

The potential of big housing data: an application to the Italian real estate market*

Banca d'Italia, Big Data Workshop, 27/03/2018

Michele Loberto¹, Andrea Luciani¹, Marco Pangallo²

1 Banca d'Italia, Directorate General for Economics, Statistics and Research

2 Institute for New Economic Thinking at the Oxford Martin School and Mathematical Institute, University of Oxford

**The views expressed in this paper are those of the authors and do not necessarily reflect the views of Banca d'Italia*

Why studying housing markets?

- Current economic situation: business activity of the **construction industry, consumption** through wealth effects (housing is the “democratic asset”)
 - Existing data available about 3 months later, limited hedonic-corrected indexes; no info on demand & supply
- Implications for **financial stability**
 - US housing bubble different across segments (Piazzesi et al., 2015, AER)
- Long-term effects in the **labor market**: misallocation of workers in the most productive areas (Hsieh and Moretti, 2017)
 - Absence of microdata

Why studying housing markets? (and current limits with Italian data)

- Current economic situation: business activity of the **construction industry, consumption** through wealth effects (housing is the “democratic asset”)
 - Existing data available about 3 months later; limited hedonic-correction; no info on demand & supply
- Implications for **financial stability**
 - US housing bubble different across segments (Piazzesi *et al.*, 2015, *AER*): very limited information on segments
- Long-term effects in the **labor market**: misallocation of workers in the most productive areas (Hsieh and Moretti, 2017)
 - Tradeoff spatial coverage-representativeness

Our data

- Housing sales advertisements (**ads**) posted on the website [immobiliare.it](https://www.immobiliare.it)
- 1.1million ads posted 2015Q1 to 2017Q2 in all province capital cities (30% of Italian population)
- Here focus on “standard” dwellings for sale with full ownership
- Ads can also refer to offices and shops, rents, foreclosure auctions, buildings under construction, partial ownership, usufruct, etc.

Our data

- Housing sales advertisements (**ads**) posted on the website [immobiliare.it](https://www.immobiliare.it)
- 1.1million ads posted 2015Q1 to 2017Q2 in all province capital cities (30% of Italian population)
- Here focus on “standard” dwellings for sale with full ownership
- Ads can also refer to offices and shops, rents, foreclosure auctions, buildings under construction, partial ownership, usufruct, etc.

Our data

- For each ad we know:
 - Date posted, date removed
 - Geolocation
 - Floor area, rooms, bathrooms, floor, energy class, etc.
 - Garden, garage, elevator
 - Textual description & figure hash codes
 - (Asking) price, number of **clicks** every Friday of each week
- Spatial matching with 2011 census and tax agency (OMI) data

Trilocale via Etruria, Roma

120 m² | 3 locali | 1 bagno | € 465.000 |

Id: 57700898



Associato
FLAIP



Dati principali

Riferimento e Data annuncio	CDC2908 (1548035) - 02/09/2016
Contratto	Vendita
Tipologia	Appartamento
Superficie	120 m ²
Locali	2 camere, 1 bagno, cucina abitabile
Piano	2° di 5 piani, con ascensore
Disponibilità	Libero
Tipo proprietà	Intera proprietà, classe immobile media

Caratteristiche

Parzialmente arredato | Cantina | Giardino comune | Balcone | Terrazze

Descrizione

Rif. CDC0308 - ZONA SAN GIOVANNI a pochi passi dalla fermata metro Re di Roma e ad.te Piazza Tuscolo, proponiamo appartamento in soluzione condominiale del 1938, con bellissimo cortile condominiale, servizio di portierato ed ascensore. L'immobile è posto al 2° piano, la metratura è di 120 mq c.a. ed è così suddiviso : corridoio, salone doppio con faretti per l'illuminazione e parquet, ampia cucina abitabile, due camere da letto, studio, bagno. Servizio di portierato. Si presenta in ottimo stato in quanto ristrutturato. Possibilità di ricavare il 2° bagno. Possibilità di cantina. Classe energetica F
 Le spese condominiali ammontano a c. a. €950 annue più €520 annue di riscaldamento
 Il prezzo richiesto è di € 465.000,00.

Rif. CDC0308

Main issue: duplicates

- A seller can entrust more than one agency for the sale of the dwelling
- The same agency might delete an ad and create a new identical one for it to seem more recent
- Other possibilities

Is this really a problem? Yes

- Volume of transactions in 2016

City	IMM	OMI	Coverage
Turin	20263	12322	164.4
Genoa	10358	6601	156.9
Milan	40342	21909	184.1
Bologna	7655	5507	139.0
Florence	12833	4786	268.1
Rome	73070	30173	242.2
Naples	9764	6650	146.8
Palermo	5504	4718	116.7

Table 2: Transactions. Comparison between ads and OMI data.

Pre-processing of the ads

Pairwise identification of duplicates

Clustering pairs of duplicates and aggregation of information



€ 196,000
70 m²
Maintenance:
Excellent

Descrizione
ROMA, quartiere APPIO LATINO, in Via Cerveteri, a pochi passi dalla centralissima e strategica Piazza Re Di Roma, proponiamo in vendita un appartamento al quinto piano, con affaccio sulla stessa via. L'appartamento, ha una superficie commerciale di 70 mq ed è composto da due camere matrimoniali, una cucina abitabile, un bagno e un ripostiglio. Gli ambienti sono ben disposti con poco "spreco" di spazi in corridoi e disimpegni. La palazzina è dotata di ascensore e dalle finestre della casa si possono quasi accarezzare le foglie dei platani della Via Cerveteri. Gli ambienti sono luminosi e in buone condizioni interne.



