

# Komparace kupních a odhadních cen bytů v ČR za rok 2019<sup>#</sup>

*Petr Sunega<sup>\*</sup> – Martin Lux<sup>\*\*</sup> – Petr Kubala<sup>\*\*\*</sup>*

## Úvod

Jako zástava poskytnutých hypotečních úvěrů (a úvěrů ze stavebního spoření nad určitou výši jistiny) slouží nemovitosti (byty, rodinné domy a pozemky), které jsou nejčastěji také předmětem transakce. Odhadní ceny těchto nemovitostí jsou stanoveny odhadci banky (nebo externími odhadci) a slouží jako podklad pro rozhodnutí o poskytnutí úvěru a jeho výše (výše úvěru musí být v přiměřené relaci k odhadní ceně nemovitosti měřené ukazatelem loan-to-value ratio, LTV), o výši úrokové sazby a současně pro stanovení výše rezerv pro případ, že by došlo k nesplacení úvěru a nucenému prodeji zástavy. Z uvedeného je zřejmé, že správné stanovení odhadní ceny ovlivňuje dostupnost úvěru určeného na pořízení bydlení (zda se banka rozhodne úvěr vůbec poskytnout) a následně i jeho cenu (vyšší LTV znamená zpravidla i vyšší úrokovou sazbu). Správné stanovení odhadní ceny přímo ovlivňuje finanční dostupnost vlastnického bydlení pro ty žadatele (domácnosti), které pro jeho pořízení musí využít cizí zdroje (úvěry) a nemohou si dovolit jej financovat pouze z vlastních zdrojů. Současně tak má vliv i na samotné rozhodování domácností o (ne)pořízení vlastnického bydlení a načasování takového rozhodnutí. Stanovení odhadní ceny ovlivňuje i výši prostředků, které banky musí držet jako rezervu pro případ realizace zástav z důvodu nesplacení úvěrů dlužníky, tj. má vliv i na výši nákladů bank a stabilitu finančního sektoru.

Výše odhadní ceny a ceny úvěru na bydlení (úroku) ovlivňuje prostřednictvím důchodového efektu výši spotřebních výdajů domácností na ostatní položky spotřebního koše, přičemž tento vliv je s ohledem na vysoký podíl výdajů na bydlení ve spotřebním koši českých domácností relativně významný. Stejně tak se výše výdajů na bydlení odráží i ve výši úspor domácností. Odhadní ceny představují informaci pro subjekty trhu o tom, jaké ceny jsou „akceptovány“ finančním sektorem, tj. mohou nepřímo ovlivnit i výši transakčních cen. Ze všech výše uvedených důvodů je proto správné stanovení odhadních cen nemovitostí a sledování jejich vývoje v čase důležité.

V říjnu 2018 došlo ke zpřísnění požadavků bank na hodnoty LTV, DTI (debt-to-income, tj. poměru výše dluhu k čistým ročním příjmům domácnosti žadatele) a DSTI (debt-service-to-income, tj. poměru výše splátky úvěru k čistému měsíčnímu příjmu domácnosti žadatele). Tato zpřísnující kritéria byla doporučena ze strany regulátora (České národní banky, ČNB) v rámci makroobezřetnostní politiky snažící se snížit potenciální systémová rizika ve finančním sektoru. Důsledkem této politiky mohl být neformální tlak na „umělý“ růst odhadních cen

<sup>#</sup> Tento článek byl zpracován jako jeden z výstupů projektu „Zvýšení transparentnosti hypotečního trhu v ČR rozšířením datové základny pro sledování a odhad cen nemovitostí“ podpořeného Technologickou agenturou ČR (TA ČR) pod číslem TL03000212.

<sup>\*</sup> Ing. Petr Sunega – vědecký pracovník; Sociologický ústav AV ČR, v.v.i., Jilská 1, 110 00 Praha 1. Kontakt: petr.sunega@soc.cas.cz.

<sup>\*\*</sup> doc. Ing. Mgr. Martin Lux, Ph.D. – vedoucí vědecký pracovník; Sociologický ústav AV ČR, v.v.i., Jilská 1, 110 00 Praha 1.

<sup>\*\*\*</sup> Mgr. Petr Kubala, Ph.D. – postdoktorand; Sociologický ústav AV ČR, v.v.i., Jilská 1, 110 00 Praha 1.

zastavovaných nemovitostí za účelem zabezpečení úvěru, tedy potenciálně větší odchylka mezi odhadními a kupními cenami zastavených nemovitostí.

Banky se z důvodu úspory nákladů spojených s poskytováním úvěrů snaží využívat pro ocenění zástav i různé cenové mapy či databáze, které většinou vychází z odhadních cen nemovitostí. Oprávněným požadavkem ČNB je však větší kontrola ocenění založených na odhadních cenách ve smyslu jejich potenciálního odchýlení od kupních cen. Finální kupní ceny jsou evidovány v katastru nemovitostí spravovaném Českým úřadem zeměměřičským a katastrálním (ČÚZK). Nabízí se tak možnost srovnat odhadní ceny bank a kupní ceny evidované ČÚZK pro analýzu odchylky a správnosti ocenění.

Důležitost správného stanovení odhadní ceny je akcentována i v zahraniční literatuře. Baum et al. (2011) zmiňují důležitost procesu oceňování pro fungování trhu s aktivy, ve finančních zprávách, pro účely měření výkonnosti (výnosnosti), úvěrování bank a jako podklad při rozhodování o pořízení, prodeji nebo správě aktiv.

Autoři studie nizozemské centrální banky z roku 2019 (van der Molen, Nijskens 2019) zdůrazňují, že systematické nadceňování nemovitostí může vést k předlužení žadatelů o úvěr (dlužníků) a podhodnocení úvěrových rizik na trhu s hypotečními úvěry. Autoři analyzovali vzorek více než 200 000 rezidenčních nemovitostí, které byly oceněny mezi roky 2012 a 2017, a pro něž byly k dispozici obě ceny – kupní i odhadní. Z výsledků jejich analýz vyplynulo, že odhadní cena nemovitosti byla vyšší nebo rovna kupní ceně v téměř 95 % případů a vyšší než kupní cena v téměř 60 % případů.

