



NARODOWY
BANK POLSKI

**Price-setting factors or revealed preferences?
How to understand the results of hedonic models and hedonic
indices of housing rental market that base on listings data?**

Michał Hebdzyński, National Bank of Poland, Regional Branch in Poznań

Recent trends in the real estate market and its analysis – 2023 edition, Warszawa, 29.11.2023

The presentation represents the personal opinions of the author and does not reflect the official position of Narodowy Bank Polski.



Agenda

1. Introduction:
 - Importance of the rental market
 - Why to study tenants' preferences?
 - Characteristics of revealed preferences
2. Research questions and contribution
3. Data, variables, methods
4. Findings & discussion – hedonic models' coefficients and hedonic rent indices
5. Conclusions
6. Limitations

Importance of housing rental market

According to Czerniak & Rubaszek (2018) and Rubaszek & Rubio (2020) a developed housing rental market may:

- contribute to the stability of the housing market,
- decrease the overall macroeconomic and real estate market volatility.

It is also considered an important factor for mobility in the labour market (Łaszek et al., 2021).

Outlook for the future

- Even though only 3,4% of Polish households rent at market price (Eurostat, 2023), the preferences of young adults for collaborative housing have been detected (Lis et al., 2023).
- Transition from „homeownership society” (Ronald 2008) to „generation rent” (Ronald & Kadi 2018).
- The Polish Private Rented Sector (PRS) is estimated to rise from 7.000 housing units in 2022 to over 60.000 in 2028 (PWC, 2022).

Why to study tenants' preferences?

1. The development of the rental market has macroeconomic and social importance.
2. The market environment is changing, also because of pandemics (Trojanek et al. 2021) and war in Ukraine (Tomal & Helbich, 2022; Trojanek & Głuszak (2022)).
3. The need to provide the knowledge and analytical methods for researchers, housing policy planners, institutional investors, developers and individuals.
4. The need to adjust the ways of gathering data and calculating price indices.

Approaches to measure preferences

stated preferences

vs.

revealed preferences

- Hasanzadeh et al. (2019) and Vasanen (2012) proved their consistency.
- Earnhart (2002) – they are in line only in the case of some apartment characteristics, but combining both information leads to the best understanding of the phenomena that drive housing decisions.

Revealed preferences (Samuelson, 1938, 1948)

Assumptions:

- Utility-maximizing consumer purchases reveal the utility assigned to goods.
- Based on the transacted prices of housing with particular characteristics, it is possible to estimate the marginal prices paid for each housing feature.
- The decomposition may be conducted using a hedonic model (Lancaster, 1966; Rosen, 1974). Then, the achieved coefficients represent the revealed preferences.

Critique:

- Transactions do not necessarily portray the structures preferred by consumers.

Revealed preferences – characteristics

- The most reliable data type for modelling – the information on micro-level transactions.
- Based on the micro-level hedonic models, researchers are constructing hedonic price indices, adjusted for quality changes between analytical periods (as suggested by the European Commission, Eurostat, Organisation for Economic Co-operation and Development, and World Bank (2013)).
- The goals of a model of revealed consumer preferences are compliant with the requirements of a model aimed to track price movements.

... **But** ...

- The available transactional data are scarce (a high share of transactions is not reported, is reported with a lag or the information is fragmented).

... **However** ...

- Listings (offer/ask) data may be considered an alternative, in some areas being superior (availability, richness, timeliness).
- Listings-based indices have been found to be correlated with transactions-based ones (Micaleff, 2022) and may be regarded as their leading indicators (Anenberg & Laufer, 2017; Lyons, 2019).

... **But** ...

- Listings represent the supply side of the market and not always may be regarded as representative (Beręsewicz, 2015; Nasreen & Ruming, 2022).

But then, how to call the results of hedonic models? Revealed preferences or price-setting factors?

The research questions studied:

- Are the coefficients of hedonic models obtained based on listings data in line with those obtained on transactional data?
- Do the hedonic price indices obtained based on listings and transactional data point at the same dynamics of rents?

Contribution of the study:

- develops the understanding of the results of hedonic methods applied to the housing rental market with the use of listings data,
- verifies the compliance of transactions- and listings-based hedonic indices and indicates the sources of differences,
- adds validity to the already published studies based on housing listings.

