Spatial Modelling of House Prices in the Dublin Area

Dr. James Sweeney

University College Dublin

james.sweeney@ucd.ie

May 17, 2016

Problems of Interest

- The property price calculators used at present are limited in terms of the factors they use to estimate the value of a property. Furthermore, typically no uncertainty is provided for price predictions.

- In areas where house (sales) turnover are relatively infrequent, providing *accurate* estimates of house prices will be difficult - borrowing weight from nearby sales across all house types may help with this.

- Sentiment Analysis: an issue of interest is whether there is bias in terms of the prices people are willing to pay for a property?

January Snapshot: 1531 observations

	Type	Price (euro)	Beds	Baths	floorArea	location	postcode	lon	lat	energy-rating
1	Type	Price (euro)	Beds	Baths	floorArea	location	postcode	lon	lat	energy-rating
2	Apartment	300000	2	1	5.3	dublin-4	dublin-4	-6.223159	53.327399	null
3	Terraced	370000	3	2	7.4	sutton	dublin-13	-6.1246955	53.39088048	BER D2
4	Semi-Detached	344950	4	2	12	clonee	dublin-15	-6.4282101	53.40535372	BER D2
5	Apartment	147500	1	1	30	christchurch	dublin-8	-6.2765622	53.34374308	BER E1
6	Apartment	130000	1	1	30.01	dublin-8	dublin-8	-6.280181	53.340944	null
7	Apartment	185000	1	1	30.658	dundrum	dublin-14	-6.2449229	53.3037748	BER D2
8	Apartment	205000	1	1	31.56	dublin-1	dublin-1	-6.2628615	53.34733627	BER E2
9	Terraced	220000	1	1	32.99	harolds-cross	dublin-6w	-6.2757863	53.32839777	BER G
10	Apartment	149950	1	1	33	dublin-8	dublin-8	-6.2802194	53.3411067	BER E2
11	Terraced	199000	3	3	33.12	dublin-8	dublin-8	-6.3193438	53.3396372	null
12	Apartment	150000	1	1	35	north-circular-roa	dublin-7	-6.2723485	53.35785792	BER C3
13	Apartment	130000	1	1	35	dublin-1	dublin-1	-6.2679116	53.3524931	null
14	Apartment	152500	1	1	35.6	dublin-8	dublin-8	-6.2841144	53.3457748	BER D1
15	Apartment	124950	1	1	36	dublin-8	dublin-8	-6.2753029	53.3334848	null
16	Apartment	124950	1	1	36	dublin-8	dublin-8	-6.2752855	53.33351587	null
17	Apartment	145000	1	1	37	dublin-1	dublin-1	-6.269507	53.35189	BER C3
18	Apartment	155000	1	1	37	dublin-1	dublin-1	-6.2558274	53.3521797	BER E2
19	Apartment	165000	1	1	37	dublin-1	dublin-1	-6.255827	53.35218	BER F
20	Apartment	225000	1	1	37	ballsbridge	dublin-4	-6.2093031	53.31814898	BER E1
21	Apartment	135000	1	1	37.2	south-circular-roa	dublin-8	-6.2921972	53.33492907	null
22	Apartment	115000	1	1	37.2	dublin-8	dublin-8	-6.2843466	53.34544556	BER E1

Complete information on: **asking** price, type, floor area, #
beds, # baths, location, postcode, local area.
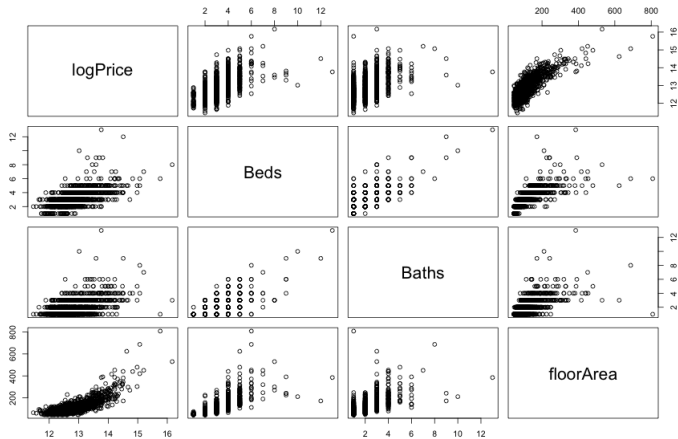**Spatial mislabelling....**

