



# Rent Imputation for Welfare Aggregates: A Methodological Framework

Cesar Cancho and Joao Pedro Azevedo  
Poverty and Equity Global Practice, The World Bank

# Summary

- Econometric/Statistical methods for rent imputation do not consider biases in collected data
- Proposed decision tree to identify sources of biases
  - Segmentation in dwelling characteristics between owners and renters
  - Systematic differences in responses between owners and renters
- Implementation shows how to:
  - Detect existence of differences in observable characteristics for renters and home-owners.
  - Identify common support in the characteristics of both groups

# Outline

- Why to impute rent?
- How to impute rent?
- What do we propose?
- How to implement it?
- Conclusions

# Why to impute rent?

# Key component of the consumption aggregate

- Utility derived from occupying dwelling
- Housing can represent a considerable share of expenses
  - OECD 2012 estimates: 14 – 25 percent of households adjustable disposable income
  - Average share increases with aggregate welfare
- Inclusion in welfare aggregate may impact
  - Ranking of households
  - Value of poverty line (CBN method)
  - Composition of poor
- Homeownership is prevalent in most regions
  - Europe: 39 percent (Switzerland) to 90 percent (Romania)
  - Latin America: 44 percent (Colombia) to 80 percent (Nicaragua)

# How to impute rent?

# Imputed rent: easier said than done

- Rent is ideal measure
- Implicit rental value must be treated with caution
  - Rents known to them may be subsidized or out of date
  - Housing and rental markets may not well enough developed to permit estimates
  - Best available solution may simply be to exclude the housing component.
- Balcazar et al (2014)
  - Comprehensive review of methodologies to impute rent

# Methods to impute rent

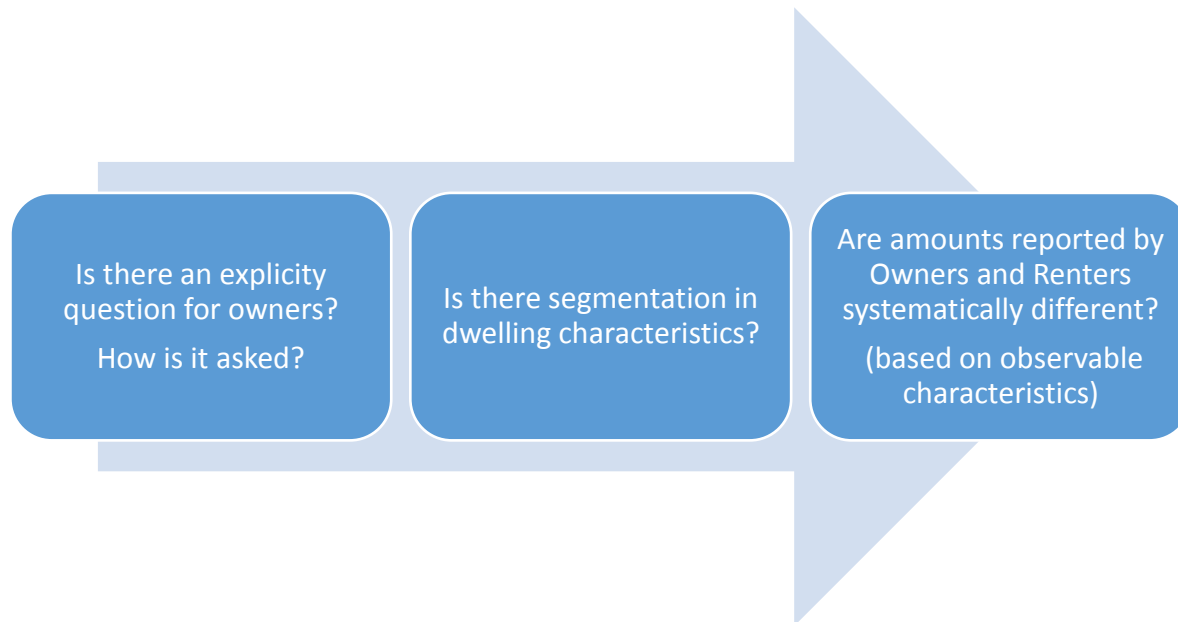
- Econometric/Statistical Approach
  - Linear, log-linear, higher-order, semi-parametric, non-parametric, quantile, spatial models, heckman selection, stratification (imputations cells)
- Capital Market Approach
  - Rent-to-value
    - Requires estimations of capitalization rate based on dwelling value or annual rent
  - User-cost approach
    - Requires information on dwelling value
- Payment Approach
- Self-reported
- External sources
  - Administrative records, newspapers, mortgage transactions



