

FINAL REPORT  
AUSTRIAN MARSHALL PLAN FOUNDATION  
BERKELEY-AUSTRIA EXCHANGE PROGRAM  
2016/2017

Estimating aggregate quantile-specific gross rental  
yields for residential housing in Sydney

Sofie R. Wabl\*

Department of Economics, University of Graz  
Universitätsstraße 15/F4, 8010 Graz, Austria  
sofie.wabl@gmail.com

January 3, 2017

**Abstract**

Rental yields are used by investors and policy makers to assess housing markets. Houses rented and sold are likely to be different in their characteristics, which is a crucial measurement issue. A framework is proposed to measure rental yields cross-sectionally and over time thus providing a comprehensive picture of housing markets. The two-step procedure based on quantile regression and propensity score matching is designed to fundamentally control for differences in house characteristics. The methodology is applied to data on house transactions and asking rents in Sydney, Australia, between 2004 and 2014. The paper finds decreasing rental yields when moving from the low end of the distribution to the top end, and detects a systematic sample selection bias when restricting the analysis to houses bought-to-let.

**Keywords:** Rental yield, Price-to-rent ratios, Hedonic models, Quantile regression, Housing markets, Economic measurement

---

\*This work was supported by the Austrian National Bank (Jubiläumsfondsprojekt 14947); the Austrian Marshallplan Foundation Fellowship (UC Berkeley Program 2016/2017); and the JungforscherInnenfonds sponsored by the council of the University of Graz (2014). I thank Australian Property Monitors for supplying the data. This article constitutes the last chapter of my dissertation at the University of Graz under the supervision of Robert J. Hill. I am grateful for valuable comments from Robert J. Hill, Daniel P. McMillen, and Michael Scholz as well as the participants in the research seminar at the Department of Economics at the University of Graz, the Young Scholars Program at the Lincoln Institute of Land Policy in Cambridge, MA, and a research seminar at the Austrian National Bank in Vienna.

# 1 Introduction

Aggregate gross rental yield,<sup>1</sup> defined as the ratio of annual rent over sales price, is an important measure for investors to base their investment decision on as well as policy makers, central bankers, and researchers to assess the state of housing markets. Investors evaluate a property’s potential return (and the real estate market in general) using the *net rental yield*, that is *gross rental yield* minus costs such as maintenance costs or interest payments. Forward-looking present value models predict that low current rental yields signal higher future capital gains as well as increasing rents and thus attractive investment opportunities (see Clark, 1995; Capozza and Seguin, 1996).

Rental yields are also used to detect disorders in housing markets (see for instance Weeken, 2004; Fox and Tulip, 2014; OECD, 2016). Aggregate measures of rental yields may be used as an ingredient in the user cost formula for housing (Poterba, 1992; Himmelberg et al., 2005). In equilibrium, the cost of owning, i.e., the user cost of housing, should equal the return on owning: the rent. As Fox and Tulip (2014) put it: “Given the supply of housing is fixed in the short run, prices are determined by how much buyers are willing to pay. Hence a comparison of the costs of home ownership with the costs of the nearest alternative [i.e., rents] seems central to a measure of overvaluation.”

By rearranging the user cost formula and plugging in rental yields, it is possible to measure expected capital gains (McCarthy and Peach, 2004; Fox and Tulip, 2014; Hill and Syed, 2016), which may assist in the early detection of housing bubbles. Stiglitz (1990) defines asset bubbles as follows: “[I]f the reason that the price is high today is only because investors believe that the selling price is high tomorrow – when *fundamental* factors do not seem to justify such a price – then a bubble exists.” Rapidly increasing expected capital gains, that are noticeably higher than long-term averages, may hence indicate irrational exuberance. For instance, the *UBS Global Real Estate Bubble Index*, which is “designed to track the risk of housing bubbles in global financial centers” (see Holzhey and Skoczek, 2016, page 4) includes reciprocal rental yields as one of five components.

Whereas *average* rental yields are a good starting point to assess investment opportunities or the state of a housing market, they may however obscure substantial cross-sectional variation. There are theoretical arguments (see section 2) why the ingredients in the user cost formula are expected to vary across the distribution and hence, in equilibrium, such variation directly translates into cross-sectional variation of rental yields. This paper therefore constructs quantile-specific rental yields which allow this variation to be measured.

In general, there are two alternative ways to look at aggregate average rental yields which also translate into two different ideas how to generalize the concept to quantile-specific rental yields. First, one may aim to calculate rental yields separately for individual houses. As there are only few houses for which price and rent information is available, one may choose to restrict the analysis on this sub-sample of observations or impute missing prices or rents. One may finally obtain an aggregate measure of the average rental yield by taking the average over all individual ratios. Quantile-specific rental yields would be obtained by evaluating the distribution of rental yields at different quantile levels.

Second, one may think of an average rental yield as the ratio of an average annual rent over an average sales price. This understanding seems to be very common among real estate agents and investment advisers. A generalization of this concept may be to match quantiles of the rent distribution to the same quantiles of the price distribution.

---

<sup>1</sup>If not stated otherwise, I always mean *gross* rental yield when referring to rental yields in this paper.

In this paper, I will follow the second approach as it has some appealing advantages in terms of interpretation: A rental yield for a high quantile level is associated with high prices *and* high rents and vice versa. Such kind of interpretation is not possible for the first approach where a high quantile level neither reflects a high price nor a high rent but only indicates a *high ratio*. When explaining quantile-specific rental yields through quantile-specific ingredients in the user cost formula the relationship between rental yields and the price distribution is however essential.

Although the basic concept of a rental yield is straight-forward, there are crucial measurement challenges. Similar as to when constructing house price indices, it is important to control for differences in house characteristics to compare like with like. In the house price index literature such “quality-adjustment” is usually performed by applying repeat-sales or hedonic methods (see de Haan and Diewert, 2013). When constructing rental yields an additional dimension of quality-adjustment should be considered: Next to quality differences *within* houses sold and *within* houses rented, there is possibly also a mismatch *across* houses sold and rented. For instance, houses sold may be on average larger than houses rented leading to biased results.

Several studies construct rental yields by comparing rental indices with house price indices (see Fu and Ng, 2001; Himmelberg et al., 2005; Gallin, 2008; Campbell et al., 2009; Duca et al., 2011). Whereas such a procedure accounts for quality differences within sales and within rents, it ignores quality differences across sales and rents as quality-adjustment is most probably performed in different ways for rental and sales price indices. Furthermore, using indices only allows changes of rental yields, but not levels, to be measured.

Alternatively, one may construct rental yields by taking the ratio of the average observed rent over the average observed sales price. Such kind of calculations ignore quality differences and are often found on real estate agents or financial advisers websites. Davis et al. (2008) calculate historic rental yields based on average sales prices and average imputed rents.

Hill and Syed (2016) use a hedonic imputation approach to account for quality differences in rental yields. If a house was sold but not rented they treat the rental price of this particular dwelling as missing and vice versa. They estimate separate hedonic models for the rental and the sales data set and use these models to impute missing rental and sales prices. Ultimately, they gain estimates for the rental and sales price for all dwellings in their data set and calculate dwelling-specific quality-adjusted price-to-rent ratios. They calculate the median over all these ratios to obtain an aggregate measure.

Fox and Tulip (2014) use a data set that consists of observed or imputed rent and sales prices for identical dwellings and construct average rental yields based on them. Imputations rely on hedonic methods or on extrapolated prices using price and rent indices.

Bracke (2015) adapts the repeat-sales idea and restricts the sample to those dwellings having been sold and rented within a short period of time. Bracke hence uses an exact matching procedure to create the sample on which he performs his analysis. Smith and Smith (2006) also use a matching approach but allow next to exact matches also pairs of observations that are, though not identical, similar in their characteristics. Both Bracke, and Smith and Smith use conservative matching approaches which come at the cost of strikingly small sample sizes that may be subject to sample selection bias. In the case of Smith and Smith, samples consist of 100 observations only.

There is a trade-off between aiming for a good match between houses sold and rented in terms of characteristics, and avoiding a sample selection bias. When relying on exact

matches, i.e., houses that were sold and rented within a short period of time, there is per construction no mismatch between houses sold and rented. However, as shown in this paper, such a conservative approach is likely to introduce a severe sample selection bias, which is conceptually similar to the well-known Akerlof-type lemons bias in repeat-sales indices (see Clapp and Giaccotto, 1992; Wallace and Meese, 1997; Steele and Goy, 1997). On the other hand, neglecting the fact, that houses sold and rented tend to be different in their characteristics, may lead to noisy estimates. This paper thus aims to find a compromise between these two competing goals by suggesting a two-step procedure.

In the first step, houses rented are matched to houses sold based on their characteristics using customized propensity score matching. The second step constructs samples from the *marginal* sales and rental price distribution net of house characteristics using penalized quantile regression in combination with a sampling algorithm proposed by Machado and Mata (2005). These samples are used to calculate quantile-specific rental yields that are fundamentally controlled for differences in house characteristics.

This paper thus contributes to the yet very sparse literature on measurement issues related to rental yields and is the first to develop a method to measure quality-adjusted rental yields cross-sectionally. The method is applied to sales and rent data from Sydney, Australia, between 2004 and 2014. Such comprehensive data on rents are very rare which probably at least partly explains the gap in the literature on techniques to accurately measure rental yields.

As predicted by theory, rental yields are consistently downwards-sloping when moving from the low end of the market to the top end. Refraining from quality-adjustment led to an average bias of roughly -2%. Even more importantly, this paper is the first one to document a systematic sample selection bias when relying on houses that are sold and rented within a short period of time. For the Sydney data, this bias is found to range on average between 6% to 8%.

The remainder of the paper is structured as follows. Section 2 elaborates on expectations about cross-sectional variation in rental yields based on the user cost formula. Section 3 develops the methodology to construct quantile-specific and quality-adjusted rental yields. Section 4 describes the data set and presents empirical results. Finally, section 5 concludes.

## 2 Expectations about cross-sectional variation

Poterba (1992) and Himmelberg et al. (2005) argue that in equilibrium the expected annual cost of owning should equal the annual cost of renting and hence compare

[...] the value of living in [a] property for a year – the ‘imputed rent’, or what it would have cost to rent an equivalent property – with the lost income that one would have received if the owner had invested the capital in an alternative investment – the ‘opportunity cost of capital’. This comparison should take into account differences in risk, tax benefits from owner-occupancy, property taxes, maintenance expenses, and any anticipated capital gains from owning the home (Himmelberg et al., 2005, page 74).

They hence equate

$$R_t = P_t u_t, \tag{1}$$



where  $R_t$  denotes the annual rent and  $P_t$  the house price in period  $t$ .  $P_t u_t$  is called the *user cost of housing* and  $u_t$  alone the *per dollar user cost* defined as

$$u_t = r_t + \omega_t + \delta_t + \gamma_t - g_{t+1}.$$

Thereby,  $r_t$  denotes an appropriate interest rate,  $\omega_t$  running average transaction costs including taxes,  $\delta_t$  the depreciation rate or maintenance costs,  $\gamma_t$  an additional risk premium compensating home owners for the higher risk of owning a property as opposed to renting and  $g_{t+1}$  the expected capital gain during the period. All these factors are given as fractions of the house price  $P_t$  and are specific for period  $t$ .<sup>2</sup>

The equilibrium condition (1) relates *the* annual rent to *the* house price. Hence,  $R_t$  and  $P_t$  may describe an average price per period but also any other specific point of the price distributions:

$$R_t(\vartheta) = P_t(\vartheta)u_t(\vartheta)$$

for all  $\vartheta \in (0, 1)$  where each  $\vartheta$  denotes a specific quantile level.

Prices as well as rents naturally vary with  $\vartheta$ . But there are also good reasons why one would expect the per dollar user cost to vary cross-sectionally and the equilibrium condition to be violated in certain price segments. I will explain reasons for both types of deviations starting with the cross-sectional variation of the user cost. Hill and Syed (2016) give an extensive list of arguments together with empirical evidence which the following is based on.

First, maintenance costs as a fraction of house prices  $\delta_t$  are likely to be lower for high-priced dwellings. Diewert et al. (2015) argue that a house price is actually a composite price for the structure and the land the structure is built on. More expensive dwellings tend to have a higher share of land and since land in contrast to structure does not depreciate, maintenance costs are expected to be lower for expensive dwellings. Bracke (2015) points out that dwellings with higher utilization rates, i.e., more people per square meter, have higher maintenance costs. It is plausible that utilization rates – and hence depreciation rates – are higher at the low-end of the market.

Second, the risk premium  $\gamma_t$  may be higher at the low-end of the distribution as homeownership may be more risky for low-income households (see Peng and Thibodeau, 2013). There is some evidence that houses belonging to a low price segment react stronger to an overall boom-bust cycle (see for instance Guerrieri et al., 2013; Walzl, 2016b).

Third, low-income households may face higher costs of borrowing and hence higher interest rates  $r_t$  thus pushing up the user cost at the low-end of the housing market.

When it comes to transaction costs and expected capital gains, the direction of cross-sectional differences is not so clear. Diewert et al. (2015) mention that property tax rates, which are part of the term  $\omega_t$ , are often different on the land and structures components of a property. As mentioned above, the share of land in a house price is likely to be larger at the high-end. However, some taxes are applicable only for properties worth more than a certain threshold.

---

<sup>2</sup>There is some debate whether  $\gamma$  should be included into the user cost (see Fox and Tulip, 2014) as it is not clear whether owning or renting is more risky. Rosen et al. (1984) argue that when estimating the user cost, one obtains a measure of the *ex post* cost of owner-occupation for a period. Tenure decisions are however based upon the *expected* costs which include uncertainty that should be accounted for. It is a challenging task to measure  $u$  accurately which is however not the focus of this paper.

