Residential Rental Yields in Melbourne

Andy Krause and Gideon Aschwanden

April 18, 2016

Introduction

Rental yields (rent-to-price ratios) are an important metric in judging both the viability of a housing investment as well as the overall health of the market. In this study, we use the property level rental yield calculations made in an earlier work[1] to analyse the variation in the Melbourne market. This study is focused on market transactions that occurred in the year 2015.

Data Preparation

Our initial study produced yields for the period 2011 to 2015. Here we isolate only those transactions and resulting calculated yields during the 2015 calendar year (~120,000 observations). As property data often contains errors, we first check for outlying observations. Based on natural breaks in the data, we then remove observations with yields below 1% or over 8%. We also eliminate all observations from suburbs which do not have at least 3 observations per quarter as these represent sparsely populated areas. Finally, a number of properties show very large number of bedrooms. We remove all units with more than 4 bedrooms and houses with more than 6.

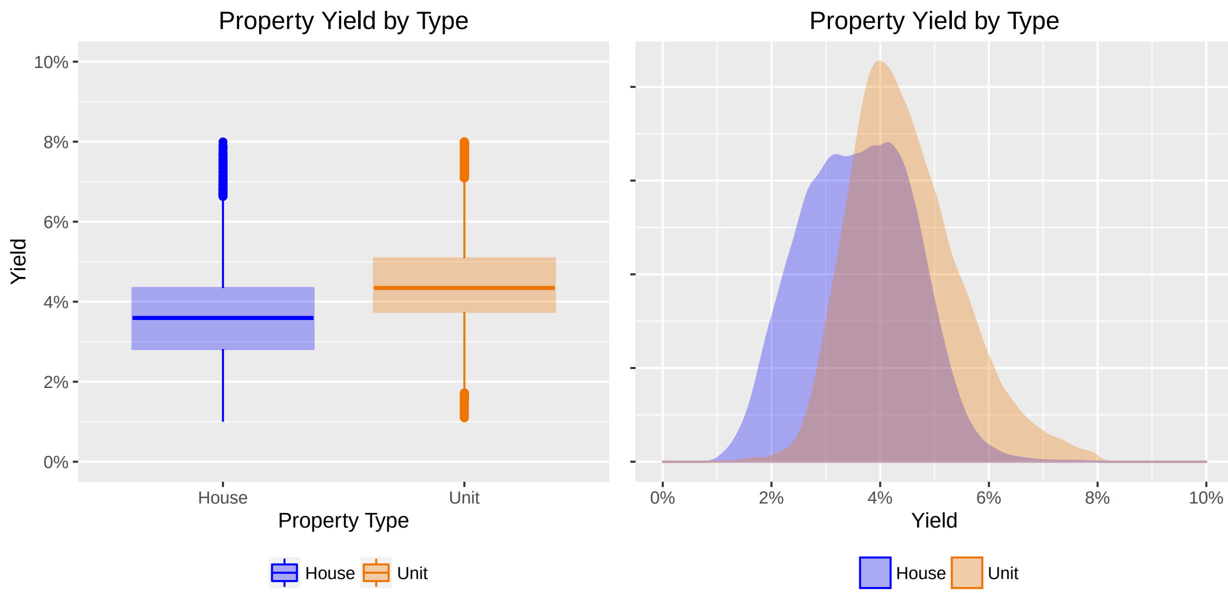
Next, we merge data on property-specific distances to train stations, tram and bus stops. At this stage we also remove any observation of a property not labeled as a 'Unit' or 'House' and then split the data into separate unit and house datasets.