Data

- Listings of apartments located in multi-family buildings in Poznań (Poland), listed online for long-term rent via Otodom.pl from Q4 2020 to Q2 2023, source: OLX Group, N = 9186
- Transactions (paired with listings), source: BaRN, National Bank of Poland, from Q1 2021 to Q2 2023, N = 197 * **

* Only those observations have been selected for which it was possible to pair the transaction with the corresponding listing with a high probability that both refer to the same apartment.

** If the listing reappeared in adjacent periods, only the last observation has been retained in the dataset. The procedure aimed to ensure us that the listed rents would be as close as possible to the transacted rents.

Variables

VARIABLE	DESCRIPTION
<i>T_RENT / L_RENT (dependent variable)</i>	transacted (T) or listed (L) rent for the apartment in Polish Złoty [PLN] (in logarithmic form)
<i>AREA</i>	floor area of the apartment (in logarithmic form)
<i>ROOM_INT</i>	room "intensity" - rooms per 1 m ² of the apartment
<i>TO_1945</i>	1 – if the building in which the apartment is located was built in 1945 or earlier, 0 – otherwise
<i>FROM_1946_TO_2004</i>	1 – if the building in which the apartment is located was built between 1946 and 2004, 0 – otherwise
<i>FROM_2005</i>	1 – if the building in which the apartment is located was built in 2005 or later, 0 – otherwise
<i>GROUND_FLOOR</i>	1 – if the apartment is located on the ground floor, 0 – otherwise
<i>GARRET</i>	1 – if the apartment is located on the highest floor of the building, 0 – otherwise
<i>PARKING_SPACE</i>	1 – if there is an access to the designated parking space, 0 – otherwise
<i>AIR_COND</i>	1 – if there is air conditioning in the apartment, 0 – otherwise
<i>DISHWASH</i>	1 – if there is a dishwasher in the apartment, 0 – otherwise
<i>DIST_CC</i>	distance to the city centre [in km]

Source: own elaboration

Hedonic methods used

1. Ordinary Least Squares (OLS)

A hedonic model with logged dependent variable and time dummies indicating the period of rent transaction or listing:

$$\ln R = \beta_0 + \sum_{j=1}^J \beta_j C_j + \sum_{i=2}^I \gamma_i D_i + \varepsilon$$

2. Quantile regression (QR) (Koenker & Bassett, 1978)

- Modelling any quantile of the conditional distribution of the dependent variable.

Analytical steps

1. Constructing an OLS model explaining transacted rents (N = 197) using all variables. Transaction dates have been used to construct time dummies. Selecting the statistically significant variables and re-estimating the model. (*MOD_1*)
2. Constructing an OLS model for the same (paired) observations and variables but selecting listed rent as the dependent variable and using listing dates to construct time dummies. (*MOD_2*)
3. Constructing an OLS model for all listings (N = 9186, dependent variable – listed rent). (*MOD_3*)
4. Comparison of the obtained variables' coefficients and hedonic rent indices.

Coefficients – findings & discussion (1/4)

	MOD_1	MOD_2	MOD_3
	PAIRED TRANSACTIONS	PAIRED LISTINGS	ALL LISTINGS
	Dependent variable: transacted rent	Dependent variable: listed rent	Dependent variable: listed rent
	Mean [PLN] = 1951,3	Mean [PLN] = 1992,2	Mean [PLN] = 2161,4
	N = 197	N = 197	N = 9186
	R ² = 0,777	R ² = 0,785	R ² = 0,689
VARIABLE	COEFFICIENT	COEFFICIENT	COEFFICIENT
In_AREA	0,538 ***	0,545 ***	0,586 ***
TO_1945	0,147 ***	0,138 **	0,124 ***
FROM_2005	0,214 ***	0,210 ***	0,198 ***
PARKING_SPACE	0,096 ***	0,077 **	0,078 ***
DIST_CC	-0,050 ***	-0,046 ***	-0,028 ***
TIME-DUMMIES	YES	YES	YES
CONSTANT	5,316 ***	5,327 ***	5,133 ***

Source: own elaboration

- The differences between the coefficients of hedonic models based on transactions and listings have been lower than 20% when the calculations have been made on the same group of apartments (= constant quality).
 - The presented structure of the compared models should be considered satisfactory based on Diewert & Shimizu (2021).