Likewise there is probably no clear answer whether expected capital gains increase or decrease over the price distribution. Hill and Syed (2016) provide some evidence that they are higher at the high-end but claim that this might hold true only in the long run.

All in all, one hence expects the user cost to be decreasing with  $\vartheta$ , i.e.,  $\frac{\partial}{\partial \vartheta} u_t(\vartheta) < 0$ .

Above that, the equilibrium condition is likely to be violated at the very top as well as the very low end of the market: Owners of expensive properties may be more selective on renters as they wish to find reliable renters that will maintain the dwelling properly and Diewert et al. (2009) argues that rents may hence be offered at a discount. Bracke (2015) finds that the duration of tenancies tends to be longer at the top-end of the market and Larsen and Sommervoll (2009) find evidence to suggest that rent discounts accrue to long-term tenants.

In contrast, at the low-end of the market – the market of interest for low-income households – there may be a higher demand for rental objects. Some low-income households may wish to buy but are not able to get a large enough mortgage and are hence unintentionally forced into the rental market. Such excess demand would push up rents and thus lower rental yield in this segment.

The latter two arguments thus suggest that the equilibrium condition is violated in the following way

$$\begin{cases} \frac{R_t(\vartheta)}{P_t(\vartheta)} \geq u_t(\vartheta), & \text{low } \vartheta, \\ \frac{R_t(\vartheta)}{P_t(\vartheta)} = u_t(\vartheta), & \text{average } \vartheta, \\ \frac{R_t(\vartheta)}{P_t(\vartheta)} \leq u_t(\vartheta), & \text{high } \vartheta. \end{cases}$$

As a consequence of all these arguments, rental yields are hence expected to be downwards-sloping when moving from the low end of the distribution to the top end. I will test this expectation in section 4.

### 3 Quality-mismatch: A twofold problem

In the house price index literature, it is known that comparing house prices without controlling for differences in house characteristics – commonly known as *quality adjustment* – leads to noisy results. When constructing house price indices, this problem is usually addressed by applying either hedonic or repeat-sales techniques. Hedonic methods assume that house prices are composed of a list of shadow prices associated with housing characteristics. Changes in prices net of these shadow prices over time are used to construct price indices. Repeat-sales methods restrict the analysis to dwellings sold multiple times to ensure that like is compared with like.

Additionally to a quality-mismatch *within* houses sold or *within* houses rented, another dimension of quality-mismatch emerges when constructing rental yields as the distribution of characteristics tend to be different across rental and sales observations. (I will refer to this issue as *across quality-mismatch*.) In the empirical section of this paper it is shown that for the Sydney data houses rented tend to have smaller land areas, less bed- and bathrooms, and are located closer to the city center than houses sold.

### 3.1 Addressing *across* quality-mismatch: Matching

Exact matching (i.e., relying on repeated sales) has a long tradition in the construction of house price indices. Constructing rental yields from houses sold and rented would thus be a natural extension of the repeat-sales methodology. Considering house price indices, several studies found evidence of an Akerlof-type lemons bias. Houses sold and rented within a short period of time may not be lemons but they are hardly a random sub-sample drawn from the population of all dwellings currently on the market. Relying on these observations only may hence induce a sample selection bias. If for instance a house is bought to let, the well-informed buyer would rather try to buy a house that fits the demand in the rental market in terms of characteristics and location. Owner-occupiers buy for themselves and may hence demand different types of houses. I use several methods to test for a sample selection bias for the Sydney data, which thoroughly confirm the existence of such a bias. Results are reported in subsection 4.5.

Moving from exact matching to a less strict matching strategy may reduce the impact of a sample selection bias as a much greater share of observations is used. Such matching strategies have, however, yet been rarely used in housing contexts.<sup>3</sup>

Ho et al. (2007) and McMillen (2012) point out that matching is also an appropriate data preprocessing technique that reduces noise arising from a changing mix of characteristics. Additionally, matching reduces sensitivity towards outliers and makes results less sensitive towards model specification. Here, I use a combination of exact matching (i.e., matching houses that were sold and rented) and matching of houses similar in their characteristics to address the across quality-mismatch problem. Matching is performed as a data-preprocessing step before hedonic models are applied, which control for within quality-mismatch as described in subsection 3.2.

In general, there is no *right* matching procedure. Matching performance can be checked directly by analyzing the balance of house characteristics across matched samples. Hence, several matching procedures may be tried out and one should choose the approach that yields the best balance. Ho et al. (2007, page 216) state that

[t]rying different matching methods is not like trying different models, some of which are right and some wrong, since balance provides a reasonably straightforward objective function to maximize and choose matching solutions. [...] one should try as many matching solutions as possible and choose the one with the best balance as the final preprocessed data set. [...] matching solutions with suboptimal balance are in fact irrelevant and should play no part in our ultimate inferences.

All procedures that analyze covariates only to find similar observations are eligible matching strategies. It is important to note that the analysis has to be done independently of the response variable, i.e., the sales or rental price, to guarantee unbiased results.

A commonly used matching procedure is *propensity score matching* (PSM)<sup>4</sup> which summarizes all covariates with a single number – the propensity score. Similar propensity scores indicate a similar set of covariates. Propensity scores are obtained by regressing the dummy

---

<sup>3</sup>McMillen (2012) and Guo et al. (2014) develop matching estimators to construct house price indices and apply them to house sales in Chicago, Illinois, USA, and new home sales in Chengdu, Sichuan Province, China, respectively. Deng et al. (2012) apply such a matching estimator to house sales in Singapore.

<sup>4</sup>See Rosenbaum and Rubin (1983).

variable *type*, that indicates whether the dwelling was sold  $type = 1$  or rented  $type = 0$ , on the vector of house characteristics using a probit or logit link function. Here, I use a logit link function.<sup>5</sup> The logit model is evaluated for each observation in the rental and sales data set yielding a specific propensity score for each observation. Then, an observation from the sales data set is drawn randomly and matched to an observation from the rental data set with most similar propensity score, which is called *nearest neighbor matching*. In case there are more possible matches with identical propensity scores, the matched observation is chosen randomly. The matching is done separately for each period. Before applying PSM, covariates are checked whether they fall outside the common support and are discarded in this case.<sup>6</sup>

---

**Procedure 1:** Matching strategy.

---

Let  $T$  be the number of periods,  $L$  the number of regions and *rent* a dummy variable indicating the type of transaction.

**For each period  $t$  perform the following steps:**

- 1 | Generate sub-samples  $S_t^{all}$  and  $R_t^{all}$  consisting of all sales and rental observations in period  $t$ ;
- 2 | Filter all observations that appear in  $S_t^{all}$  and  $R_t^{all}$ , i.e., identify all exact matches yielding exact matches samples  $S_t^e$  and  $R_t^e$ ;
- 3 | Stratify the remaining observations  $S_t = S_t^{all} \setminus S_t^e$  and  $R_t = R_t^{all} \setminus R_t^e$  by regions yielding  $2 \cdot L$  stratified samples  $S_t^l$  and  $R_t^l$ ,  $l = 1 \dots, L$ ;

**For each region  $l$  perform the following steps:**

- 4 | Estimate logit models for structural covariates  $x_{il}$  belonging to observations in  $S_{il} \cup R_{il}$ 

$$logit(type_t | x_{il}) = x_{il} \beta_{il}$$
and calculate propensity scores  $\pi_{il}^r$  for the rental observations in  $R_{il}$  and  $\pi_{il}^s$  for the sales observations in  $S_{il}$ ;
- 5 | Perform nearest neighbor PSM with random ordering to obtain matched samples  $R_{il}^*$  and  $S_{il}^*$ ;

6 | **End**

- 7 | Combine matched samples  $R_t^h = R_t^e \cup \left( \bigcup_{l=1}^L R_{tl}^* \right)$  and  $S_t^h = S_t^e \cup \left( \bigcup_{l=1}^L S_{tl}^* \right)$ .

8 | **End**

---

In this paper, PSM is the core element of the matching strategy. Since PSM can be performed on *observed* house characteristics only, omitted variables may play a role. The list of possibly important house characteristics is endless and it is practically impossible to obtain all the characteristics that may affect house prices or rents. For instance, the specific layout of rooms most probably determines part of the price or rent but it is almost impossible to measure it in a processable way. Therefore, I include two additional steps that help to reduce such influences.

Procedure 1 summarizes the matching strategy. First, I identify all dwellings that were sold and rented in the same period and include these exact matches in the final sample

---

<sup>5</sup>A robustness check finds that for the data used in this paper probit models yield very similar but slightly worse results in terms of balance. A distance measure based on the Mahalanobis distance was also tried. Results are again similar but slightly worse.

<sup>6</sup>The PSM is conducted using the R package `MatchIt` by Ho et al. (2011).

(steps 1 and 2). The Sydney data set provides a unique identifier for each house which is used to filter exact matches. Further matching is performed on all observations that are not exact matches. As a fraction of all matched observations is of equal quality and is thus not at all affected by omitted variables, the potential of an overall omitted variables bias is reduced.

Location is the most important price-determining factor. Hence, good matching results are particularly important for locational characteristics. Ho et al. (2007) recommend to treat covariates differently if they are known to be of distinct importance. Therefore, I stratify all observations geographically by creating sub-samples of observations belonging to the same region (step 3).<sup>7</sup> PSM is then performed on the stratified samples using only structural characteristics<sup>8</sup> in the logit model (steps 4 and 5). Neighborhoods tend to develop at the same time (see Basu and Thibodeau, 1998; Hill, 2013) which yields houses similar in their characteristics within short distances. The stratified PSM searches for houses with similar *observed characteristics* within each region. As houses within a region tend to be of similar quality, this approach yields in fact houses of similar *observed and unobserved characteristics* and thus a potential omitted variables bias is further reduced.

Ultimately, exact and PSM pairs are included into the final samples denoted by  $R_t^h$  and  $S_t^h$  which I refer to as *hybrid matches samples* (step 6). Samples including exact matches only are denoted by  $R_t^e$  and  $S_t^e$  and the original data sets, i.e., the full samples, by  $R_t^{all}$  and  $S_t^{all}$ . Note that generally  $\#R_t^{all} \neq \#S_t^{all}$ , whereas per construction  $\#R_t^e = \#S_t^e$  and  $\#R_t^h = \#S_t^h$ .<sup>9</sup>

### 3.2 Addressing *within* quality-mismatch: Marginal densities

I rely on a hedonic approach to address quality differences within the sales and rental sample. A hedonic equation writes the house price or rent written as a function of house characteristics whose associated parameters are interpreted as shadow prices. Here, I estimate hedonic equations using quantile regression models which allows different conditional quantiles, i.e., points in the price or rent distribution conditional on house characteristics, to be estimated. I use these conditional quantiles to construct samples from the marginal rental and sales price distribution *net of quality differences*.<sup>10</sup>

Let  $n$  be the number of observations,  $x \in \mathbb{R}^{n \times d}$  a matrix of the values of  $d$  house characteristics including an intercept,  $p = (p_1, \dots, p_n)^\top$  a vector of observed rental or sales prices and  $\beta = (\beta_1, \dots, \beta_d)^\top$  a vector of unknown shadow prices associated with the  $d$  characteristics. A standard hedonic semi-log model<sup>11</sup> estimates the conditional mean

<sup>7</sup>Residex, an Australian provider of property information, divides Sydney into 16 regions: Campbelltown, Canterbury-Bankstown, Cronulla-Sutherland, Eastern Suburbs, Fairfield-Liverpool, Inner Sydney, Inner West, Lower North Shore, Manly-Warringah, Mosman-Cremorne, North Western, Parramatta Hills, Penrith-Windsor, St Georges, Upper North Shore and Western Suburbs.

<sup>8</sup>In the logit model, the number of bed- and bathrooms is treated as metric variable rather than categorical allowing comparisons across categories, e.g., a house with one bedroom is more similar to a house with two bedrooms than to a house with three bedrooms.

<sup>9</sup>The symbol  $\#$  denotes the cardinality of each set.

<sup>10</sup>Quantile regression models date back to Koenker and Bassett Jr (1978) but have yet been rarely applied in housing contexts (see Waltl, 2016b, for a literature review). McMillen (2008) applies a similar technique as used here to analyze temporal changes in house price distributions.

<sup>11</sup>The semi-log functional form is the standard specification of hedonic models in housing contexts. Predicting prices requires a back-transformation to the original scale. In case of a standard linear model, this requires a reliable estimate of the variance to gain an estimate of the conditional mean (see Kennedy, 1981;

function

$$\mathbb{E}(\log p|x) = x\beta.$$

Quantile regression allows the same model to be estimated at different points of the price or rent distribution, i.e., for different quantile levels  $\vartheta \in (0, 1)$

$$Q_{\vartheta}(\log p|x) = x\beta(\vartheta),$$

where

$$Q_{\vartheta}(\log p|x) := F_{\log p|x}^{-1}(\vartheta) = \inf\{z \in \mathbb{R} | F_{\log p|x}(z) \geq \vartheta\}$$

denotes the  $\vartheta$ th quantile of the conditional distribution  $f(\log p|x)$  with cumulative distribution function  $F_{\log p|x}(\cdot)$ .

For each quantile level specific shadow prices  $\hat{\beta}(\vartheta)$  are estimated. Implicitly, hedonic quantile regression models hence assume that buyers and renters value certain characteristics differently across the distribution which is a beneficial feature and makes the model more flexible.