Spatially Referenced:
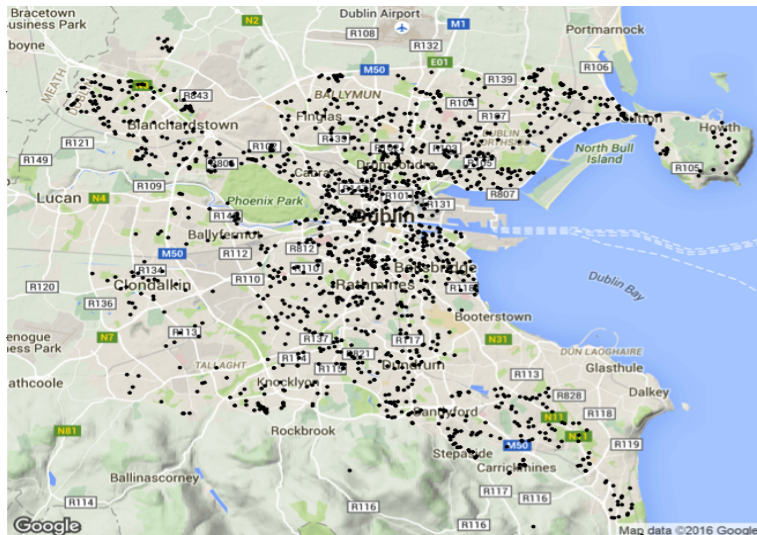
Data: log (price per m^2)

log(price) per M2
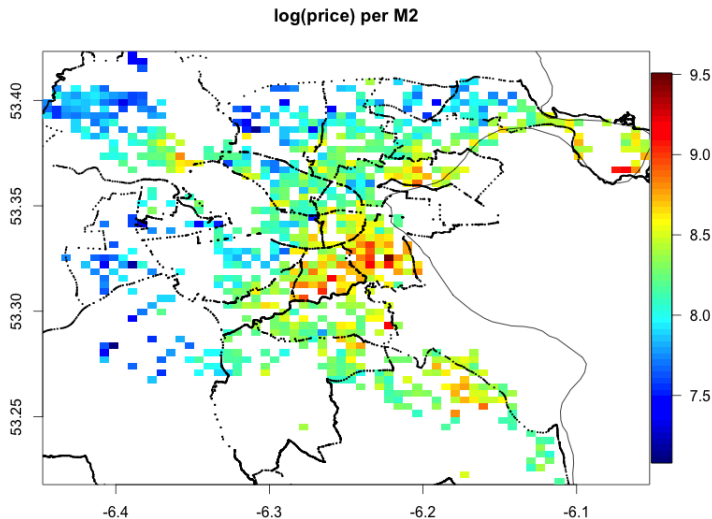
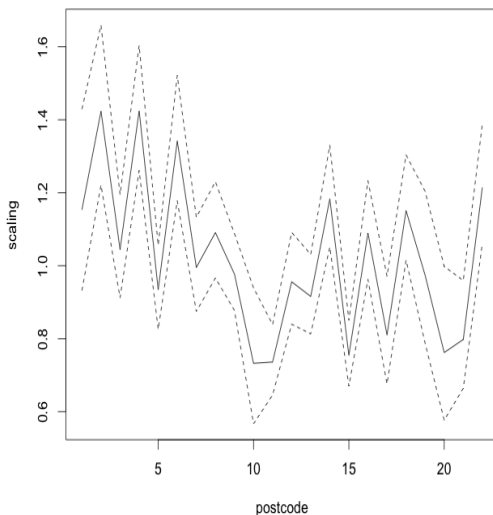# Data: Relationship between variables

(Typical) Hedonic Price (Linear) Models

Let $y_i^* = \log(\text{Price/Floor Area})$, $X_i = (\text{size}_i, \text{beds}_i, \text{baths}_i)$ then:

$$
\begin{aligned}
y_i^* &= \mu + X_i\beta + \text{postcode}_i + \text{area}_i + \text{type}_i + \epsilon_i \\
\epsilon_i &\sim \mathcal{N}(0, \sigma^2) \\
\text{postcode}_i &\sim \mathcal{N}(0, \tau_1^2) \\
\text{area}_i &\sim \mathcal{N}(0, \tau_2^2) \\
\text{type}_i &\sim \mathcal{N}(0, \tau_3^2)
\end{aligned}
$$

Note, this model implies that:

$$E(y_i) = e^{\mu + \beta_1 \text{size}_i + \beta_2 \text{beds}_i + \beta_3 \text{baths}_i + \text{postcode}_i + \text{area}_i + \text{type}_i}$$

