Some Transaction Prices can be Detrimental to your House Price Index

Robert J. Hill[2], Norbert Pfeifer[2], **Miriam Steurer[2],** Radoslaw Trojanek[3]

[1] Graz University, Austria
[2] Poznan University, Poland

May 15, 2022

Table of Contents

Introduction

Motivation

- House Price Indices (HPIs) inform the public, banks, and government institutions on the price developments in housing markets.
- Housing downturns are a good predictor of recessions (Leamer, 2007).
- Economic crises that involve the housing market tend to be more severe and prolonged.

  $\Rightarrow$ timely information on what is going on in the housing market is important to establish optimal macroprudential and monetary policies.

The Importance of Transaction Data

- Constant-quality House Price Indices need micro-level price data.
- Different data sources for micro-level house-price data, e.g., transaction data, appraisal data, online housing platforms.
- Transaction prices are the "gold standard" for compiling house price indices (Eurostat, 2013, IMF,2020).
  - Appraisal data (e.g., from banks, real estate agents, or property tax departments) tend to be downward biased and backward looking.
  - Prices from online platforms reflect the wishes of the seller side, but typically not the realised purchase price.

The Problems with Transaction Data

- Transaction data are typically based on the purchase contracts between sellers and buyers.
- In some countries there are long time lags before these prices are recorded and made available to index compilers.
- These lags undermine the timeliness and usefulness of the HPI itself as well as the broader CPI (if the HPI is a sub-component). (Shimizu et al. 2016)
- We concentrate on a second – less appreciated – problem with transaction data: namely that transaction prices for new-build properties often take years to be recorded in transaction data.

New-build properties

- New-builds are often sold before the building process is complete, and sometimes even before the building process started (i.e., “off the plan”). Prices are set when these preliminary agreements are signed.
- However, the transaction is only completed and registered once the building is finished and ownership transferred.
- Depending on when during the building stage thee property is sold, this (additional) time lag can consist of many months or years.
- This additional lag in transaction data is particularly problematic for countries with large new-build markets (e.g. 50 percent of transacted properties in Poland).

Data

Our Dataset

- We concentrate on apartment market in two large Polish cities: Warsaw and Poznan.
- Transaction prices and dates are recorded by the Property registry offices in Warsaw and Poznan. These data contain information on the individual apartment (i.e., size, parking arrangements, story, and transaction date) as well as the building (e.g., age and exact location). If it exists, the date of the preliminary contract is recorded as well.
- We link these datasets with the cadastre dataset to obtain additional information on each building (e.g., how many floors it has and the age of the building). Where this information was not available, we used Google Street view.

Hedonic Price Indices

Establishing Hedonic Price Indices

- We use the Rolling-Time-Dummy (RTD) hedonic method. This estimates a hedonic model with a fixed window length (for example, $m+1$ periods).

$$\ln p_{\tau n} = \sum_{c=1}^{C} \beta_c z_{\tau cn} + \sum_{s=t+1}^{t+m} \delta_s d_{\tau sn} + \varepsilon_{\tau n},$$

  where t is the first period in the window and n indexes the housing transactions that fall in the rolling window.

- The characteristics of the properties are denoted by $z_{\tau cn}$, while $d_{\tau sn}$ is a time-dummy variable that equals 1 when $\tau = s$, and zero otherwise.

- The change in the price index from period $t+m-1$ to period $t+m$ is then calculated as follows:

$$\frac{P_{t+m}}{P_{t+m-1}} = \frac{\exp(\hat{\delta}^t_{t+m})}{\exp(\hat{\delta}^t_{t+m-1})}.$$

- A superscript t is included on the estimated δ coefficients to indicate that they are obtained from the hedonic model with period t as the base (i.e., $P_t = 1$).