Data Analysis

Houses vs Units

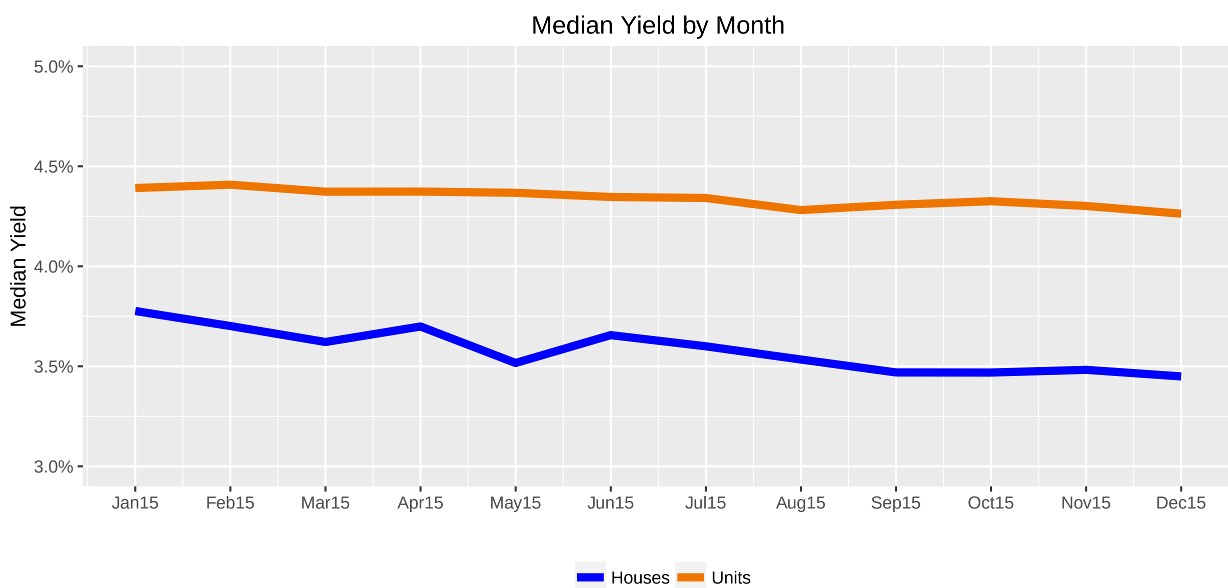
We will start by looking into the difference in yields between houses and units. On the left, box plots show that unit yields are generally higher than those of houses, and similarly dispersed (at least in terms of the interquartile ranges). The density plots on the right present a like story, with rental yields being higher and a bit more tightly distributed than the lower house yields. As result of these differences and the differences in general market dynamics between houses and units we will analyze the two separately throughout the remainder of this analysis.

[1] See Krause, A. and Aschwanden, G. 'Deriving a rent-to-price ratio in Residential Markets', Working Paper, 2016. Available at http://www.andykrause.com/current-research



Over Time

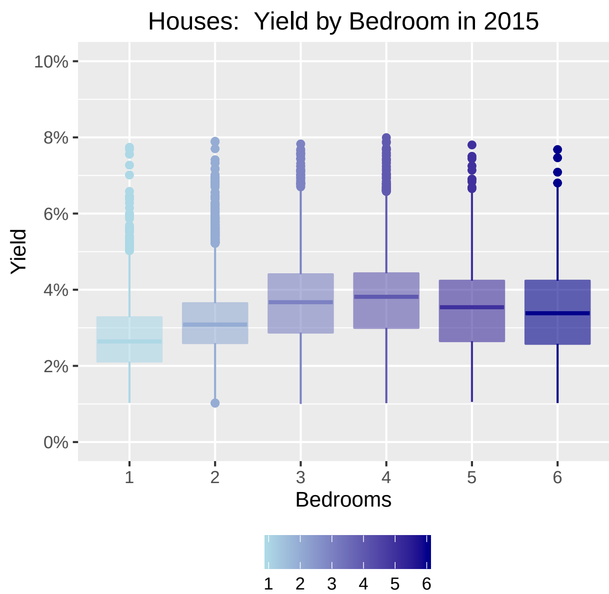
Next, we will check to see if yields have changed over 2015. From the plot below we see that unit yields (orange) have varied slightly over the year analyzed here, whereas the house yields show a more sustained decrease (with some variation between). For the time being, we'll ignore the impacts of time and return to this aspect in the multivariate analysis later on.