Coefficients – findings & discussion (2/4)

	MOD_1	MOD_2	MOD_3
	PAIRED TRANSACTIONS	PAIRED LISTINGS	ALL LISTINGS
	Dependent variable: transacted rent	Dependent variable: listed rent	Dependent variable: listed rent
	Mean [PLN] = 1951,3	Mean [PLN] = 1992,2	Mean [PLN] = 2161,4
	N = 197	N = 197	N = 9186
	R ² = 0,777	R ² = 0,785	R ² = 0,689
VARIABLE	COEFFICIENT	COEFFICIENT	COEFFICIENT
In_AREA	0,538 ***	0,545 ***	0,586 ***
TO_1945	0,147 ***	0,138 **	0,124 ***
FROM_2005	0,214 ***	0,210 ***	0,198 ***
PARKING_SPACE	0,096 ***	0,077 **	0,078 ***
DIST_CC	-0,050 ***	-0,046 ***	-0,028 ***
TIME-DUMMIES	YES	YES	YES
CONSTANT	5,316 ***	5,327 ***	5,133 ***

Source: own elaboration

2. The final, transacted rents in Poznań were very close to the listed rents.

Coefficients – findings & discussion (3/4)

	MOD_1	MOD_2	MOD_3
	PAIRED TRANSACTIONS	PAIRED LISTINGS	ALL LISTINGS
	Dependent variable: transacted rent	Dependent variable: listed rent	Dependent variable: listed rent
	Mean [PLN] = 1951,3	Mean [PLN] = 1992,2	Mean [PLN] = 2161,4
	N = 197	N = 197	N = 9186
	R ² = 0,777	R ² = 0,785	R ² = 0,689
VARIABLE	COEFFICIENT	COEFFICIENT	COEFFICIENT
In_AREA	0,538 ***	0,545 ***	0,586 ***
TO_1945	0,147 ***	0,138 **	0,124 ***
FROM_2005	0,214 ***	0,210 ***	0,198 ***
PARKING_SPACE	0,096 ***	0,077 **	0,078 ***
DIST_CC	-0,050 ***	-0,046 ***	-0,028 ***
TIME-DUMMIES	YES	YES	YES
CONSTANT	5,316 ***	5,327 ***	5,133 ***

Source: own elaboration

- The differences between the coefficients of the transactions-based model and the all-listings-based model have been larger (albeit not crucial), especially for the distance variable.

Where does the discrepancy come from?

1. We have ensured that the discrepancy has not been rooted in the fact that some listed apartments have been transacted faster than others because, in the analysis, we have included only the last listing of each apartment.
2. We have shown that the problem has not originated from the difference between the height of the listed and transacted rent.

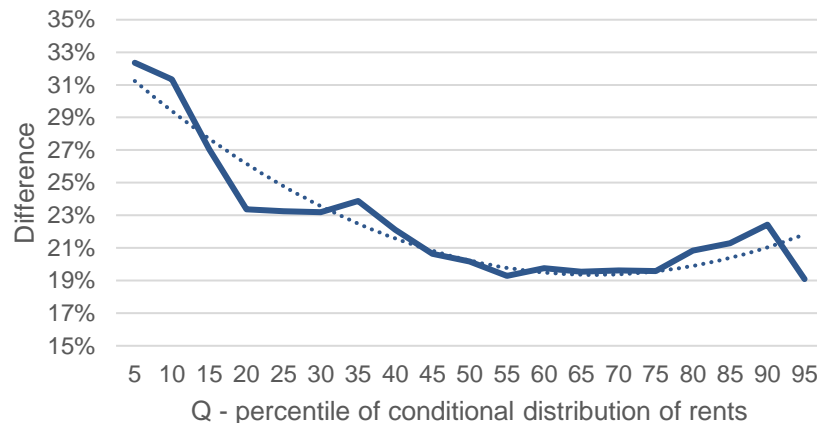
Auxiliary hypothesis 1: the discrepancy between the transactions- and listings-based hedonic models' coefficients stems from the difference in the quality structure of the datasets.

We suspect that the listings data underrepresent the lower segment of the market (advertised via different channels).