€ 196,000
70 m²
Maintenance:
Excellent

Descrizione
ROMA, quartiere APPIO LATINO, in Via Cerveteri, a pochi passi dalla centralissima e strategica Piazza Re Di Roma, proponiamo in vendita un appartamento al quinto piano, con affaccio sulla stessa via. L'appartamento, ha una superficie commerciale di 70 mq ed è composto da due camere matrimoniali, una cucina abitabile, un bagno e un ripostiglio. Gli ambienti sono ben disposti con poco "spreco" di spazi in corridoi e disimpegni. La palazzina è dotata di ascensore e dalle finestre della casa si possono quasi accarezzare le foglie dei platani della Via Cerveteri. Gli ambienti sono luminosi e in buone condizioni interne.



€ 196,000
70 m²
Maintenance:
Excellent

Descrizione
ROMA, quartiere APPIO LATINO, in Via Cerveteri, a pochi passi dalla centralissima e strategica Piazza Re Di Roma, proponiamo in vendita un appartamento al quinto piano, con affaccio sulla stessa via. L'appartamento, ha una superficie commerciale di 70 mq ed è composto da due camere matrimoniali, una cucina abitabile, un bagno e un ripostiglio. Gli ambienti sono ben disposti con poco "spreco" di spazi in corridoi e disimpegni. La palazzina è dotata di ascensore e dalle finestre della casa si possono quasi accarezzare le foglie dei platani della Via Cerveteri. Gli ambienti sono luminosi e in buone condizioni interne.

(2.2; 5; -1.8; 9; -6)
doc2vec
(2.5; 7; -2.2; 8; -2)

$$\mathbb{P}(dup) = 0.62$$

C5.0 classification tree

€ 194,000 60 m²



€ 192,000
50 m²
Maintenance:
Good

Descrizione
ROMA, quartiere APPIO LATINO, in Via Cerveteri, a pochi passi dalla centralissima e strategica Piazza Re Di Roma, proponiamo in vendita un appartamento al quinto piano, con affaccio sulla stessa via. L'appartamento, ha una superficie commerciale di 70 mq ed è composto da due camere matrimoniali, una cucina abitabile, un bagno e un ripostiglio. Gli ambienti sono ben disposti con poco "spreco" di spazi in corridoi e disimpegni. La palazzina è dotata di ascensore e dalle finestre della casa si possono quasi accarezzare le foglie dei platani della Via Cerveteri. Gli ambienti sono luminosi e in buone condizioni interne.



€ 192,000
50 m²
Maintenance:
Good

Descrizione
ROMA, quartiere APPIO LATINO, in Via Cerveteri, a pochi passi dalla centralissima e strategica Piazza Re Di Roma, proponiamo in vendita un appartamento al quinto piano, con affaccio sulla stessa via. L'appartamento, ha una superficie commerciale di 70 mq ed è composto da due camere matrimoniali, una cucina abitabile, un bagno e un ripostiglio. Gli ambienti sono ben disposti con poco "spreco" di spazi in corridoi e disimpegni. La palazzina è dotata di ascensore e dalle finestre della casa si possono quasi accarezzare le foglie dei platani della Via Cerveteri. Gli ambienti sono luminosi e in buone condizioni interne.



€ 192,000
50 m²
Maintenance:
Good

Descrizione
ROMA, quartiere APPIO LATINO, in Via Cerveteri, a pochi passi dalla centralissima e strategica Piazza Re Di Roma, proponiamo in vendita un appartamento al quinto piano, con affaccio sulla stessa via. L'appartamento, ha una superficie commerciale di 70 mq ed è composto da due camere matrimoniali, una cucina abitabile, un bagno e un ripostiglio. Gli ambienti sono ben disposti con poco "spreco" di spazi in corridoi e disimpegni. La palazzina è dotata di ascensore e dalle finestre della casa si possono quasi accarezzare le foglie dei platani della Via Cerveteri. Gli ambienti sono luminosi e in buone condizioni interne.

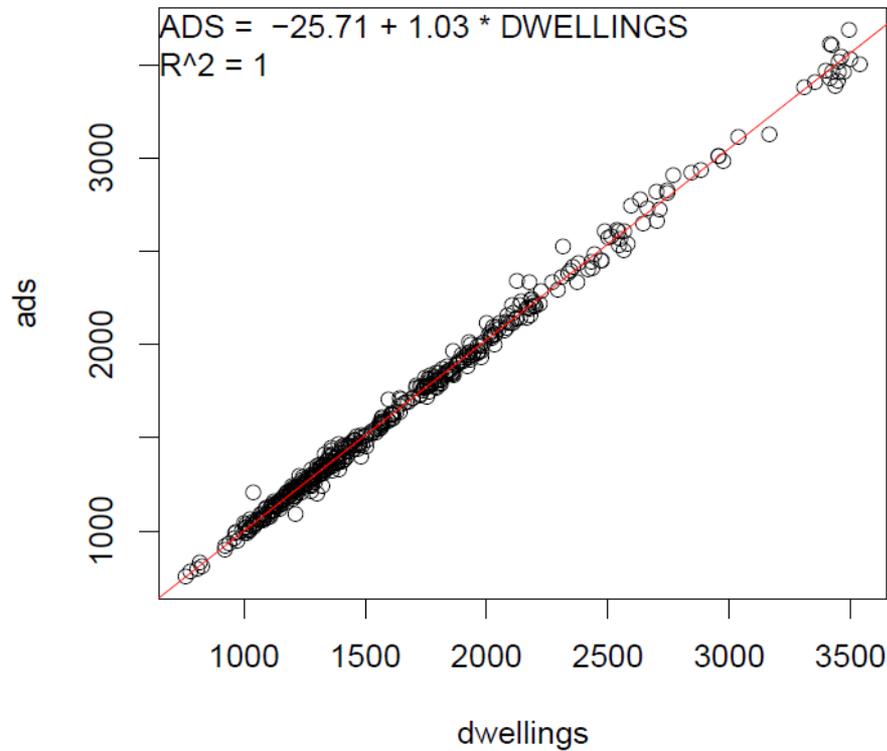
Did it work?