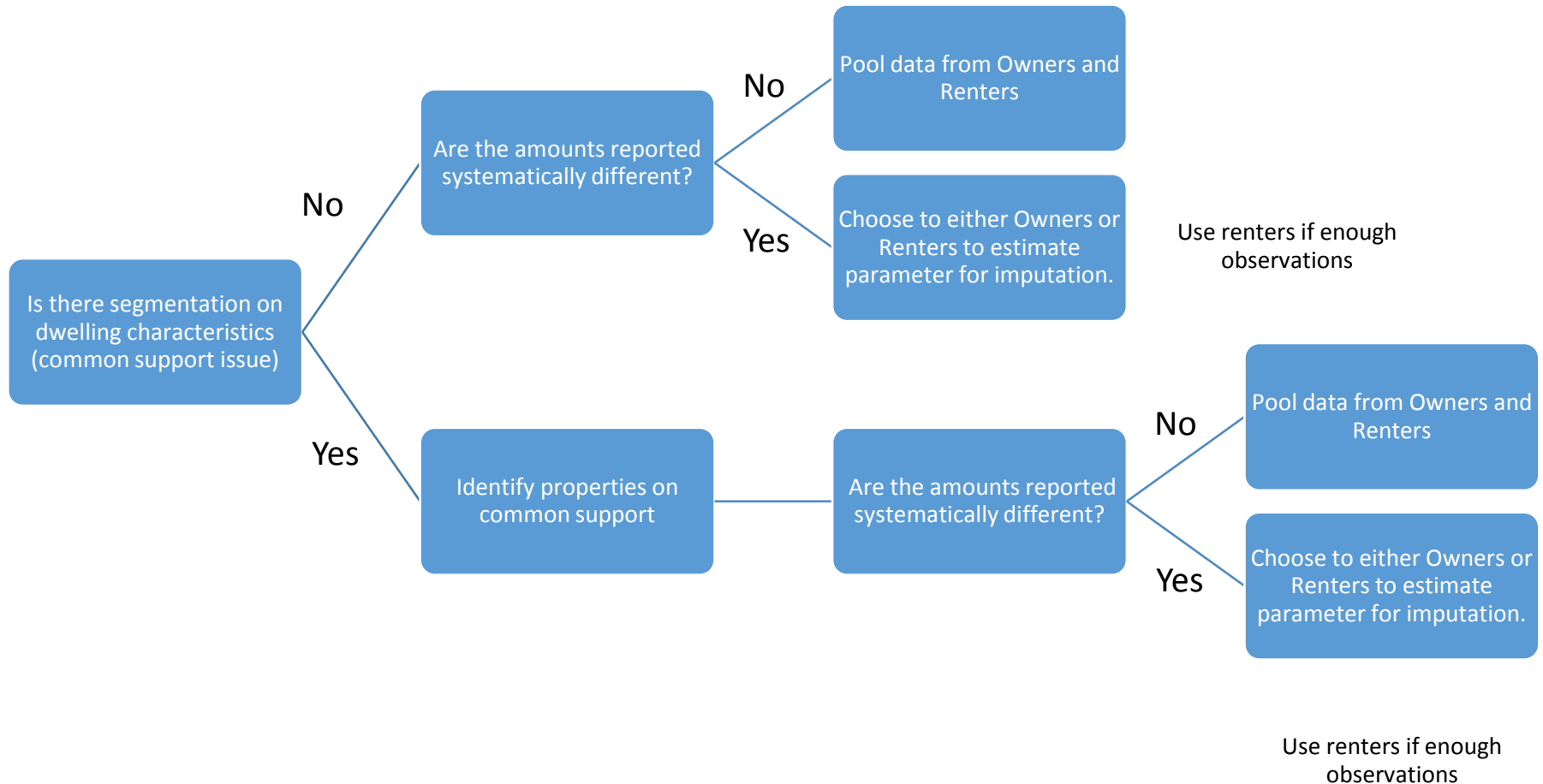
# What do we propose?