Predicted house prices or rents for houses with characteristics  $x$  are obtained by

$$\hat{p}(\vartheta|x) = \exp\left(x\hat{\beta}(\vartheta)\right).$$

These predictions are per construction *conditional on house characteristics*. Put differently, quantile regression delivers a sample from the conditional distribution  $f(p|x)$  but one is more interested in a sample from the marginal distribution  $f(p)$ . These two densities are related to each other via

$$f(p) = \int f(p|x)f(x) dx = \int f(p|x) dF(x). \quad (2)$$

To obtain quality-adjusted distributions, effects of house characteristics have to be integrated out.

Machado and Mata (2005)<sup>12</sup> give an intuitive algorithm to gain a marginal sample from quantile regression models, i.e., a procedure to numerically integrate out the effects of covariates following formula (2). Procedure 2 summarizes the algorithm adapted to suit this particular case: Quantile regression models are estimated for randomly drawn quantile-levels (steps 1 and 2) which – as a consequence of the inverse probability integral transformation – may be used to construct the conditional distribution  $f(\log p|x)$ . House characteristics are “integrated out” by re-sampling from the covariates distribution (step 3) and evaluating the quantile regression models for these randomly drawn sets of characteristics (step 4). Taking exponents of the resulting predictions ultimately yields a random sample from the marginal distribution  $f(p)$ .

From these samples one obtains quantile-specific, quality-adjusted transaction prices and rents by calculating a series of empirical quantiles for a narrowly spaced sequence  $\{\vartheta_1, \dots, \vartheta_K\}$  of quantile levels yielding

$$\{\hat{p}^r(\vartheta_1), \dots, \hat{p}^r(\vartheta_K)\} \quad \text{and} \quad \{\hat{p}^s(\vartheta_1), \dots, \hat{p}^s(\vartheta_K)\}.$$

---

Waltl, 2016a). For quantiles, this back-transformation is straight-forward as  $Q_{\vartheta}(\log p|x) = \log[Q_{\vartheta}(p|x)]$  which implies  $Q_{\vartheta}(p|x) = \exp[Q_{\vartheta}(\log p|x)]$ , i.e., predicted prices are obtained by just taking the exponent. This is not possible for linear models since  $\mathbb{E}(\log p|x) \neq \log[\mathbb{E}(p|x)]$ .

<sup>12</sup>Firpo et al. (2009) address a related issue and develop a method to estimate unconditional quantile effects.

---

**Procedure 2:** Calculating samples from the marginal price distribution.

---

Let  $T$  be the number of periods,  $p^r$  rental and  $p^s$  sales prices.

- 1 Generate a random sample from  $U(0, 1)$  of size  $J$ :  $\vartheta_1, \dots, \vartheta_J$ ;

**For each period  $t$  perform the following steps:**

- 2 For each quantile level  $\vartheta_j$  estimate quantile regression models separately for the rental and sales data set

$$Q_{\vartheta_j}^r(\log p_t^r | x_t^r) = x_t^r \beta_t^r(\vartheta_j) \quad \text{and} \quad Q_{\vartheta_j}^s(\log p_t^s | x_t^s) = x_t^s \beta_t^s(\vartheta_j)$$

yielding  $J$  sets of coefficients  $(\hat{\beta}_t^r(\vartheta_j)^\top, \hat{\beta}_t^s(\vartheta_j)^\top)$ ;

- 3 Generate a random sample of size  $J$  with replacement from the rows of the rental covariates  $x_t^r$  denoted by  $x_{jt}^{r*}$  and the sales covariates  $x_t^s$  denoted by  $x_{jt}^{s*}$  for  $j = 1, \dots, J$ ;

- 4 Predict prices for the re-sampled covariates to obtain

$$\exp\left(x_{jt}^{r*} \hat{\beta}_t^r(\vartheta_j)\right) \quad \text{and} \quad \exp\left(x_{jt}^{s*} \hat{\beta}_t^s(\vartheta_j)\right)$$

for  $j = 1, \dots, J$ , which form random samples from  $f_t(p^r)$  and  $f_t(p^s)$ .

- 5 **End**
- 

These prices are eventually used to construct quantile-specific, quality-adjusted rental yields<sup>13</sup>

$$\frac{\hat{p}^r(\vartheta_1)}{\hat{p}^s(\vartheta_1)}, \dots, \frac{\hat{p}^r(\vartheta_K)}{\hat{p}^s(\vartheta_K)}.$$

Constructing marginal densities from quantile regression models has some distinct advantages. First, the approach directly yields ratios for a list of quantiles and thus provides a more complete picture of the interaction of rental and housing markets. Second, as a by-product one obtains samples from the marginal rent and price distribution which allows changes in the distribution – changes in means, medians, quantiles, dispersion, etc. – to be analyzed. Finally, the marginal densities approach relies on a hedonic equation, which is a widely-used and accepted concept in housing and urban economics.

There are two main disadvantages. First, a large number of models have to be estimated to gain stable results.<sup>14</sup> Depending on the complexity of the hedonic models and the number of observations per period this may be a challenging task. Second, as with all hedonic approaches, the marginal densities approach may be prone to an omitted variables bias. The marginal densities method integrates out the effects of observed characteristics but fails to adjust prices for unobserved characteristics. Put differently, one observes prices from the distribution  $f(p|x, z)$ , where  $x$  denotes observed and  $z$  unobserved house characteristics, and the marginal densities approach generates samples from  $f(p|z)$ . If  $x$  and  $z$  are independent,

---

<sup>13</sup>Note that for rental yields the *annual* rent is compared to the transaction price. Usually, rents are paid in monthly or weekly installments and must therefore be multiplied by 12 or 52, respectively.

<sup>14</sup>Here, I use  $J = 1,000$  which implies in total the estimation of  $2,000 \cdot T$  models, where  $T$  denotes the number of periods.

one obtains

$$\begin{aligned} f(p|z) &= \frac{f(p, z)}{f(z)} = \frac{\int f(p, x, z) dx}{f(z)} = \frac{\int f(p|x, z)f(x, z) dx}{f(z)} \\ &= \int f(p|x, z) dF(x). \end{aligned}$$

If the list of observed variables includes the most important price-determining characteristics, omitted variables are less important. Generally, the marginal densities approach reduces within quality differences but probably cannot totally eliminate them.

### 3.3 The marginal densities matching approach

The matching strategy proposed in subsection 3.1 addresses across quality-mismatch and the marginal densities approach described in subsection 3.2 within quality-mismatch. Combining these two procedures yields the *marginal densities matching* (MDM) approach, which fundamentally controls for quality-mismatch of both types.

Both, the matching and marginal densities approach, may however also be applied independently of each other or one may rely on exact matches rather than hybrid matches. Using exact matches naturally eliminates across quality-mismatch but may be prone to within quality-mismatch and most importantly sample selection bias. Relying on the full sample in combination with the marginal densities approach may reduce within quality-mismatch but does not address across quality-mismatch.

Table 1 summarizes all possible combinations of the methods discussed in the previous sections and indicates which approach may be subject to which error type. The MDM approach turns out to find an ideal compromise between all possible error sources and is hence the recommended method.

## 4 Empirical application

### 4.1 Data

For the empirical analysis, I use a data set created by Australian Property Monitors<sup>15</sup> which includes sales and rental observations for Sydney, Australia, between 2004 and 2014. In the sales data set, the exact transaction date as well as the sales price in AUD is included. The rental data set includes weekly asking rents collected from newspapers and web advertisements, and the publishing date. Usually, there is little bargaining with rents indicating that asking rents are expected to reflect the market well.

Both data sets include the exact address of the property which was geo-coded to obtain exact longitudes and latitudes. Additionally, there is the number of bedrooms, the number of bathrooms, and land area in square meters. Data has been restricted to observations with land areas less than  $5,000m^2$ , and with a maximum of six bed- or bathrooms. Above that, observations are only included if they fall into a rectangle spanned by longitudes and latitudes that covers the greater Sydney area. Longitudes lie within  $[150.60, 151.35]$  and latitudes within  $[-34.20, -33.40]$ . Observations without price or rent information as well as observations lacking address information are excluded. Obvious erroneous observations

---

<sup>15</sup>See <http://apm.com.au> in order to obtain access to their data sets.



Table 1: Error types.

	Method applied		Error source		
	Matching	Marginal densities	Within quality-mismatch	Across quality-mismatch	Sample selection bias
Exact unadjusted	exact	no	✓	✗	✓
Exact adjusted	exact	yes	~	✗	✓
Hybrid unadjusted	hybrid	no	✓	~	✗
<b>MDM</b>	hybrid	yes	~	~	✗
Full unadjusted	no	no	✓	✓	✗
Full adjusted	no	yes	~	✓	✗

*Note:* The table summarizes all possible combinations of the marginal densities approach and matching procedures. If the marginal densities approach is applied, the method is called *adjusted* and otherwise *unadjusted*. *Exact*, *hybrid* and *full* refer to the data set (and hence the type of matching) which the analysis is based on. Using this terminology, the MDM approach may also be labelled *hybrid adjusted*. All approaches may be subject to *within quality-mismatch*, *across quality-mismatch* and *sample selection bias*. The symbol ✓ indicates that the approach is subject to a bias, ✗ that it is not subject to a bias and ~ that a bias may not be totally eliminated but definitely reduced.

(e.g., a house price of 1 AUD) are eliminated. Altogether there are 341,202 complete sales and 311,674 complete rental observations. Table 2 provides summary statistics and Appendix A reports the PSM results.

## 4.2 Reconstruction strategy

In the sales data set 75.4% and in the rental data set 93.7% of all observations are completely observed, i.e., all characteristics are available. Some properties appear more than once in the data sets as they are sold and rented, or sold (rented) multiple times. It is possible to use these repeated observations to reconstruct some incomplete observations.<sup>16</sup>

The reconstruction algorithm consists of three steps: Missing observations are, if possible, refilled separately: first, within the sales and, second, within the rental data set. In a third step, sales and rental observations are pooled and the reconstruction algorithm is applied on the pooled sample.

A missing characteristic is refilled using information of a completely observed observation of the same property, subject to certain constraints. First, if there are several completely observed values, the algorithm checks whether the observed values differ and refills only if the same value is observed all the times. For instance, if a dwelling appears three times in the data set and a characteristic is completely observed two out of three times, refilling is permitted only when the same value is observed for both complete observations. Second, if a dwelling appears twice within a period of six months, this may be

<sup>16</sup>The algorithm applied in this paper is similar as in Waltl (2016a,b) but extended to cross-refilling between sales and rental observations. In contrast to Waltl (2016a,b), the reconstruction algorithm is not only applied to the number of bed- and bathrooms but also to the variable land area and there is no price restriction as this is not meaningful when comparing rents and sales prices.

Table 2: Summary statistics.

	Rental data			Sales data		
	Full	Hybrid	Exact	Full	Hybrid	Exact
<b>Rental / sales price in AUD</b>						
1st quartile	360	360	390	450,000	448,000	430,000
Median	470	470	500	650,000	646,000	625,000
Mean	569	574	609	826,300	829,200	785,700
3rd quartile	650	650	685	950,000	950,000	912,000
<b>Land area in <math>m^2</math></b>						
1st quartile	405	439	459	465	451	459
Median	573	579	582	590	582	582
Mean	659	665	613	638	624	610
3rd quartile	716	721	704	722	712	704
<b>Number of bedrooms in %</b>						
1	2.35	1.37	0.46	0.31	0.31	0.29
2	16.23	13.27	10.30	8.75	9.24	10.46
3	51.56	52.97	53.17	45.61	46.11	52.83
4	24.52	26.57	28.94	34.43	33.74	28.67
5	4.68	5.11	6.27	9.38	9.13	6.61
6	0.66	0.72	0.87	1.52	1.47	1.14
<b>Number of bathrooms in %</b>						
1	60.29	57.61	56.49	44.28	45.21	54.48
2	31.32	33.30	33.12	39.53	39.08	34.37
3	7.23	7.83	8.87	13.61	13.18	9.41
$\geq 4$	1.16	1.26	1.52	2.59	2.53	1.74
<b>Number observations</b>						
All	330,102			427,211		
Complete	311,674	284,184	22,851	341,202	284,184	22,851
in % of all	94.4%			79.9%		
in % of full	100%	91.2%	7.3%	100%	83.3%	6.7%

*Note:* The table reports summary statistics for the full, hybrid matches, and exact matches sample separately for rental and sales data. Values for the characteristics of the exact matches sample are slightly different due to erroneous data entries (this is only the case for the number of bedrooms and bathrooms). In the last section *all* refers to all observations, *complete* to all fully observed or fully reconstructed observations. *Complete in % of all* gives the share of complete observations per sample and *complete in % of full* the share of complete observations in the matched samples in percent of all complete observations. As the hybrid matching algorithm is applied to the complete data set, there is no value for *all observations*.

a signal for renovations. Repeat-sales indices usually discard such observations from their calculations for this very reason.

The reconstruction is successful: The share of complete observations in the sales data set is increased from 75.4% to 79.9% and in the rental data set from 93.7% to 94.4%. The empirical analysis is then performed on all completely observed or successfully refilled

observations.

### 4.3 Hedonic models

The MDM approach requires the estimation of hedonic models for various quantile levels  $\vartheta$ . All models are of the structure

$$Q_{\vartheta}(\log p|X) = \beta_1(\vartheta) + \beta_2(\vartheta) \log(\text{area}) + \sum_{j=2}^6 \beta_j^{\text{bed}}(\vartheta) \mathbb{1}_{\{j\}}(\text{bed}) \\ + \sum_{j=2}^4 \beta_j^{\text{bath}}(\vartheta) \mathbb{1}_{\{j\}}(\text{bath}) + f^{\vartheta}(\text{long}, \text{lat}),$$

where  $X$  contains all covariates. Models are estimated semester-wise.<sup>17</sup> The variable  $p$  denotes either the sales price or the rent,  $\text{area}$  the land area,  $\text{bed}$  the number of bedrooms,  $\text{bath}$  the number of bathrooms, and  $\beta$  associated shadow prices. The symbol  $\mathbb{1}_{\{j\}}(x)$  denotes the indicator function

$$\mathbb{1}_{\{j\}}(x) = \begin{cases} 1, & x = j, \\ 0, & x \neq j. \end{cases}$$

There is only a small number of observations with five or six bathrooms. Therefore, the categories four, five, and six bathrooms are merged to a single category.