- 1.1million ads -> 650thousand housing units
- Volume of transactions in 2016

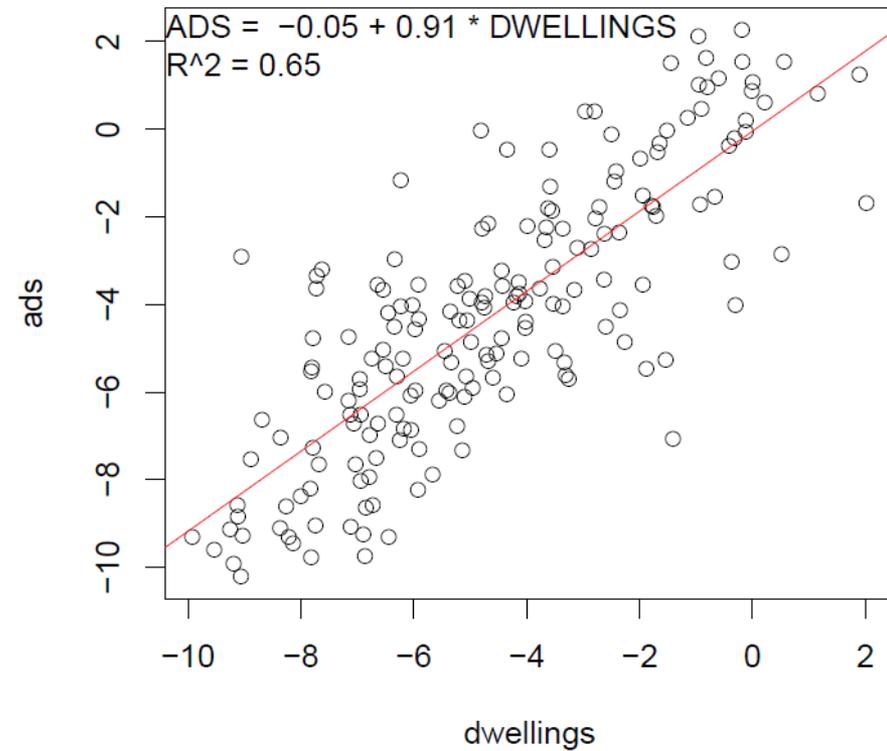
City	IMM	OMI	Coverage
Turin	7615	12322	61.8
Genoa	4816	6601	73.0
Milan	12570	21909	57.4
Bologna	2866	5507	52.0
Florence	4122	4786	86.1
Rome	22103	30173	73.3
Naples	3832	6650	57.6
Palermo	2427	4718	51.4