# Improvement to econometric methods

- Identify and correct potential biases in estimation
  - Guidance on how to test and validate estimates
- Layers of information:



# Decision Tree



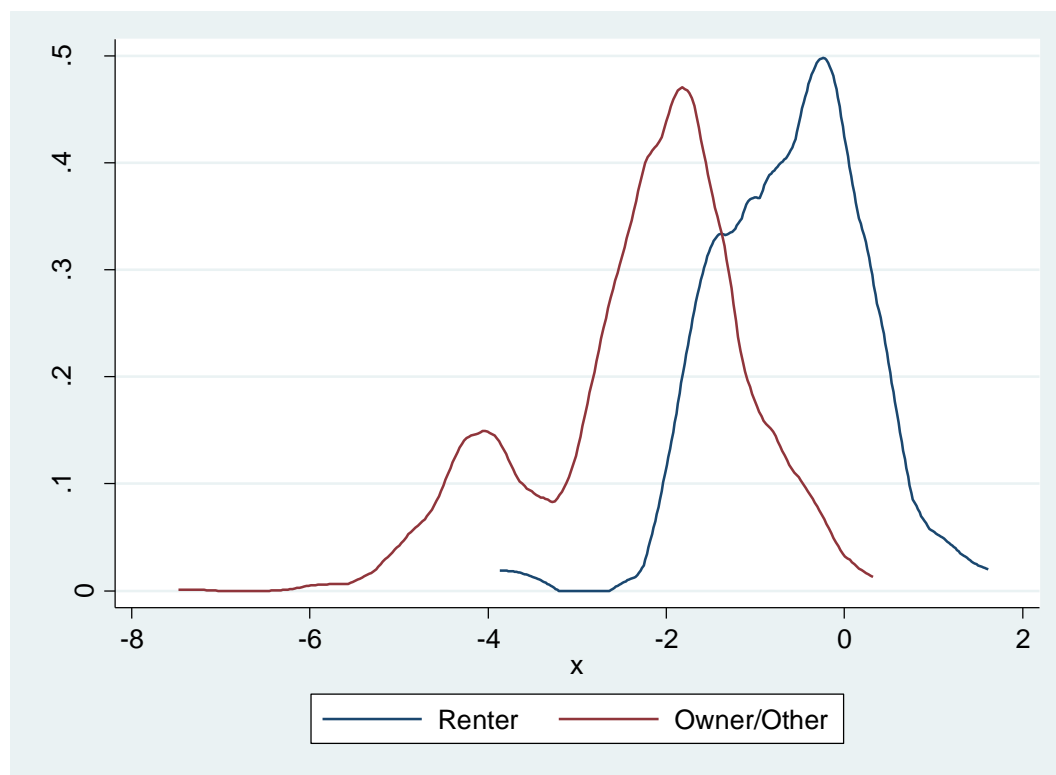
# What to impute?

- Value
  - Rental value on the owners
  - Implicit rent of owners on the renters

# How to implement it?

# Step 1: Segmentation owners-renters?

Linear projection (XB) from probit model based on observable characteristics (*psmatch2*)

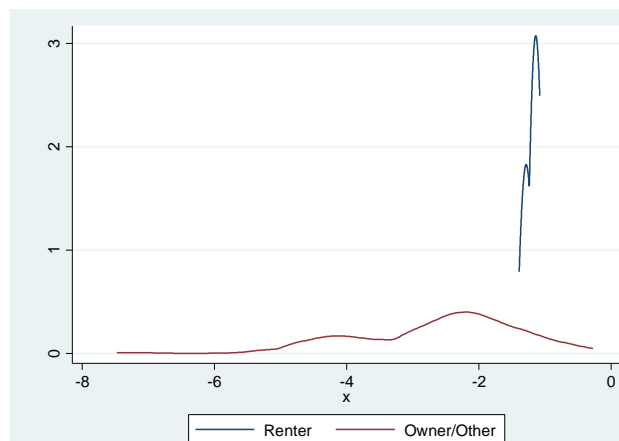


- Concentration of masses of the distributions in different places indicates different observable characteristics
- Overlapping indicates common support