A dwelling's particular location is a most crucial price-determining characteristic and it is hence most important to model locational effects precisely. Location comprises a wide variety of effects such as for instance crime rates, air pollution, and distance to amenities such as parks, schools, hospitals, or public transportation. A geo-spatial spline  $f(\text{long}, \text{lat})$  defined on exact longitudes and latitudes smoothly links neighboring dwellings.<sup>18</sup> The spline ultimately leads to a precise price map (see Figure 9 in Appendix B) that models locational variation within the city.

The MDM approach implicitly assumes that renters and buyers value house characteristics differently as models are estimated separately for rental and sales observations.<sup>19</sup> Heterogeneous preferences are plausible for various reasons and also supported by estimated

<sup>17</sup>In this paper, I call the period between January and June the *first* and between July and December the *second semester* of a particular year. In general, results are more precise when using shorter time periods. However, the rental data used in this analysis is not suited for shorter periods.

<sup>18</sup>Using geo-spatial splines has been suggested by Hill and Scholz (2014) for hedonic imputation house price indices and adapted for hedonic quantile imputation house price indices by Waltl (2016b). As in Waltl (2016b), I use penalized quantile regression together with the triogram method developed by Hansen et al. (1998) and Koenker and Mizera (2004) to estimate quantile-specific geo-spatial splines. The smoothing parameter is chosen using an adapted Schwartz Information Criterion as suggested by Koenker et al. (1994).

<sup>19</sup>Homogeneous tastes across renters and buyers may be estimated using a single model for price and rent observations. The common price variable  $p$  equals the sales price if the respective house was sold and otherwise the annual rent. The dummy variable  $\text{type}$  indicates the type of observation: If the house was sold  $\text{type} = 1$  and otherwise  $\text{type} = 0$ . A hedonic model that regresses  $\log p$  on house characteristics  $x$  and the dummy variable  $\text{type}$ ,  $Q_{\vartheta}(\log p|x, \text{type}) = x\beta(\vartheta) + \delta(\vartheta) \cdot \text{type}$ , directly generates quality-adjusted rental yields:

$$\frac{\hat{R}(\vartheta)}{\hat{P}(\vartheta)} = \frac{\exp(x\hat{\beta}(\vartheta))}{\exp(x\hat{\beta}(\vartheta) + \hat{\delta}(\vartheta))} = \exp(-\hat{\delta}(\vartheta)).$$

Whereas  $\hat{R}$  and  $\hat{P}$  are generally unbiased when using quantile regression, for OLS estimates this holds true only in the case of homoscedastic errors. (In that case the variance term cancels out.)

shadow prices in this analysis (see Appendix B): For instance, someone might choose to rent because she expects to stay in the house only temporarily. Hence, she might be more willing to accept a bad location or low quality of the structure compared to someone planning to stay in the area permanently. However, also the reverse scenario is plausible: A homebuyer may be more willing to invest into renovation and is hence more likely to settle for a lower quality structure. Renters in contrast may not be able to “reap the full benefits of improvements they make to the property inside and out” as Smith and Smith (2006) argue. Renters and buyers may represent different age and income groups, which on average have different preferences.

The MDM approach additionally allows shadow prices to vary cross-sectionally thus adding another dimension of heterogeneity in preferences. Quantiles refer to house prices and rents, respectively, and hence are expected to be linked to the income or wealth distribution. Households choosing very expensive houses tend to be wealthier than households looking for cheaper ones and it is plausible that households do not value house characteristics uniformly across the income or wealth distribution.

## 4.4 Rental yields

Estimated rental yields vary strongly over time and cross-sectionally. Figure 1 shows rental yields as a function of quantile levels for three selected periods and as a function of time for three selected quantile levels.<sup>20</sup>

As discussed in section 2, the user cost formula suggests declining rental yields when moving from the low end of the distribution to the top end, i.e., a negative relationship between quantile levels and rental yields. This pattern is formally tested using a simple linear regression model:

$$\frac{\hat{R}_t(\vartheta)}{\hat{P}_t(\vartheta)} = \alpha_t + \beta_t \cdot \vartheta + \varepsilon_t,$$

where  $\hat{R}_t(\vartheta)/\hat{P}_t(\vartheta)$  denotes the estimated quantile-specific rental yield,  $\vartheta$  the respective quantile-level, and  $\varepsilon_t$  an independently and normally distributed error term. This regression is run separately for each semester. Estimated slope parameters  $\hat{\beta}_t$  are significantly negative ( $H_0 : \beta_t = 0$  versus  $H_1 : \beta_t < 0$  using a one-sided  $t$ -test) in all semesters thus confirming the hypothesis of decreasing rental yields.

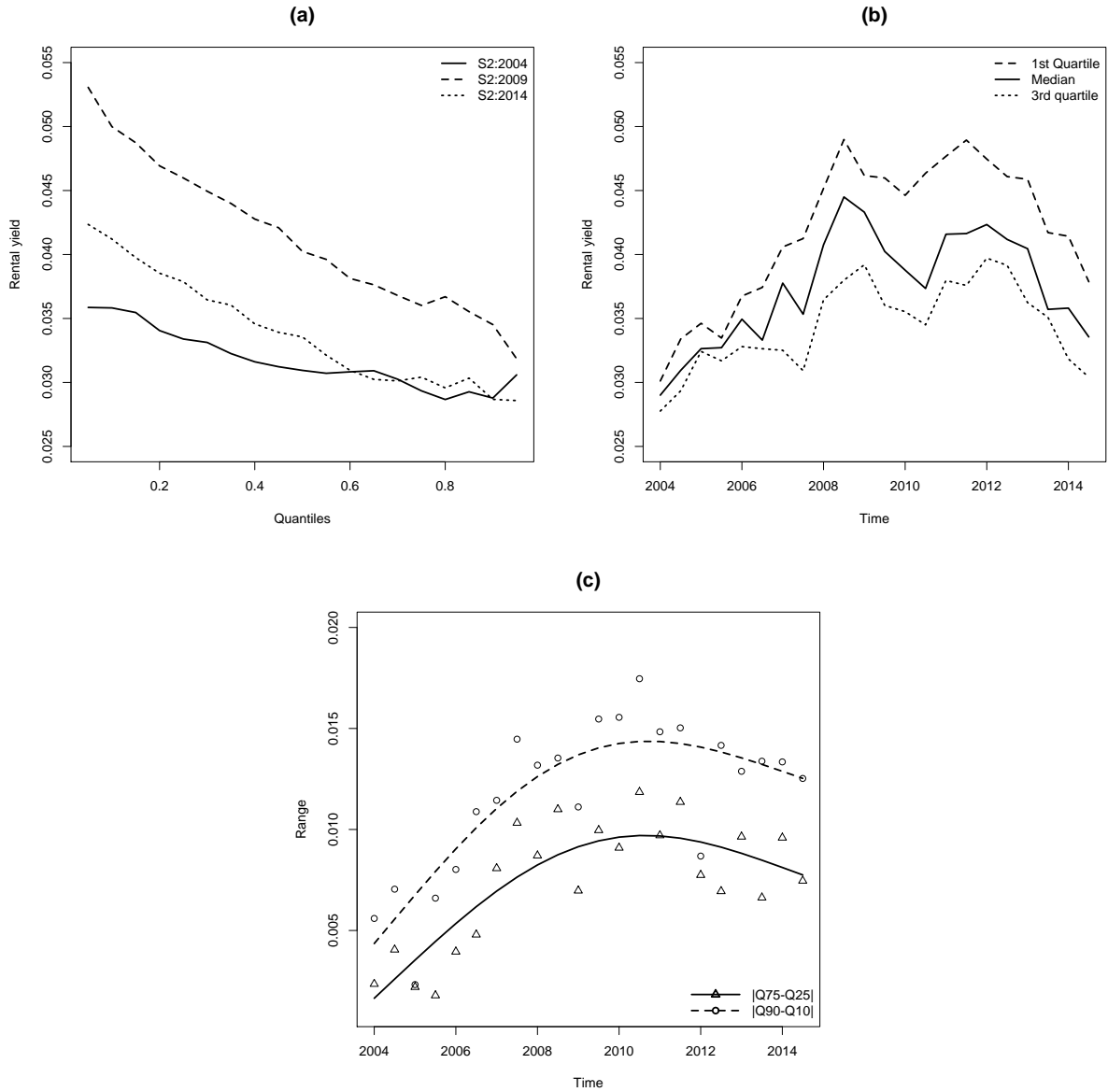
The temporal development of rental yields is strongly linked to changes in house prices (see Figure 2), which might be a hint that rents are more sticky than prices: 2004 constitutes the peak of a great housing boom in Sydney and prices fell thereafter. This period is linked to rising rental yields. The same is true after the (in Australia generally less pronounced) peak associated with the global financial crisis around 2008. During periods of rapidly increasing prices (2009–2010 and from 2012 onwards) rental yields are falling.

Dispersion in rental yields varied considerably over time: The spread increased until 2009 (coinciding with increasing rental yields) and moderately fell thereafter. (See panel (c) in Figure 1.)

---

<sup>20</sup>Appendix C provides more empirical results including a table reporting rental yields per period for the first, second (median) and third quartile (Table 6) and a figure summarizing all results simultaneously in a three-dimensional surface plot spanned by time and quantiles (Figure 10).

Figure 1: Quantile-specific rental yields over time.



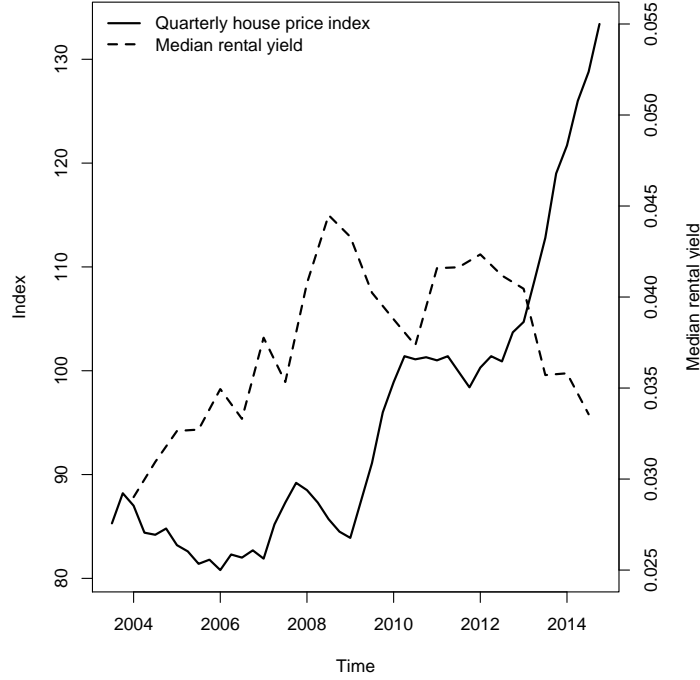
*Note:* Panel (a) shows MDM rental yields as a function of quantile levels for three semesters. Panel (b) shows rental yields over time for three quantile levels. Panel (c) shows the interquartile range  $|Q_{75} - Q_{25}|$  and the range between the 0.9- and 0.1-quantile  $|Q_{90} - Q_{10}|$ . To ease visualization, I included a smooth regression curve obtained from an additive model using thin-plate regression splines.

## 4.5 Sample selection bias

Relying on exact matches, i.e., houses sold and rented within the same period, may introduce a sample selection bias similar to the well-known Akerlof-type lemons bias in repeat-sales indices. In the Sydney data set, median rents calculated from exact matches are higher than in the overall sample whereas median prices are lower (see Table 2).<sup>21</sup> This

<sup>21</sup>Bracke (2015) uses exact matches to analyze rental yields in London, UK. For this data set too, median rents are higher and median sales prices lower when relying on exact matches only.

Figure 2: House price index compared to median rental yields.



*Note:* The figure plots a quarterly house price index based on stratification methods constructed by the Australian Bureau of Statistics together with median rental yields.

implies that the rental yield may be pushed up by both the numerator and denominator and gives a first hint for a systematically positive sample selection bias.

I will test for a sample selection bias more formally in two ways. First, I directly take advantage of the MDM methodology developed in this paper. As seen in Table 1, a sample selection bias may be estimated when comparing median rental yields calculated according to the MDM approach with median rental yields that are based on quality-adjusted exact matches (referred to by *exact adjusted* in the table and by the superscript *EA* in the formula below). The bias is approximated<sup>22</sup> by

$$100 \cdot \frac{1}{T} \sum_{t=1}^T \left( \frac{\text{rental yield}(\vartheta = 0.5)_t^{EA}}{\text{rental yield}(\vartheta = 0.5)_t^{MDM}} - 1 \right) = 8.14\%.$$

The average bias as well as all but one period-specific biases are positive indicating that rental yields calculated from exact matches only are likely to be systematically upwards biased. The largest period-specific bias equals 17.12% and is found in S1:2013.