Comparison and validation

Aggregation level: city



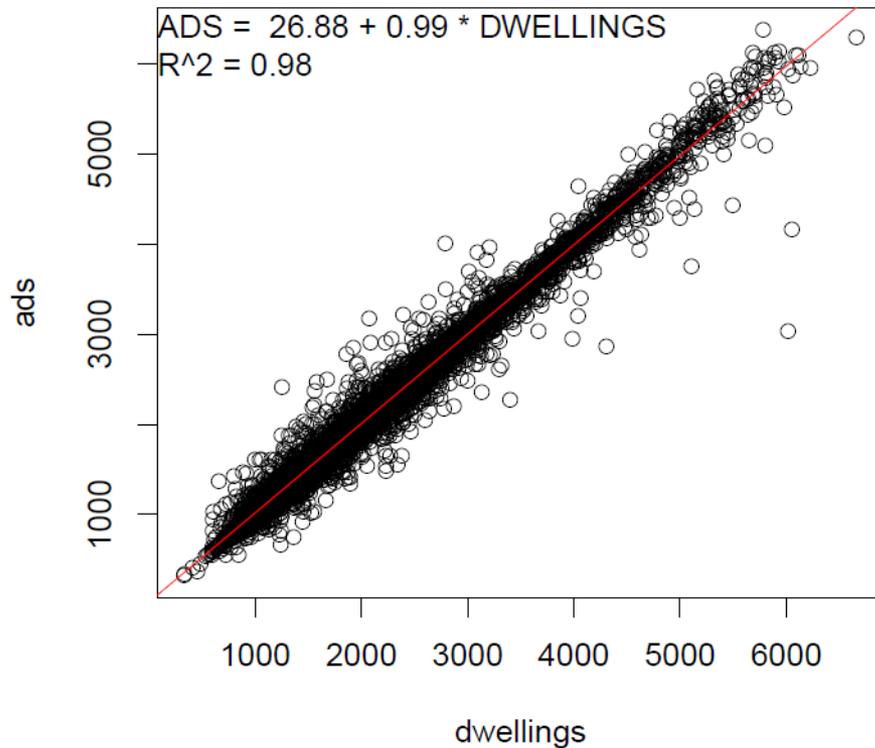
(e) Prices - Levels



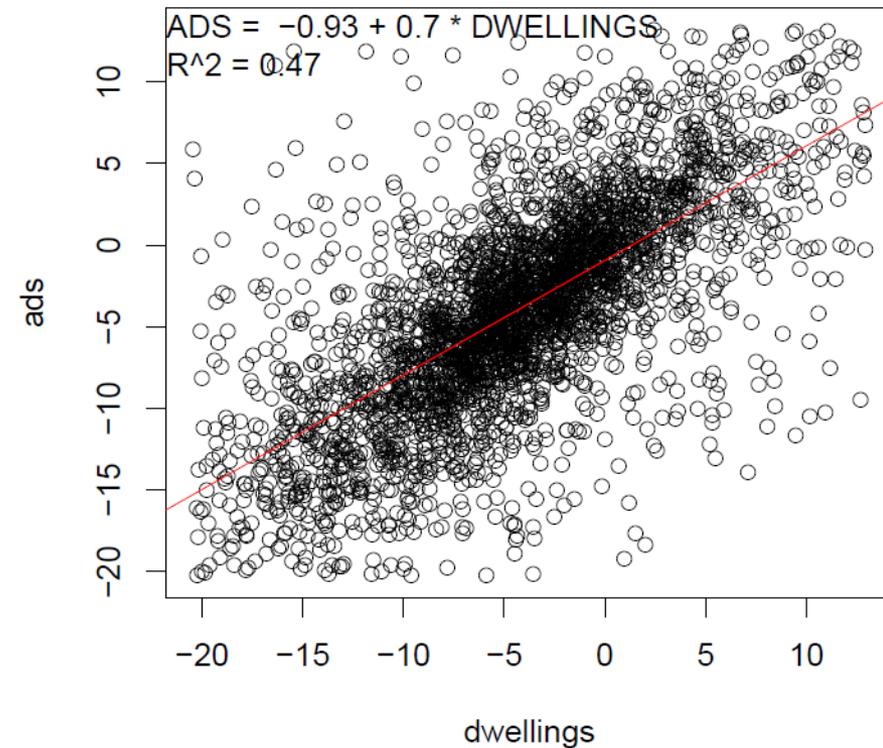
(f) Prices - Growth rates

Comparison and validation

