BANQUE DE FRANCE
EUROSYSTÈME

APPORTS DU *WEB-SCRAPING* POUR UN SUIVI À HAUTE FRÉQUENCE DU MARCHÉ IMMOBILIER : LE CAS DU ROYAUME-UNI

CNIS – COMMISSION SYSTÈME FINANCIER ET FINANCEMENT DE L'ÉCONOMIE

27 MAI 2021

J-C. BRICONGNE

BANQUE DE FRANCE

BANQUE DE FRANCE

Messages principaux :

- Les données issues du *web-scraping* (téléchargement massif de données sur Internet) permettent un suivi granulaire [tout en veillant à la représentativité] et plus régulier des marchés immobiliers, et offrent un angle complémentaire, avec le point de vue des vendeurs

- Dans le cas du Royaume-Uni, le marché a été gelé pendant le premier confinement, avec un effondrement temporaire des offres postées, et une relative stabilité des prix (absolus) postés sauf sur Londres

- Les données de *scraping* sont cohérentes avec les sources officielles et permettent d'obtenir des prix en niveau, qui apportent de l'information par rapport aux indices, et permettent de mieux prévoir les retournements et d'estimer les stocks de richesse immobilière

Nombreuses informations publiques en temps réel sur les sites immobiliers au UK : 92 % des agences immobilières publient des annonces sur Internet

Exemples d'utilisation de données alternatives pour suivre le marché immobilier : Kulkarni *et al.*, 2009 avec les *Google Trends* ou *web-scraping* (par exemple Hanson et Santas, 2014 ou Bricongne *et al.*, 2019)

Le *web-scraping* offre le point de vue des vendeurs (par opposition avec les transactions finales qui résultent de l'interaction entre vendeurs et acheteurs)

Couverture du marché UK dans cette période particulière de Covid-19 & Brexit, sachant que l'approche peut être étendue à d'autres zones géographiques

Couverture des sites utilisés variable selon les régions, mais assez élevée dans tous les cas pour assurer une bonne représentativité au niveau infra-national

Données récupérées par *web-scraping* des sites majeurs d'annonces immobilières au UK : Rightmove, Zoopla et OnTheMarket.

Pour améliorer la couverture sur des régions plus spécifiques, PropertyPal – un site immobilier spécialisé dans l'Irlande du Nord – et S1Homes, son *alter ego* pour l'Écosse, sont également couverts.

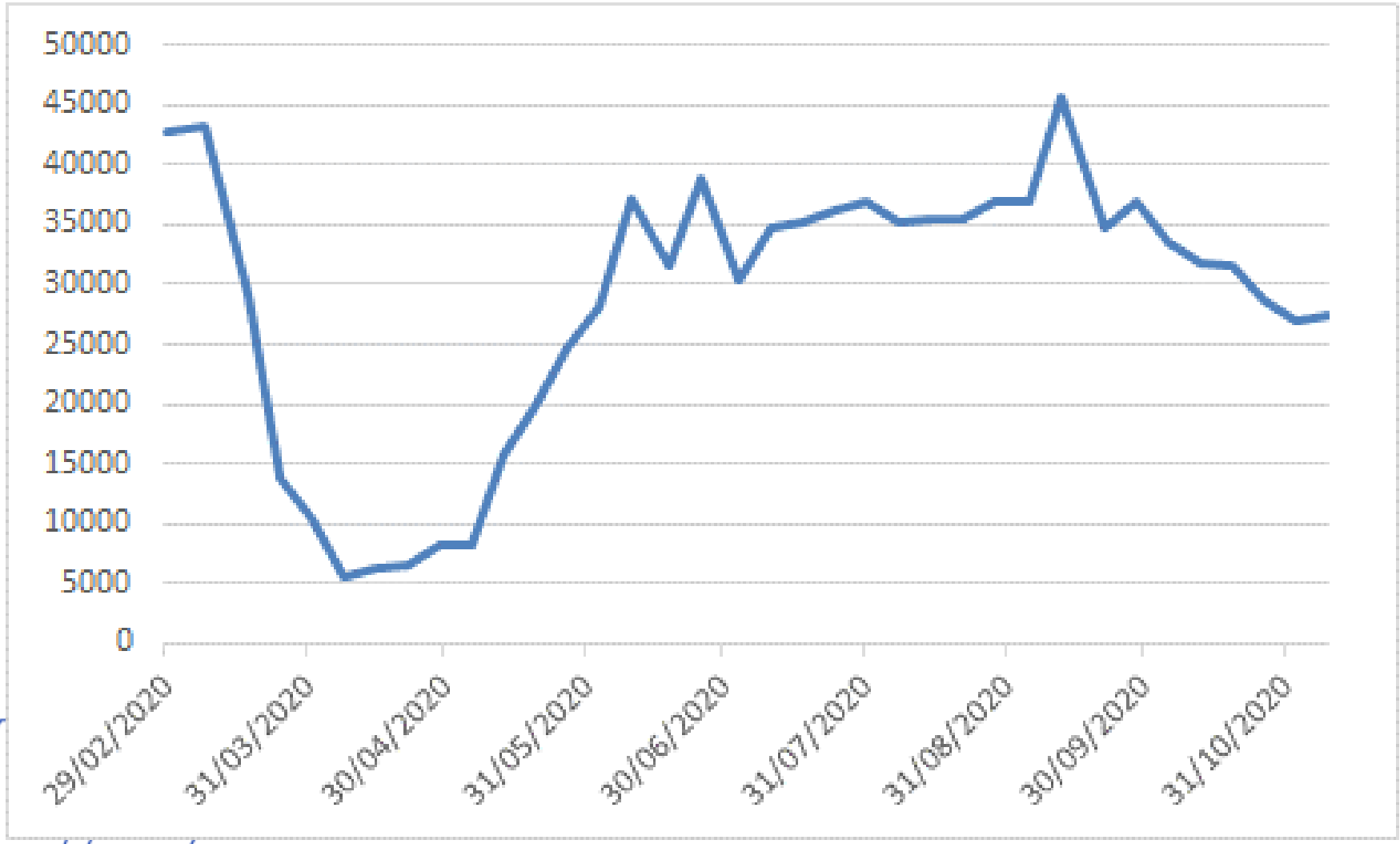
Grâce à ces sites *Web*, plus de 1,5 million d'offres immobilières sont téléchargées en moyenne chaque jour. Environ les deux tiers de ces dernières sont des offres de vente (cf. tableau). Les données pour Zoopla sont téléchargées (et nettoyées) depuis début mars 2020 alors que les données d'autres sites *Web* le sont depuis juillet 2020.

BANQUE DE FRANCE
EUROSYSTÈME

Nombre moyen d'offres *web-scrapées* par jour (par URL unique)