Cannon a Cole (2011) mimo jiné zjistili, že výše odchylky kupních a odhadních cen se v USA lišila podle fáze vývoje trhu bydlení – v době růstu cen nemovitostí byly odchylky vyšší (a kupní ceny vyšší než odhadní), zatímco v době krize a poklesu cen byly nižší (a odhadní ceny vyšší než kupní). Uvedené zjištění podporuje hypotézu, že odhadní ceny jsou zpožděným odrazem vývoje kupních cen.

Společnost MSCI (viz např. Moloney 2019) vydává studie porovnávající kupní a odhadní ceny ve Velké Británii již řadu let. Z nich mimo jiné vyplynulo, že v době boomu odhadní ceny zaostávaly za cenami kupními, zatímco v době korekce klesaly odhadní ceny pomalejším tempem než ceny kupní. MSCI zpracovalo i komparativní analýzu pro 12 zemí, za něž byly k dispozici časové řady za období 2000 až 2018: Austrálii, Kanadu, Francii, Německo, Itálii, Japonsko, Nizozemí, Jižní Afriku, Švédsko, Švýcarsko, Velkou Británii a USA. Jedním ze závěrů bylo, že v řadě zemí byla nejvyšší průměrná odchylka (s kupními cenami převyšujícími ceny odhadní) dosažena v letech 2006 a 2007, tj. v době vrcholu růstu cen bydlení před poslední globální finanční krizí. Časování se přitom mírně lišilo mezi skupinami zemí. Po obratu převýšily odhadní ceny ty kupní ve Velké Británii, USA a Švédsku už v roce 2008, zatímco v Německu, Japonsku, Nizozemí, Kanadě a Austrálii tomu bylo až na konci roku 2009.

Parker (1998) mimo jiné zmiňuje důvody nepřesnosti ocenění – nedokonalosti trhu s nemovitostmi (informační bariéry, heterogenita obchodovaných nemovitostí, omezená mobilita aktérů i aktiv, náklady spojené se získáním úvěru a jeho dostupnost, daňové faktory, a jiné), chybějící centrální databáze, důvěrnost informací o uzavřených obchodech, omezená nabídka srovnatelných nemovitostí, postavení odhadců (jejich práce může být hodnocena s ohledem na výsledek jejich odhadu, tedy podléhájí zájmu zadavatele).

Řada autorů (např. Levy a Schuck 2005, Achu 2013) se ve svých textech zabývala otázkou vlivu klientů (nebo obecně aktérů participujících na procesu oceňování) na odhadce a na výši

odhadních cen. Podle nich jsou odhady jsou nepřesnými aproximacemi tržních cen, což je mimo jiné důsledkem efektu zpoždění (odhadní ceny zaostávají za tržními cenami, se zpožděním následují trendy na trhu), sériové korelace (korelace mezi aktuální a o jedno období zpožděnou hodnotou cen; tj. aktuální hodnoty jsou ovlivněny minulými hodnotami), statistických chyb odhadu a náchylnosti odhadců k ovlivnění ze strany dalších subjektů (např. klientů). Příčinou je vysoká konkurence v oblasti oceňování, které vyžaduje poměrně těsnou interakci mezi odhadcem a klientem. Ta s sebou přináší riziko morálního hazardu, kdy odhadce nejedná nezávisle, ale zohledňuje potřeby klienta.

Cílem tohoto článku je pokusit se odpovědět na otázku, zda v roce 2019 byly v ČR nadhodnoceny nebo podhodnoceny odhadní ceny v porovnání s kupními cenami, a pokud ano, v jakých segmentech. K jejímu zodpovězení budou využita data s cenovými údaji Českého úřadu zeměměřičského a katastrálního (ČÚZK) a data s odhadními cenami České spořitelny, a.s. (ČS). Struktura článku je následující. Ve druhé kapitole budou stručně popsány datové zdroje použité pro komparaci kupních a odhadních cen bytů. Třetí kapitola se věnuje vysvětlení metodologie použité pro účely srovnání. Ve čtvrté kapitole jsou prezentovány výsledky vybraných analýz, následuje závěr.

## Data

Soubor s kupními cenami ČÚZK byl tvořen vzorkem dat nakoupených od ČÚZK za rok 2019. Nakoupena byla data za 984 katastrálních území, výsledkem byl soubor s více než 254 000 cenovými záznamy, z toho bylo 68 % záznamů o parcelách, 1 % o budovách a 31 % o jednotkách. Z něj byly pro účely dalších analýz vyřazeny záznamy o vícenásobných prodejkách (prodejkách, kdy byla prodávána jednotka a další nemovitost), protože u tohoto typu záznamů nelze spolehlivě stanovit cenu pro jednotlivé prodávané nemovitosti. Jednalo se o situace, kdy bylo pod jedním číslem řízení evidováno více nemovitostí s cenovým údajem (který byl však pro všechny nemovitosti stejný, tj. jednalo se o celkovou cenu za všechny prodané položky). Zhruba ve 42 % případů se s jednotkou prodávala jedna nebo více dalších nemovitostí, nejčastěji se jednalo o jednu nebo více parcel.

Dále byly ze souboru vyřazeny záznamy s jiným typem využití, než je bydlení,<sup>1</sup> a v neposlední řadě jednotky, kde byl v poli popis nemovitosti uveden jiný typ nemovitosti než byt.<sup>2</sup> Důvodem bylo, že cílem bylo porovnat odhadní a kupní ceny pouze pro segment bytů, tj. nemovitostí určených k bydlení. Výsledkem byl datový soubor s více než 44 000 byty.

Dalším krokem čištění dat bylo určení segmentace obou datových souborů na malá (ale současně dostatečně početně zastoupená z hlediska počtu bytů) území, kde by byly jednotkové ceny bytů dostatečně homogenní pro účely srovnání odhadních a kupních cen. V rámci těchto segmentů pak probíhalo detailní čištění dat. Konkrétně byla hledána extrémní a odlehlá pozorování, která vybočovala ze srovnání s podobnými byty ve stejné lokalitě z hlediska celkové ceny, plochy bytu nebo jednotkové ceny (ceny na m<sup>2</sup>). Jednalo se o časově náročný proces, jehož cílem bylo odfiltrovat z dat maximum pozorování, která by mohla nežádoucím způsobem vychylovat průměrné a mediánové hodnoty, a v konečném důsledku zkreslit

<sup>1</sup> Konkrétně jednotky, kde způsob využití byl "garáž", stavba pro výrobu a skladování, stavba občanského vybavení, stavba pro obchod, rodinný dům, stavba pro administrativu, jiná stavba, víceúčelová stavba, stavba ubytovacího zařízení, objekt občanské vybavenosti a stavba bez uvedeného způsobu využití.