Aggregation level: OMI microzone



(e) Prices - Levels

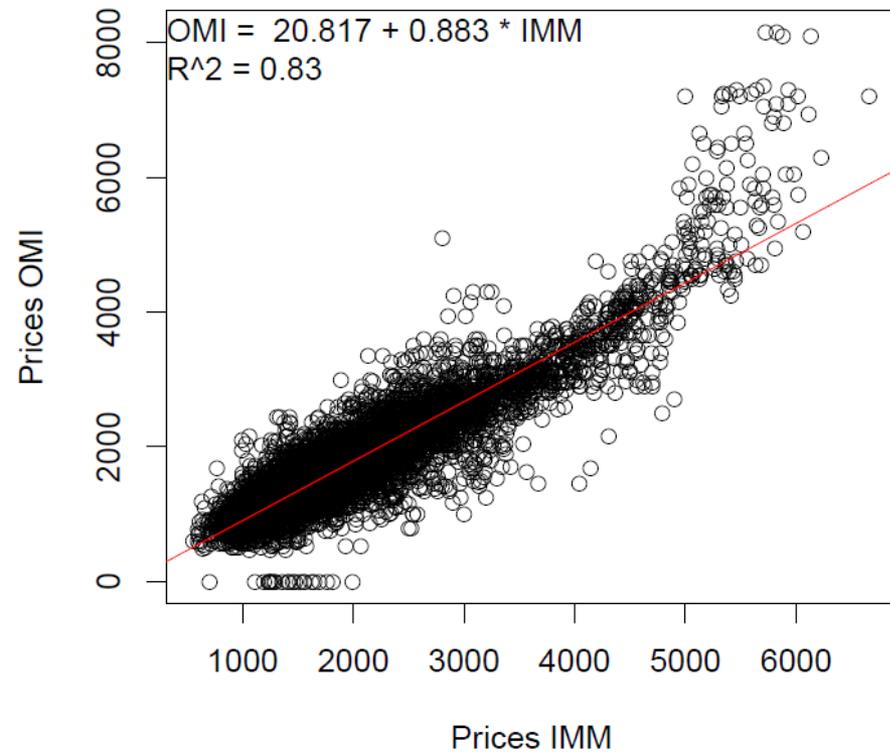


(f) Prices - Growth rates

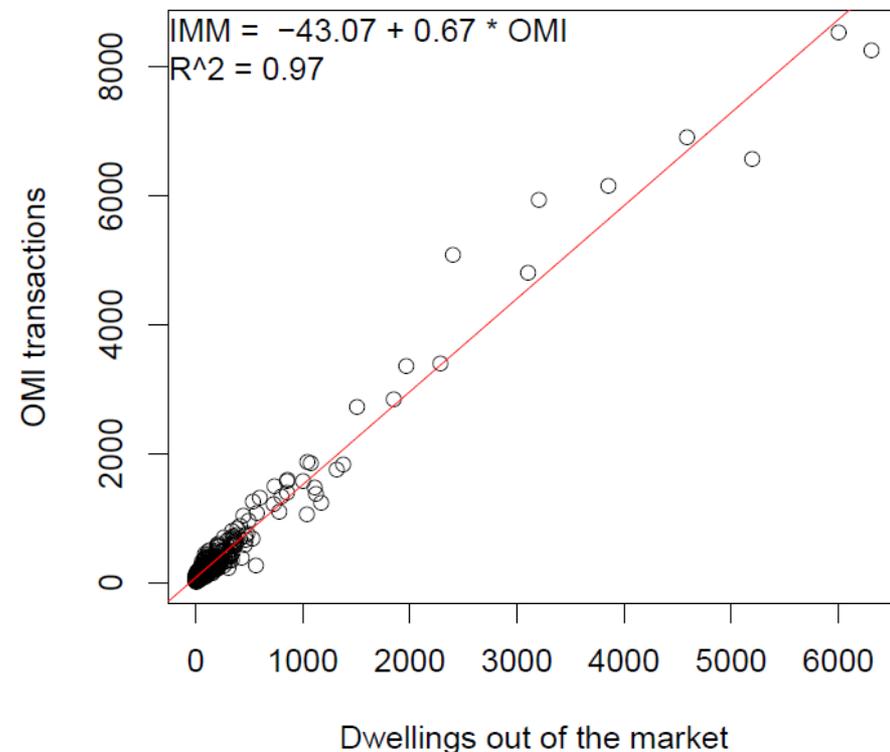
Same pattern for volume of transactions and stock of dwellings for sale