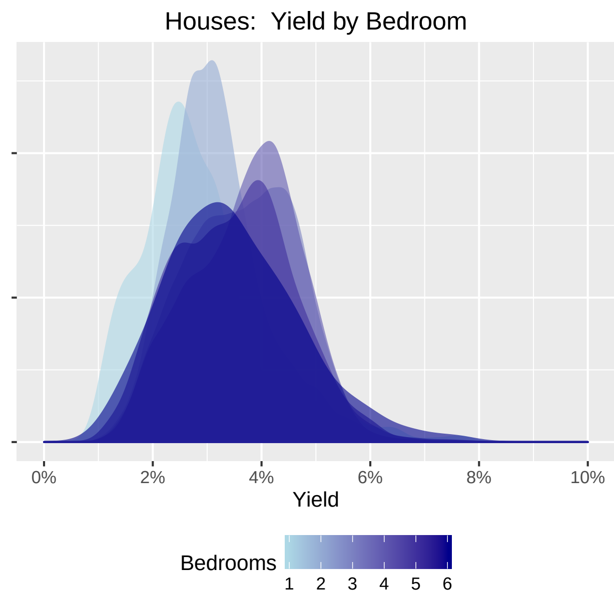


By Bedrooms

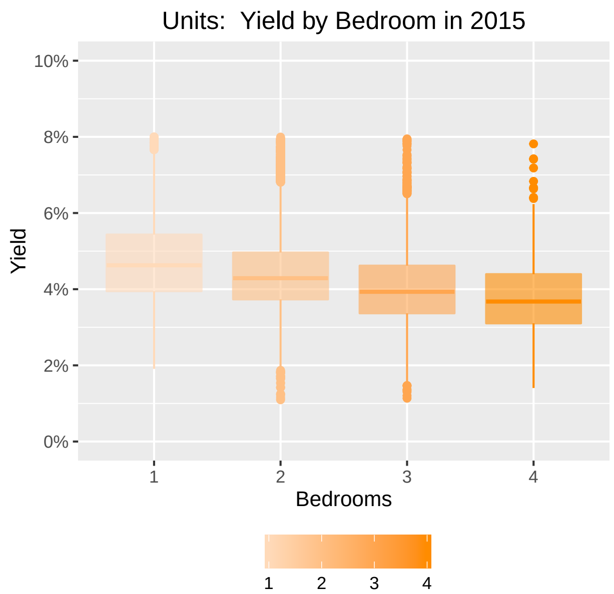
It is likely that the characteristics of the housing units have some impact on the yields. Bedrooms are a particularly likely candidate to have an impact given that most homes are advertised by their bedroom count and bedroom count often determines the number of people that can feasibily occupy a structure. To examine this potential impact, we look at the median rental yield by bedroom count for both houses and units during this periods; beginning with houses.

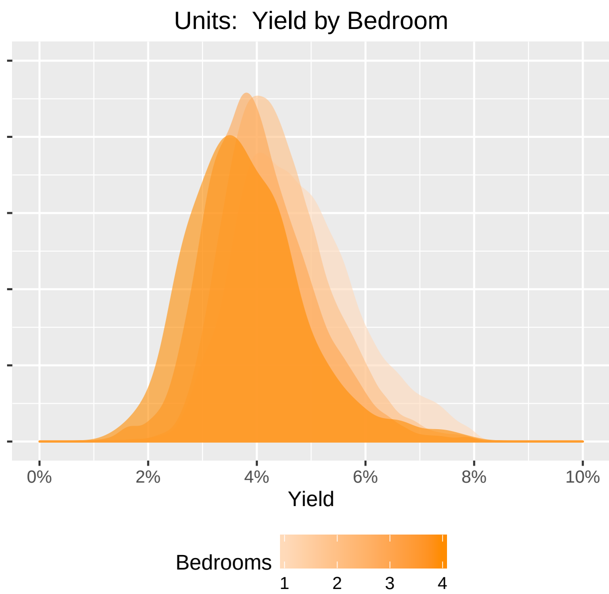
Houses As the box plot on the left shows, 3 and 4 bedroom homes produce the best yields, with the lowest yields for 1 bedroom houses. Note that the overall sample size of 1 and 6 bedrooms homes are small (these are rare) and therefore the median values of each should be interpreted cautiously. The combined density plot (right) and the separated one (below) highlights the differences between the full distribution of yields for each bedroom category. Here we see the consistency of yields in the 2 through 4 bedroom categories with much more variation from the others.