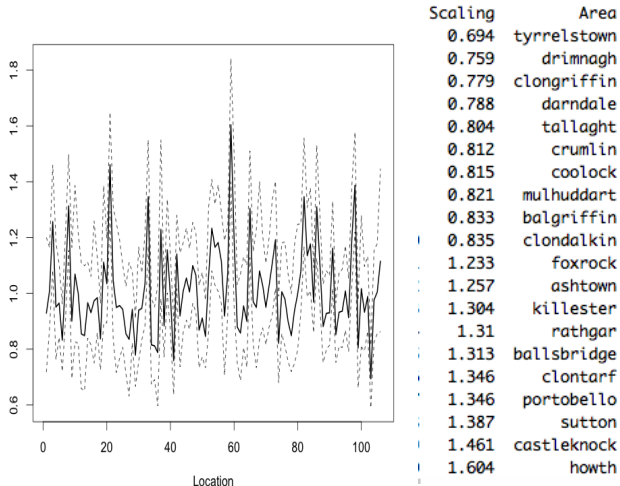# Postcode Effects

Note:
19 Dublin 20
20 Dublin 22
21 Dublin 24
22 Dublin 6W

Location Effects

Scaling	Area
0.694	tyrrelstown
0.759	drimnagh
0.779	clongriffin
0.788	darndale
0.804	tallaght
0.812	crumlin
0.815	coolock
0.821	mulhuddart
0.833	balgriffin
0.835	clondalkin
1.233	foxrock
1.257	ashtown
1.304	killester
1.31	rathgar
1.313	ballsbridge
1.346	clontarf
1.346	portobello
1.387	sutton
1.461	castleknock
1.604	howth

A More General Model including Spatial Variation

$$
\begin{aligned}
y_i^* &= \mu + X(s_i) + f(\text{size}_i) + f(\text{beds}_i) + f(\text{baths}_i) + \text{postcode}_i + \\
&\quad \text{area}_i + \text{type}_i + \epsilon_i
\end{aligned}
$$

$$
\begin{aligned}
\epsilon_i &\sim \mathcal{N}(0, \sigma^2) \\
\text{postcode}_i &\sim \mathcal{N}(0, \tau_1^2) \\
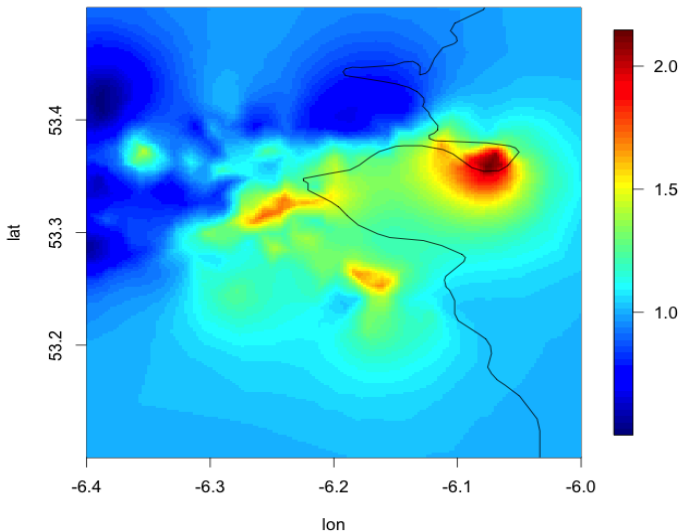\text{area}_i &\sim \mathcal{N}(0, \tau_2^2) \\
\text{type}_i &\sim \mathcal{N}(0, \tau_3^2) \\
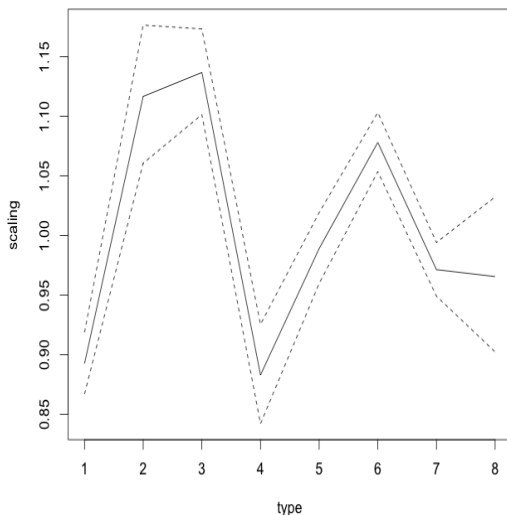f_k() &\sim \textit{smooth} - \Delta^2 f \sim \mathcal{N}(0, \tau_k^2) \\
X(\mathbf{s}) &\sim \textit{GP}(0, \Sigma), \text{ stationary, isotropic......}
\end{aligned}
$$