Comparison and validation

Validation OMI-Immobiliare.it

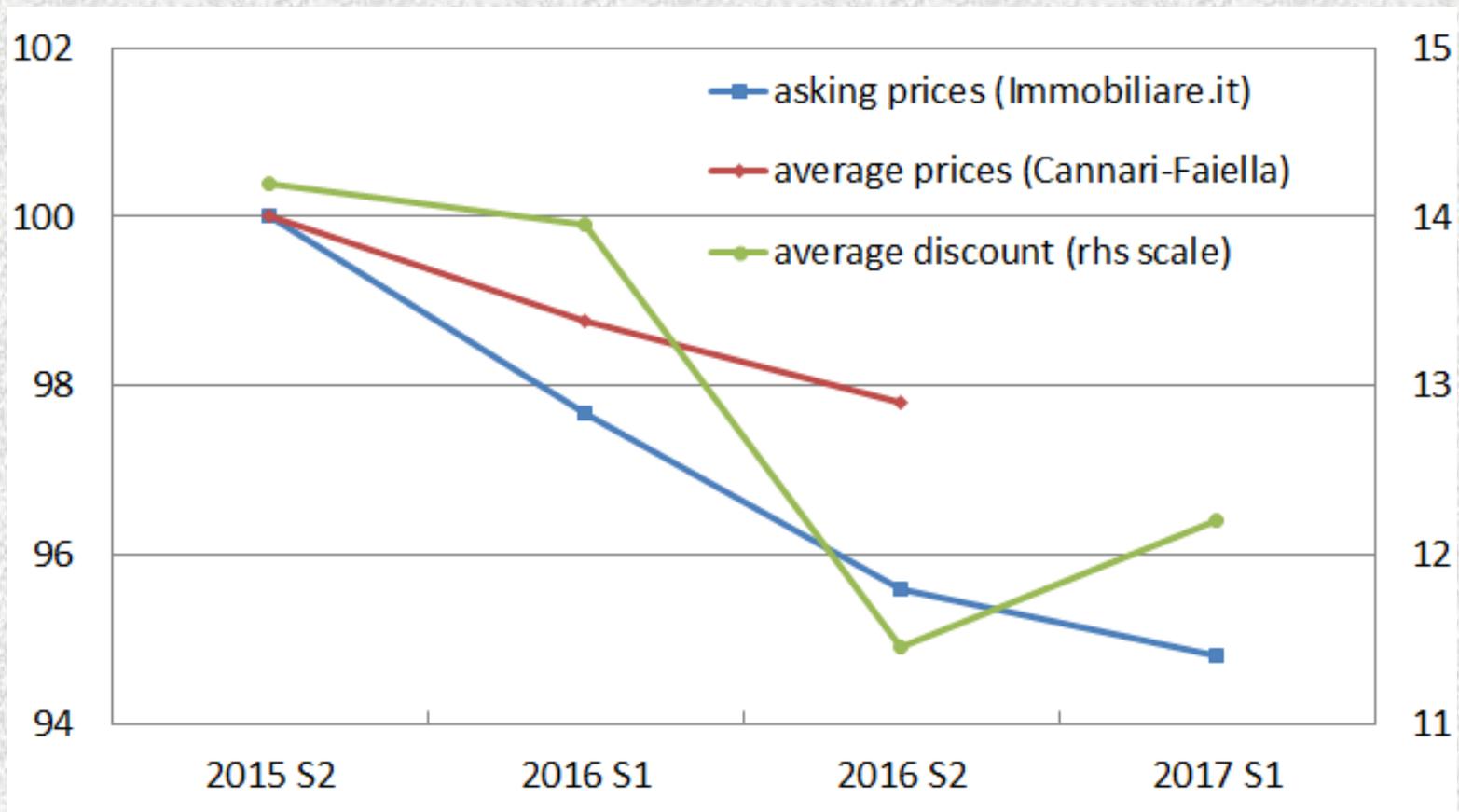


(a) Prices

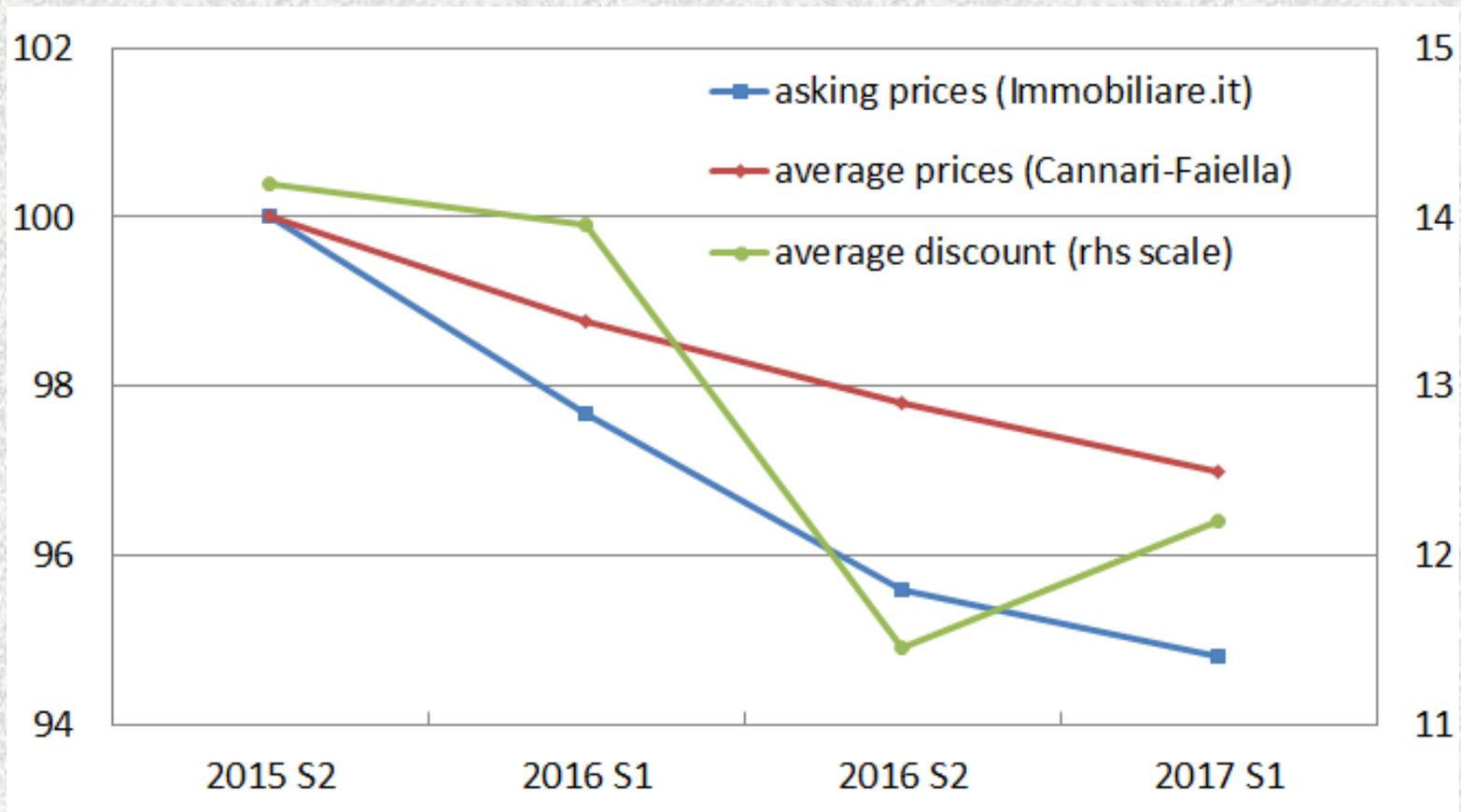


(b) Transactions

Applications: nowcasting



Applications: nowcasting



Applications: hedonic regression

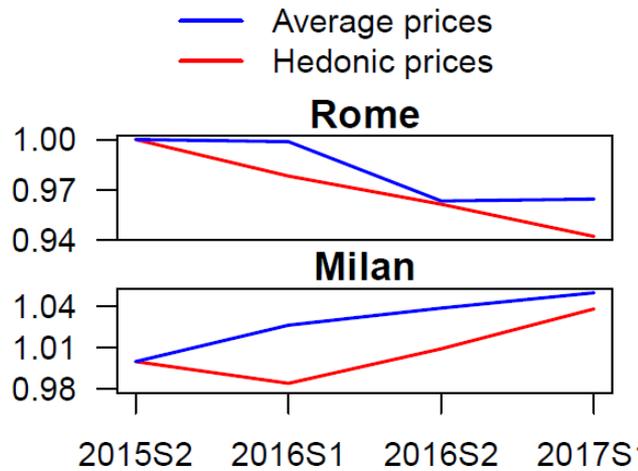
Table 10: Hedonic regression

	<i>Dependent variable:</i>		
		Price per m2	
	Italy (1)	Rome (2)	Milan (3)
Number of bathrooms	154.577*** (3.743)	72.178*** (8.034)	208.084*** (12.031)
Floor of the apartment	32.363*** (0.926)	51.505*** (2.077)	48.876*** (2.338)
Floor area [m ²]	-1.281*** (0.048)	-2.811*** (0.102)	-0.141 (0.174)
Number of rooms	-22.093*** (2.103)	-48.384*** (4.904)	-21.860*** (8.086)
Air conditioning	106.661*** (3.499)	126.300*** (7.545)	122.863*** (10.211)
Balcony	-30.398*** (3.693)	11.483 (7.912)	0.764 (10.935)
Elevator	168.574*** (3.969)	144.780*** (9.708)	223.752*** (12.614)
Energy class; Ref: EG; Level: AB	360.102*** (7.560)	267.614*** (18.907)	393.328*** (21.635)
Energy class; Ref: EG; Level: CD	114.052*** (5.742)	92.140*** (20.624)	135.590*** (16.620)
Garage; Ref: No; Level: Single	120.817*** (5.838)	127.548*** (11.158)	108.439*** (27.483)
Garage; Ref: No; Level: Double	205.992*** (4.330)	319.284*** (10.700)	149.507*** (12.491)
