Pre-Analysis
10/03/19

**Specifications to Run**

- ~~Linear~~
- Semi-log (Taylor 2003, Ottensman et al 2008, Netusil 2005)
- Log (Mahan et al 2000) (is this logged ind and non-logged dep)
- Double-log (Netusil 2005)
- Quadratic
- Linear Box-Cox
    - Cropper et al (1988) showed that when all housing attributes are observed without error, more complex functional forms (quadratic, quadratic box-cox) can be used to estimate marginal prices; produce biased estimates when some vars are proxied or missing
    - Linear box-cox provides accurate prices when all attributes measured correctly and accommodates some misspecification of the hedonic functional form (Cropper 1988)
    - Note: absolute value of t-statistics estimated from Box-Cox can be biased upward (Blackley et al, 1984)
    - STEPS:
        - Assume model of the form $y^{(\lambda)} = \beta_1 Z_1{}^{(\theta_1)} + \beta_2 Z_2{}^{(\theta_2)} + \cdots + \beta_n Z_n{}^{(\theta_n)}$ where each $X_i = Z_i{}^{(\theta_i)}$ is an independent variable.
        - Let lambda be flexible (typically ranges between -2 and 2)
        - Iterate over several OLS models, varying lambda until mean squared error is minimized
        - Recover MWTP by transforming the estimated $\beta_i$ s (Blackley et al, 1984)
- Quadratic Box-Cox

**Dependent Variable**

- Sales price (of residences in Portland from assessor data)
    - Arms length transactions: “types of sales assert that both parties act in their own self-interest and are not subject to pressure from the other party; furthermore, it assures others that there is no collusion between the buyer and seller”
    - Adjust sales price using price index for multnomah residential housing market[1]
        - from June 2016 to July 2019? What is most recent month/year available?
        - Base year?

**Independent Variables: see table below**

**Independent Variables -- Questions**

[1] Mahan et al. (2000)

Pre-Analysis
10/03/19

- Structural
    - Lot size greater than 1 acre or other specification? Use dummy or continuous measure?
    - Square footage of structure and lot (possibly non-linear, could add quadratic term)
    - “Capacity” variable: interaction or ratio of structure to lot sq footage?
- Neighborhood
    - Dummy for spatial fixed effects in neighborhood -- is this allowed? If so, should we take out most other neighborhood independent vars? How coarse/fine should spatial variables be?
        - Census tracts
        - Neighborhood association (e.g. Eastmoreland)
        - Quadrant
    - Which distance measure is best?
        - Straight-line distance
        - Travel time
        - Distance by major highway
    - Distance/Travel time to CBD
        - Free-flow
        - Congested
    - Distance to airport— consider as employment center OR residential amenity/disamenity
    - Distance to major highway
    - Significant Cultural Resources: this adds value to the house, but increases costs of redevelopment. Should we consider redevelopment costs or focus on marginal cost? Is this being incorporated somewhere else in the proforma? Reverse problem for hazardous constraints
- Environmental
    - Open space zone
        - Interaction of open space with diversity of uses within 1 km radius, measured with buffer zones (Acharya et al, 2001)
        - Use a few different radii (200ft, ¼ mi, ½ mi, etc.)
        - How should we set up our buffer variables if we choose to include spatial fixed effects?
    - What distance away should we do for these? E.g. 200ft, ¼ mi, ½ mi

| Type | Name | Expected | Source(s) |
|---|---|---|---|

Pre-Analysis
10/03/19

| | | Sign | |
|---|---|---|---|
| Structural | No. Bathrooms | positive | Ottensman et al. (2008) |
| Structural | Fireplace Dummy | | |
| Structural | Age of property | negative | Netusil (2005), Donovan |
| Structural | Structure size (sqft or sqft + sqft^2 or log sqft or $sqft^2$) | | |
| Structural | Lot Size (sqft) | | |
| Structural | Number of stories | | |
| Structural | Garage square footage | | |
| Structural | “Capacity” - ratio of structure to lot size (interaction) - base zoning capacity | | |
| Structural | Oversized Lot Dummy | positive | Ottensman et al. (2008) |
| Structural | Basement Dummy (crawlspaces, basement+crawlspace, etc) | positive | Ottensman et al. (2008) |
| Structural | Attic | | |
| Structural | Elevation (ft) | positive | Johnston et al. (2019) |
| Neighborhood | % Vacant Lots/Houses Nearby | negative | Ottensman et al. (2008) |
| Neighborhood | District SAT (school performance score) | positive | Ottensman et al. (2008) |
| Neighborhood | Per student expenditures | positive | Turner et al. (2014) |
| Neighborhood | Distance/Travel to Central Business District (free flow) | positive | Ottensman et al. (2008), Mahan et al (2000) |
| Neighborhood | School boundary zones | | |
| Neighborhood | Distance/Travel to CBD (congested) | positive | Ottensman et al. (2008), Mahan et al (2000) |
| Neighborhood | Distance/Travel to multi employ centers | | Ottensman et al. (2008) |
| Neighborhood | Walkscore | | Ottensman et al. (2008) |
| Neighborhood | Population density per census tract [proxy for congestion] | negative | Irwin et al. (2001) |
| Neighborhood | Distance to Airport | mixed | Ottensman et al. (2008) |

Pre-Analysis
10/03/19

| | | | |
|---|---|---|---|
| Neighborhood | Market Segment Dummy (quadrant) | mixed | Mahan et al (2000) |
| Neighborhood | Zoning designation | | |
| Neighborhood | Racial composition of neighborhood | | |
| Neighborhood | Non-overlapping neighborhood associations (spatial fixed effects) | IMPORTANT!!!!!! | Kuminoff (forthcoming) |
| Neighborhood | Average education level | | |
| Neighborhood | Median neighborhood income | positive | Turner et al (2014) |
| Environmental | Distance to Wetland (Lake vs stream) | negative | Mahan et al (2000) |
| Environmental | Size of Wetland (lake vs stream) | positive | Mahan et al (2000) |
| Environmental | Open space | | |
| Environmental | Slope + Stream | | Netusil (2005) |
| Environmental | Proximity to trees | positive | Netusil (2005), Donovan |
| Environmental | Proximity to stream | positive | Netusil (2005), Donovan |
| Environmental | P-zone + quadrant | mixed | Netusil (2005) |
| Environmental | C-zone + quadrant | mixed | Netusil (2005) |
| Environmental | Tree canopy coverage on lot | | Donovan (2010) |
| Environmental | Proximity to Willamette (bridge effect) | | |
| Structural | Area-perimeter ratio/number of vertices (capture odd shapes?) | | |
| Neighborhood | High school catchment area (but there's some flexibility--how many students go outside their catchment area?) | | |
| Structural | Property taxes (assessment in year house sold) paid in year home transacted & across time | | |
| Neighborhood | Traffic (layer from PBOT) | | |
| Neighborhood | Street classification (major arteries, biking only, etc) | | |