Alternatively, one may approximate the bias by estimating hedonic models including a dummy variable marking exact matches: *exact* = 1 when a specific transaction is an exact match and *exact* = 0 otherwise. I hence estimate the following models separately for the

<sup>22</sup>As the MDM approach does not eliminate across and within quality-mismatch biases but only reduces them, the result here is an approximation.

sales and rental data:<sup>23</sup>

$$Q_{0.5}(\log p^{s,r}|X, exact) = \delta^{s,r} \cdot exact + X\beta,$$

whereas  $X$  denotes a matrix of house characteristics including a locational spline and structural characteristics, and  $p^{s,r}$  either the sales or rental price. I estimate these models using penalized quantile regression to obtain an estimate for the median.<sup>24</sup>

The exponent of the estimated coefficients  $\hat{\delta}^s$  and  $\hat{\delta}^r$  may be used to measure the magnitude of the sample selection bias. Let  $R$  denote the median annual rent and  $P$  the median sales price representing the entire market. For exact matches the rental yield is expected to be biased following

$$\frac{R \cdot \exp(\delta^r)}{P \cdot \exp(\delta^s)} = \text{rental yield} \cdot \underbrace{\exp(\delta^r - \delta^s)}_{\text{bias term}}.$$

Insignificant estimates for  $\delta^r$  and  $\delta^s$  would indicate that exact matches are a perfect sub-sample of the overall set of observations, i.e.,  $\delta^r \approx \delta^s \approx 0$ . However, the coefficients  $\hat{\delta}_t^r$  are significantly different from zero at the 0.01-level in all but one period and consistently positive. The coefficients  $\hat{\delta}_t^s$  are significant in 11 out of 22 periods at the 0.01-level and in all but one period negative. To estimate the average bias, I take the mean over all period-wise estimated biases:

$$\frac{1}{T} \sum_{t=1}^T \left( \exp(\hat{\delta}_t^r - \hat{\delta}_t^s) - 1 \right) = 5.75\%.$$

Again a large positive bias is estimated.

Almost all  $\hat{\delta}_t^r$  are significant whereas just half of all  $\hat{\delta}_t^s$  are. This indicates that it is rather distortion in the rent that pushes the sample selection bias than distortion in sales prices.

A sample selection bias is likely not to affect the median only but also other quantiles. The methods described above may directly be extended to a quantile-specific measure of sample selection bias.<sup>25</sup> Also quantile-specific rental yields based on exact matches only are found to be severely and systematically upwards biased. The magnitudes range between 5.60% and 8.89%. Table 7 in Appendix C reports more results.

## 4.6 Comparison to alternative methods

The MDM approach is designed to construct aggregate quality-adjusted and quantile-specific rental yields. To measure the success of quality-adjustment, I compare the MDM results to rental yields calculated from empirical quantiles without any control of quality differences (this method is referred to by *full unadjusted* in Table 1 and by *FU* in the formula below) in a similar way as in the calculation of the sample selection bias in

<sup>23</sup>The regressions are estimated on the hybrid matches sample to guarantee comparability to the calculations above.

<sup>24</sup>Alternatively, the models may also be estimated using penalized least squares to obtain an estimate for the mean. For this data sets, results for the mean and median are almost identical.

<sup>25</sup>For the second approach, the hedonic models are simply estimated for other quantile levels than the median.

subsection 4.5:

$$100 \cdot \frac{1}{T} \sum_{t=1}^T \left( \frac{\text{rental yield}(\vartheta)_t^{FU}}{\text{rental yield}(\vartheta)_t^{MDM}} - 1 \right) = \begin{cases} -2.36\%, & \vartheta = 0.25, \\ -1.69\%, & \vartheta = 0.5, \\ 0.38\%, & \vartheta = 0.75. \end{cases}$$

There is a tendency towards a negative bias when neglecting quality-control. However, it is not clear whether this is a systematic bias as the sign of the bias is not consistently negative in all periods or all price segments. (For the median, the bias is negative in 15 out of 22 periods.)

To obtain quantile-specific rental yields, the MDM approach compares a quantile of the quality-adjusted rental price distribution to the corresponding quantile of the quality-adjusted sales price distribution. Alternatively, one may calculate quantiles from the distribution of rental yields obtained from individual dwellings directly. Individual rental yields may be obtained from exact or hybrid matches. However, the interpretation is fundamentally different: A high quantile level does neither refer to a high rent nor to a high sales price as it is the case for the MDM approach. It rather reports the amount of variation in rental yields within a market. Due to this different interpretations, I only compare median rental yields.

Figure 3 shows the results: The levels of MDM rental yields are consistently lower than those from aggregated individual ratios. The temporal patterns are, however, very similar. When comparing rental yields aggregated from exact and hybrid matches, respectively, it turns out that exact rental yields are consistently larger supporting again a systematically positive sample selection bias.

## 4.7 Robustness check

The MDM approach compares quantile  $\vartheta$  of the rental distribution with quantile  $\vartheta$  of the sales price distribution. However, one would hope that results change only little when evaluating either distribution not exactly for the quantile  $\vartheta$  but in a neighborhood of  $\vartheta$ , i.e.,

$$\frac{R([\vartheta - \varepsilon, \vartheta + \varepsilon])}{P(\vartheta)} \quad \text{or} \quad \frac{R(\vartheta)}{P([\vartheta - \varepsilon, \vartheta + \varepsilon])}$$

for a small  $\varepsilon > 0$ .

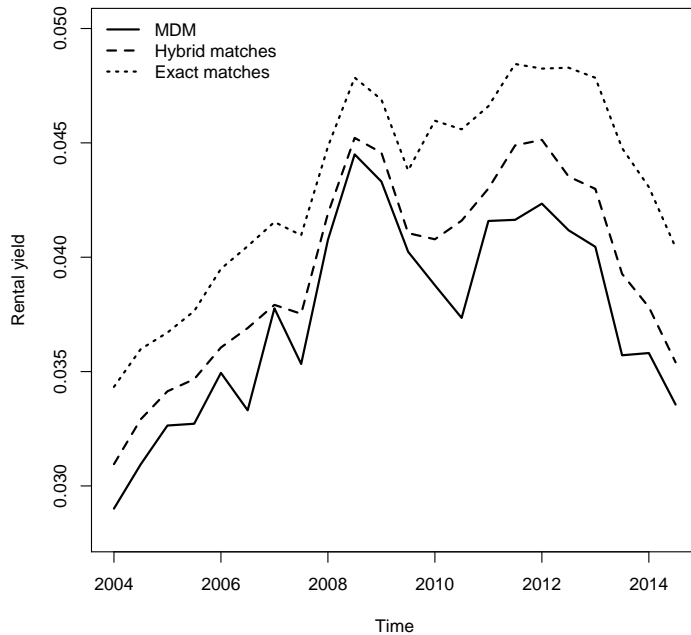
I analyze the range of ratios for an  $\varepsilon$ -neighborhood of 1% for three quantile levels  $\vartheta \in \{0.25, 0.5, 0.75\}$  and therefore measure the distances between minimum and maximum ratio for each period, i.e.,

$$\text{distance}_t(\vartheta) = \max_{\varepsilon \in [0,1]} \left\{ \frac{R_t([\vartheta - \varepsilon, \vartheta + \varepsilon])}{P_t(\vartheta)}, \frac{R_t(\vartheta)}{P_t([\vartheta - \varepsilon, \vartheta + \varepsilon])} \right\} \\ - \min_{\varepsilon \in [0,1]} \left\{ \frac{R_t([\vartheta - \varepsilon, \vartheta + \varepsilon])}{P_t(\vartheta)}, \frac{R_t(\vartheta)}{P_t([\vartheta - \varepsilon, \vartheta + \varepsilon])} \right\}.$$

It turns out that deviations are very small thus increasing confidence in the MDM results. Table 3 reports the minimum, maximum, and average distance over all periods. Figure 11 in Appendix C shows results for three time periods.



Figure 3: Aggregating individual rental yields.



*Note:* The figure compares MDM results with results obtained from aggregating individual rental yields based on hybrid and exact matches, respectively.

Table 3: Robustness check.

	Q1	Median	Q3
Average distance	0.22%	0.18%	0.25%
Minimum distance	0.17%	0.13%	0.18%
Maximum distance	0.33%	0.29%	0.37%

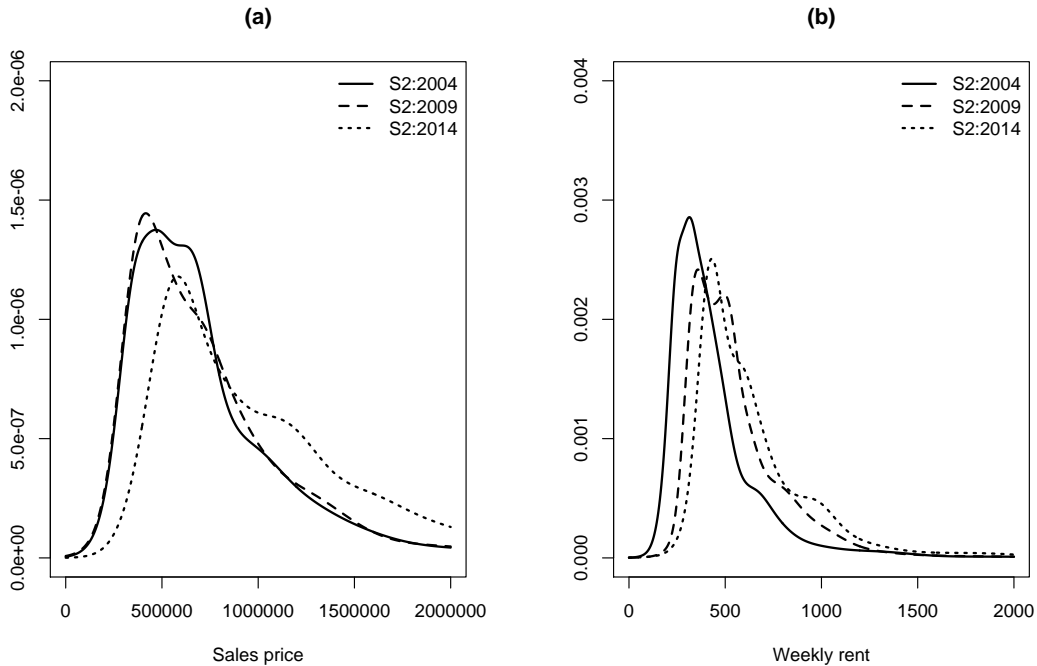
*Note:* The table reports deviations of the  $\vartheta$  MDM rental yield when evaluating either the rent or sales price distribution not for  $\vartheta$  exactly but for a small environment  $[\vartheta - 1\%, \vartheta + 1\%]$ . Distances are calculated for each period and the table reports the average, minimum and maximum distance for the first (Q1), second (median) and third (Q3) quartile.

## 4.8 Marginal distributions

As a by-product of the MDM approach, one obtains random samples from quality-adjusted rental and sales price distributions. Figure 4 shows these densities for three selected periods.

Distributions change over time not only in the mean or median but also in other aspects such as variance or skewness. Over the whole period, median prices increased by roughly 40% and median rents by 62% (see Figure 5). Relative variation within sales prices – measured by the coefficient of variation – is generally larger than within rents. Both distributions are consistently right-skewed at a similar degree and are also very similar in

Figure 4: Marginal densities.



*Note:* The figures shows marginal densities resulting from the MDM approach for three selected periods. Panel (a) shows densities for sales prices and panel (b) for rental prices. Densities are constructed using an adaptive kernel approach with local bandwidths as suggested by Portnoy and Koenker (1989) in combination with a normal kernel, Silverman’s rule of thumb, and a sensitivity parameter equal to 0.5.

terms of kurtosis, which is a measure of a distribution’s peakedness. Both distributions are consistently leptocurtic.

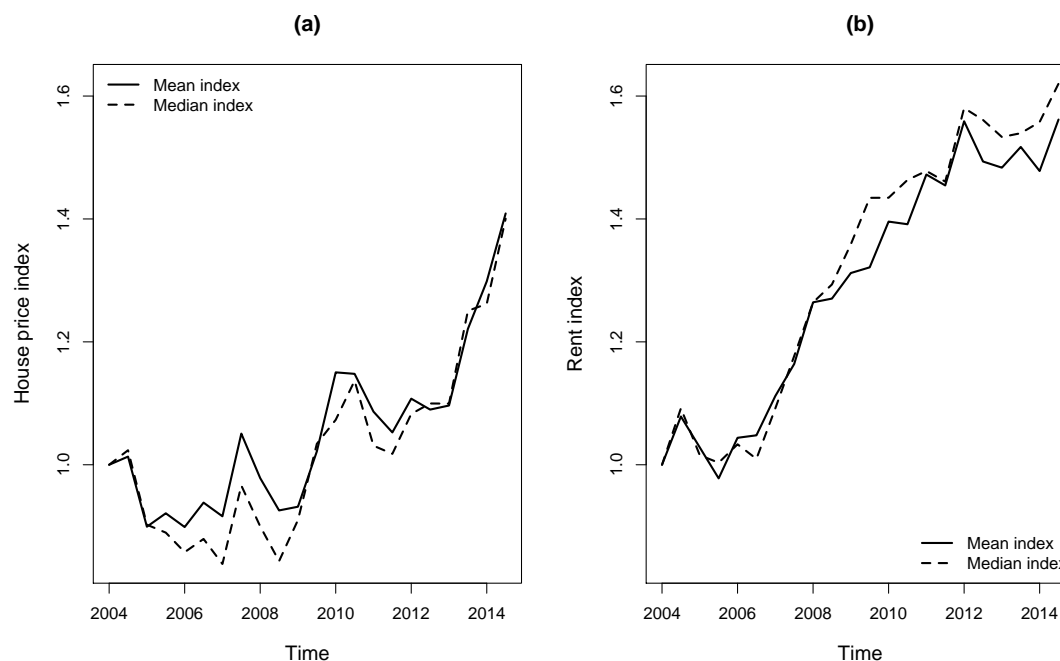
## 5 Summary and conclusions

This paper develops a methodology to construct quality-adjusted and quantile-specific rental yields. The method consists of two steps: The first step is based on propensity score matching and addresses quality-mismatch across sales and rental observations. The standard propensity score matching technique is adapted to minimize a potential omitted variable bias. The second step uses hedonic quantile regression techniques to account for quality-mismatch within sales and within rental observations. Quantile-specific hedonic models are estimated which include a geographical spline that precisely models locational effects. Predictions from these models are utilized to construct marginal sales and rental price distributions net of house characteristics which are ultimately used to construct rental yields.