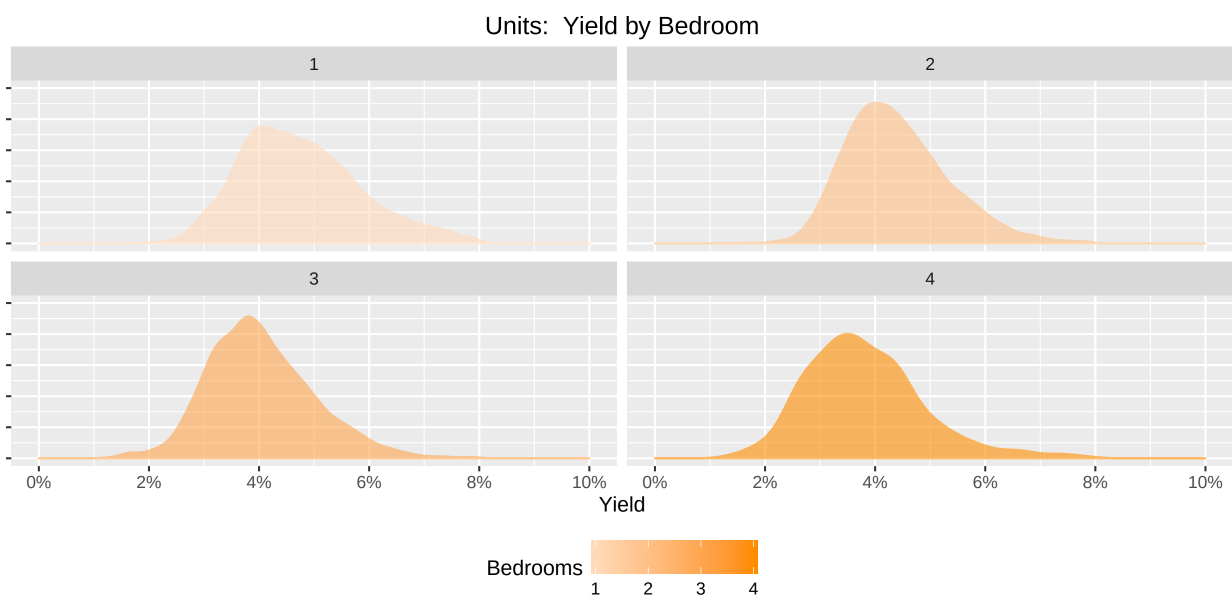




Units For units we see a different trend. Smaller units (1 and 2 bedroom) generate the highest yields, with yield increasing steadily with the addition of extra bedrooms. Again, the distribution of yields is most consistent in 2 and 3 bedroom units.







Multivariate Approach

It is evident from the above that bedrooms have an impact on yield (both for houses and units) and that time may also have an impact, especially with houses. It is also likely that other factors such as bathrooms, presence of a pool, lot size, etc. may impac the results as well. To analyze the impact that a number of structural and temporal factors have on yields we regress yields on a host of variables via a multivariate regression analysis.

Houses Running a multivariate regression model of yields on a collection of structural and time variables shows us a number of things. First, the model of housing yields only explains 5.31 percent of the variation in housing yields. In other words, structural attributes and time are not the key determinants of yields. Note that this dataset does not contain information on property condition, quality, age, views or water frontage;

characteristics that may influence yields. Overall, this suggests that either: A) Location is very important in determining house yields; B) The omitted spatial variables noted above are important in determining yields; C) Some part of A and B; or D) Rental yields are somewhat random. We will explore these considerations later in this analysis.

Looking more closely at the model results, we see that the number of bathrooms, presence of a pool and presence of a garage have a positive relationship with yields, whereas larger lot size has a negative one. Yields have gone down steadily over time as indicated by the coefficients on the month variables. The coefficients on the bedrooms variables shows a similar pattern (1 bedroom homes are the holdout) with 3 and 4 bedrooms homes showing the best yields.

	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	0.0291	0.0004	74.7222	0.0000
Bedrooms - 2	0.0033	0.0004	8.7987	0.0000
Bedrooms - 3	0.0071	0.0004	19.2716	0.0000
Bedrooms - 4	0.0071	0.0004	18.8932	0.0000
Bedrooms - 5	0.0044	0.0004	10.3396	0.0000
Bedrooms - 6	0.0038	0.0007	5.6626	0.0000
Nbr of Baths	0.0012	0.0001	16.4520	0.0000
Plot Size (100m)	-0.0000	0.0000	-10.0715	0.0000
Has A Pool	-0.0009	0.0003	-2.8253	0.0047
Has a Garage	0.0024	0.0001	23.2291	0.0000
Feb	-0.0009	0.0002	-5.4729	0.0000
Mar	-0.0015	0.0002	-9.7207	0.0000
Apr	-0.0012	0.0002	-7.0668	0.0000
May	-0.0025	0.0002	-16.0187	0.0000
Jun	-0.0015	0.0002	-7.6428	0.0000
Jul	-0.0015	0.0002	-8.6564	0.0000
Aug	-0.0019	0.0002	-10.0503	0.0000
Sep	-0.0023	0.0002	-12.8868	0.0000
Oct	-0.0026	0.0002	-14.0088	0.0000
Nov	-0.0025	0.0002	-13.1642	0.0000
Dec	-0.0027	0.0002	-15.1879	0.0000

Units Much like houses, structural and temporal variables do not explain much of the variation in unit yields; 7.94 % in this case. Again this suggest that either space, omitted variables, some combination of the two explain yield variation, if explainable at all.