- The window is then rolled forward by one period, and the hedonic model is re-estimated. The price index over multiple periods is computed by chaining these bilateral comparisons together as follows:

$$\frac{P_{t+m+1}}{P_t} = \left[\frac{\exp(\hat{\delta}^{t-m}_{t+1})}{\exp(\hat{\delta}^{t-m}_{t})}\right]\left[\frac{\exp(\hat{\delta}^{t-m+1}_{t+2})}{\exp(\hat{\delta}^{t-m+1}_{t+1})}\right] \times \cdots \times \left[\frac{\exp(\hat{\delta}^{t+1}_{t+m+1})}{\exp(\hat{\delta}^{t+1}_{t+m})}\right].$$

Hedonic Price Indices for Warsaw

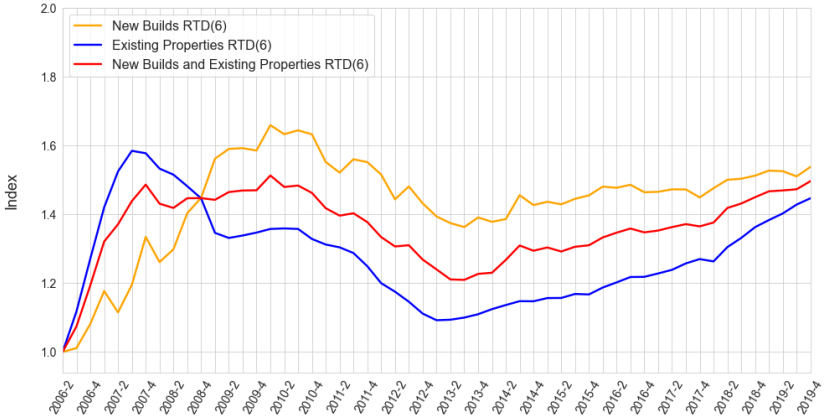

Figure: Real House Price Indices for Warsaw, derived with RTD method with a 6-quarter data window

The Time Lag of a New-Builds Index

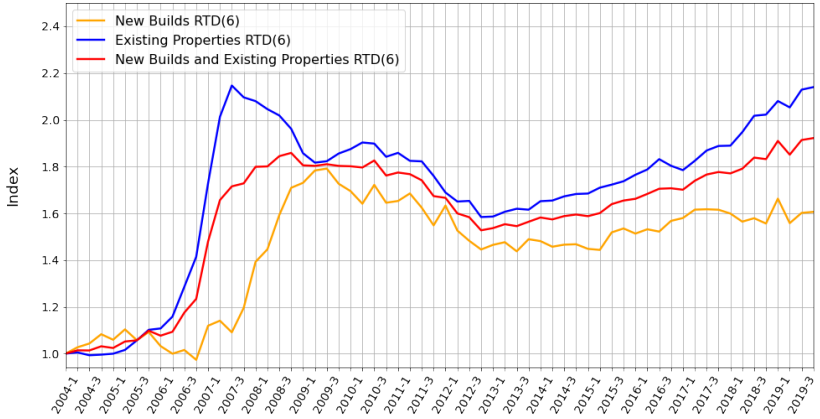

Figure: Real House Price Indices for Poznan, derived with RTD method with a 6-quarter data window

Distance Metrics

Measuring Lags using Distance Metrics

- The previous figures showed that the indices for new-build properties lag behind the index for existing properties.
- We want to estimate how long this lag is.
- We do this using Distance Metrics.
- Various distance metrics exist. We find that the estimated lag does not depend on which method is used.
- We like the metrics proposed by Diewert (2009).

Diewert Metric 1 (DM1):

$$DM1(k) = \frac{1}{T-k-1} \sum_{t=1}^{T-k-1} \left[\frac{P^p_{t+1}}{P^p_t} \Big/ \frac{P^s_{t+k+1}}{P^s_{t+k}} + \frac{P^s_{t+k+1}}{P^s_{t+k}} \Big/ \frac{P^p_{t+1}}{P^p_t} - 2 \right]$$