The methodology is applied to house sales and rents in Sydney, Australia, between 2004 and 2014. Rental yields are found to be downwards-sloping when moving from the low end of the distribution to the top end as predicted by the user cost formula. The negative trend is statistically significant in all periods.

Rental yields also show strong temporal variation which matches developments in house

Figure 5: House price and rent indices.



*Note:* The figure depicts changes in quality-adjusted sales prices and rents obtained from the MDM approach. Panel (a) shows mean and median house price indices and panel (b) rent indices.

prices. This may indicate sticky rents.

A major finding suggests that relying on houses sold and rented within a short period of time – i.e., a repeat-sales type of approach – is likely to introduce a severe sample selection bias. For the Sydney data, I find a systematic positive bias of roughly 6%–8%. A sample selection bias is not only found for the median rental yield but throughout the entire distribution.

Neglecting quality-adjustment introduces errors of on average -2% per period in the median rental yield.

## References

- Basu, S. and Thibodeau, T. G. (1998). Analysis of spatial autocorrelation in house prices. *The Journal of Real Estate Finance and Economics*, 17(1):61–85.
- Bracke, P. (2015). House prices and rents: Microevidence from a matched data set in Central London. *Real Estate Economics*, 43(2):403–431.
- Campbell, S. D., Davis, M. A., Gallin, J., and Martin, R. F. (2009). What moves housing markets: A variance decomposition of the rent-price ratio. *Journal of Urban Economics*, 66(2):90–102.
- Capozza, D. R. and Seguin, P. J. (1996). Expectations, efficiency, and euphoria in the housing market. *Regional Science and Urban Economics*, 26(3-4):369–386.

- Clapp, J. M. and Giaccotto, C. (1992). Estimating price indices for residential property: A comparison of repeat sales and assessed value methods. *Journal of the American Statistical Association*, 87(418):300–306.
- Clark, T. E. (1995). Rents and prices of housing across areas of the United States. A cross-section examination of the present value model. *Regional Science and Urban Economics*, 25(2):237–247.
- Davis, M. A., Lehnert, A., and Martin, R. F. (2008). The rent-price ratio for the aggregate stock of owner-occupied housing. *Review of Income and Wealth*, 54(2):279–284.
- de Haan, J. and Diewert, W. E., editors (2013). *Handbook on Residential Property Prices Indices (RPPIs)*. Methodologies and Working Papers. Eurostat, Luxembourg.
- Dehejia, R. H. and Wahba, S. (1999). Causal effects in nonexperimental studies: Reevaluating the evaluation of training programs. *Journal of the American Statistical Association*, 94(448):1053–1062.
- Deng, Y., McMillen, D. P., and Sing, T. F. (2012). Private residential price indices in Singapore: A matching approach. *Regional Science and Urban Economics*, 42(3):485–494.
- Diewert, W. E., Haan, J. d., and Hendriks, R. (2015). Hedonic regressions and the decomposition of a house price index into land and structure components. *Econometric Reviews*, 34(1-2):106–126.
- Diewert, W. E., Nakamura, A. O., and Nakamura, L. I. (2009). The housing bubble and a new approach to accounting for housing in a CPI. *Journal of Housing Economics*, 18(3):156–171.
- Duca, J. V., Muellbauer, J., and Murphy, A. (2011). House prices and credit constraints: Making sense of the US experience. *The Economic Journal*, 121(552):533–551.
- Firpo, S., Fortin, N. M., and Lemieux, T. (2009). Unconditional quantile regressions. *Econometrica*, 77(3):953–973.
- Fox, R. and Tulip, P. (2014). Is housing overvalued? *Research Discussion Paper Reserve Bank of Australia*, RDP 2014-06.
- Fu, Y. and Ng, L. K. (2001). Market efficiency and return statistics: Evidence from real estate and stock markets using a present-value approach. *Real Estate Economics*, 29(2):227–250.
- Gallin, J. (2008). The long-run relationship between house prices and rents. *Real Estate Economics*, 36(4):635–658.
- Guerrieri, V., Hartley, D., and Hurst, E. (2013). Endogenous gentrification and housing price dynamics. *Journal of Public Economics*, 100:45–60.
- Guo, X., Zheng, S., Geltner, D., and Liu, H. (2014). A new approach for constructing home price indices: The pseudo repeat sales model and its application in China. *Journal of Housing Economics*, 25:20–38.

- Hansen, M., Kooperberg, C., and Sardy, S. (1998). Triogram models. *Journal of the American Statistical Association*, 93(441):101–119.
- Hill, R. J. (2013). Hedonic price indexes for residential housing: A survey, evaluation and taxonomy. *Journal of Economic Surveys*, 27(3):879–914.
- Hill, R. J. and Scholz, M. (2014). Incorporating geospatial data in house price indexes: A hedonic imputation approach with splines. *Graz Economics Papers*, 2014-05.
- Hill, R. J. and Syed, I. A. (2016). Hedonic price-rent ratios, user cost, and departures from equilibrium in the housing market. *Regional Science and Urban Economics*, 56:60–72.
- Himmelberg, C., Mayer, C., and Sinai, T. (2005). Assessing high house prices: Bubbles, fundamentals and misperceptions. *Journal of Economic Perspectives*, 19(4):67–92.
- Ho, D. E., Imai, K., King, G., and Stuart, E. A. (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis*, 15(3):199–236.
- Ho, D. E., Imai, K., King, G., and Stuart, E. A. (2011). MatchIt: Nonparametric preprocessing for parametric causal inference. *Journal of Statistical Software*, 42(8):1–28.
- Holzhey, M. and Skoczek, M. (2016). UBS Global Real Estate Bubble Index. Available at: [https://www.ubs.com/global/en/wealth\\_management/wealth\\_management\\_research/global-bubble-index.html](https://www.ubs.com/global/en/wealth_management/wealth_management_research/global-bubble-index.html) (accessed 21 December 2016).
- Kennedy, P. E. (1981). Estimation with correctly interpreted dummy variables in semilogarithmic equations. *The American Economic Review*, 71(4):801.
- Koenker, R. and Bassett Jr, G. (1978). Regression quantiles. *Econometrica*, 46(1):33–50.
- Koenker, R. and Mizera, I. (2004). Penalized triograms: total variation regularization for bivariate smoothing. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 66(1):145–163.
- Koenker, R., Ng, P., and Portnoy, S. (1994). Quantile smoothing splines. *Biometrika*, 81(4):673–680.
- Larsen, E. R. and Sommervoll, D. E. (2009). The impact on rent from tenant and landlord characteristics and interaction. *Regional Science and Urban Economics*, 39(3):316–322.
- Machado, J. A. F. and Mata, J. (2005). Counterfactual decomposition of changes in wage distributions using quantile regression. *Journal of Applied Econometrics*, 20(4):445–465.
- McCarthy, J. and Peach, R. W. (2004). Are home prices the next bubble? *Economic Policy Review*, 10(3):1–17.
- McMillen, D. P. (2008). Changes in the distribution of house prices over time: Structural characteristics, neighborhood, or coefficients? *Journal of Urban Economics*, 64(3):573–589.
- McMillen, D. P. (2012). Repeat sales as a matching estimator. *Real Estate Economics*, 40(4):745–773.

- OECD (2016). *OECD Economic Outlook*, volume 2016 issue 2. OECD Publishing, Paris.
- Peng, L. and Thibodeau, T. (2013). Risk segmentation of American homes: Evidence from Denver. *Real Estate Economics*, 41(3):569–599.
- Portnoy, S. and Koenker, R. (1989). Adaptive  $L$ -estimation for linear models. *The Annals of Statistics*, 17(1):362–381.
- Poterba, J. (1992). Taxation and housing: Old questions, new answers. *American Economic Review*, 82(2):237–242.
- Rosen, H. S., Rosen, K. T., and Holtz-Eakin, D. (1984). Housing tenure, uncertainty, and taxation. *The Review of Economics and Statistics*, 66(3):405–416.
- Rosenbaum, P. R. and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55.
- Smith, M. H. and Smith, G. (2006). Bubble, bubble, where’s the housing bubble? *Brookings Papers on Economic Activity*, 2006(1):1–67.
- Steele, M. and Goy, R. (1997). Short holds, the distributions of first and second sales, and bias in the repeat-sales price index. *The Journal of Real Estate Finance and Economics*, 14(1/2):133–154.
- Stiglitz, J. E. (1990). Symposium on bubbles. *Journal of Economic Perspectives*, 4(2):13–18.
- Wallace, N. E. and Meese, R. A. (1997). The construction of residential housing price indices: A comparison of repeat-sales, hedonic-regression, and hybrid approaches. *The Journal of Real Estate Finance and Economics*, 14(1/2):51–73.
- Waltl, S. R. (2016a). A hedonic house price index in continuous time. *International Journal of Housing Markets and Analysis*, 9(4):648–670.
- Waltl, S. R. (2016b). Variation across price segments and locations: A comprehensive quantile regression analysis of the Sydney housing market. *Real Estate Economics*, doi:10.1111/1540-6229.12177.
- Weeken, O. (2004). Asset pricing and the housing market. *Bank of England Quarterly Bulletin*, pages 32–41.

# Supplemental Appendix

## A Matching strategy

Houses rented tend to have smaller land areas, a lower number of bedrooms and a lower number of bathrooms than houses sold (see Table 2). The hybrid matching approach works well to balance house characteristics across sales and rental observations. Rented houses in the hybrid matches sample have larger land areas and sold houses lower land areas, thus increasing balance from two directions. The same effect is observed for the number of bedrooms and the number of bathrooms.

In terms of prices, houses in the hybrid matches sample and houses in the full sample are almost identical. The median (mean) rental price is 1.2% (0.8%) lower in the hybrid matches sample than in the full sample and the median (mean) sales price is 3.0% (2.5%) lower. The average and maximum *absolute* deviations per semester are comparably small (see Table 4).

Per construction, characteristics are balanced across rental and sales observations in the exact matches sample. Average characteristics lie well between average characteristics in the full rental and full sales sample. The median (mean) rental price is however 4.7% (7.2%) higher in the exact matches sample than in the full sample and the median (mean) sales price is 4.6% (4.5%) lower. The average and maximum absolute deviations per semester are even larger (see Table 4). In S2:2004 the mean absolute deviation between rents in the full and exact matches sample is almost 30%.

Median prices and rents obtained from the exact matches sample differ strongly from prices in the full sample, whereas prices and rents in the hybrid matches sample follow them closely (see also Figure 6) indicating that relying on exact matches only is expected to induce a sample selection bias. In subsection 4.5 I formally test and find strong evidence for such a sample selection bias. Above that, only roughly 7% of all observations are exact matches yielding small sample sizes per semester ranging between 187 and 2,029.

Table 4: Average and maximum absolute deviations from full sample prices.

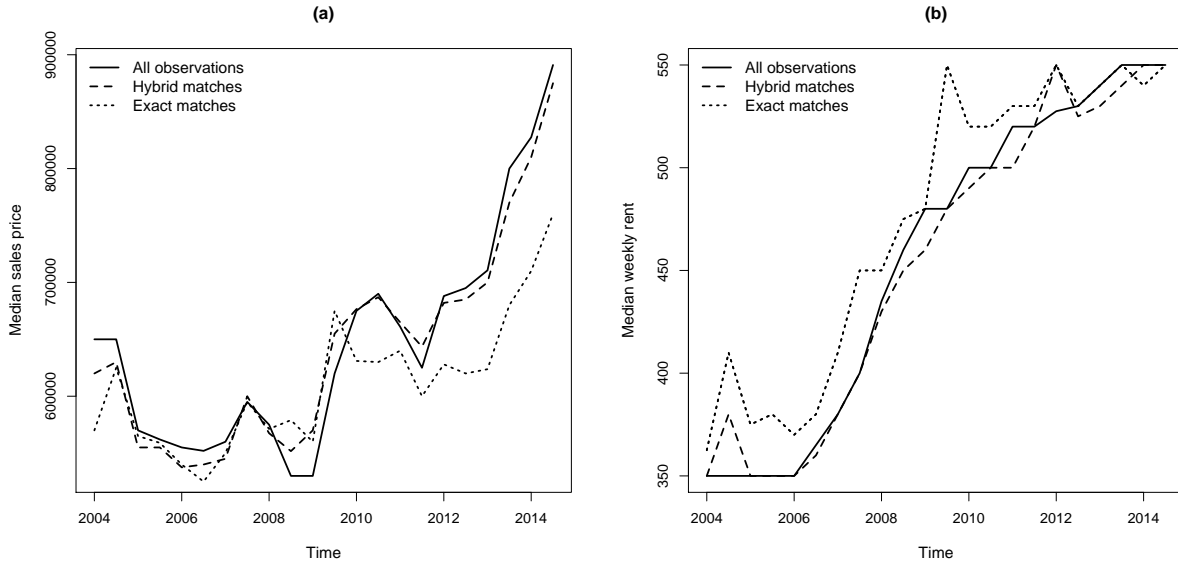
	Hybrid		Exact	
	average	maximum	average	maximum
Median rent	1.46%	8.57%	4.81%	17.14%
Mean rent	1.44%	10.41%	7.37%	29.44%
Median sales price	2.49%	7.55%	6.79%	15.00%
Mean sales price	2.09%	6.95%	6.74%	17.66%

*Note:* The table reports average absolute deviations per semester of mean / median prices in the hybrid and exact matches sample from mean / median prices in the full sample. For instance, the average and absolute absolute deviation of mean prices resulting from the hybrid matches sample are given by

$$\frac{1}{T} \sum_{t=1}^T \left| \frac{\bar{p}_t^{hybrid} - \bar{p}_t^{full}}{\bar{p}_t^{full}} \right| \quad \text{and} \quad \max_{t \in \{1, \dots, T\}} \left| \frac{\bar{p}_t^{hybrid} - \bar{p}_t^{full}}{\bar{p}_t^{full}} \right|,$$

where  $\bar{p}_t^{hybrid}$  ( $\bar{p}_t^{full}$ ) denotes the mean price in period  $t$  in the hybrid matches (full) sample.