Heating type; Ref: No; Level: Centralized	47.737*** (10.439)	156.469*** (29.439)	225.680** (88.304)
Heating type; Ref: No; Level: Autonomous	135.250*** (9.878)	222.994*** (29.050)	281.075*** (88.417)
Kitchen type; Ref: Kitchenette; Level: Small eat-in kitchen	-49.458*** (5.350)	-23.742** (11.736)	-89.975*** (14.268)
Kitchen type; Ref: Kitchenette; Level: Large eat-in kitchen	-77.403*** (4.477)	-86.146*** (10.131)	-57.177*** (12.522)
Status; Ref: To renovate; Level: Good	192.238*** (5.661)	164.210*** (11.707)	186.031*** (16.280)
Status; Ref: To renovate; Level: Excellent	480.591*** (5.874)	462.987*** (12.252)	535.513*** (16.606)
Status; Ref: To renovate; Level: New	610.984*** (8.808)	479.285*** (21.080)	584.823*** (25.053)
Terrace	143.539*** (3.761)	229.414*** (8.255)	217.474*** (12.834)
Utility room	-56.913*** (3.516)	-84.680*** (8.234)	-77.461*** (10.396)
Basement	11.039*** (3.660)	57.360*** (8.388)	24.695** (9.814)
Porter	99.019*** (5.557)	73.887*** (10.685)	62.285*** (10.476)
Constant	718.810*** (241.357)	2,026.932*** (724.520)	457.625 (749.341)
Observations	174,143	43,945	27,212
R ²	0.783	0.675	0.731
Adjusted R ²	0.781	0.673	0.730

Note:

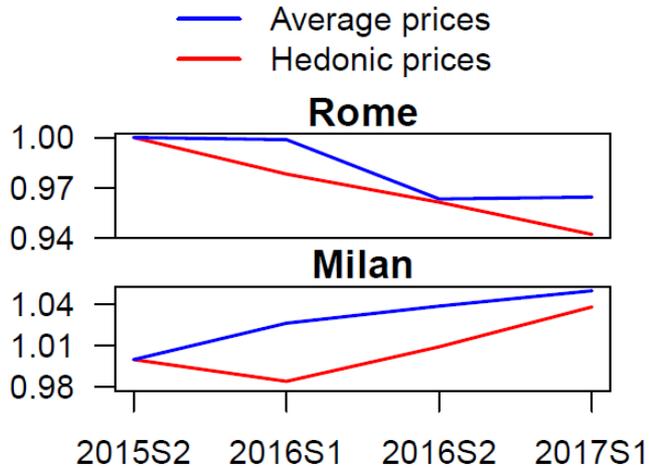
*p<0.1; **p<0.05; ***p<0.01
OMI micro-zone and quarter dummies