Auxiliary hypothesis → new analytical steps

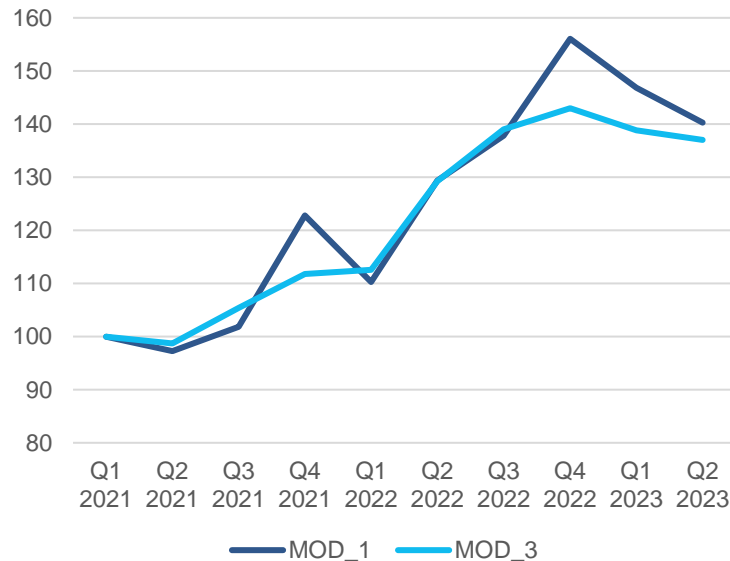
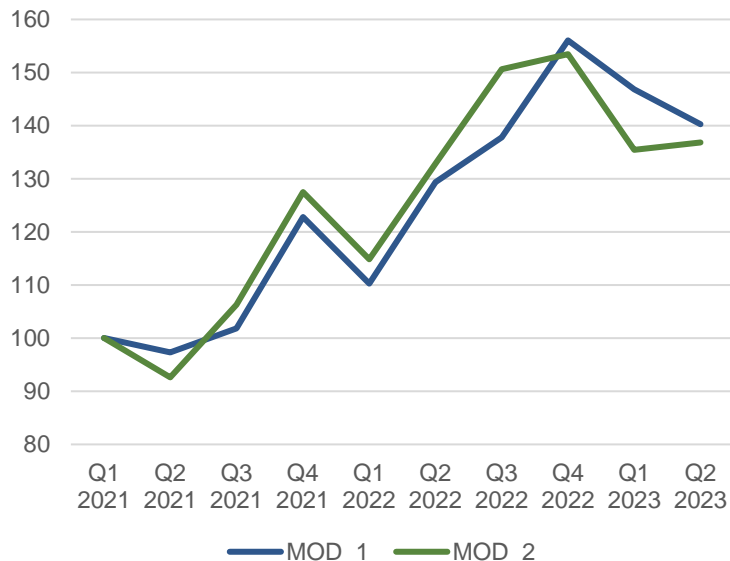
1. Constructing quantile regression models for every fifth conditional quantile of distribution of the dependent variable, based on all listings ($N = 9186$).
2. Verifying, which quantile of the listings-based model deviates least from the transactions-based model (in terms of the obtained variables' coefficients).

Coefficients – findings & discussion (4/4)



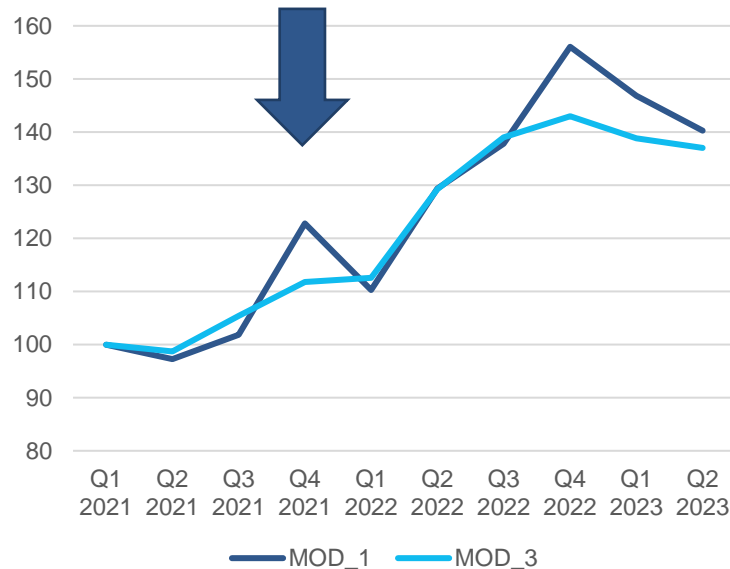
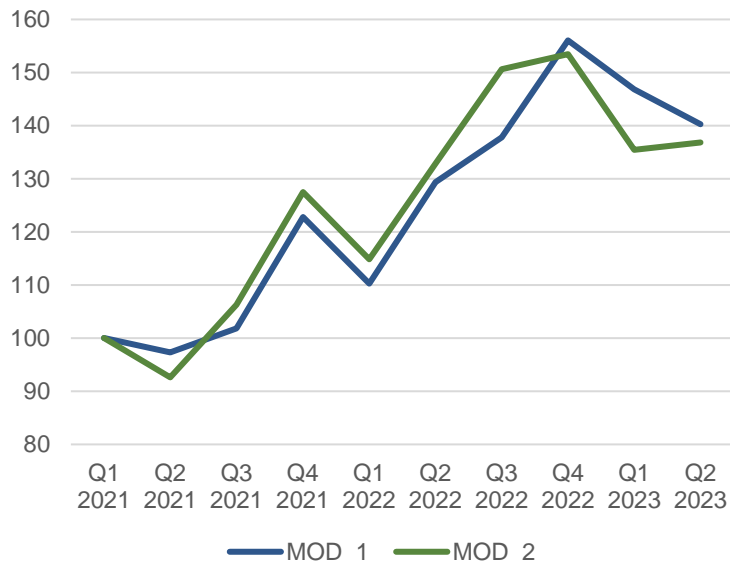
- The highest compliance of coefficients of transactions-based and listings-based models has been reached for the 55th – 75th percentile of the conditional distribution of listed rents. → Our transactional data may represent an even higher market segment than the listings data.

