

This document describes the replication files for “**Initial Property Offering: Underpricing and Learning Behavior in the Presale Housing Market**”, by Changwei Zhan.

Structure:

Each Stata Do file contains codes that are necessary to reproduce a table or a figure in the paper.

Variable descriptions:

The “condotxn.dta” file is the sample of property transactions to reproduce results. The descriptions of variables are shown below:

- **areasqm**: Flat area in square meters.
- **unitpricesm**: Unit price per square meter.
- **postalcode**: The 6-digit postal code.
- **saleyear**: Year in which the property was sold.
- **ym**: Year and month in which the property was sold.
- **lnunitpricesm**: Natural logarithm of the unit price per square meter.
- **level**: Floor level.
- **storey**: Storey range (1 for level ≤ 10 ; 2 for $10 < \text{level} \leq 20$; and 3 for level > 20).
- **stack**: Identifier for the stack in the building.
- **purchasertype**: Type of purchaser.
- **projectid**: Unique identifier for the project.
- **dist_mrtstn**: Distance to the nearest MRT (Mass Rapid Transit) station.
- **dist_busstop**: Distance to the nearest bus stop.
- **dist_psh30**: Distance to the nearest top 30 primary school.
- **dist_cityhall**: Distance to CBD (the City Hall MRT station).
- **lnareasqm**: Natural logarithm of the area in square meters.
- **lnindist_mrtstn**: Natural logarithm of the distance to the nearest MRT station.
- **lnindist_busstop**: Natural logarithm of the distance to the nearest bus stop.
- **lnindist_psh30**: Natural logarithm of the distance to the nearest top 30 primary school.
- **lnindist_cityhall**: Natural logarithm of the distance to CBD.
- **firstsaledate**: Date of the first sale of the project.
- **daysincelaunch**: Number of days since the property launch.
- **developer_amended**: Identifier for the amended developer name (I use the common words in companies' names to manually classify them as the same developer).
- **registdate**: Registration date of the developer.
- **daysincelaunch_100**: Days since launch normalized to a scale of 100.
- **salepercent**: Percentage of the property sold.
- **topdeveloper**: Indicator for whether the developer is classified as a top developer.
- **publicdeveloper**: Indicator for whether the developer is a public developer.
- **oldfirm**: Indicator for whether the developer is a mature firm (registration date earlier than the average registration date in the sample).
- **sora**: The Singapore Overnight Rate Average (SORA) value.
- **projectsequence_2km**: Sequence of the project within a 2 km radius.