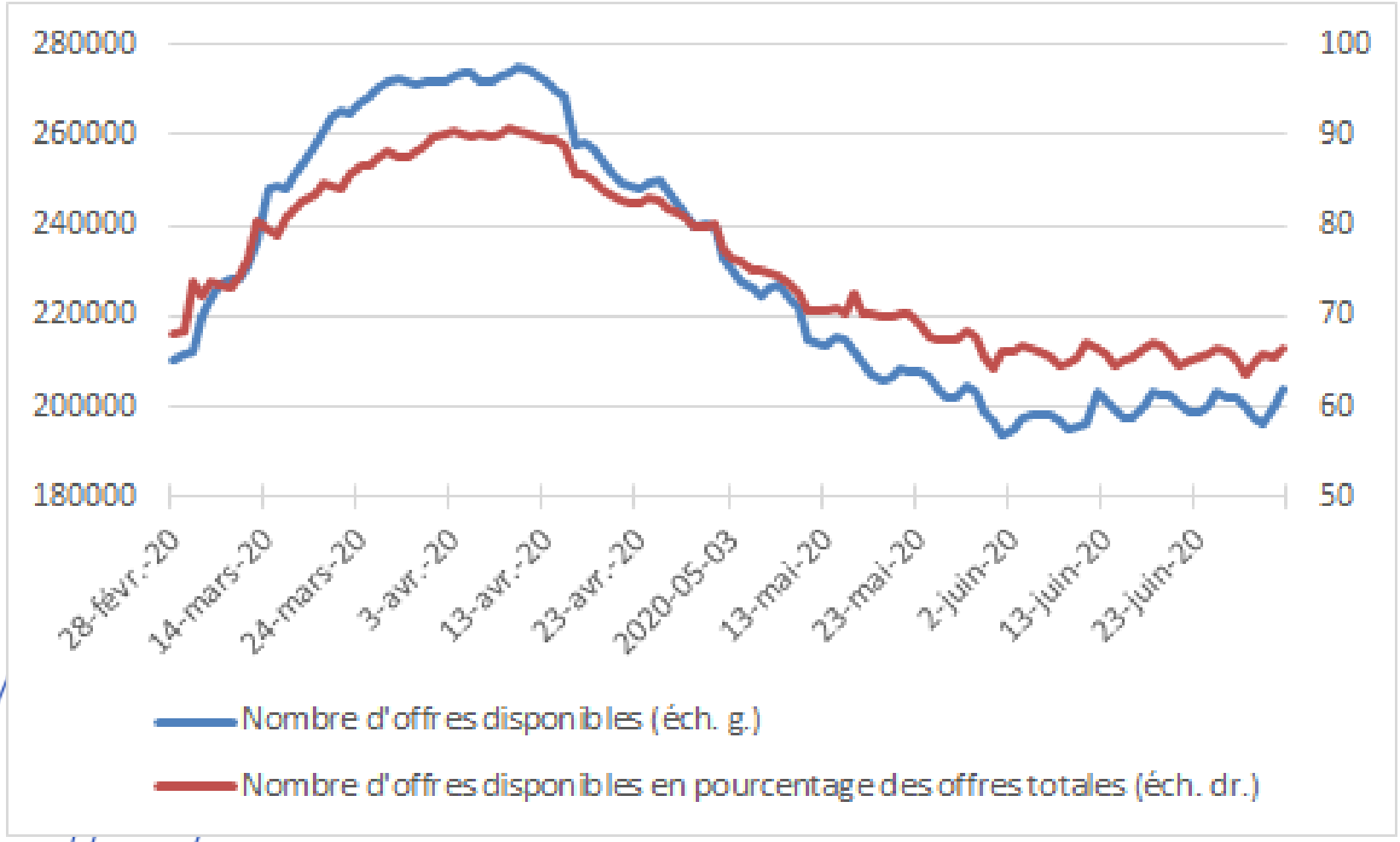
Nom du site	Vente (résidentiel)	Location (résidentiel)	Vente (commercial)	Location (commercial)
Zoopla	300 000	200 000	15 000	35 000
Rightmove	500 000	200 000	20 000	25 000
OnTheMarket	250 000	100 000	5 000	10 000

Des nouvelles offres immobilières hebdomadaires en baisse très forte lors du premier confinement et moindre lors du deuxième (en unités/semaine)



Sources : Zoopla calculs des auteurs.

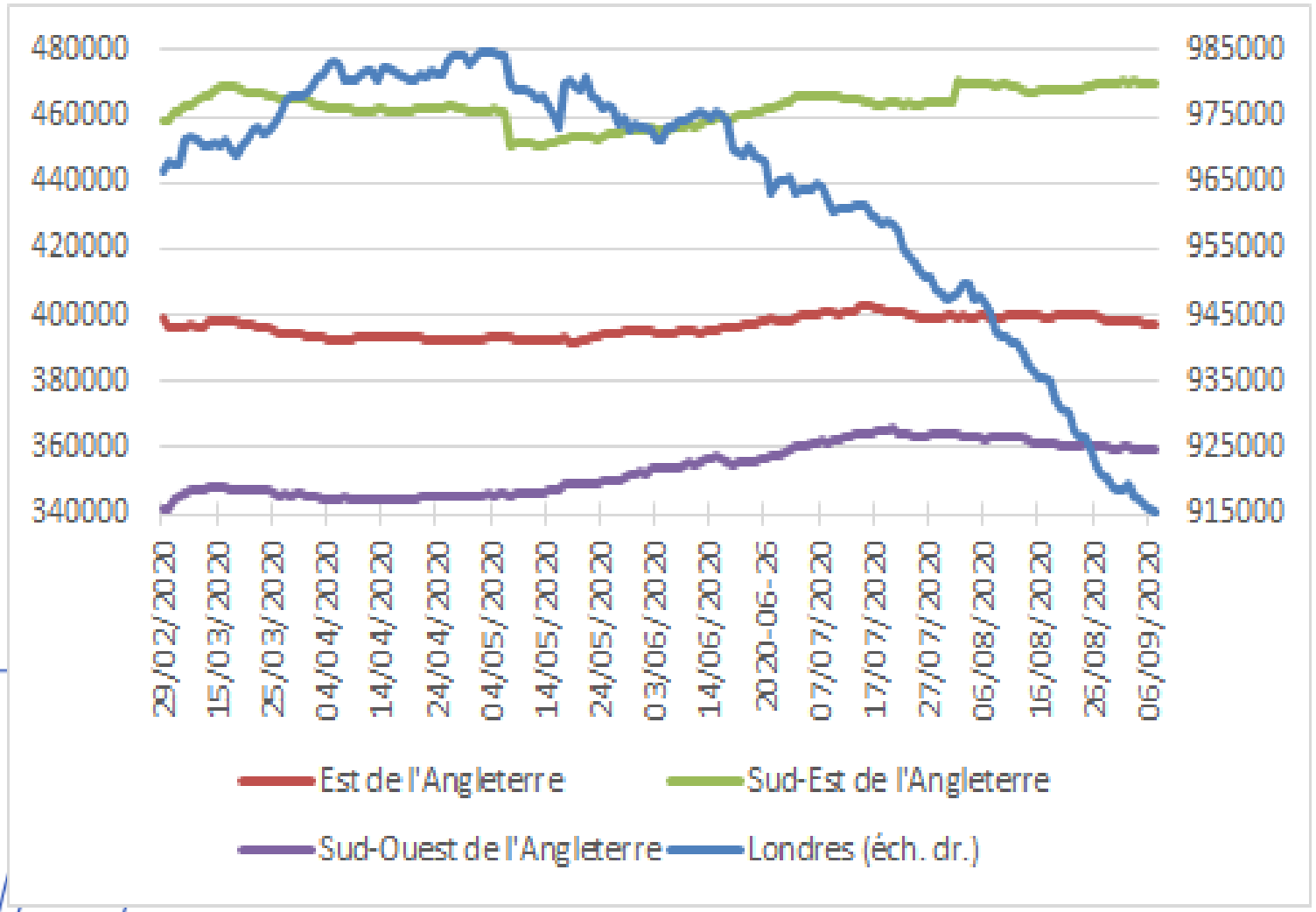
Un marché quasi gelé : offres encore disponibles après un mois sur Internet (à rapprocher d'une chute de 95% des transactions) (en unités)



Sources : Zoopla, calculs des auteurs.

BANQUE DE FRANCE

Prix de vente moyen affiché par région
(en livres sterling par logement)



Sources : Zoopla et calculs des auteurs.

BANQUE DE FRANCE
EUROSYSTÈME

Les statistiques issues du *web-scraping* sont complémentaires des statistiques officielles et permettent un suivi en temps réel et granulaire.

De façon plus structurelle, elles permettent aussi un calcul de prix (ou de loyers) en niveaux.