In terms of individual coefficients, we see that pools have a positive correlation with yields, whereas garages and # of baths do not. These may appear to be somewhat counterintuitive, but remember that we are measure yields, not rents or prices themselves. In this case, the negative value of garages on yields might signify that owner-occupiers value garage space more than renters (who might not own a vehicle). On the other hand, these negatively signs could be the result of omitted variable biases (newer structures may have less garage space, therefore the lack of a garage is a proxy for newer, better condition units.) Yields have varied a bit over time for units, but with a slow decrease. This trend was indicated above in the univariate analysis as well. Finally, for units, each extra bedroom decreases yields; another finding echoing the previous univariate analysis.

	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	0.0485	0.0002	251.0658	0.0000
Bedrooms - 2	-0.0033	0.0001	-33.4184	0.0000
Bedrooms - 3	-0.0065	0.0002	-37.8251	0.0000
Bedrooms - 4	-0.0082	0.0006	-14.3982	0.0000
Nbr of Baths	-0.0007	0.0001	-5.3361	0.0000
Has A Pool	0.0088	0.0002	35.9891	0.0000
Has a Garage	-0.0023	0.0004	-6.2643	0.0000
Feb	0.0001	0.0002	0.2982	0.7655
Mar	-0.0004	0.0002	-1.8528	0.0639
Apr	-0.0003	0.0002	-1.5553	0.1199
May	-0.0006	0.0002	-3.1424	0.0017
Jun	-0.0007	0.0002	-2.9665	0.0030
Jul	-0.0007	0.0002	-3.5261	0.0004
Aug	-0.0012	0.0002	-5.4623	0.0000
Sep	-0.0013	0.0002	-6.3172	0.0000
Oct	-0.0008	0.0002	-3.6434	0.0003
Nov	-0.0009	0.0002	-4.3688	0.0000
Dec	-0.0013	0.0002	-6.9325	0.0000

Accounting for Space

As we saw above, leaving out spatial factors produced models with very low explanatory power. Location is obviously an important factor in determining rents as well as prices. Here we test if the relationship between those two (the rental yield) also has definitive spatial patterns.

We begin by following a spatially aggregated approach to defining location, that of differentiating yields by suburb. This is the most common spatial scale at which yields are analyzed in industry practice, likely due to the simplicity of doing so. The next question to ask is how much adding suburbs helps explain variation in yields. We re-estimate the multivariate equations from before, adding suburb as fixed effects (indicator or dummy variables).

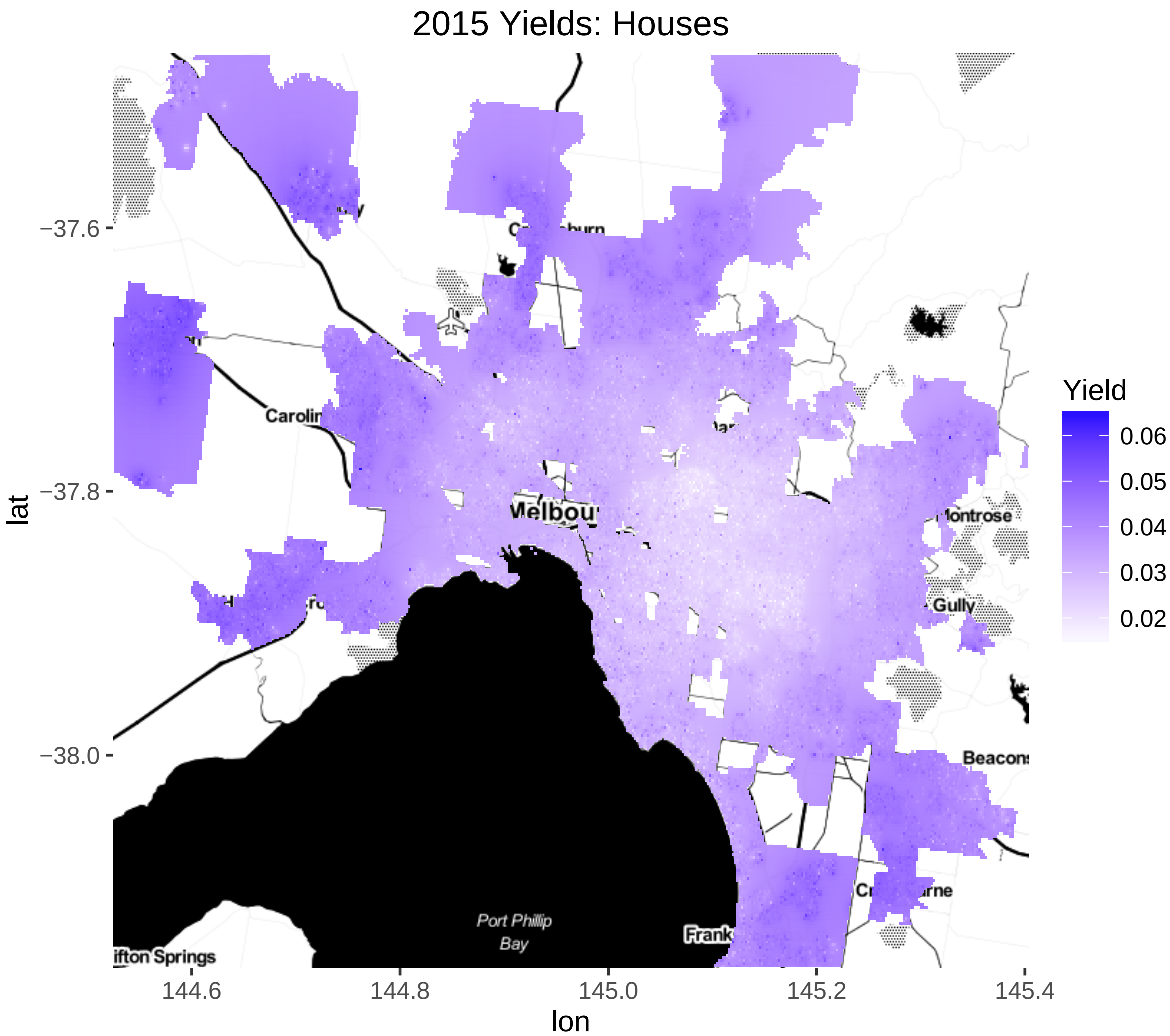
By doing so, the model of house yields now explains 66% of the variation in yields, a market improvement over the aspatial model estimated earlier. Slightly smaller gains are found when adding suburbs to the model specification for units; an increase from 7.94% to 43.63% of variation explained.