<sup>2</sup> Jednalo se o případy, kdy byly v poli popis nemovitosti uvedeny následující hodnoty: ateliér v budově, dílna nebo provozovna v budově, garáž v budově, jiný nebytový prostor v budově, rozestavěná jednotka v budově, skupina nebytových prostorů.

výsledky porovnání kupních a odhadních cen. K tomuto účelu byla využita procedura EXAMINE ve statistickém paketu SPSS (verze 27).

Zvlášť pak probíhalo čištění ve městech s více než 50 000 obyvateli, kde byl obvykle vysoký počet bytů s cenovými údaji, a jejich zahrnutí do některého ze segmentů by tak bylo příliš hrubé. V rámci čištění byly identifikovány i byty s nesmyslnými hodnotami spoluvlastnického podílu (a tudíž i z něj odvozené podlahové plochy), které byly z výsledného datového souboru vymazány. Kromě toho byly identifikovány i byty se špatně dopočtenou „nejpravděpodobnější“ hodnotou podlahové plochy – takové byty vybočovaly z porovnání s ostatními obvykle podezřelou nízkou nebo vysokou jednotkovou cenou. Obvykle u takových bytů stačilo upravit hodnotu podlahové plochy (posunout o desetinné místo) a následně přepočítat jednotkovou cenu. Pokud opravená jednotková cena nevybočovala ze srovnání s ostatními byty, byl byt v datech ponechán. Výsledkem byla sada filtrů pro každý segment, tj. „přípustných“ intervalů pro hodnoty celkové kupní/odhadní ceny, plochy bytu a jednotkové ceny (cena na m<sup>2</sup>). Výsledkem čištění dat byl soubor s cenovými údaji ČÚZK (kupními cenami) zahrnujícími necelých 32 000 bytů.

Soubor s odhadními cenami ČS obsahoval daleko více informací nejen o samotném bytu a bytovém domě, ale i o okolí nemovitosti. V případě bytů jde například o informaci o technickém stavu bytu (zda se jedná o novostavbu, výborně udržovaný byt, dobře udržovaný byt, neudržovaný byt k částečné rekonstrukci nebo neudržovaný byt k celkové rekonstrukci) i technickém stavu celého bytového domu (jelikož ten se může od technického stavu bytu lišit), opotřebení bytu, podlaží, kde se byt nachází, i celkovém počtu podlaží v bytovém domě, o roce, kdy proběhla celková a/nebo dílčí rekonstrukce bytu apod. V porovnání s daty ČÚZK je mnohem detailněji sledována plocha bytu, a to jak podlahová plocha jednotlivých příslušenství k bytu (sklep, balkón, terasa, WC, koupelna atd.), tak celková započitatelná plocha bytu. Pokud jde o širší okolí nemovitosti, je zde sledována například poloha domu v rámci obce a typ zástavby, dostupnost veřejné dopravy (vlaku, autobusu, dálnice, MHD) a veřejných komunikací (dálnice/silnice I. třídy). Evidována jsou i rizika spojená s lokalitou, kde se byt nachází, a rizika spojená s bytem jako takovým (např. věcné břemeno, riziko špatného přístupu bytu, sociální riziko v lokalitě apod.).

Data o jednotlivých attributech nemovitostí jsou sbírána odhadci banky, kteří odhad provádějí přímo na místě ocenění, tedy v bytě samotném. K zápisu odhadu využívají i specializovaný software obsahující vestavěné kontroly chyb nebo nevyplněných polí, který jim neumožní odeslat záznam bez zadání klíčových údajů o nemovitosti a bez případných oprav chybných nebo inkonzistentních údajů, což přispívá k minimalizaci zadání chybných údajů a k úplnosti databáze.

Čištění dat souboru s odhadními cenami ČS bylo realizováno s využitím stejných metod a mechanismů jako v případě souboru s cenovými údaji ČÚZK. Na rozdíl od souboru s cenovými údaji ČÚZK nebylo v případě dat ČS nutno řešit problém s nespolehlivou plochou bytu, stejně tak se v datech nevyskytovaly případy bytů na stejné adrese s výrazně odlišnými odhadními cenami. Na druhou stranu, zatímco součástí datového souboru s cenovými údaji ČÚZK jsou i geografické souřadnice polohy domu (ve formátu JTSK), v souboru s odhadními cenami ČS původně tyto souřadnice k dispozici nebyly a byly do souboru dodatečně přidány; u zhruba 22 % bytů však v adrese chybělo číslo popisné, proto byly GPS souřadnice přiřazeny pouze na základě názvu ulice v dané obci a poloha bytu tak stanovena ve středu dané ulice. Po očištění dat zahrnoval výsledný datový soubor s odhadními cenami ČS celkem 6 706 záznamů o bytech. V porovnání se souborem s cenovými údaji ČÚZK se tak jednalo o soubor se skoro pětikrát menším počtem bytů.

K oběma souborům byla připojena externí data z jiných (převážně veřejně dostupných) datových zdrojů. V případě dat ČÚZK se jednalo zejména o data z databáze RUIAN.<sup>3</sup> Určitým problémem těchto dat je jejich neúplnost: například druh konstrukce domu chyběl u více než 5 % bytových jednotek, rok dokončení stavby u 62 % bytových jednotek, podlahová plocha domu u 79 % záznamů; informace o napojení na inženýrské sítě (plyn, kanalizace, vodovod) a o způsobu vytápění byla naopak poměrně dobrá a chyběla jen u 2 až 5 % jednotek.

Dalším externím datovým zdrojem byl seznam katastrálních území s vazbou na vyšší územně správní celky (obec, okres, kraj), který ČÚZK poskytuje volně ke stažení (ČÚZK 2021). Jiným externím datovým zdrojem byl soubor s počtem obyvatel v obcích a průměrným věkem obyvatel v obcích publikovaný každoročně ČSÚ (ČSÚ 2019).