Les deux sources sont utiles notamment pour les banques centrales : risques de retournement, stabilité financière…

LE *WEB-SCRAPING* PERMET UN CALCUL DE PRIX EN NIVEAU

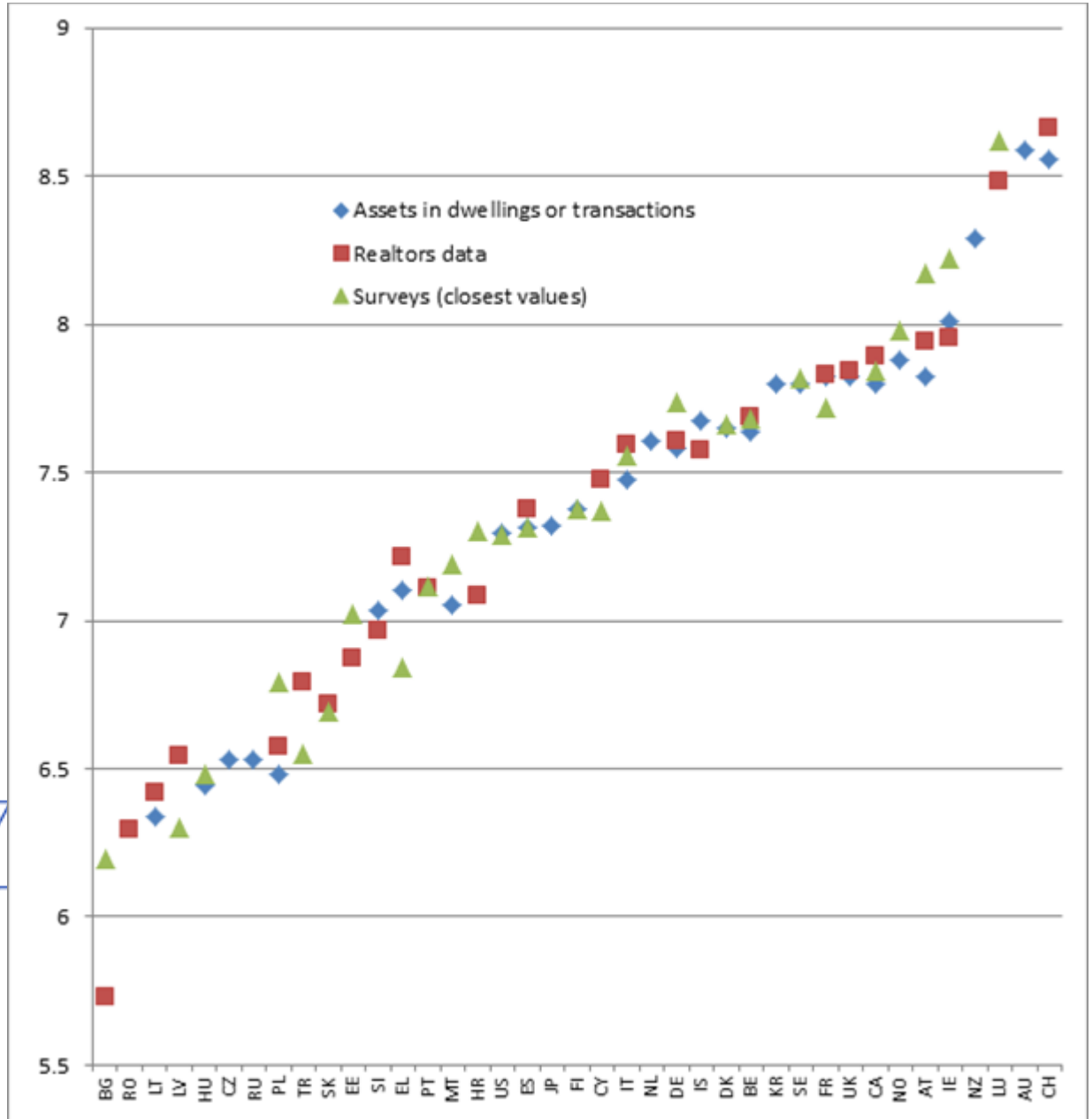
Méthode basée sur données individuelles (*web-scraping*, voire transactions : IE, FR ou MT), au niveau infra-national i (région, municipalité…) avec w_i le stock de m^2 :

$$p = \frac{\sum_{i=1}^{n} w_i * prix\ moyen\ pondéré\ au\ m^2 i}{\sum_{i=1}^{n} w_i} \quad (2)$$

Avec $\text{prix moyen pondéré au m}^2_{\text{i}}$ égal à la somme des prix divisée par la somme des surfaces, equivalent à :

$$\text{prix moyen pondéré au m}^2_{\text{i}} = \frac{\sum_{j=1}^{n} surface(j) * prix\ au\ m^2\ (j)}{\sum_{j=1}^{n} surface(j)}$$

WEB-SCRAPING : DES ORDRES DE GRANDEUR COHÉRENTS



BANQUE DE FRANCE

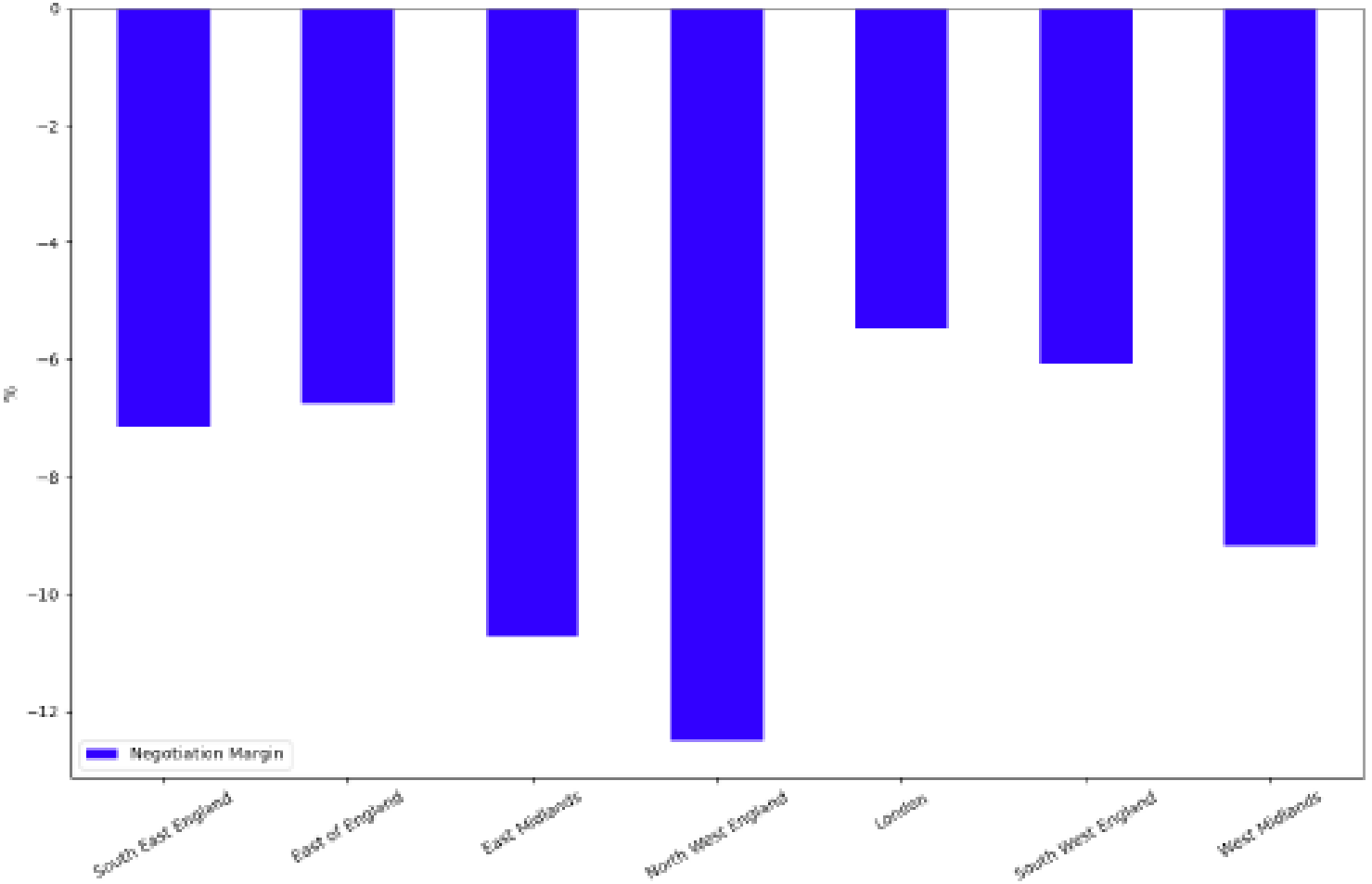
Quelques applications possibles :