The question thus remains, is the remaining variation due to omitted variables and randomness, or are more microspatial factors (factors operating at a finer scale than suburbs) influencing yields. To begin to understand this we first look at the residuals from the regression models including the suburb fixed effects. If the residuals from these models exhibit high spatial autocorrelation, this suggests that some spatial impacts remaining unaccounted for. To test for this we estimate the Moran's I of each set of residuals; houses and units. We use a spatial weights matrix of the 10 nearest neighbors, distance weighted.

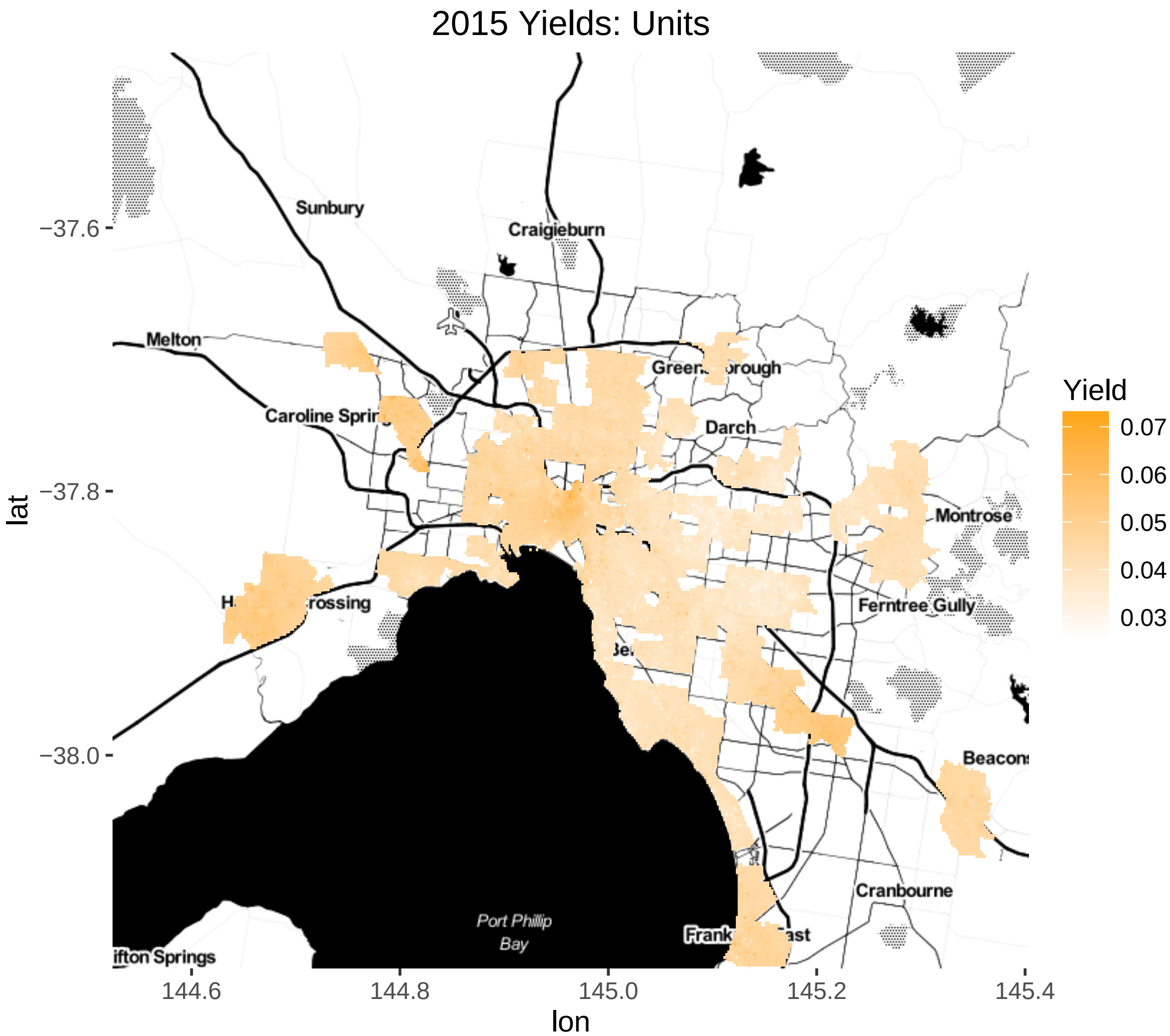
The Moran's I tests show highly statistically significant levels of spatial autocorrelation for both houses and units. This suggest that there is at least some remaining variation in yields that can be explained by additional spatial considerations. The next step then involves determining which type of spatial dependence exists, lagged or in the error terms.