## Metodologie

Metodologicky nejčistší a nejméně problematické by bylo porovnání kupních a odhadních cen na vzorku identických bytů. Bohužel použitá data neumožňovala takové porovnání realizovat. Proto bylo porovnání kupních a odhadních cen provedeno s využitím tří alternativních metodologických přístupů tak, aby naše výsledky byly co nejvěrohodnější a ověřené. První přístup byl založen na podrobném třídění obou datových souborů podle vybraných proměnných zastoupených v obou souborech v dostatečně malých geografických celcích (za okresy a okresní města). V těchto tržních segmentech (tříděních) byly zjištěny průměrné a mediánové kupní a odhadní ceny, resp. rozdíly mezi nimi, a ty byly dále analyzovány a hodnoceny. Druhý přístup byl založen na porovnání kupních a odhadních cen na spojených datových souborech se zachováním těch proměnných (atributů), které se vyskytovaly v obou datových souborech; jednotková cena bytu pak byla na spojeném souboru vysvětlována pomocí regresních modelů, a to včetně využití interakčních proměnných. S ohledem na omezený prostor neuvádíme analýzy ani výsledky založené na dvou výše zmíněných přístupech, pro zájemce jsou k dispozici v Sunega et al. 2021. V tomto článku se zaměříme pouze na třetí přístup založený na vytvoření dvojic geograficky nejbližších bytů z obou datových souborů.

Cílem bylo pro byty z jednoho souboru najít identický nebo alespoň nejpodobnější byt z druhého souboru a v souboru tvořeném takovými dvojicemi bytů porovnat kupní a odhadní ceny (tj. srovnávat srovnatelné). Propojení obou datových souborů bylo realizováno na základě znalosti geografické polohy bytů v obou datových souborech. S využitím volně dostupného geografického software QGIS (verze 3.20.3) byl ke každému bytu v souboru s odhadními cenami ČS přiřazen jeden nebo více bytů ze souboru s kupními cenami ČÚZK, a to na základě jejich geografické vzdálenosti (určené podle GPS souřadnic), která nesměla přesáhnout 500 metrů. 500 metrů bylo zvoleno s ohledem na možnou nepřesnost v určení polohy některých bytů bez uvedení úplné adresy v datech ČS. Jiné varianty než 500 metrů nebyly testovány s ohledem na dlouhý výpočetní čas, který spojení obou souborů vyžadovalo (vzhledem k rozsahu obou souborů se jednalo skoro o jeden a půl dne). Vzhledem k tomu, že k jednomu bytu z dat ČS mohlo být přiřazeno více bytů z dat ČÚZK (pokud se jich více nacházelo ve vzdálenosti do 500 metrů), následně bylo nutno vybrat z těchto více bytů pouze jeden „nejpodobnější“. Výběr byl proveden už nejen podle geografické vzdálenosti, ale i podle dalších proměnných – plochy bytu, typu konstrukce a data zplacení/provedení odhadu. Jinými

<sup>3</sup> Jde o Registr územní identifikace, adres a nemovitostí, jehož správcem je Český úřad zeměměřický a katastrální (ČÚZK), spuštěn byl 1.7.2012 (viz <https://www.cuzk.cz/ruian/>). Obsahem RUIAN jsou popisné a lokalizační údaje o územních prvcích, územně evidenčních jednotkách, účelových prvcích, adresách a jejich vzájemných vazbách.

slovy, spočítána byla geografická vzdálenost v metrech mezi bytem z dat ČS a jedním nebo více byty z dat ČÚZK, rozdíl v plochách bytů, typu konstrukce a datech (ve dnech), následně byly tyto hodnoty rozdílů převedeny na jednotnou škálu v rozmezí 0 až 1 (prostřednictvím standardizace s využitím minimální a maximální hodnoty) a spočítána vážená celková „vzdálenost“ mezi byty (jako vážený součet jednotlivých dílčích vzdáleností). Testována byla tři různá nastavení vah. Nejvyšší váhu měla ve všech případech geografická vzdálenost mezi byty, nejnižší datum zplatnění (záznamu v katastru) nebo provedení odhadu (bytu ze souboru s odhadními cenami ČS).

Po určení dvojic „nejbližších“ bytů v datech výše popsaných způsobem byly v souboru vypočteny ukazatele rozsahu nesouladu mezi kupními a odhadními cenami a míry nesouladu mezi kupními a odhadními cenami a ty byly analyzovány s využitím různých statistických metod. Rozsah nesouladu mezi kupními a odhadními cenami a zjednodušeně řečeno udává podíl segmentů (v dané kategorii) s vyšší hodnotou odhadní než kupní ceny. Míra nesouladu byla definována jako rozdíl mezi odhadní a kupní cenou na  $m^2$ . Míra nesouladu byla zjištěna jak v korunách (jako prostý rozdíl mezi oběma cenami na  $m^2$ ), tak v procentech.

## Výsledky

Tabulka 1 ukazuje kvantifikaci rozsahu nesouladu pro celý datový soubor, tj. 5 495 bytů (ne pro všechny byty ze souboru ČS se podařilo najít alespoň jeden byt ve vzdálenosti do 500 metrů ze souboru ČÚZK). Z tabulky je zřejmé, že rozsah nesouladu se v této variantě týkal 60 % bytů s jednotkovou odhadní cenou vyšší než jednotkovou kupní cenou.

**Tab. 1: Rozsah nesouladu mezi kupními a odhadními cenami**

	Počet bytů	Podíl bytů (%)
Kupní cena na $m^2$ vyšší nebo rovna odhadní ceně na $m^2$	2 181	39,7
Odhadní cena na $m^2$ vyšší než kupní cena na $m^2$	3 314	60,3
<b>Celkem</b>	<b>5 495</b>	<b>100,0</b>

Zdroj: vlastní výpočty

S cílem provést vyhodnocení rozsahu nesouladu při kontrole všech výše uvedených faktorů současně byly odhadnuty parametry logistického regresního modelu, kde závislou proměnnou byla dichotomická proměnná indikující rozsah nesouladu (resp. v situaci, kdy byla odhadní cena na  $m^2$  vyšší než kupní cena na  $m^2$ , závislá proměnná nabývala hodnoty 1, v ostatních případech 0) a nezávislými proměnnými pak výše uvedené faktory. S ohledem na skutečnost, že v datech ČS je v datech více informací o bytech, bylo možno sadu vysvětlujících proměnných rozšířit o další proměnné (například technický stav bytu a způsob vytápění). Současně je však nutné upozornit, že tyto informace se vztahují pouze k bytu ze souboru ČS. V případě jeho protějšku ze souboru s cenovými údaji ČÚZK vůbec nemusí platit, že jde o byt ve stejném technickém stavu nebo se stejným typem vytápění jako u bytu ze souboru ČS, protože tyto informace prostě nemáme k dispozici. Parametry jednotlivých proměnných odhadnutého logistického regresního modelu jsou uvedeny v tabulce 2.