- Estimer la richesse immobilière (nationale ou régionale), surtout dans les pays qui n'en publient pas
- Prévoir les retournements (information additionnelle par rapport aux indices) : prix sur revenu, capacité d'achat, marges de négociation si données de transactions sont publiées et les prix postés

BANQUE DE FRANCE
EUROSYSTÈME

Figure 8: Negotiation margin per region

Source: Real-estate websites, notarial data, and authors' calculation



BANQUE DE FRANCE
EUROSYSTÈME

Prolongement : le projet Alter Hous en liaison avec le réseau REFINE :

- Proposer des statistiques alternatives (non officielles mais présentant des principes communs de fiabilité / représentativité) issues du *web-scraping* ou d'autres sources (satellites…) et des articles méthodologiques sur l'immobilier
- Projet commun avec la Commission européenne, l'OCDE, la Banque mondiale, l'AMSE, le LIEPP, l'ESRI irlandais…

CONCLUSIONS / PERSPECTIVES / ENJEUX

- Approfondir la comparaison avec les statistiques officielles, en gardant en tête les biais possibles (marges de négociation : quels écarts avec les derniers prix postés ? Travaux en cours)

- Extension aux loyers, au prix des terrains et à l'immobilier commercial (en cours)

- Utilisation de la granularité pour étudier les effets de certaines politiques

- Extension de l'approche à d'autres pays : travaux en cours avec la Banque mondiale

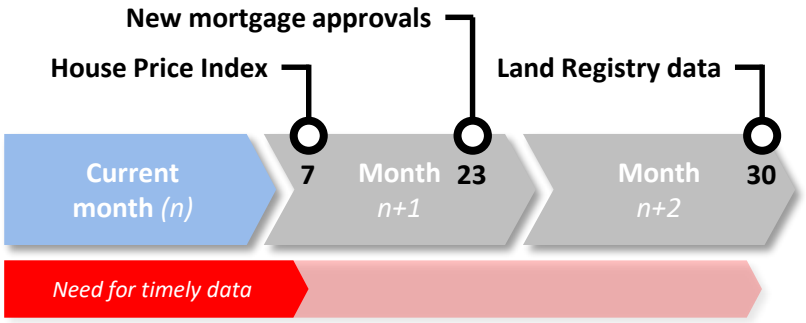
- Questions juridiques liées au *web-craping* et de sécurisation des sources, qui peuvent changer au cours du temps, voire disparaître…

ANNEXES

MOTIVATIONS
INFORMATION AVAILABLE PUBLICLY AND IN REAL-TIME

Lags in official statistics …

- Official statistics published only **after month end** and often only aggregated at **national level** or at large region level (Wales, Scotland, Northern Ireland, and England – with potentially a rough decomposition of the latter)



… but information publicly available

- Price information available **publicly** on real-estate website, in **real-time**, and with a high **granularity**

- Evidence that a **large share of offers are posted on the web** (92% of realtors post their ads on those websites)

Real-time housing price index (possibly more granular too)

DATA
WEB-SCRAPING

On each of the 5 websites (Zoopla, Rightmove, OnTheMarket, S1homes, and PropertyPal)

Request all offers corresponding to a ZIP code

Loop over all ZIP codes

Scan the "results" web page to retrieve information for all available offers

ZIP code	City	Address	Surface	# of bedrooms	# of bathrooms	# of living rooms	Transaction type	Building type	Type of good	Price	General description
E1 7A E	London	-	45	2	1	1	Auction	New	Flat	525,000	"Lovely flat with large terrace"
E1 7A E	London	-	130	5	2	1	Sale	Ancient	House	4,500,000	"Garage and safe neighbourhood"
...											
E1 7A E	London	-	-	2	2	2	Sale	Ancient	Duplex	850,000	"Balcony and view on Westminster"

Around **1.5 million** offers in total per day

Figure 13: Median selling prices by region
Source: Zoopla and authors' calculation

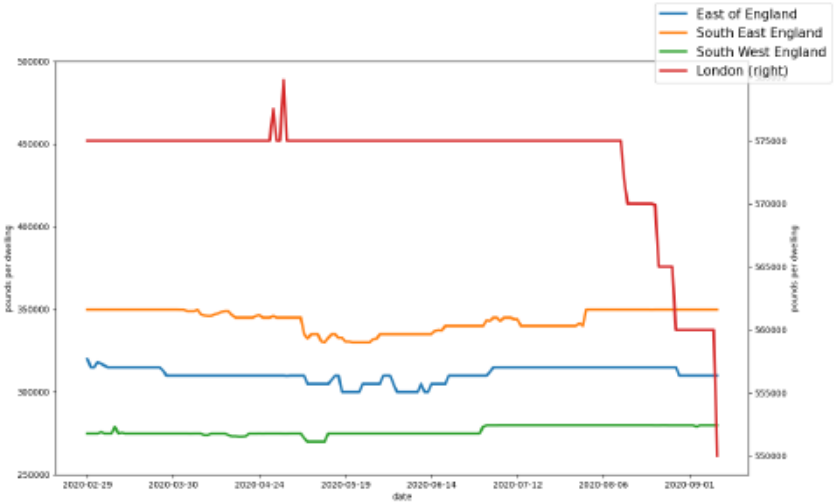
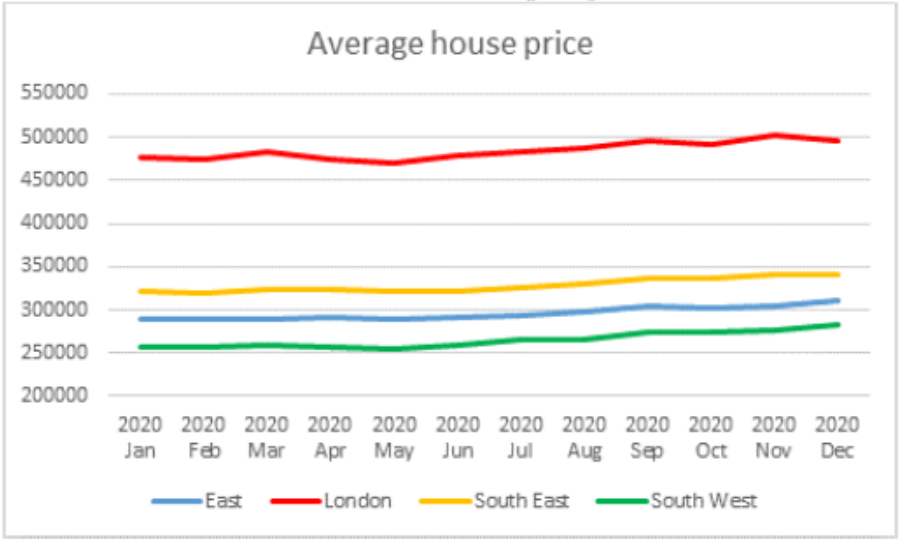


Figure 14: Statistically-adjusted geometrical average selling prices by region
Source: Land Registry



Price level (€/m², from BIS, OECD, Eurostat, WidWorld, NSIs, central banks, censuses, other

