsilon_{i,t} \\ \beta_{i,t} &= -0.11 - 0.01 \cdot Z_{i,t} + \omega_{i,t} \end{aligned}$$

with the following distributional parameters

	t=1	t=2
$Z_{i,t}$	Poisson(3), upper truncation at 7	$Z_{i,1}$ + N(0,1), upper truncation at 9
$\epsilon_{i,t}$	N(0, 0.03)	N(0, 0.03)
$\omega_{i,t}$	N(0, 0.01)	N(0, 0.01)
$\eta_{i,t}$	N(0, 0.001)	N(0, 0.001)

Back