**Tab. 2: Logistický regresní model vysvětlující rozsah nesouladu mezi kupními a odhadními cenami na úrovni ČR**

Proměnná	B	S.E.	Wald	df	Sig.	Exp(B)
Konstanta	1,751	0,523	11,223	1	0,001	5,761
Započitatelná plocha bytu v m <sup>2</sup>	-0,007	0,001	42,139	1	0,000	0,993
Typ konstrukce (cihla x panel), cihla = 1	0,284	0,073	15,118	1	0,000	1,328
Počet obyvatel v obci	0,000	0,000	2,241	1	0,134	1,000
Technický stav bytu			104,163	4	0,000	
<i>Novostavba - referenční kategorie</i>						
Výborně udržovaný	-0,391	0,095	17,073	1	0,000	0,676
Dobře udržovaný	-0,857	0,093	84,362	1	0,000	0,425
Neudržovaný k částečné rekonstrukci	-0,876	0,162	29,403	1	0,000	0,416
Neudržovaný k celkové rekonstrukci	-0,256	0,255	1,012	1	0,314	0,774
Způsob vytápění			14,026	2	0,001	
<i>Dálkové (zdroj mimo budovu) - referenční kategorie</i>						
Ústřední (včetně tepelného čerpadla)	0,253	0,071	12,549	1	0,000	1,288
Lokální	-0,008	0,135	0,004	1	0,950	0,992
Kraj			65,778	13	0,000	
<i>Praha - referenční kategorie</i>						
Středočeský	-0,603	0,502	1,442	1	0,230	0,547
Jihočeský	-0,601	0,506	1,410	1	0,235	0,548
Plzeňský	-0,712	0,482	2,184	1	0,139	0,491
Karlovarský	-0,303	0,515	0,347	1	0,556	0,738
Ústecký	-0,146	0,499	0,085	1	0,771	0,865
Liberecký	-0,429	0,504	0,725	1	0,395	0,651
Královéhradecký	-0,209	0,511	0,168	1	0,682	0,811
Pardubický	-0,453	0,505	0,805	1	0,370	0,635
Vysočina	-0,579	0,516	1,261	1	0,261	0,560
Jihomoravský	-0,594	0,438	1,836	1	0,175	0,552
Olomoucký	0,027	0,499	0,003	1	0,957	1,027
Zlínský	0,201	0,517	0,151	1	0,698	1,222
Moravskoslezský	-0,051	0,466	0,012	1	0,912	0,950
-2 Log likelihood	7078,302 (Sig. = 0,000)					
Cox & Snell R Square	0,052					
Nagelkerke R Square	0,071					
Hosmer and Lemeshow test	Chi square = 9,754, df = 8, Sig. = 0,283					
N	5 495					

Zdroj: vlastní výpočty

Poznámka: B ve druhém sloupci tabulky 2 jsou nestandardizované koeficienty jednotlivých vysvětlujících proměnných, které udávají, jak se změni logit závislé proměnné v případě změny hodnoty nezávislé proměnné o jednotku a fixovaných hodnotách ostatních nezávislých proměnných. Hodnoty ve sloupci S.E.

udávají standardní chybu odhadu parametrů  $B$ , což je variabilita v odhadech hodnot  $B$ , kterou bychom získali, pokud bychom opakovaně vybírali vzorek ze stejné populace a opakovaně odhadovali hodnoty  $B$ . Hodnoty ve sloupci Wald jsou hodnoty Waldovy statistiky, která testuje, nakolik je zařazení jednotlivých proměnných do modelu přínosné (nakolik jsou významné). Testuje se nulová hypotéza, že  $B=0$ . Pokud je dosažená hladina významnosti (Sig.) menší než například 0,05, zamítá se nulová hypotéza. Hodnoty ve sloupci  $df$  jsou stupně volnosti odpovídající počtu kategorií dané proměnné zařazených do regresního modelu. Hodnoty ve sloupci  $Exp(B)$  udávají násobky, o které se změní šance (že odhadní cena je vyšší než kupní), jestliže se hodnota nezávislé proměnné změní o jednotku a hodnoty ostatních proměnných zůstanou nezměněné. V posledních pěti řádcích tabulky jsou uvedeny statistiky pro ohodnocení regresního modelu a na posledním řádku tabulky ( $N$ ) je počet pozorování (bytů), na nichž byl regresní model odhadnut. Čím vyšší hodnoty statistiky  $-2 \text{ Log likelihood}$ , tím horší predikce závislé proměnné (tj. nižší je lepší). V našem případě je dosažená významnost 0,000, takže zahrnutím nezávislých proměnných do modelu umožňuje lepší predikci závislé proměnné než jejich vypuštění z modelu. Hodnota  $R^2$  Coxové a Snella činí 0,052 a hodnota Nagelkerkova  $R^2$  činí 0,071, tj. lze konstatovat, že model vysvětluje 7,1 % „variability“ závislé proměnné. Test dobré shody Hosmera a Lemeshova říká (Sig. 0,283 > 0,05), že nezamítáme nulovou hypotézu, že mezi pozorovanými a modelem predikovanými hodnotami není žádný rozdíl.