Figure 6: Observed median sales prices and weekly rents over time.



*Note:* Panel (a) shows observed median sales prices and panel (b) observed median weekly rental prices per semester. The figures compare results obtained from the full, hybrid matches and exact matches samples.

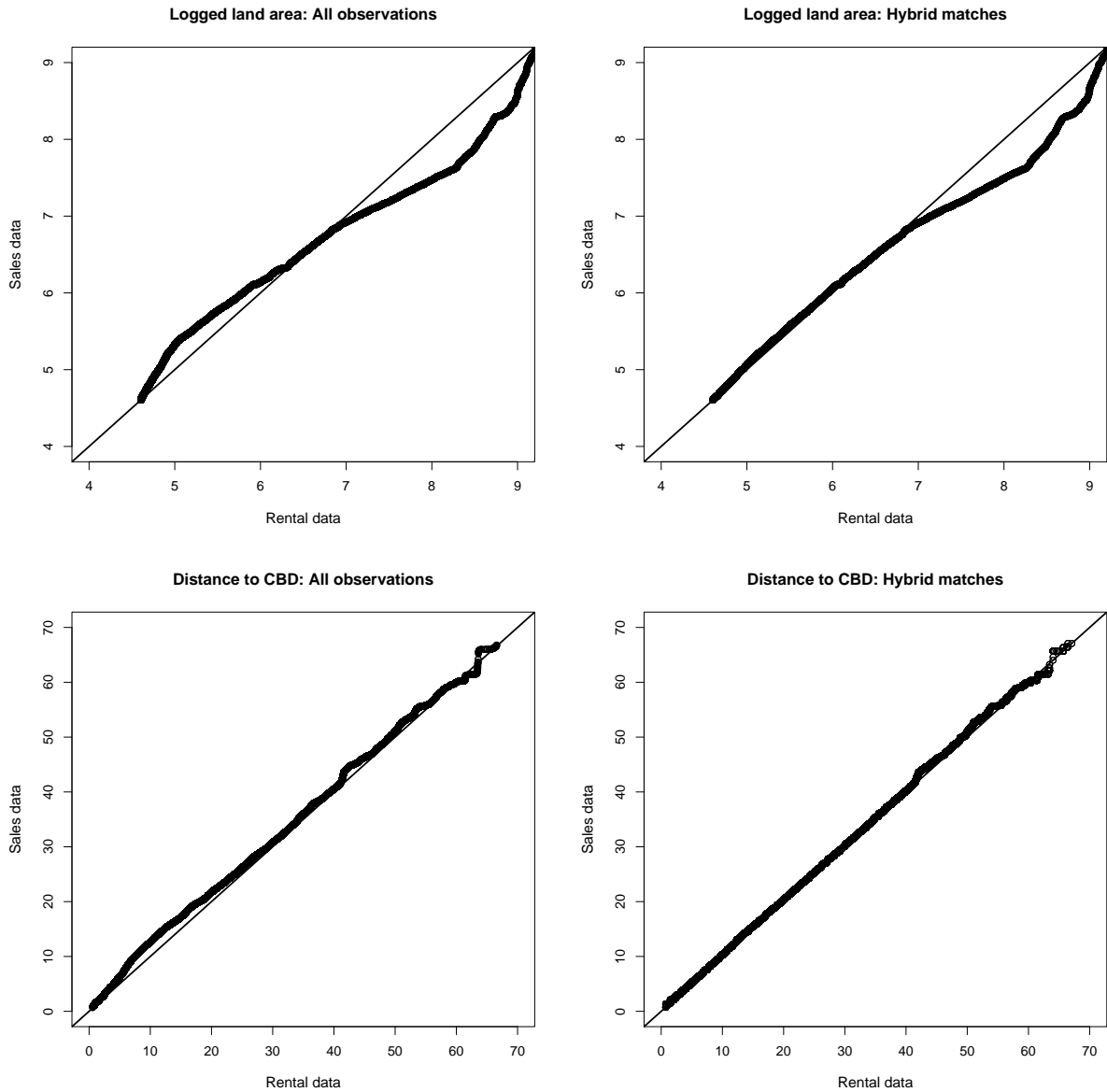
Figure 7 and Figure 8 provide more balance checks. As the exact matches sample is per construction perfectly balanced, results are presented for the full and hybrid matches sample only. Figure 7 compares the distributions of the covariate land area and as a proxy for location the distance to the central business district between rented and sold houses using empirical quantile-quantile plots. If there were perfect balance, it would appear as a straight 45 degrees line. In the full sample, there are systematic deviations between land area for rented and sold houses. The matching procedure eliminates deviations in the interval  $[0m^2; 1,000m^2]$  almost completely. As more than 90% of all observations (in each sample) have smaller land areas than  $1,000m^2$ , this is a great success in increasing balance.

In terms of location, the full sample is already quite well balanced, however balance is even improved in the hybrid matches sample. Quantile-quantile plots are not well suited for analyzing the integer variables number of bed- and number of bathrooms. Therefore, Figure 8 shows bar plots reporting the share of observations with a particular number of bed- or bathrooms. Matching is again successful in increasing balance.

There are consistently more sales observations than rental observations in each period. The matching approach presented here does *not* allow more than one rental observation to be matched to a sales observation, which restricts the maximum number of matches to the number of rental observations. Alternatively, one may relax the restriction to use each observation only once and perform matching *with replacement* (see Dehejia and Wahba, 1999; Ho et al., 2007). The main benefit of this approach is that more sales observations enter the analysis and the balance is increased significantly. However, the sales price and rent distribution is shifted quite substantially (i.e., deviations as reported in Table 4). The more conservative matching process *without replacement* is hence applied here.



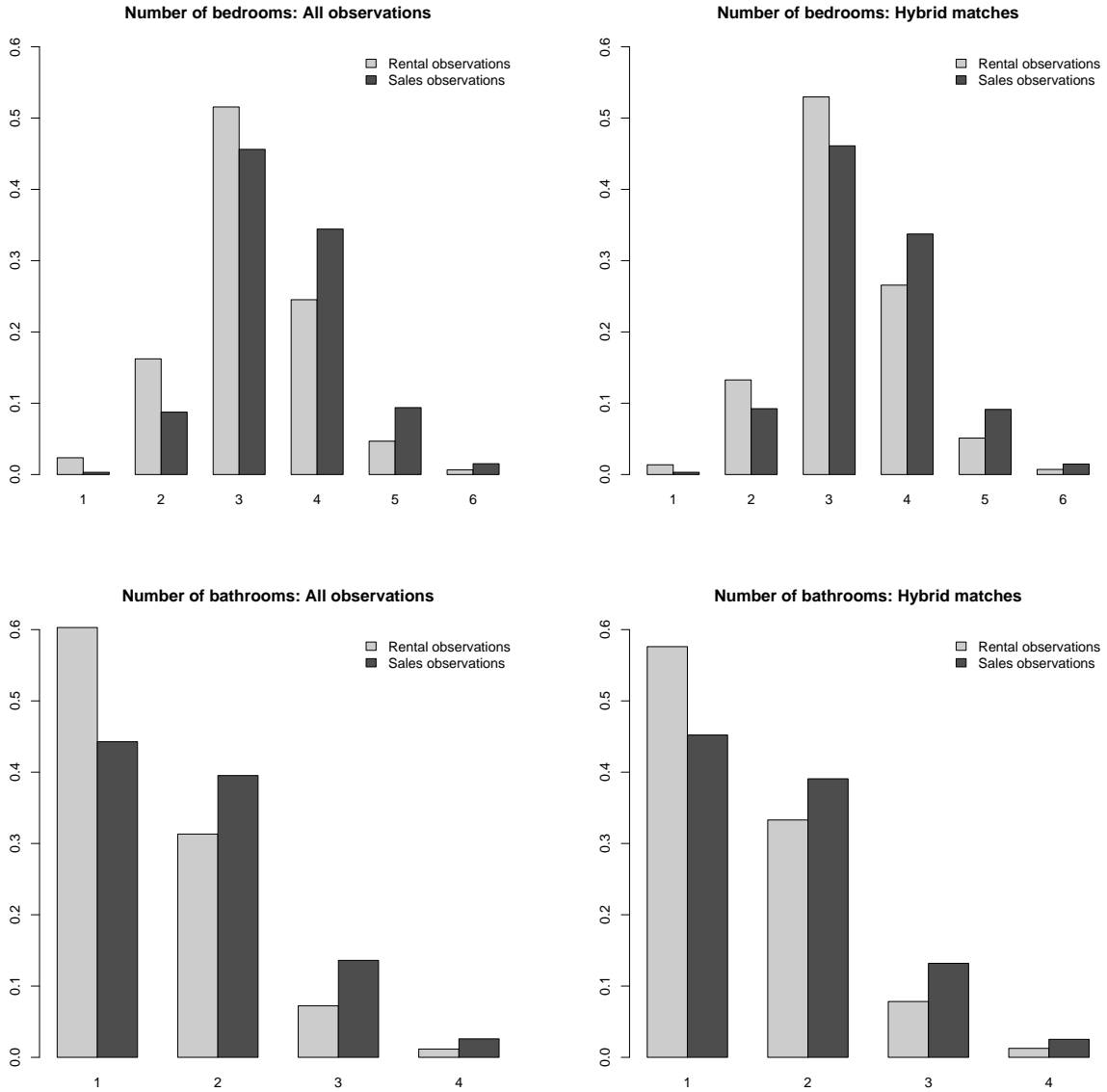
Figure 7: Balance checks for logged land area and distance to CBD.



*Note:* Empirical quantile-quantile plots comparing characteristics in the sales and rental data set. Left panels depict results using all observations and right panels those using the hybrid matches sample. Additionally, the 45 degrees line indicates the case of perfect balance.

Applying the marginal densities approach on matches obtained from matching with replacement generally yields very similar results. Rental yields are slightly higher which is a consequence of the fact that more large homes from the sales data set can be matched and thus enter the calculations.

Figure 8: Balance checks for the number of bed- and bathrooms.

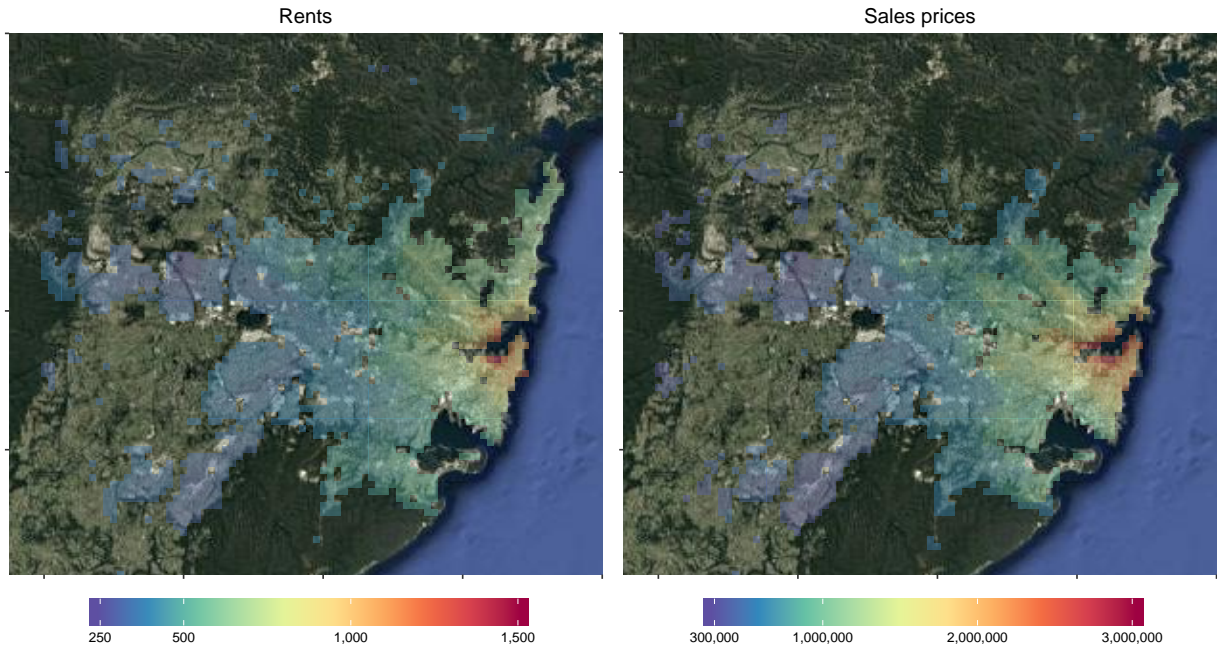


*Note:* The figures show the percentage of observations having a certain number of bed- or bathrooms in each respective sample. The figures on the left refer to all observations and the figures on the right to the hybrid matches sample.