Diewert Metric 2 (DM2):

$$DM2(k) = \frac{1}{T-k-1} \sum_{t=1}^{T-k-1} \left[\left(\frac{P^p_{t+1}}{P^p_t} \Big/ \frac{P^s_{t+k+1}}{P^s_{t+k}} - 1 \right)^2 + \left(\frac{P^s_{t+k+1}}{P^s_{t+k}} \Big/ \frac{P^p_{t+1}}{P^p_t} - 1 \right)^2 \right]$$

Diewert Metric 3 (DM3):

$$DM3(k) = \frac{1}{T-k-1} \sum_{t=1}^{T-k-1} \left[\ln \left(\frac{P^p_{t+1}}{P^p_t} \Big/ \frac{P^s_{t+k+1}}{P^s_{t+k}} \right) \right]^2$$

- P^p_t and P^s_t denote the levels of the price indices for the new-built and existing property markets in period t.
- DM(k) denotes a modified-Diewert metric with the market for new-builds lagging the market for existing properties by k quarters.

Results of Dissimilarity Measures

- All metrics find a new-builds index lags an existing property index by 8 quarters in Warsaw, although the lag is much less pronounced in the second half of our sample.
- For Poznan the lag is 6 quarters for the first half of the sample and 5 quarters for the second half.
- Dissimilarity Results are shown on the next two slides.

Results for Warsaw

Time	Metric	Quarters Lagged 0	1	2	3	4	5	6	7	8	9	10	Δ
2006-Q2 -2019-Q4	ED	0.322291	0.301150	0.316026	0.315085	0.286432	0.285711	0.278631	0.225813	**0.207451**	0.232671	0.259395	0.1148
	DM1	0.001835	0.001636	0.001843	0.001882	0.001592	0.001589	0.001547	0.001039	**0.000883**	0.001138	0.001454	0.0010
	DM2	0.003690	0.003292	0.003716	0.003788	0.003207	0.003206	0.003115	0.002082	**0.001771**	0.002281	0.002918	0.0019
	DM3	0.001833	0.001635	0.001841	0.001880	0.001590	0.001587	0.001545	0.001039	**0.000883**	0.001138	0.001453	0.0010
	HD	0.112542	0.105226	0.110518	0.110400	0.100446	0.099760	0.097350	0.078936	**0.072262**	0.081085	0.090498	0.0403
	1-PCC	0.893741	0.815482	0.746859	0.661109	0.551372	0.465249	0.382675	0.297460	**0.286216**	0.368467	0.497925	0.6075
2006-Q2 -2012-Q4	ED	0.310755	0.286076	0.304319	0.298690	0.268366	0.264212	0.256014	0.188852	**0.174555**	0.202955	0.227166	0.1362
	DM1	0.003535	0.003121	0.003696	0.003741	0.003167	0.003150	0.003110	0.001764	**0.001561**	0.002253	0.003017	0.0020
	DM2	0.007113	0.006284	0.007456	0.007534	0.006386	0.006363	0.006269	0.003535	**0.003131**	0.004516	0.006061	0.0040
	DM3	0.003531	0.003118	0.003691	0.003737	0.003163	0.003145	0.003106	0.001764	**0.001561**	0.002252	0.003015	0.0020
	HD	0.108449	0.099880	0.106367	0.104587	0.094039	0.092088	0.089267	0.065695	**0.060444**	0.070419	0.078931	0.0480
	1-PCC	0.911577	0.815349	0.749268	0.657993	0.511637	0.436830	0.334184	0.147936	**0.108369**	0.212241	0.417553	0.8032
2013-Q1 -2019-Q4	ED	0.085456	0.094019	0.083612	0.097119	0.083546	0.092245	0.086325	0.057868	**0.049172**	0.054252	0.057758	0.0363
	DM1	0.000256	0.000322	0.000264	0.000371	0.000284	0.000363	0.000332	0.000159	**0.000121**	0.000155	0.000185	0.0001
	DM2	0.000512	0.000643	0.000528	0.000743	0.000569	0.000726	0.000664	0.000318	**0.000242**	0.000309	0.000370	0.0003
	DM3	0.000256	0.000322	0.000264	0.000371	0.000284	0.000363	0.000332	0.000159	**0.000121**	0.000155	0.000185	0.0001
	HD	0.030075	0.033090	0.029422	0.034196	0.029371	0.032456	0.030359	0.020438	**0.017386**	0.019170	0.020406	0.0126
	1-PCC	0.100483	0.102373	0.113530	0.132201	0.127686	0.147156	0.132947	0.092091	**0.090600**	0.118081	0.166613	0.0098