Z tabulky je patrné, že rozsah nesouladu kupních a odhadních cen se snižoval s rostoucí velikostí (plochou) bytu, byl významně vyšší u cihlových bytů (v porovnání s panelovými), snižoval se se zhoršujícím se technickým stavem bytu v porovnání s novostavbami. Rovněž byl významně vyšší v bytech s ústředním vytápěním v porovnání s byty s dálkovým vytápěním. Velikost obce měřená počtem obyvatel stejně jako region (kraj) nebyly statisticky významné. Z celkového specifikace modelu je rovněž zřejmé, že model dokázal vysvětlit jen velmi malou část variability závislé proměnné (tj. skutečnosti, že odhadní ceny jsou vyšší než kupní ceny). Pravděpodobně je to dáno tím, že vysvětlující proměnné nedokážou dostatečně postihnout důvody spojené s rozdílnou výší cen. Aby se podíl vysvětlené variability zvýšil, do modelu bychom museli zařadit další „správné“ proměnné, které nemáme k dispozici. Současně je možné, že důvody nesouladu mezi oběma typy cen mohou být i na straně odhadců, tj. nemusí nutně souviset s charakteristikami bytů.

Mediánová míra nesouladu, tedy medián rozdílu jednotkových odhadních a kupních cen v jednotlivých dyádách vyjádřený v procentech, se pohybovala kolem 9 %.<sup>4</sup> Podobně jako v případě rozsahu nesouladu jsme se i v případě míry nesouladu pokusili zjistit, zda se míra nesouladu liší mezi kraji, velikostními kategoriemi bytů, typy konstrukce a velikostmi obcí a podle dalších faktorů. Proto byly odhadnuty parametry lineárního regresního modelu, kde závislou proměnnou byla spojitá proměnná v podobě rozdílu mezi odhadními a kupními cenami na  $m^2$  (v korunách) a nezávislými proměnnými pak jednotlivé výše zmíněné faktory. V tabulce 3 jsou uvedeny parametry jednotlivých proměnných odhadnutého lineárního regresního modelu.

---

<sup>4</sup> V tomto článku neuvádíme výši průměrné míry nesouladu, jelikož (1) primárním cílem je zejména nalezení hlavních faktorů v pozadí míry nesouladu (2) a oba zdroje dat jsou natolik odlišné co do velikosti i struktury, že by konkrétní výše míry nesouladu mohla být nepřesná a zavádějící.



**Tab. 3: Lineární regresní model vysvětlující míru nesouladu mezi kupními a odhadními cenami v korunách na úrovni ČR**

	Nestand. koeficienty		Stand. koeficienty	t	Sig.	VIF
	B	Std. Error	Beta			
Konstanta	-1313,661	1125,086		-1,168	0,243	
Započitatelná plocha bytu v m <sup>2</sup>	-57,691	9,639	-0,081	-5,985	0,000	1,062
Typ konstrukce (cihla x panel), cihla = 1	2911,297	624,674	0,078	4,661	0,000	1,641
Obce do 4 999 obyvatel	414,863	1153,347	0,007	0,360	0,719	2,523
Obce s 5 000 - 19 999 obyvateli	870,335	1115,840	0,019	0,780	0,435	3,534
Obce s 20 000 - 49 999 obyvateli	1320,873	1172,632	0,025	1,126	0,260	2,985
Obce s 50 000 - 99 999 obyvateli	1543,426	1115,362	0,028	1,384	0,166	2,413
<i>Obce se 100 000 a více obyvateli - referenční kategorie</i>						
Novostavba	6758,837	755,010	0,139	8,952	0,000	1,408
Výborně udržovaný byt	3538,642	582,167	0,088	6,078	0,000	1,229
<i>Dobře udržovaný byt - referenční kategorie</i>						
Neudržovaný byt k částečné rekonstrukci	-654,098	1251,747	-0,007	-0,523	0,601	1,057
Neudržovaný byt k celkové rekonstrukci	6819,210	1993,488	0,045	3,421	0,001	1,033
<i>Dálkové topení (zdroj mimo budovu) - referenční kategorie</i>						
Ústřední topení (včetně tepelného čerpadla)	3061,181	607,836	0,081	5,036	0,000	1,522
Lokální topení	-214,421	1147,393	-0,003	-0,187	0,852	1,195
Praha	6439,779	986,551	0,151	6,528	0,000	3,129
Středočeský	1648,950	1116,274	0,028	1,477	0,140	2,176
Jihočeský	-644,992	1423,867	-0,007	-0,453	0,651	1,556
Plzeňský	1058,984	1177,067	0,015	0,900	0,368	1,569
Karlovarský	3051,678	1534,303	0,052	1,989	0,047	4,011
Ústecký	306,721	1619,225	0,004	0,189	0,850	3,063
Liberecký	1428,538	1513,291	0,014	0,944	0,345	1,295
Královéhradecký	148,967	1510,457	0,002	0,099	0,921	1,478
Pardubický	514,299	1479,101	0,006	0,348	0,728	1,546
Vysočina	-60,869	1596,417	-0,001	-0,038	0,970	1,403
<i>Jihomoravský kraj - referenční kategorie</i>						
Olomoucký	4430,727	1253,187	0,056	3,536	0,000	1,447

	Nestand. koeficienty		Stand. koeficienty	t	Sig.	VIF
	B	Std. Error	Beta			
Zlínský	4888,067	1517,720	0,051	3,221	0,001	1,445
Moravskoslezský	3884,132	1254,156	0,054	3,097	0,002	1,799
Nejkratší dojezdová vzdálenost do krajského města v km	-2,981	16,607	-0,004	-0,180	0,858	2,865
R	0,262					
R square	0,069					
Adjusted R square	0,064					
Std. Error of the Estimate	17734,816					
Durbin-Watson	1,863					
F	15,520 (Sig. < 0,001)					
N	5 488					

Zdroj: vlastní výpočty.