Running LaGrange Multiplier tests shows that spatial errors are the dominant for of dependence in each case. What this means is that omitted spatial variables (unaccounted for local factors) are likely to blame. To attempt to explain these micro-spatial impacts, we use a two-step approach: 1) Map yields at sub-suburb level; and 2) test the influence of some local accessibility factors.

The map of house yields in 2015 show a range of about 2% to 6% over the Melbourne metropolitan region. House yields are lowest in the near eastern suburb and highest in a many of the outer suburbs.

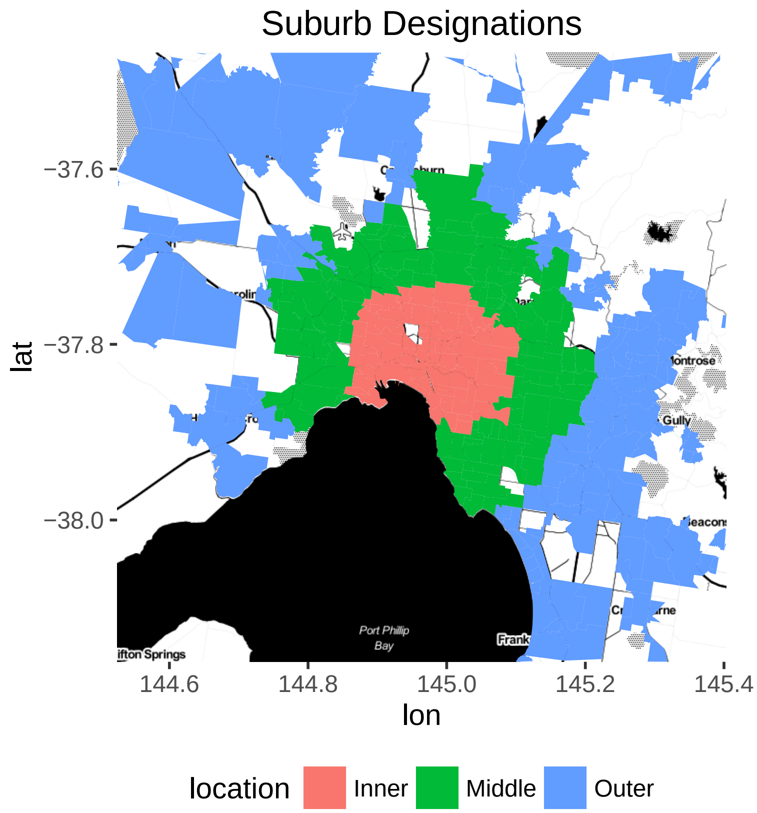


Unit yields look somewhat similar, however, there are many less areas that have enough transactions from which to develop yield estimates. The notable exception here is that Carlton and the other areas around the University of Melbourne and RMIT have high yields.