Applications: segmentation

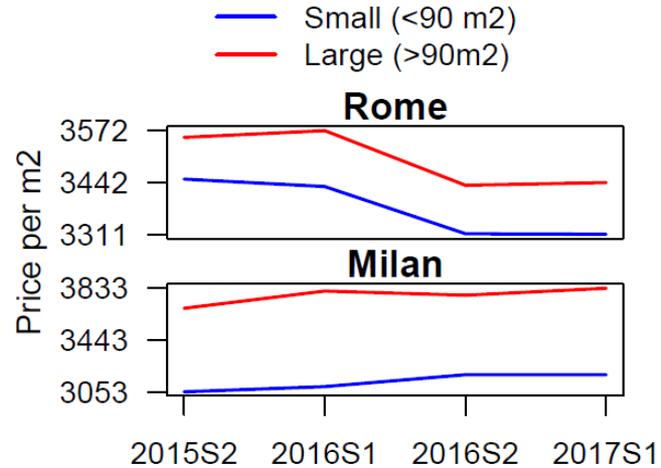


(a) Comparison

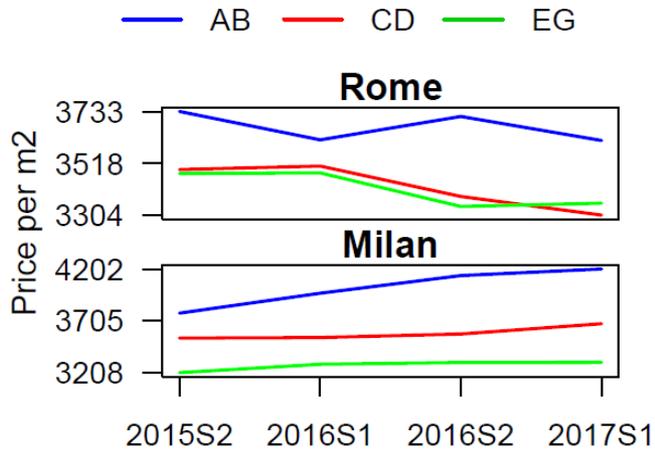
Applications: segmentation



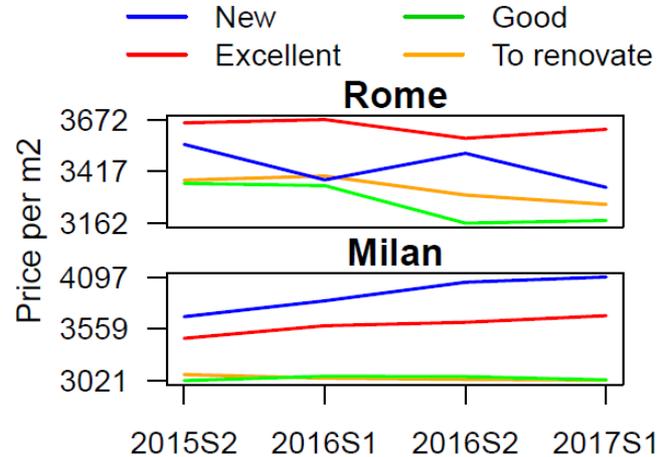
(a) Comparison



(b) Floor area

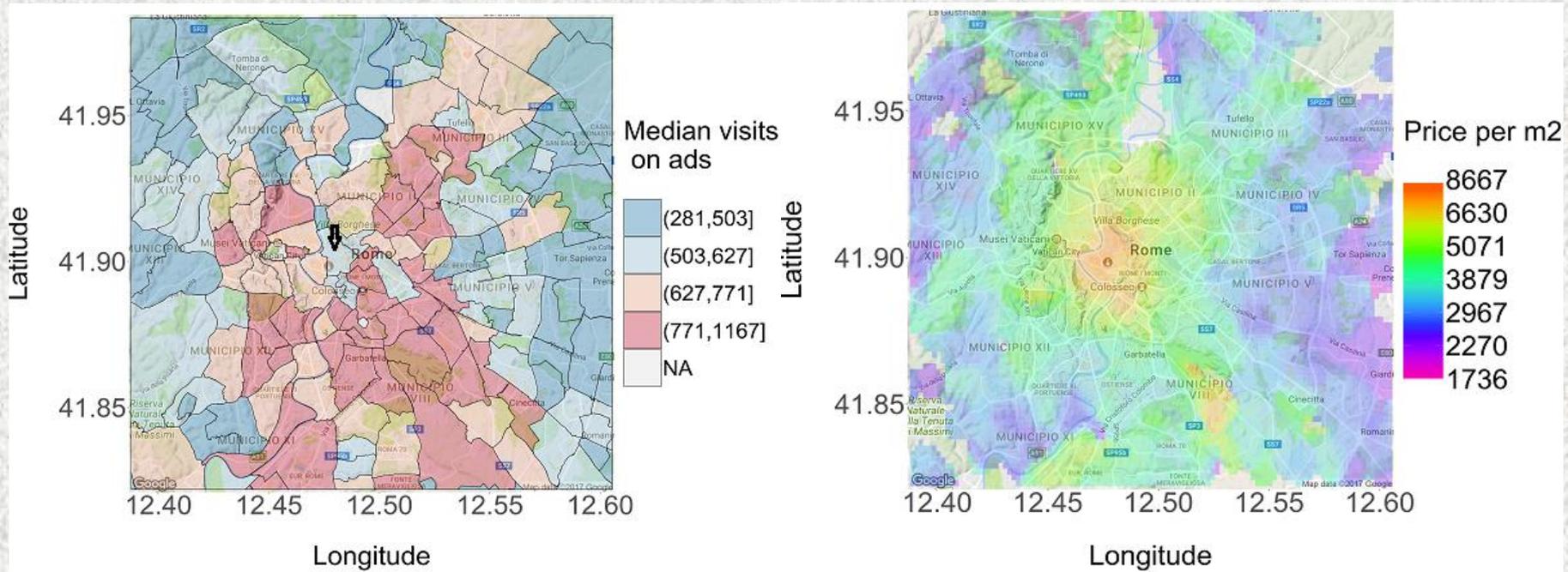


(c) Energy class



(d) Maintenance status

Applications: clicks and demand tightness



Tightness as leading indicator of prices

$$\log(P_{i,t}) = \alpha + \beta_1 \log(D_{i,t-1}) + \beta_2 \log(D_{i,t-2}) + \gamma T + \delta_0 \log(P_{i,t-1}) + \delta X_i + \epsilon_{i,t}$$

- i is an OMI micro-zone (neighborhood)
- t is a quarter
- D is tightness (number of clicks per ad)
- T is quarter fixed effects
- X_i is OMI micro-zone characteristics (from census). No chance to use OMI fixed effects (so far).

<i>Dependent variable:</i>	
Price per m2 (t) [log]	
Price per m2 (t-1) [log] (δ_0)	0.506*** (0.010)
Tightness (t-1) [log] (β_1)	0.044*** (0.013)
Tightness (t-2) [log] (β_2)	0.048*** (0.013)
Constant (α)	3.211*** (0.120)
Observations	6,288
R ²	0.811
Adjusted R ²	0.807

Conclusion

- Online granular housing data have **large potential** both for policy and for research, for the analysis of the current economic situation and for long run economic outlook. They address many shortcomings of existing data sources (at least for Italy).
- **Next steps:** matching with notary deeds data to obtain transaction prices and with mortgage data to obtain loan-to-value
- At INET we build **Agent-Based Models (ABMs)**. Using these datasets would allow ABMs to deliver their full potential in quantitative forecasting. (Geanakoplos, Axtell, Farmer, Howitt *et al.*, 2012, *AER P&P*; Haldane and Turrell, 2017, *Oxrep*)

Thank you for your attention!