Poznámka: hodnoty B ve druhém sloupci tabulky 3 jsou nestandardizované koeficienty jednotlivých vysvětlujících proměnných, které udávají, jak se změní hodnota závislé proměnné v případě změny hodnoty nezávislé proměnné o jednotku a fixovaných hodnotách ostatních nezávislých proměnných. Hodnoty ve sloupci Std.Error udávají standardní chybu odhadu parametrů B, což je variabilita v odhadech hodnot B, kterou bychom získali, pokud bychom opakovaně vybírali vzorek ze stejné populace a opakovaně odhadovali hodnoty B. Hodnoty ve sloupci Beta jsou standardizované regresní koeficienty, které umožňují porovnat sílu vlivu jednotlivých nezávislých proměnných na závislou proměnnou (v případě nestandardizovaných koeficientů to možné není, protože jsou vyjádřeny v různých jednotkách). Čím vyšší je absolutní hodnota beta koeficientu, tím silnější je vliv dané nezávislé proměnné na závislou proměnnou. Hodnoty ve sloupci t jsou hodnoty t-statistiky, která je určena jako podíl nestandardizovaného koeficientu (B) a směrodatné chyby (Std.Error) a říká, s jakou pravděpodobností (Sig.) můžeme zamítnout hypotézu, že  $B = 0$ . Jinými slovy, testujeme nulovou hypotézu, že nezávislá proměnná nemá vliv na vysvětlení závislé proměnné. VIF je variation inflation factor, který měří rozsah multikolinearity v modelu (riziko, že jedna nebo více proměnných v modelu jsou vzájemně silně korelovány). Vysoké hodnoty VIF (doporučení se liší, ale zpravidla vyšší než 5) indikují multikolinearitu v modelu. V posledních sedmi řádcích tabulky jsou uvedeny statistiky pro ohodnocení regresního modelu a na posledním řádku tabulky (N) je počet pozorování (bytů), na nichž byl regresní model odhadnut. Čím vyšší hodnoty statistiky  $R^2$ , tím větší část variability závislé proměnné model vysvětluje (tj. vyšší je lepší). V našem případě lze konstatovat, že model vysvětluje pouze 6,9 % „variability“ závislé proměnné (resp. 6,4 % při zohlednění „penalizace“ za počet proměnných v modelu podle Adjusted  $R^2$ ). Durbin-Watsonova statistika se používá k odhalení autokorelace reziduí z regrese, konkrétně reziduí a jejich zpoždění v čase t-1. Přítomnost autokorelace reziduí má za následek, že jsou podhodnoceny standardní chyby odhadu parametrů B, takže B se mohou jevit jako statisticky významné, což ve skutečnosti nejsou. Hodnoty blízké 2 (zpravidla v rozmezí 1,5 až 2,5) indikují nulovou autokorelaci reziduí, hodnoty nižší než 2 pozitivní autokorelaci a hodnoty vyšší než 2 negativní autokorelaci. F-statistika (F) testuje celkovou významnost modelu, tj. zda model s vysvětlujícími (nezávislými) proměnnými lépe odpovídá datům než model bez vysvětlujících proměnných. Nulová hypotéza zní, že model s nezávislými proměnnými a model bez nezávislých proměnných odpovídají datům stejně. V našem případě můžeme nulovou hypotézu zamítnout (protože  $\text{Sig.} < 0,001$ , což je menší než 0,05).

Výsledky regresního modelu uvedené v tabulce 3 ukazují, že model vysvětluje jen velmi nízké procento variability míry nesouladu v korunách, což může být způsobeno více faktory – nejdůležitější je zřejmě ten, že neznáme správné proměnné, které by dokázaly lépe vysvětlit variabilitu závislé proměnné. Významnou vysvětlující proměnnou se ukázala být plocha bytu – s rostoucí plochou bytu se míra nesouladu v korunách snižuje. Převažující typ konstrukce

domu byl rovněž statisticky významný – vyšší míru nesouladu v korunách vykazují byty v cihlových domech v porovnání s panelovými byty. Velikost obce měřená počtem obyvatel nebyla statisticky významná. Významným faktorem se ukázal být technický stav bytu<sup>5</sup> – z modelu je zřejmé, že míra nesouladu v korunách je vyšší u bytů v novostavbách, výborně udržovaných bytů a u neudržovaných bytů vhodných k celkové rekonstrukci v porovnání s dobře udržovanými byty. Významným faktorem se ukázal být i typ vytápění bytu<sup>6</sup>, kdy byty s ústředním vytápěním vykazují vyšší míru nesouladu v korunách než byty s dálkovým vytápěním. V neposlední řadě je významnou vysvětlující proměnnou kraj (region), kdy se významně vyšší míra nesouladu v korunách ukázala v Praze (v průměru o 6 440 Kč na m<sup>2</sup>), Karlovarském kraji, Olomouckém kraji, Zlínském kraji a v Moravskoslezském kraji v porovnání s Jihomoravským krajem. Nejkratší dojezdová vzdálenost do nejbližšího krajského města (v kilometrech ani v minutách), stejně jako dojezdová vzdálenost do centra zaměstnanosti (nejbližší obce s více než 20 000 obyvateli) nebyly statisticky významné, tj. neměly statisticky významný vliv na míru nesouladu v korunách.

## Závěr

Z provedených analýz celkem jednoznačně vyplývá, že odhadní ceny byly v roce 2019 vyšší než kupní ceny, podle způsobu (varianty) hodnocení se tato situace týkala zhruba 60 % bytů v námi analyzovaném souboru. Pokud jde o segmenty trhu, kde se nesoulad mezi kupními a odhadními cenami nejvíce projevil, šlo především o byty v cihlové zástavbě (často novostavby nebo výborně udržované). Z výše uvedených analýz vyplývá, že by se mělo jednat rovněž hlavně o segment menších bytů, ale tento závěr se nepotvrdil podle výsledků dalších dvou metodologických přístupů (viz metodologie). Naopak, podle nich se ukázalo, že rozdíly mezi kupními a odhadními cenami se projeví zejména ve velkoměstech (obcích se 100 tisíci a více obyvateli).

Jakkoliv se výše uvedená zjištění ohledně nadhodnocení odhadních cen mohou zdát jednoznačné, je třeba upozornit, že jsou v rozporu se závěry některých zahraničních studií prezentovaných v úvodu, které spíše potvrzovaly hypotézu, že odhadní ceny jsou zpožděným odrazem vývoje kupních cen. Důvodem může být několik zásadních omezení, která je nutno vzít v potaz. Především byly rozdíly mezi kupními a odhadními cenami hodnoceny pouze za období jednoho roku (2019) a pouze pro jeden segment (byty). Je tedy možné, že situace v jiných letech nebo jiných segmentech trhu může být zcela odlišná. Další značná omezení vyplývají z charakteru dat, na nichž bylo srovnání prováděno. V případě obou datových souborů se nejednalo o úplný vzorek dat, ale určitý výběr, který mohl být v některých směrech zkreslený, což se mohlo projevit i na výsledcích. V datech s cenovými údaji ČÚZK chyběly některé zásadní informace (např. o technickém stavu bytu), které by umožnily lepší výběr vzájemně srovnatelných bytů z obou datových zdrojů. Zmínit je potřeba i odlišnosti ve způsobu definice plochy bytu – zatímco bankovní data používají započitatelnou plochu, ČÚZK eviduje podlahové plochy (resp. výše spoluvlastnických podílů je odvozena z podlahových ploch). Samotné odvození podlahové plochy ze spoluvlastnických podílů v datech ČÚZK je problematické. V neposlední řadě může být důvodem zjištěných rozdílů mezi kupními a odhadními cenami i značný časový nesoulad mezi datem provedení ocenění a zápisem