	BE	BG	CZ	DK	DE	EE	IE	EL	ES	FR	HR	IT	CY	LV	LT	LU	HU	MT	NL	AT	PL	PT	RO	SI	SK	FI	SE	UK	EA19
1970	124.0				321.8		141.2					150.7														141.6	193.2	92.2	
1971	125.2			157.4	361.4		154.5		70.3			153.6														142.7	202.3	100.5	
1972	140.5			188.5	405.4		156.2		74.8			157.9														150.4	217.3	123.9	
1973	164.5			225.5	467.1		158.9		92.3			172.6														192.2	235.2	160.2	
1974	195.2			243.4	525.1		182.5		124.5	263.1		240.1				216.8			211.2							236.4	267.1	163.3	
1975	222.3			284.3	531.5		207.8		135.8	314.8		293.9				195.9			246.6							249.3	314.5	160.7	
1976	305.1			342.8	621.6		215.5		143.8	339.0		274.7				288.4			354.5							279.6	399.6	152.0	
1977	357.9			358.9	678.7		265.6		153.3	380.9		283.3				348.2			489.2							246.9	377.8	168.5	
1978	406.8			421.6	748.8		337.4		209.0	425.7		304.5				389.8			544.5							233.4	418.3	186.3	
1979	433.8			419.8	820.7		424.4		221.6	478.5		358.0				419.2			511.2							249.2	456.2	251.6	
1980	426.6			404.6	857.0		462.5		214.2	559.3		454.0				445.7			460.6							303.2	501.7	358.4	472.9
1981	399.4			382.9	961.8		564.2		217.9	594.4		567.8				458.6			431.0							370.1	481.4	367.3	536.3
1982	354.4			366.2	1032.5		599.5		195.5	596.9		623.3				408.0			404.7							405.5	409.2	354.9	551.7
1983	342.8			441.2	1054.7		613.9		218.6	592.8		649.7				403.9			412.7							521.9	438.9	417.3	580.9
1984	360.5			521.5	1043.8		652.6		248.3	622.4		667.7				438.7			412.9							606.8	467.9	426.8	590.5
1985	366.5			612.7	1055.0		661.2		245.9	655.0		626.3				464.4			423.2							608.8	460.9	459.9	602.4
1986	396.1			692.7	1112.8		634.9		274.3	668.5		663.5				534.0			465.3							594.9	455.7	443.4	626.8
1987	422.2			634.7	1111.9		618.0		379.4	704.9		667.0				577.6			493.0							661.7	493.7	538.8	678.0
1988	451.8			635.1	1121.5		672.7		501.8	768.5		748.3				662.1			509.5			489.5				940.7	614.0	738.4	739.3
1989	526.3			642.4	1188.5		735.6		627.2	883.2		889.9				810.9			557.6			554.4				1159.0	706.3	769.7	824.6
1990	587.7			594.0	1270.8		838.7		727.8	956.9		1028.2				934.2			564.4			624.5				1074.1	762.6	797.9	918.1
1991	616.4			598.2	1332.3		864.1		838.0	1006.1		1189.5				1015.0			581.4			757.7				823.3	839.7	777.4	994.5
1992	686.2			616.0	1461.9		907.9		773.2	1024.2		1179.6				1168.5			657.9			862.1				598.7	662.7	669.7	1044.5
1993	720.7			611.8	1548.1		870.8		671.9	1023.4		1156.7				1149.6			722.4			787.5				535.1	542.1	696.0	1072.9
1994	788.9			693.3	1637.8		907.5		663.5	1021.8		1092.4				1203.8			806.6			810.3				628.5	574.2	685.4	1094.9
1995	834.5			765.1	1680.9		934.6		697.3	1033.8		1055.9				1221.6			860.1			820.1				617.6	607.6	641.1	1108.4
1996	822.7			829.6	1605.1		1117.1		690.3	1023.5		1193.3				1257.1			919.6			835.8				640.3	617.5	763.7	1124.9
1997	828.7	97.2		915.3	1552.9		1237.2	867.6	687.8	1011.0		1215.9				1281.0			1009.1			839.7				730.8	650.5	918.7	1139.2
1998	890.7	111.8		1008.1	1555.0		1505.3	939.0	724.1	1040.7		1219.0				1340.8			1131.6			884.3				811.1	655.6	967.8	1161.6
1999	954.0	112.7		1076.9	1556.6		1829.0	1020.7	799.2	1111.3		1231.8			211.4	1427.1			1316.0			963.7				883.6	794.5	1217.7	1209.4
2000	1005.8	112.0	247.2	1143.9	1563.8		2206.6	1093.9	867.8	1209.5	788.5	1280.0		205.9	206.0	1530.6		374.3	1555.7	1389.5		1038.4				934.2	856.8	1394.0	1274.4
2001	1054.2	112.8	296.8	1215.2	1565.3		2539.3	1251.4	952.6	1303.7	828.0	1354.5		213.9	245.1	1746.4		418.5	1704.2	1400.1		1094.9				919.4	878.1	1545.9	1339.8
2002	1121.3	102.3	338.4	1260.3	1543.6		2652.3	1424.7	1099.2	1417.7	859.1	1515.8	1472.7	284.4	264.1	1893.8		450.1	1812.4	1409.1		1101.1				977.7	949.3	1679.3	1427.1
2003	1199.5	114.1	363.2	1297.4	1549.7		3032.4	1501.6	1288.1	1586.2	867.2	1609.0	1457.5	318.1	289.8	2106.2		507.6	1857.0	1412.6		1113.4		743.3		1039.0	1019.6	1794.4	1510.2
2004	1303.6	168.1	388.0	1414.8	1526.3	476.9	3371.0	1536.4	1508.0	1826.7	960.4	1708.2	1626.3	314.9	313.2	2400.6		589.6	1931.3	1386.1		1120.8		804.7		1122.6	1122.6	2006.4	1609.9
2005	1468.8	229.6	413.9	1658.8	1544.3	632.8	3697.0	1703.8	1712.3	2106.7	1111.2	1838.5	1747.7	391.3	393.6	2674.5		642.3	2023.3	1455.1		1146.6		921.7	401.3	1214.1	1182.9	2306.0	1726.9
2006	1612.0	263.4	470.2	2059.2	1538.8	946.1	4246.7	1929.3	1941.9	2362.0	1312.1	1944.0	1939.1	606.3	505.2	2972.4		770.0	2109.7	1515.2		1169.9		1076.1	515.3	1298.7	1381.3	2537.9	1840.5
2007	1737.6	339.5	579.2	2113.8	1505.4	1142.6	4563.7	2044.1	2132.4	2498.3	1474.1	2041.0	2140.6	827.2	638.2	3180.0	667.9	931.7	2211.8	1585.9		1185.9		1330.3	681.4	1375.3	1488.4	2554.2	1925.1
2008	1814.2	424.3	647.6	2006.6	1525.8	1032.6	4247.0	2078.5	2101.8	2518.7	1511.7	2094.9	2261.5	822.3	695.5	3286.4	650.5	1034.9	2260.4	1601.8	768.8	1230.8	979.1	1423.5	895.6	1386.2	1306.8	1877.9	1962.9
2009	1805.7	337.7	632.0	1768.4	1538.8	648.5	3433.5	2001.2	1963.0	2362.8	1447.0	2084.6	2113.1	514.8	487.3	3248.4	607.9	990.1	2159.6	1665.3	756.2	1219.5	708.5	1288.6	781.4	1406.1	1426.8	1835.1	1893.2
2010	1862.4	303.4	655.9	1815.1	1555.4	685.3	2970.2	1908.1	1928.3	2476.3	1340.2	2067.9	1992.0	458.3	451.3	3424.5	577.3	1000.8	2122.3	1768.0	751.6	1228.8	651.2	1290.3	750.1	1494.8	1762.1	2001.8	1906.4
2011	1937.2	286.7	638.1	1788.9	1609.2	743.4	2462.8	1803.8	1781.0	2620.0	1315.1	2083.9	1960.1	513.3	481.2	3550.5	492.6	987.2	2080.4	1878.5	670.7	1168.5	562.9	1325.3	738.9	1542.4	1817.0	2032.8	1926.2
2012	1980.6	281.2	644.7	1734.2	1664.8	797.6	2131.5	1593.5	1517.7	2605.9	1291.0	2025.2	1900.3	529.8	480.0	3699.7	510.5	1017.3	1940.6	2015.9	708.3	1085.9	519.8	1233.9	718.9	1579.7	1909.1	2089.0	1894.7
2013	2003.9	275.0	591.2	1802.3	1716.7	882.5	2158.1	1420.6	1379.3	2555.6	1228.8	1909.6	1822.5	562.0	485.9	3883.8	489.5	1013.1	1824.2	2120.2	664.2	1065.5	515.4	1169.2	725.4	1597.9	1950.8	2097.4	1858.2
2014	1992.7	278.9	598.9	1873.8	1770.5	1003.6	2514.9	1314.5	1383.6	2515.7	1204.5	1826.1	1790.4	595.8	517.1	4054.3	480.1	1039.0	1839.1	2194.6	652.0	1110.7	503.5	1092.0	735.6	1592.3	2012.7	2425.1	1863.3
2015	2026.1	286.7	639.0	1999.6	1853.9	1072.4	2803.1	1248.9	1433.1	2478.5	1172.7	1778.1	1764.4	575.7	536.0	4273.1	542.2	1099.3	1904.6	2302.1	663.4	1144.6	513.2	1100.8	775.1	1592.3	2327.2	2726.9	1894.0
2016	2079.6	306.8	685.0	2101.1	1965.2	1123.4	3012.2	1218.4	1499.4	2503.5	1195.4	1763.8	1769.2	624.6	565.0	4529.9	626.9	1159.2	2005.0	2498.5	653.3	1226.1	541.9	1136.6	827.0	1602.0	2432.0	2500.8	1954.7
2017	2156.0	333.4	809.4	2190.8	2043.0	1185.1	3340.7	1206.2	1592.2	2585.6	1261.1	1756.7	1808.9	679.8	615.3	4785.0	679.3	1215.7	2166.7	2630.4	716.4	1339.4	559.9	1228.1	875.8	1625.2	2511.0	2526.0	2034.8