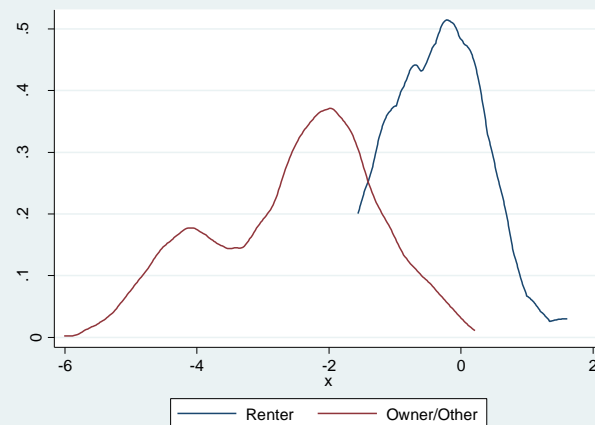
# Differences in observables across income levels?

Linear projection (XB) from probit model based on observable characteristics (*psmatch2*)

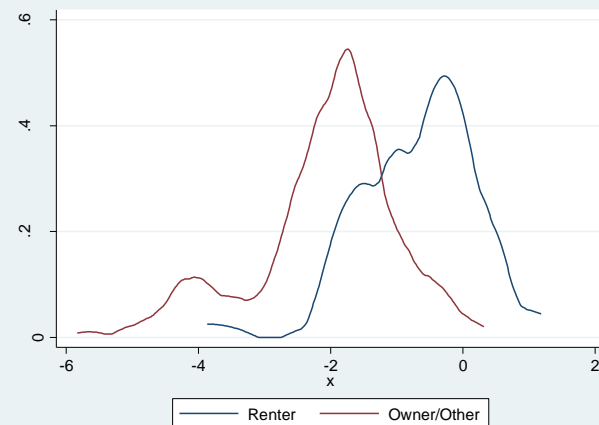
Lowest income tercile



Middle income tercile



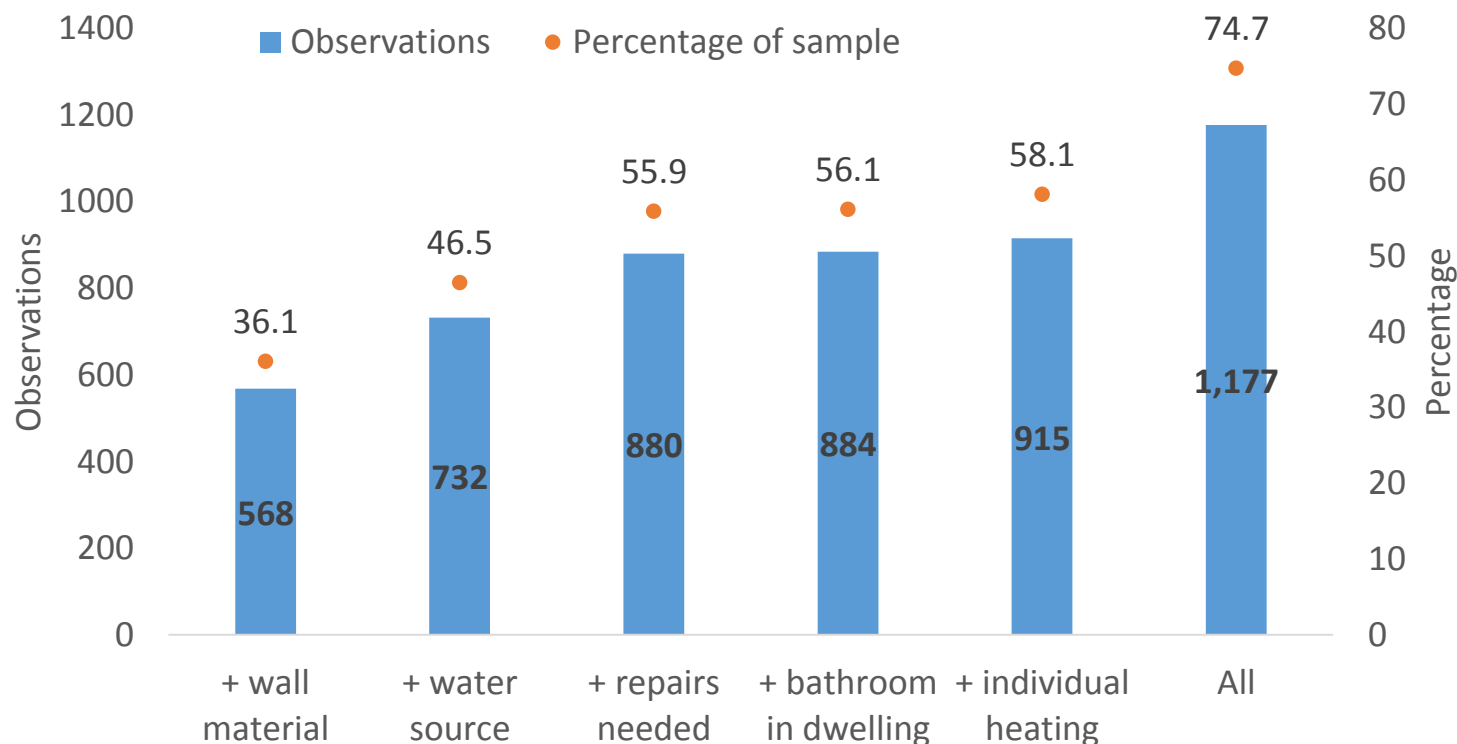
Upper income tercile



- Minimal presence of renters at the bottom of the distribution
- In middle and upper terciles there are different characteristics, but also common support on the characteristics.

# Step 2: Identifying Common support

Non Parametric method: Observations with owner self-reported implicit rent but no observations with actual rent information with **exact** same characteristics



- Large share of observations out of common support (75 percent), very restrictive.
- All variables: Repairs needed, gas connection, hot water, electricity, individual heating, phone, internet, year home build, walls material, floor material, ceiling material, area, kitchen, garage, water source, type of toilet, bedrooms, bathroom, region, location.



# Parametric method for identifying common support

Trimming method: Observations with predicted Propensity Score Matching in cells with both actual renter renter and owner self-reporting implicit rent

Cells	Owner/Other	Renter	Total
Lowest	1	0	1
2	0	0	0
3	0	0	0
4	5	0	5
5	3	0	3
6	32	0	32
7	51	0	51
8	109	2	111
9	65	0	65
10	57	0	57
11	167	0	167
12	242	0	242
13	324	13	337
14	171	21	192
15	104	12	116
16	68	33	101
17	20	28	48
18	2	7	9
19	0	2	2
Highest	0	3	3