<sup>5</sup> V tomto případě je však potřeba zdůraznit, že informace o technickém stavu bytu byla k dispozici pouze v datech ČS, nikoliv v datech ČÚZK. Jinými slovy nemůžeme zaručit, že byly porovnávány byty se stejným technickým stavem, byť s ohledem na způsob vytvoření dyád bytů bylo možno předpokládat, že jde o byty ve stejném (nebo podobném) technickém stavu.

<sup>6</sup> Platí totéž, co v případě technického stavu bytu v předchozí poznámce pod čarou.

v katastru nemovitostí. Rovněž je třeba zmínit, že do hodnocení byly zahrnuty odhadní ceny bytů bez ohledu na účel úvěru, tj. nejen byty s účelem úvěru koupě.

## Literatura:

- [1] Baum, A. E. - N. Crosby, S. Devaney (2021). *Property Investment Appraisal*. Oxford: Wiley Blackwell (fourth edition).
- [2] Cannon, S. - Cole (2011). *How Accurate are Commercial Real Estate Appraisals? Evidence from 25 Years of NCREIF Sales Data*. Journal of Portfolio Management 35 (5): 1-44. Available at SSRN: <https://ssrn.com/abstract=1824807> or <http://dx.doi.org/10.2139/ssrn.1824807>
- [3] ČSÚ (2019). *Počet obyvatel v obcích - k 1.1.2019*. Praha: Český statistický úřad. <https://www.czso.cz/csu/czso/pocet-obyvatel-v-obcich-za0wri436p>. Navštíveno dne 10.11.2020.
- [4] ČSÚ (2021). *Registr sčítacích obvodů a budov*. [https://www.czso.cz/csu/rso/registr\\_scitacich\\_obvodu](https://www.czso.cz/csu/rso/registr_scitacich_obvodu). Navštíveno dne 21.12.2021.
- [5] ČÚZK (2021). Číselníky katastrálních území a pracovišť resortu. <https://cuzk.cz/Katastr-nemovitosti/Poskytovani-udaju-z-KN/Ciselniky-ISKN/Ciselniky-katastralnich-uzemi-a-pracovist-resortu.aspx>. Navštíveno dne 21.10.2020.
- [6] Achu, K. (2013). *Client Influence on Property Valuation: A Literature Review*. International Journal of Real Estate Studies 8 (2): 24-47.
- [7] Levy, D. – Schuck, E. (2005). *The influence of clients on valuations: the clients' perspective*. Journal of Property Investment & Finance 23 (2): 182-201. <https://doi.org/10.1108/14635780510584364>.
- [8] Parker, D. (1998). *Valuation accuracy-an Australian perspective*. Pp. 19-21 in 4th Pacific Real Estate Society Conference, Perth.
- [9] Sunega, P. – P. Kubala, M. Lux, M. Macek 2021. *Faktory nadhodnocení odhadních cen: analýza rozdílů mezi kupními a odhadními cenami bytů v roce 2019*. Praha: Sociologický ústav AV ČR.
- [10] van der Molen, R. - Nijskens, R. (2019). *The quality and independence of residential property appraisals*. Occasional Studies Volume 17-1. Amsterdam: De Nederlandsche Bank N.V. [https://www.dnb.nl/media/el3ppobp/201903\\_nr\\_1\\_-2019-the\\_quality\\_and\\_independence\\_of\\_residential\\_property\\_appraisals.pdf](https://www.dnb.nl/media/el3ppobp/201903_nr_1_-2019-the_quality_and_independence_of_residential_property_appraisals.pdf) Navštíveno dne 28.11.2021.
- [11] Moloney, S. (ed.) (2019). *Valuation and sale price*. London: Royal Institution of Chartered Surveyors (RICS).

## **Komparace kupních a odhadních cen bytů v ČR za rok 2019**

*Petr Sunega - Martin Lux - Petr Kubala*

### **ABSTRAKT**

Význam správného stanovení odhadních cen je akcentován v české i zahraniční literatuře. Provedli jsme porovnání kupních cen (cenových údajů Českého úřadu zeměměřičského a katastrálního) a odhadních cen (České spořitelny, a.s.) bytů. Párování bytů bylo provedeno primárně na základě geografické polohy (s využitím GPS souřadnic), dále podle plochy bytu, typu konstrukce a data zplatnění záznamu, resp. provedení odhadu. Výsledkem je zjištění, že u 60 % bytů byly odhadní ceny vyšší než kupní. Nesoulad mezi kupními a odhadními cenami se týkal zejména bytů v cihlové zástavbě a spíše menších bytů. Uvedené výsledky je však třeba brát s rezervou s ohledem na omezení daná použitými datovými zdroji.

**Klíčová slova:** Kupní ceny; odhadní ceny; byty; banky.

## **Comparison of purchase and appraisal prices of flats in the Czech Republic in 2019**

### **ABSTRACT**

The importance of accuracy in appraisal prices is emphasized in both Czech and foreign literature. We have compared the purchase prices (price data of the State Administration of Land Surveying and Cadastre) and appraisal prices (of Česká spořitelna, a.s.) of flats. The pairing of flats was done primarily on the basis of geographic location (using GPS coordinates), then by flat area, type of construction and date of the record / estimate. As a result, it was found that for 60% of the flats the appraisal prices were higher than the purchase prices. The discrepancy between the purchase and appraisal prices was mainly for flats in brick buildings and rather smaller flats. However, these results should be taken with caution in view of the limitations of the data sources used.

**Key words:** Purchase prices; appraisal prices; flats; banks.

**JEL classification:** G30