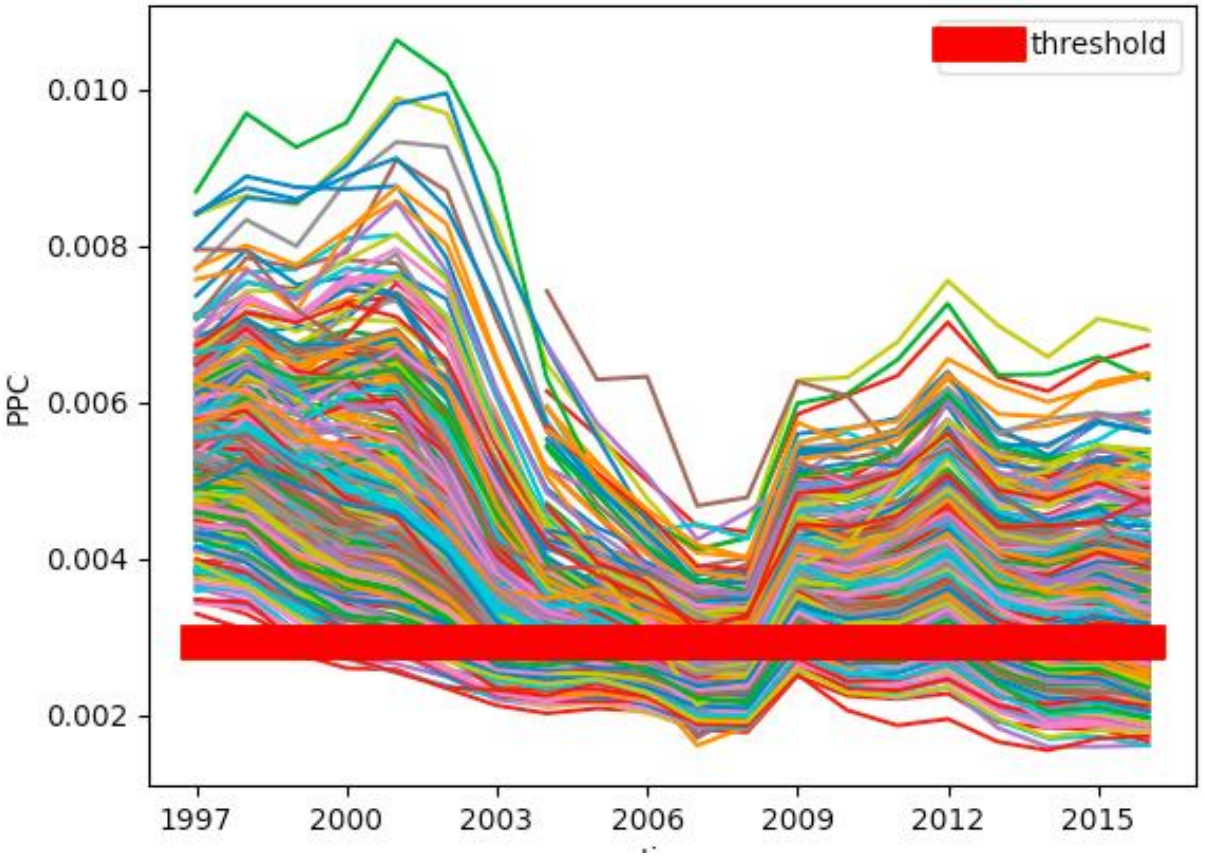
	IS	NO	CH	TR	AU	CA	HK	JP	NZ	RU	KR	US
1970		129.7						158.2				157.6
1971		133.3						191.3				162.5
1972		143.2						234.7				160.6
1973		165.8						312.0				174.4
1974		200.2	650.3					340.3				183.5
1975		226.2	619.4			284.0		323.1				211.4
1976		272.7	655.8			329.9		354.4				234.0
1977		277.9	742.0			294.8		417.3				241.5
1978		266.0	853.9			254.9		486.3				245.2
1979		269.7	897.1			272.2		410.3				266.5
1980		311.6	978.3			331.6		602.4				316.1
1981		368.2	1278.2		747.1	460.4		744.5	307.6			400.5
1982		371.3	1356.0		754.3	469.4		842.2	402.1			464.1
1983		424.5	1511.1		854.9	574.3		1045.8	463.7			563.9
1984		479.7	1566.6		1021.9	628.7		1162.6	443.6			686.5
1985		553.8	1641.6		742.8	499.7		1190.0	425.6			576.2
1986		617.7	1844.6		642.1	488.0		1272.2	409.4			510.3
1987		646.6	2052.8		606.1	493.3		1477.5	492.0			446.1
1988		709.2	2231.9		977.1	703.4		1690.3	590.9			522.3
1989		646.2	2404.1		1103.7	801.2		1552.4	580.3			538.7
1990		589.6	2537.7		960.3	682.9		1639.5	531.4			487.7
1991		553.7	2390.7		987.9	727.0		1887.2	485.2			502.6
1992		502.7	2357.1		1005.2	753.3		2012.7	515.1			570.5
1993		507.2	2384.9		1102.9	797.5		2333.3	631.8			634.2
1994		578.8	2440.7		1185.5	707.7		2317.5	746.8			594.6
1995		621.7	2501.9		1077.7	650.8		2063.0	779.9			570.9
1996		698.8	2119.6		1218.8	678.7		1882.5	977.2			616.4
1997		777.1	2155.3		1179.6	754.4		1884.2	968.3		933.1	719.5
1998		789.7	2133.5		1127.4	652.5		2004.1	816.2		1130.2	713.9
1999		963.9	2134.1		1489.2	828.9	10923.7	2509.6	951.5		1376.4	879.8
2000	999.7	1096.0	2249.8		1483.2	904.2	10524.0	2320.6	868.2		1354.2	1013.8
2001	912.2	1213.1	2332.1		1601.1	939.4	9760.0	2056.6	880.2		1426.4	1144.5
2002	1036.5	1392.0	2377.0		1770.5	863.4	7290.2	1805.0	1029.9	258.8	1553.7	1029.7
2003	1098.6	1223.7	2247.2		2308.0	953.6	5353.1	1559.7	1278.9	287.4	1398.8	920.6
2004	1302.7	1377.3	2351.7		2359.8	1020.5	6280.4	1416.3	1524.6	359.5	1504.5	934.3
2005	1895.0	1537.9	2419.3		2604.4	1318.1	8574.5	1354.6	1907.6	494.2	1809.9	1191.5
2006	1772.5	1694.8	2486.0		2687.2	1322.4	7717.2	1163.3	1944.2	716.2	1867.1	1130.3
2007	1965.0	1974.8	2555.8		2958.5	1560.2	7688.9	1095.8	2122.5	997.1	1825.4	1011.5
2008	1128.5	1594.6	2955.7		2541.8	1398.6	9528.9	1442.3	1595.6	1081.2	1464.6	983.4
2009	962.9	1909.0	2972.0		3349.5	1527.7	9257.0	1287.0	1917.7	1028.2	1630.2	894.1
2010	1092.0	2198.1	3623.8	748.1	4558.5	1889.2	12389.1	1600.1	2250.8	1177.5	1852.7	935.1
2011	1106.3	2388.2	3964.9	680.0	4605.1	1999.1	15443.3	1736.3	2341.9	911.3	1943.8	925.5
2012	1106.4	2690.9	4186.4	771.9	4594.2	2107.8	17196.1	1517.8	2558.4	1085.6	2097.1	934.4
2013	1253.6	2459.6	4239.8	678.5	4037.5	1935.9	19324.3	1211.1	2670.7	1002.7	2010.5	958.2
2014	1396.4	2336.9	4446.1	796.1	4579.4	2126.3	23256.2	1225.6	3071.0	636.8	2236.4	1144.5
2015	1650.0	2334.6	5062.5	812.7	4970.0	2084.8	29986.0	1390.7	3347.0	578.0	2378.2	1345.6
2016	2148.6	2640.7	5190.0	781.5	5351.5	2442.1	29836.3	1510.0	3977.1	686.7	2438.5	1474.0
2017	2446.5	2559.7	4855.7	704.0	5514.9	2582.9	30381.4	1416.0		613.3	2448.3	1382.3