Total	1,421	121	1,542

Identified as  
common support

- Entire range of PSM divided in 20 steps
- Additional criteria: at least 3 observations in cell both for owners and renters
- Result: 51 percent of sample identified in common support (794 observations).

# Step 3: Systematic differences in reported rent within the common support?

OLS regressions of  $\ln(\text{rent or self-reported rent})$  on observable characteristics

	Only common support (trimming method)				All observations with rent or self-reported information			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable: $\ln(\text{Rent or self-reported rent})$								
Renter	-0.109	-0.331***	-0.061	-0.073	0.11	-0.357***	-0.078	-0.095**
	(0.069)	(0.051)	(0.047)	(0.045)	(0.072)	(0.054)	(0.050)	(0.047)
Regional dummies		X	X	X		X	X	X
Urban/rural dummies		X	X	X		X	X	X
Dwelling characteristics			X	X			X	X
Consumption tercile				X				X
R-squared	0.0031	0.4854	0.7204	0.7479	0.0015	0.4798	0.6978	0.7328
N	794	794	794	794	1575	1575	1575	1575
* $p < 0.10$ , ** $p < 0.05$ , *** $p < 0.01$								

- Within the common support there are no systematic differences in rent values and self-reported implicit rent

# Conclusions

- Proposed decision tree useful tool to identify sources of biases when using reported data for rent imputation.
- Implementation shows how to:
  - Detect existence of differences in dwellings observable characteristics for renters and home-owner who self-report implicit rent
  - Identify common support in the characteristics of both groups



Thanks.