Results for Poznan

Time	Metric	0	1	2	3	4	5	6	7	8	9	10	Δ
2004-Q1 -2019-Q4	ED	0.456320	0.501780	0.437231	0.381388	0.375772	0.361538	**0.324406**	0.396296	0.426765	0.421899	0.465360	0.1319
	DM1	0.003092	0.003737	0.002884	0.002276	0.002255	0.002075	**0.001698**	0.002541	0.002982	0.002956	0.003668	0.0013
	DM2	0.006220	0.007549	0.005838	0.004571	0.004533	0.004170	**0.003405**	0.005107	0.006012	0.005953	0.007408	0.0028
	DM3	0.003089	0.003731	0.002878	0.002274	0.002253	0.002074	**0.001697**	0.002539	0.002978	0.002953	0.003662	0.0013
	HD	0.158550	0.173574	0.151215	0.132609	0.130791	0.125066	**0.112188**	0.136556	0.146826	0.145008	0.159981	0.0463
	1-PCC	0.241634	0.193087	0.136907	0.086207	0.048586	**0.030465**	0.033312	0.068995	0.125948	0.202168	0.301799	0.2111
2004-Q1 -2012-Q4	ED	0.420380	0.458756	0.406837	0.333443	0.321598	0.306302	**0.279027**	0.348294	0.377786	0.381501	0.423856	0.1413
	DM1	0.005271	0.006352	0.005181	0.003649	0.003524	0.003205	**0.002774**	0.004428	0.005408	0.005775	0.007487	0.0024
	DM2	0.010610	0.012854	0.010508	0.007333	0.007091	0.006449	**0.005566**	0.008907	0.010921	0.011644	0.015152	0.0050
	DM3	0.005265	0.006340	0.005169	0.003646	0.003520	0.003202	**0.002772**	0.004423	0.005399	0.005767	0.007472	0.0024
	HD	0.145627	0.158042	0.140216	0.115313	0.111257	0.104996	**0.095720**	0.119011	0.128915	0.130246	0.144797	0.0499
	1-PCC	0.271137	0.213816	0.150065	0.089720	0.044338	0.020508	**0.018431**	0.047339	0.097537	0.164895	0.249426	0.2527
2013-Q1 -2019-Q4	ED	0.149870	0.181066	0.146282	0.156658	0.172340	**0.137540**	0.138451	0.148472	0.150808	0.134935	0.125721	0.0123
	DM1	0.000725	0.001088	0.000739	0.000875	0.001098	**0.000720**	0.000766	0.000920	0.000985	0.000818	0.000748	0.0000
	DM2	0.001451	0.002181	0.001480	0.001755	0.002202	**0.001443**	0.001534	0.001845	0.001975	0.001639	0.001498	0.0000
	DM3	0.000725	0.001088	0.000739	0.000875	0.001098	**0.000720**	0.000766	0.000920	0.000985	0.000818	0.000748	0.0000
	HD	0.052985	0.063939	0.051737	0.055355	0.060890	0.048497	0.048926	0.052508	0.053254	0.047555	**0.044377**	0.0086
	1-PCC	0.243511	0.188139	**0.172815**	0.191483	0.198553	0.177997	0.188387	0.220142	0.277500	0.313184	0.366894	0.0707