Spatial Effect

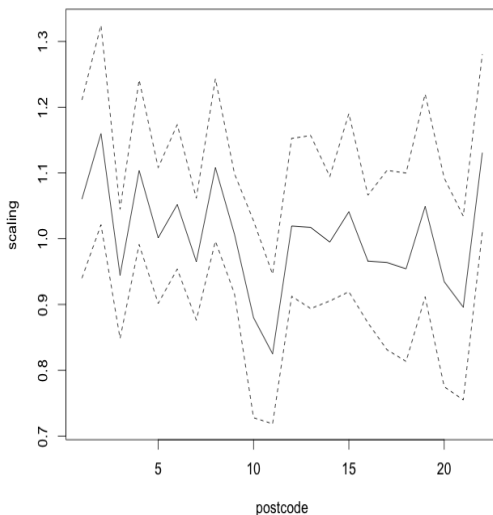**posterior mean scaling**

Effect of Type

1. Apartment
2. Bungalow
3. Detached
4. Duplex
5. End of Terrace
6. Semi-Detached
7. Terraced
8. Townhouse
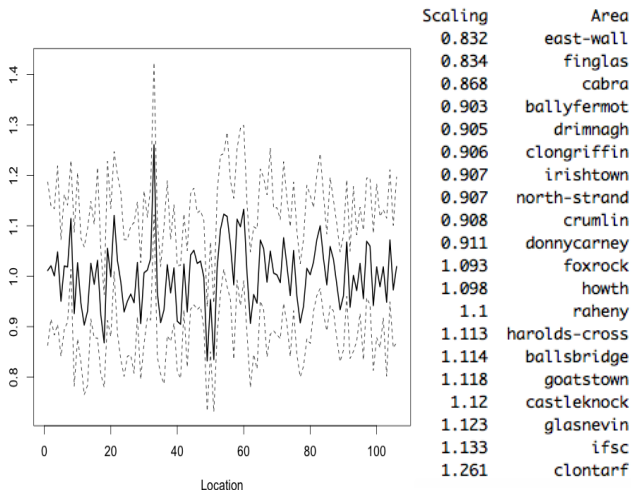
Postcode Effects

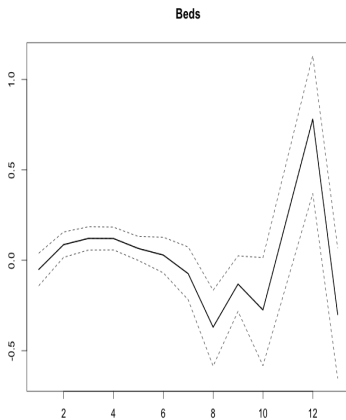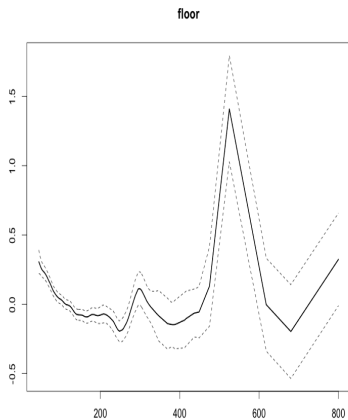
Note:
19 Dublin 20
20 Dublin 22
21 Dublin 24
22 Dublin 6W

Location Effects

Scaling	Area
0.832	east-wall
0.834	finglas
0.868	cabra
0.903	ballyfermot
0.905	drimnagh
0.906	clongriffin
0.907	irishtown
0.907	north-strand
0.908	crumlin
0.911	donnycarney
1.093	foxrock
1.098	howth
1.1	raheny
1.113	harolds-cross
1.114	ballsbridge
1.118	goatstown
1.12	castleknock
1.123	glasnevin
1.133	ifsc
1.261	clontarf

Effect Of Floor Size and Beds

Model Comparison

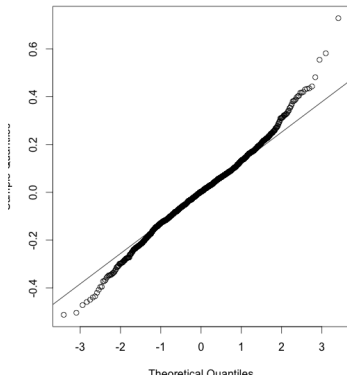| Model	*Relative DIC
Linear Model in size, beds, baths, no spatial	500.88
Smoothly varying size, beds, baths, no spatial	319.69
Smoothly varying size, spatial	90.32
Smoothly varying size, spatial, beds, baths	16.48
Smoothly varying size, beds, spatial	0.0

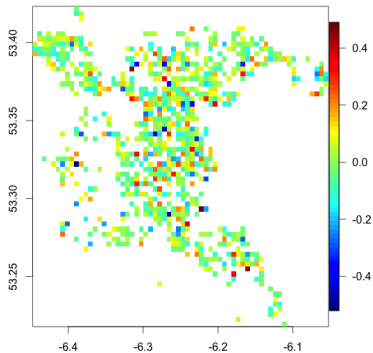Global median absolute error for best spatial model $\approx$ 29k vs 39k for linear model

Residuals

Normal Q-Q Plot

log price per M2

Conclusion

- **Existing price calculators do not incorporate available covariates**
  - Identify # baths and BER rating as being unimportant.
  - Psychological effect of postcode?

- **Spatial effect is perhaps a proxy for**
  - Transport links
  - Schools
  - Plot size
  - Provides a more parsimonius, plausible model

- **Extensions to spatio-temporal settings**
  - Not possible using continuous space methods.......