INDICATORS

PRICE-TO-INCOME RATIO & PURCHASING-POWER- CAPACITY

Purchasing-power-capacity compared with threshold signalling an housing bubble – by locality

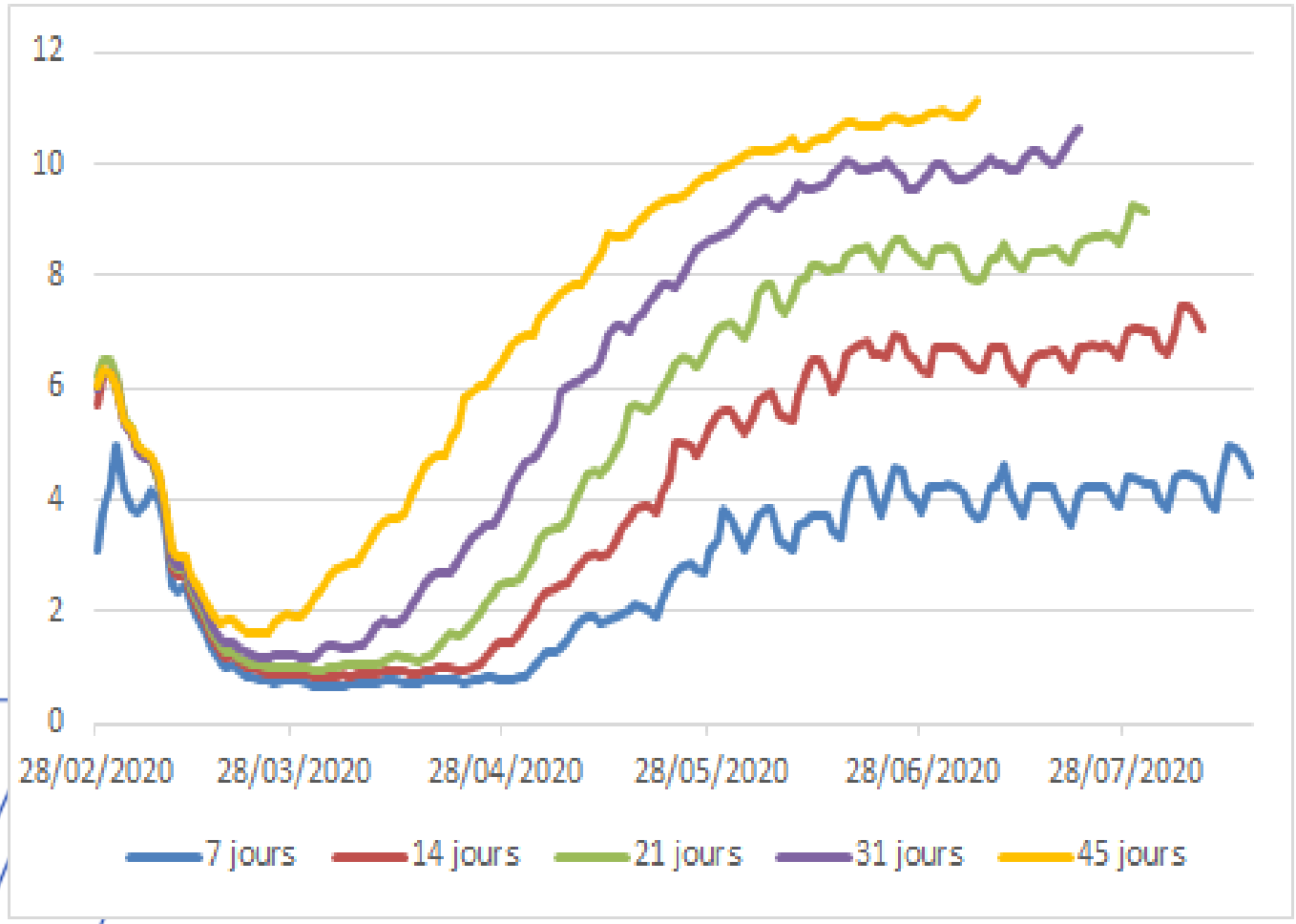
Sources: ONS and authors' calculation



BANQUE DE FRANCE
EUROSYSTÈME

BANQUE DE FRANCE

Part des offres dont le prix est révisé, selon le nombre de jours (en %)



Sources : Zoopla, calculs des auteurs

INDICATORS
RENT-TO-INCOME RATIO

- Compare each **rental offer** with its K closest **sale offers** using a K-nearest neighbor algorithm based on similar characteristics (location, surface, number of bedrooms, number of bathrooms, number of living rooms)

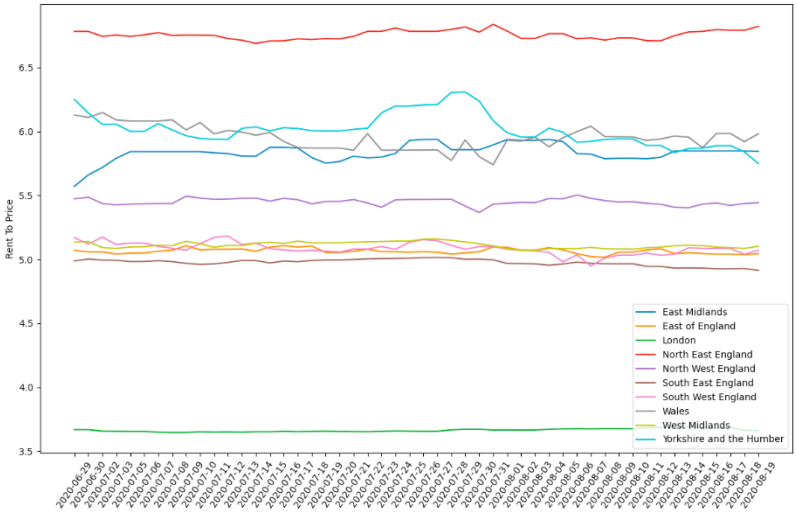
$$RTI = \frac{1}{N}\sum_{n}\frac{Annual\_Rent_n}{\frac{1}{K}\sum_k House\_Price_{n,k}}$$

- We **fix K=10** as a result of a trade-off: increasing K results in comparing the rental offer with less and less similar dwellings, but a low K introduces a large variability in the index

- To reflect the unequal coverage of our data, the ratio is estimated at local level then aggregated at national level **weighing each regional index by the stock of dwellings in that region** (source: ONS)

Rent-to-price ratios – by region

Sources: Zoopla and authors' calculation



- Little variation during the crisis but **large discrepancies across regions**. In particular, very low ratio in London – might signal over-valuation to some extent (in line with Marsden, 2015 or Petris et al., *in press*)