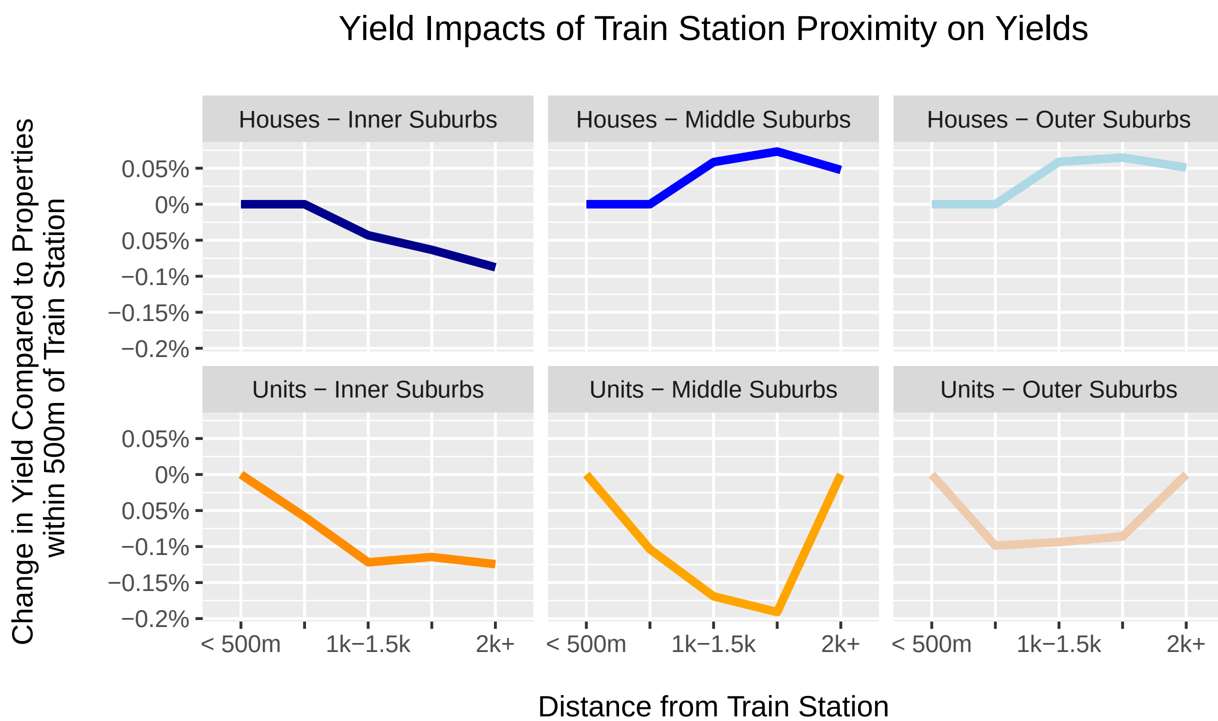


Accessibility Variables Accessibility to transit can be a key driver of rents and prices, and therefore may impact rental yields. To test this, we look for any impacts on yields due to distance from train stations[2]. In this case, we treat distance from trains as a set of distance bands (concentric rings) at distances of 500m, 1000m, 1500m, and 2000m. From this analysis, we see that train station access matters in inner suburb houses, and in inner and middle suburb units.

[2]Impacts from tram stops and other forms of transit will be analysed in future research



Below we see the differences in yields as we move outwards from a train station. At the bottom we see that units in the inner suburb lose the most yield by being far from a train station. Houses in the inner suburbs and units in the middle suburbs also lose yield with distance. Interestingly, houses in the middle and outer suburbs gain yield with distance from train stations, suggesting that areas surrounding suburban train stations are less desirable areas (or that most people drive to work anyways so the train station has little appeal or utility).



Influence of Universities Finally, we look at the impact of universities on yields. (during 2015). The panels show the yield premium or discount due to being located less than 1km or 1km to 2km from university (compared to being located 2km to 5km from university.) From this wee see that units, near universities show a moderatly positive impact, wheres negative impacts are found for houses. University of Melbourne and RMIT (analyzed together due to their proximity) show the largest premiums for unit location.