Hedonic rent indices – findings & discussion (1/2)



1. As long as the hedonic rent index calculation has been conducted on the same group of apartments, the differences between transactions-based (*MOD_1*) and listings-based (*MOD_2*) indices have been very small.

Hedonic rent indices – findings & discussion (2/2)



2. The transactions-based index (*MOD_1*) has revealed two short-term peaks, which the all-listings-based index (*MOD_3*) has not detected.

* Quarter – a three-month period ending in February (Q1), May (Q2), August (Q3) and November (Q4)

Where does the discrepancy come from?

Auxiliary hypothesis 2: the opposite short-term representation of quality in the databases of transactions and listings.

Example – negative demand shock:

1. Demand for low-quality apartments rises → more transacted low-quality apartments → more low-quality apartments in the transactional database.
2. Demand for low-quality apartments rises → less available low-quality apartments listed for rent → less low-quality apartments in the (periodically collected) listings database.
3. The transactions-based hedonic models would be better suited to the lower-quality segment of the market; the listings-based models would be worse suited to the lower-quality segment.

If the above logic is correct:

- To capture the short-term market changes using listings data we should prepare separate models for quality segments; otherwise, the listings-based indices are expected to flatten the market dynamics.

Conclusions (1/2)

1. The issue of concern in utilising listings data for the analysis of the Polish rental market should be not the difference between the listed and transacted rents but the inequality of the quality structure of analysed types of data.
2. In the medium and long-term the transactions-based and listings-based indices prove to show the same market dynamics.
3. It is needed to approximate the currently unknown quality structure of the market. Then, it would be possible to construct quality-weighted listings-based indices that would be more reliable in the short-term. Alternatively, one may construct separable indices for quality-related market segments or pay more attention to the issue of soft-quality of housing.
4. The scarce transactional data may represent the segments of the market that are even further from the real market structure than listings data. Then, relying on them for analysing preferences or studying rent indices may introduce an even bigger bias.

Conclusions (2/2)

Price-setting factors or revealed preferences?

5. The listings-based rent-setting factors have proven to be similar to the transactions-based revealed preferences. As these have not been revealed but are close to the revealed ones, when referring to them we suggest using the phrase “proxied preferences”.

Limitations

- Scarcity of the analysed transactional data.
- No golden model to compare with.
- The small number of observations has forced the use of relatively simple econometric methods to maintain the comparability of the results of the transactions-based and listings-based models.

Therefore, the study and its conclusions should be considered introductory to the topic and require further testing.