- At national level, index around **5.1%**

INDICATORS
PRICE-TO-INCOME RATIO & PURCHASING-POWER- CAPACITY

- Combine our dataset with **ONS' data on disposable income** – also available to some extent at local level – to derive a price-to-income ratio = house price divided by disposable income

- However, as defined in the literature, price-to-income ratio does not consider the **evolution of interest rates.** Propose an alternative index (the **purchasing-power-capacity**) computed as:

$$PPC = \frac{\omega}{r_t} \cdot PTI_t \cdot \left(1 - \frac{1}{(1+r_t)^d}\right)$$

ω = maximum effort rate (33%)

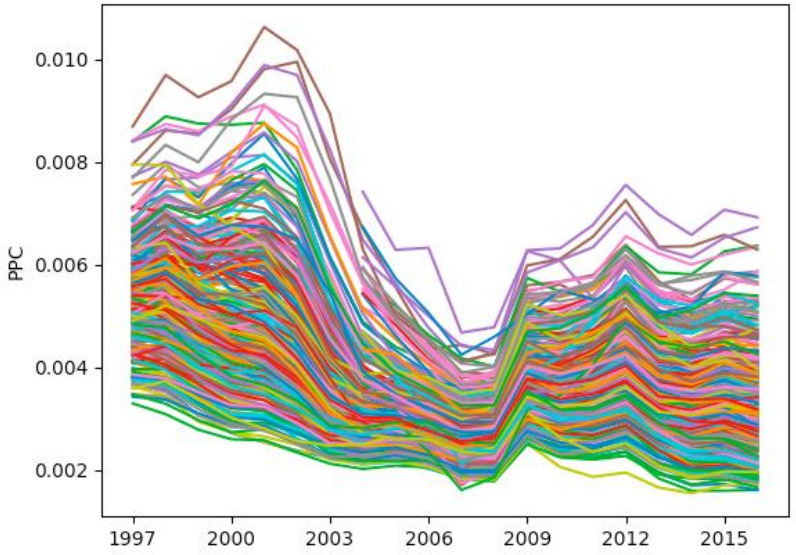
r_t = mortgage rate (source: BoE)

PTI_t = price-to-income ratio

d = median duration (15 years)

Purchasing-power-capacity – by locality

Sources: ONS and authors' calculation



- Large discrepancies but historical evidence for a **reversal in trend after the GFC**

- Empirically, find a **higher signalling power** (share of true detection of a forthcoming housing crisis) for the PPC than for the price-to-income ratio generally used in the literature

BANQUE DE FRANCE
EUROSYSTÈME

DATA
DATA CLEANING

1. Ensure **consistency** across units (price in GBP, area in squared meters, rent as monthly amount), numeric (surface/price can given be as a range), and text data (lowering characters, lemmatisation) in order to enhance comparability across offers

2. Remove **duplicates –** which can arise even on the same website

3. Exclude **commercial real-estate, miscellaneous offers** (garage, land, mobile home, bungalow, etc.) and **auctions** (since the price displayed would be the reserve price)

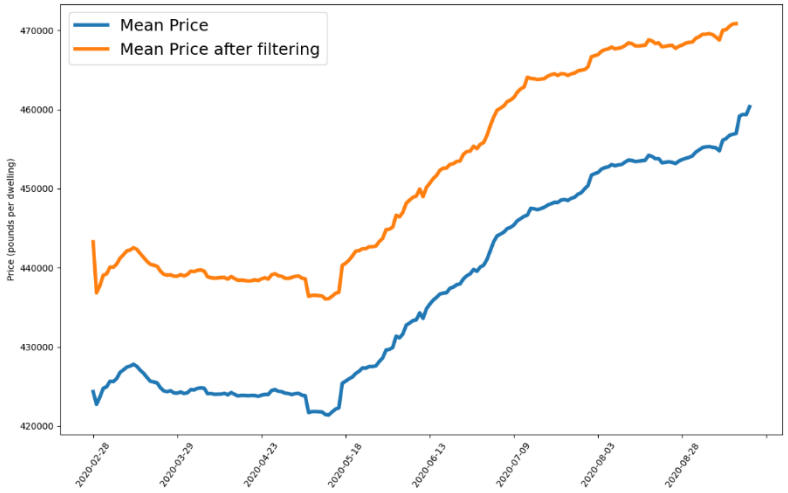
4. Remove **outliers** by winsorising at 1%

5. Using natural language processing on the general description of the offer, add **dummies to account for the presence of additional facilities** (e.g. presence of a garage, a garden, or a terrace) that might a premium to the price

REAL-TIME MONITORING
UK DURING COVID-19 (3/4)

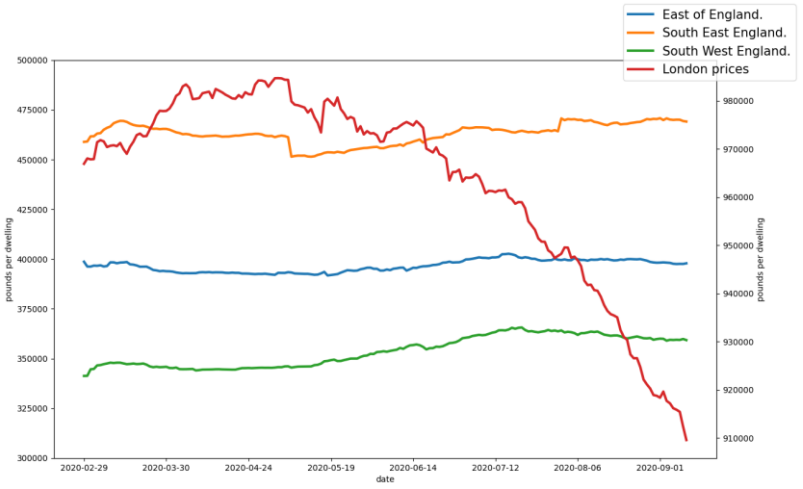
Selling price evolution – all UK

Sources: Zoopla and authors' calculation



Selling price evolution – by region

Sources: Zoopla and authors' calculation



- At national level: **slight decrease** from mid-February to the beginning of May (first lockdown), followed by a steady increase afterwards – with no apparent effect of the second lockdown

- At local level: **steady decline in London since the end of the lockdown** – in contrast with an increase in other regions (East England, South- East England, and South-West England).

- Looking at evolutions by quantiles shows that these **patterns are shared across all categories of dwellings**

BANQUE DE FRANCE
EUROSYSTÈME

Table 5. Signalling: possible cases

	No crisis episodes (NCE)	**Crisis episodes (CE)**
Crisis signal	False alert (FA)	True positive signal
No crisis signal	True negative signal	Missed crisis (MC)

	Threshold	Signal power	% False alerts (type 1 error)	% Missed crises (type 2 error)	# Crisisy ears	Size of the sample	2SD lowerbound	2SD upperbound
	(1)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Full sample								
PTI in level	10.21	0.32	0.28	0.4	40	722	9.93	10.48
Max PTI over 3 previous years	10.95	0.44	0.21	0.35	34	680	9.84	12.06
PTI/whole period mean PTI	1.03	0.26	0.39	0.35	40	722	0.93	1.13
PTI/real time mean PTI	1.20	0.06	0.19	0.75	40	722	0.87	1.53
Sample where the PTI of the first five years of each country are deleted								
PTI in level	10.20	0.33	0.27	0.4	30	593	9.94	10.47
Max PTI over 3 previous years	10.95	0.49	0.18	0.33	27	538	10.44	11.46
PTI/whole period mean PTI	0.96	0.22	0.55	0.23	30	593	0.81	1.11
PTI /real time mean PTI	1.33	0.07	0.10	0.83	30	593	0.96	1.70
Sample where the PTI of the first ten years of each country are deleted								
PTI in level	10.19	0.39	0.24	0.37	19	452	9.93	10.44
Max PTI over 3 previous years	10.02	0.49	0.26	0.25	16	408	9.06	10.98
PTI/whole period mean PTI	1.14	0.36	0.22	0.42	19	452	1.03	1.25
PTI /real time mean PTI	1.25	0.08	0.18	0.74	19	452	0.94	1.55