Using Preliminary Agreement Dates

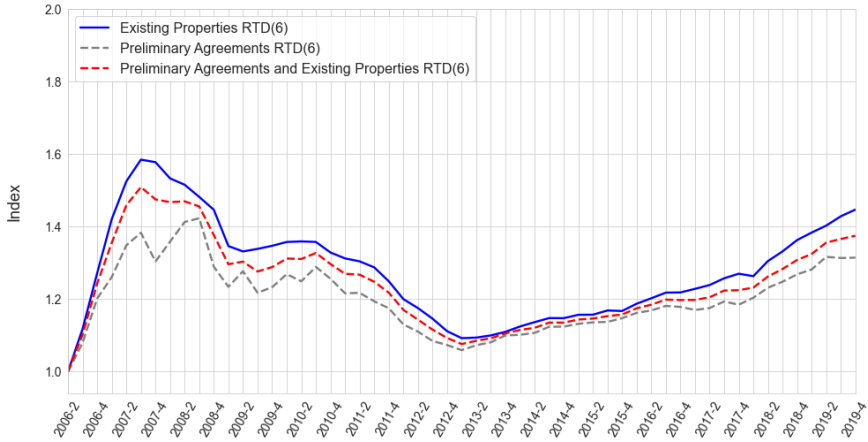

Figure: Real House Price Indices including Preliminary Agreement (dashed) for Warsaw

Explaining the Time Lag

Explaining the Time Lag

- The time lag in price indices for new-builds depends on at least three factors:
  - How long it takes to build an apartment block.
  - How far into the building process are preliminary agreements signed.
  - How well buyers and sellers anticipate future price movements.
- Average time to build an apartment complex: in Warsaw about 25 months, in Poznan around 23 months.
- Do buyers commit sooner during a boom? We do not see this in our results.
- When the market is calmer, future price developments are easier to predict. Hence the lag between the price indices for new and existing properties should be less pronounced. This is consistent with our findings.

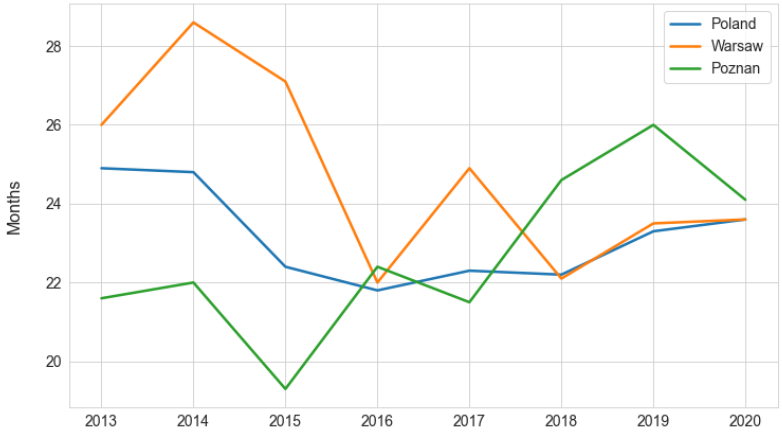

Figure: Average Building Time in Months

Implications

Implications for House Price Indices (HPIs)

- For countries in which a significant part of the property market consists of new builds, including these stale transaction prices can undermine the timeliness of the HPI.
- This is particularly problematic when the HPI is used for macroprudential supervision or monetary policy decisions.
- How to deal with this?
  1. Exclude new-builds in HPI?
     - But then the index ignores an important part of the market.
  2. Replace transactions for new-builds with preliminary agreements?
     - Then no time lag, but such data are not currently collected in most EU Land Registries (at least not independently of transaction data).
     - Undermines strong preference for transaction prices for HPI in statistical circles (e.g., Eurostat and the IMF).

Implications for the HICP

- Owner-occupied housing (OOH) is currently excluded from the HICP.
- The preferred method of Eurostat/ECB for including OOH uses a price index that only tracks new builds.
- Such an index will lag behind current market conditions and undermine the timeliness of the HICP.

Thank you for your attention!

For questions please contact

miriam.steurer@uni-graz.at

Acknowledgement:
This research received funding from the Polish National Agency for Academic Exchange under the Bekker programme (PPN/BEK/2019/1/00292/U/00001) and the Austrian Research Promotion Agency (FFG), grant #10991131.