Bibliography

- Anenberg, E., & Laufer, S. (2017). A more timely house price index. *Review of Economics and Statistics*, 99(4), 722–734.
- Beręsewicz, M. (2015). On the representativeness of internet data sources for the real estate market in Poland. *Austrian Journal of Statistics*, 44(2), 45–57.
- Czerniak, A., & Rubaszek, M. (2018). The Size of the Rental Market and Housing Market Fluctuations. *Open Economies Review*, 29(2), 261–281.
- Diewert, W. E., & Shimizu, C. (2021). Residential Property Price Indexes: Spatial Coordinates Versus Neighborhood Dummy Variables. *Review of Income and Wealth*, 3, 770–796
- Earnhart, D. (2002). Combining revealed and stated data to examine housing decisions using discrete choice analysis. *Journal of Urban Economics*, 51(1), 143–169.
- European Commission, Eurostat, Organisation for Economic Co-operation and Development, World Bank (2013). *Handbook on Residential Property Prices (RPPIs): 2013 edition*. OECD Publishing.
- Eurostat. (2023). *EU-SILC user database description*.
- Hasanzadeh, K., Kytälä, M., & Brown, G. (2019). Beyond Housing Preferences: Urban Structure and Actualisation of Residential Area Preferences. *Urban Science*, 3(1), 21.
- Koenker, R., & Bassett, G. (1978). Regression Quantiles. *Econometrica*, 46(1), 33–50.
- Lancaster, K. J. (1966). A New Approach to Consumer Theory. *Journal of Political Economy*, 74(2), 132–157.
- Łaszek, J., Augustyniak, H., & Olszewski, K. (2021). The development of the rental market in Poland, in R. Sobiecki, J. Łaszek and K. Olszewski (Eds), *Real Estate at Exposure. New Challenges, Old Problems*, pp. 263–274
- Lis, P., Rataj, Z., & Suszyńska, K. (2023). Limitations in the diffusion of collaborative housing in Poland: Expectations and beliefs of young generation as potential users. *Journal of Urban Affairs*, 1–18.
- Lyons, R. C. (2019). Can list prices accurately capture housing price trends? Insights from extreme markets conditions. *Finance Research Letters*, 30, 228–232.
- Micallef, B. (2022). Hedonic indices for registered rents in Malta. *The Annual Malta Residential Rental Study Second Edition*.
- Nasreen, Z., & Ruming, K. (2022). Struggles and opportunities at the platform interface: tenants' experiences of navigating shared room housing using digital platforms in Sydney. *Journal of Housing and the Built Environment*, 37(3), 1537–1554.
- PWC. (2022). *Przegląd rynku PRS w Polsce - I poł. 2022 r.*
- Ronald, R. 2008. The Ideology of Home Ownership: Homeowner Societies and the Role of Housing.
- Ronald, R., Kadi, J., 2018. The Revival of Private Landlords in Britain's Post-Homeownership Society. *New Political Economy* 23(6):786–803.
- Rosen, S. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy*, 82(1), 34–55
- Rubaszek, M., & Rubio, M. (2020). Does the rental housing market stabilize the economy? A micro and macro perspective. *Empirical Economics*, 59(1), 233–257.
- Samuelson, P. A. (1938). A Note on the Pure Theory of Consumer's Behaviour. *Economica*, 5(17), 61.
- Samuelson, P. A. (1948). Consumption Theory in Terms of Revealed Preference. *Economica*, 15(60), 243.
- Tomal, M., & Helbich, M. (2022a). The private rental housing market before and during the COVID-19 pandemic: A submarket analysis in Cracow, Poland. *Environment and Planning B: Urban Analytics and City Science*, 49(6), 1646–1662.
- Trojanek, R., & Gluszak, M. (2022). Short-run impact of the Ukrainian refugee crisis on the housing market in Poland. *Finance Research Letters*, 50(103236).
- Trojanek, R., Gluszak, M., Tanas, J., & Hebdzyski, M. (2021). The COVID-19 pandemic, Airbnb and housing market dynamics in Warsaw. *Critical Housing Analysis*, 8(1), 72–84.
- Vasanen, A. (2012). Beyond stated and revealed preferences: The relationship between residential preferences and housing choices in the urban region of Turku, Finland. *Journal of Housing and the Built Environment*, 27(3), 301–315.