## B Shadow prices

Figure 9 shows estimated geo-spatial splines from a median model in S2:2014. Prices and rents are highest in the inner-city and lowest in sub-urban regions. Table 5 reports average estimated shadow prices for the first, second (median) and third quartile. Estimates are very similar for models based on the full and the hybrid matches sample but differ greatly for models based on the exact matches sample.

Figure 9: Median imputed rental and sales prices across the city in S2:2014.



*Note:* The figure shows imputed median rental (left panel) and sales prices (right panel) in AUD. Prices are for S2:2014 and a two-bedrooms and one-bathroom house with a land area of  $500m^2$ . Results are projected on a map provided by Google, TerraMetrics 2016.

As expected, shadow prices differ strongly for rental and sales observations. Whereas additional bed- or bathrooms are valued higher for rented houses, land area is more important for sold houses. Shadow prices resulting from the exact matches sample also show this tendency, but differences are less pronounced. The very distinct shadow prices for these types of houses suggest, that the exact matches sample describes a sub-market that fundamentally differs from the overall market. Hence, relying on exact matches only, may lead to biased conclusions.

Parameters associated with land area are insignificant for the rental observations but highly significant (except for the exact matches sample) for sales observations. Besides that all shadow prices are highly significant in the models based on the hybrid matches and full samples. For exact matches, shadow prices associated with the number of bathrooms are generally also significant (at least at the 0.05 significance level) whereas those associated with the number of bedrooms are not.

## C Additional empirical results

This section provides more detailed empirical results. Table 6 reports period-wise MDM rental yields for the first (Q1), second (median) and third (Q3) quartile as well as average rental yields over the whole time span. Figure 10 shows the entire set of MDM rental yields as a surface plot spanned by quantiles and time. Table 7 reports more comprehensive results of the sample selection bias analysis and finally Figure 11 shows selected results from the

Table 5: Average estimated shadow prices over time for three quantile levels.

	Rental data			Sales data		
	Full	Hybrid	Exact	Full	Hybrid	Exact
<b>1st quartile</b> – $\vartheta = 0.25$						
<i>bed</i> = 2	0.419	0.401	0.365	0.161	0.161	0.183
<i>bed</i> = 3	0.596	0.577	0.512	0.263	0.262	0.270
<i>bed</i> = 4	0.714	0.697	0.630	0.344	0.345	0.352
<i>bed</i> = 5	0.799	0.782	0.715	0.387	0.387	0.381
<i>bed</i> = 6	0.817	0.796	0.751	0.375	0.374	0.406
<i>bath</i> = 2	0.132	0.130	0.137	0.091	0.094	0.106
<i>bath</i> = 3	0.299	0.296	0.304	0.213	0.217	0.260
<i>bath</i> $\geq$ 4	0.520	0.513	0.680	0.395	0.400	0.490
$\log(\textit{area})$	-0.012	-0.010	-0.010	0.166	0.171	0.149
<b>Median</b> – $\vartheta = 0.5$						
<i>bed</i> = 2	0.345	0.338	0.333	0.134	0.141	0.187
<i>bed</i> = 3	0.514	0.506	0.485	0.230	0.238	0.281
<i>bed</i> = 4	0.642	0.633	0.604	0.317	0.327	0.370
<i>bed</i> = 5	0.736	0.727	0.706	0.369	0.379	0.421
<i>bed</i> = 6	0.762	0.756	0.709	0.359	0.369	0.459
<i>bath</i> = 2	0.136	0.134	0.144	0.098	0.101	0.109
<i>bath</i> = 3	0.328	0.324	0.353	0.238	0.241	0.294
<i>bath</i> $\geq$ 4	0.612	0.608	0.817	0.472	0.474	0.631
$\log(\textit{area})$	0.001	0.002	0.016	0.213	0.215	0.182
<b>3rd quartile</b> – $\vartheta = 0.75$						
<i>bed</i> = 2	0.308	0.301	0.250	0.075	0.087	0.139
<i>bed</i> = 3	0.478	0.470	0.409	0.169	0.183	0.237
<i>bed</i> = 4	0.617	0.609	0.541	0.266	0.281	0.335
<i>bed</i> = 5	0.725	0.716	0.644	0.326	0.341	0.400
<i>bed</i> = 6	0.750	0.739	0.666	0.327	0.338	0.441
<i>bath</i> = 2	0.148	0.145	0.175	0.112	0.115	0.135
<i>bath</i> = 3	0.383	0.379	0.453	0.275	0.280	0.344
<i>bath</i> $\geq$ 4	0.728	0.716	0.881	0.561	0.563	0.670
$\log(\textit{area})$	0.011	0.015	0.038	0.248	0.251	0.221
Models include a geo-spatial spline $f(\textit{long}, \textit{lat})$ .						

*Note:* The table reports shadow prices associated with house characteristics averaged over all time periods for the first, second (median) and third quartile.

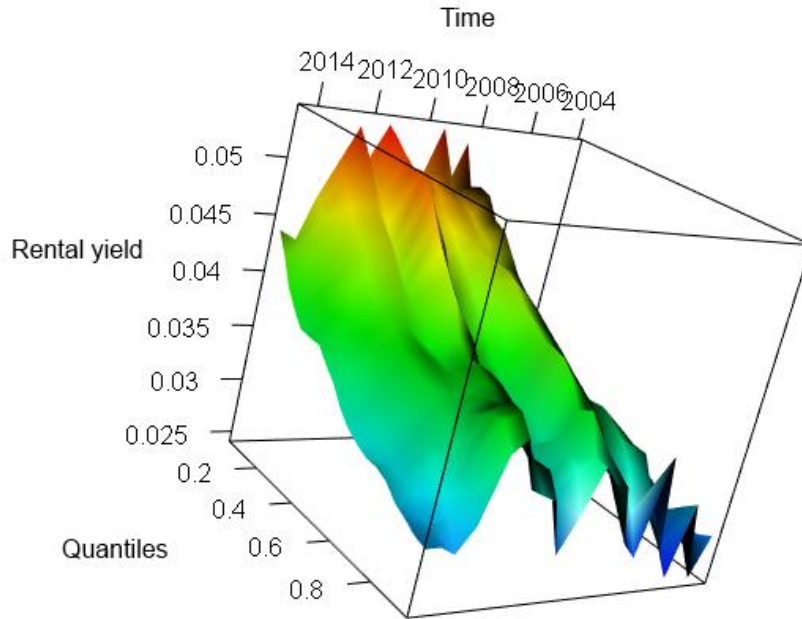
robustness check. More details regarding the supplementary results presented here are found in section 4.

Table 6: Quantile-specific rental yields over time.

Time	Q1	Median	Q3
S1:2004	3.012	2.901	2.777
S2:2004	3.339	3.094	2.934
S1:2005	3.462	3.264	3.242
S2:2005	3.347	3.272	3.169
S1:2006	3.675	3.494	3.280
S2:2006	3.743	3.331	3.264
S1:2007	4.059	3.776	3.251
S2:2007	4.124	3.533	3.092
S1:2008	4.517	4.076	3.646
S2:2008	4.899	4.450	3.799
S1:2009	4.617	4.331	3.920
S2:2009	4.598	4.024	3.602
S1:2010	4.463	3.878	3.553
S2:2010	4.635	3.735	3.450
S1:2011	4.768	4.158	3.797
S2:2011	4.894	4.164	3.758
S1:2012	4.746	4.234	3.972
S2:2012	4.609	4.118	3.915
S1:2013	4.588	4.045	3.624
S2:2013	4.171	3.571	3.509
S1:2014	4.143	3.581	3.184
S2:2014	3.787	3.356	3.041
Mean	4.191	3.745	3.445

*Note:* The table reports rental yields over time in % for the first (Q1), second (median) and third (Q3) quartile. Quality-adjustment is performed using the marginal densities approach. The bottom row reports mean ratios over all time periods.

Figure 10: Rental yield surface spanned by time and quantiles.



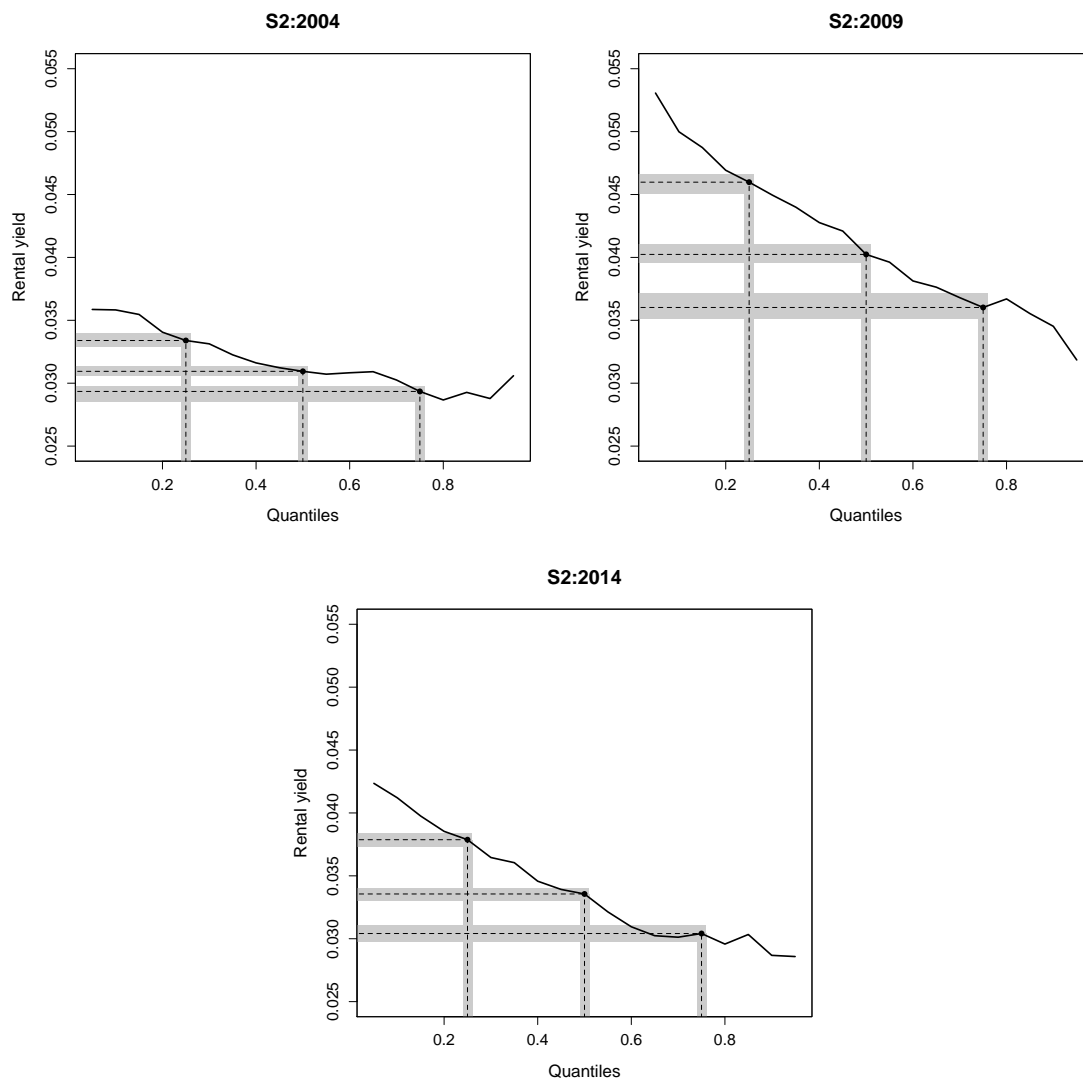
*Note:* The figure shows the full set of MDM rental yields results as a surface spanned by time and quantiles.

Table 7: Estimated sample selection bias.

	MDM methodology			Dummy approach		
	Q1	Median	Q3	Q1	Median	Q3
Average bias	8.34%	8.14%	8.89%	5.60%	5.75%	5.70%
Maximum bias	14.94%	17.12%	19.36%	9.39%	8.02%	8.43%
No. of negative biases	2	1	0			
No. of insignificant $\hat{\delta}_t^s$				15	11	12
No. of insignificant $\hat{\delta}_t^r$				0	1	0
No. of positive $\hat{\delta}_t^s$				2	1	0
No. of negative $\hat{\delta}_t^r$				0	0	0

*Note:* The table summarizes estimated sample selection biases for the median as well as the first and third quartile. *MDM methodology* and *dummy approach* refer to the two types of methods used to estimate a sample selection bias as described in subsection 4.5. *No. of negative biases* gives the number of periods where negative biases were estimated. *No. of insignificant  $\hat{\delta}_t^{s,r}$*  reports the number of insignificant dummy variables indicating exact matches and *No. of negative  $\hat{\delta}_t^s$  (positive  $\hat{\delta}_t^r$ )* the number of periods where  $\hat{\delta}_t^s$  ( $\hat{\delta}_t^r$ ) was positive (negative). In total, there are 22 periods.

Figure 11: Robustness check.



*Note:* The figure shows selected results from the robustness check. Each panel shows MDM rental yields as a function of the quantile level for a different time period as indicated in the panel caption. The horizontal parts of the gray shaded areas show the variation of rental yields obtained when increasing / decreasing the quantile level at which either the quality-adjusted rental or sales price distribution is evaluated. The quantile levels vary within  $[\vartheta - 1\%, \vartheta + 1\%]$  as indicated by the vertical parts of the gray shaded areas. Results are depicted for  $\vartheta \in \{0.25, 0.5, 0.75\}$ .