NARODOWY
BANK POLSKI

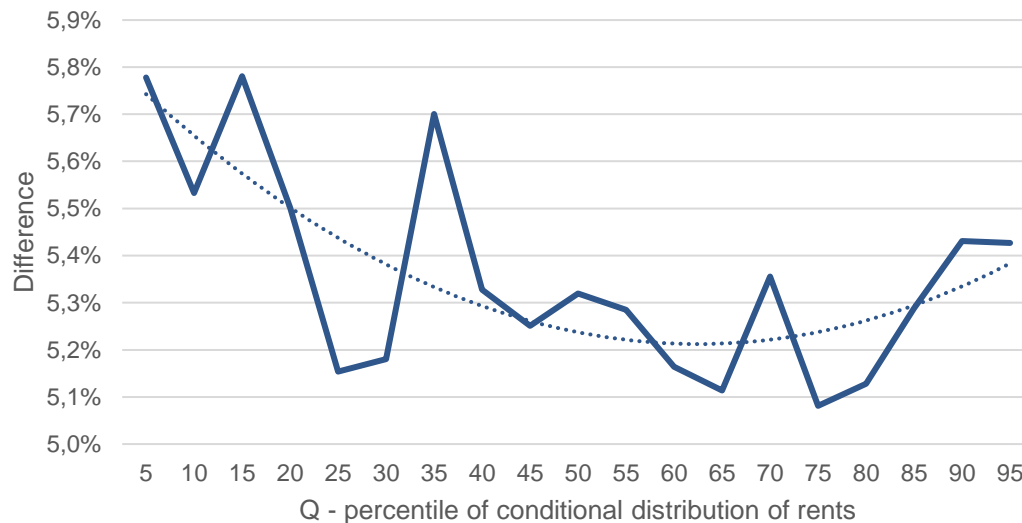
**Price-setting factors or revealed preferences?
How to understand the results of hedonic models and hedonic indices of
housing rental market that base on listings data?**

Michał Hebdzyński, National Bank of Poland, Regional Branch in Poznań
michal.hebdzynski@nbp.pl

Recent trends in the real estate market and its analysis – 2023 edition, Warszawa, 29.11.2023

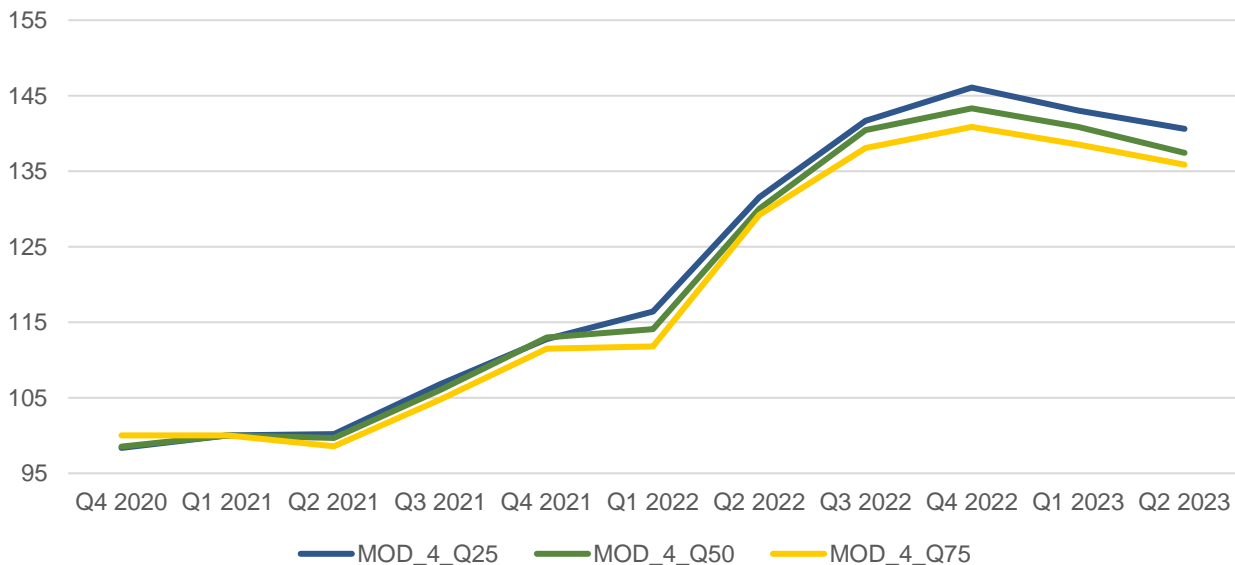


Appendix 1 - Hedonic rent indices – discussion



- The obtained transactional hedonic rent index has been closest to the all-listings-based indices that have represented the 60th – 80th percentile of the conditional distribution of rents.

Appendix 2 - Hedonic rent indices – discussion



- The listings-based models constructed for price-related market segments have shown only little, but noticeable differences.

Appendix 3 – Shares of observations from selected periods

