Josh Klein
Faculty Advisor: Professor
César Castillo-García

Gender, Nationality, Property Ownership, & Wealth in 20th Century Peru

Wesleyan University

Is there a significant relationship between sex and property value (measured in rental rate)?
Is there a significant relationship between nationality and property value (measured in rental rate)?

Introduction

- A more diverse economic elite came to prominence between the alate 1800's and 1960's (Suárez, 2009).
- Scholars have not investigated this period of economic history or observed ample data (Suárez, 2009).
- Before this era the economic elite were mainly involved in agriculture.
- The new generation of economic elite was diversified both by identity markers and means of income.
- Alfonzo Quiroz, in his 1988 study, found that the economic elite pooled in investment groups that financed projects geared towards industrialization.
- Felipe Portocarrero Suárez also reached similar conclusions in his 2009 study.
- Both studies provide valuable insight into the economic elite of Peru in this era, but neither places much emphasis on sex or nationality.
- Examining how sex and nationality shaped economic status in Peru will help to provide a clearer picture of Peru's economic landscape.

Data & Methodology

- Analysis was conducted using archival tax record data that Professor Castillo-García gathered from the Peruvian archives (Castillo-García 2026).
- Python for Image Pre-Processing, AI OCR (Optical Character Recognition), AI Structured Data Extraction, and Data Storage/Processing to assist in the digitization of the tax records was used to attempt to digitize data archives; however, it has proven to be inaccurate thus far.
- The archives document the key variables in Spanish as sexo (binary nominal categorical), nacionalidad (nominal categorical), and renta anual (continuous quantitative).
- The dataset currently stands with 752 observations from the Lima district.
- Of only 166 valid observations of nationality, Peruvians account for over 65% of the sample, while the rest of the sample is distributed over Germany, Argentina, Austria, Spain, France, the United Kingdom, Italy, and Switzerland, so a binary variable (notperuv) is better suited for running a model.
- Stata was used for data management and modeling.
- The final model is constructed using a log-linear OLS (Ordinary Least Squares) regression.

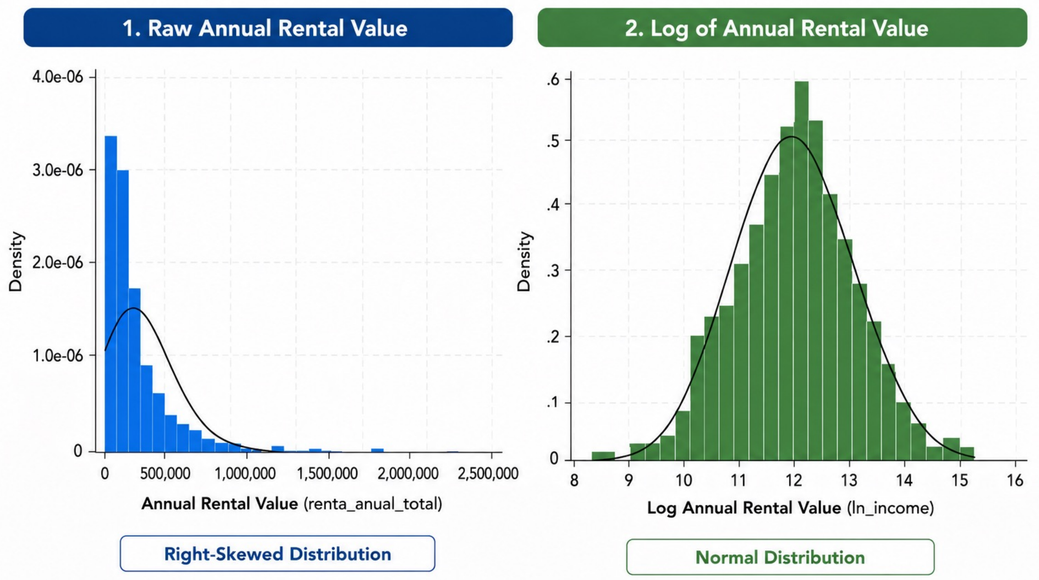
OLS Regression Results (Robust Standard Errors)

Number of observations	129	R-squared	0.0225
F-statistic (2, 126)	1.62	Root MSE	0.89863
Prob > F	0.2023		

Variable	Coefficient	Robust Std. Error	t	P>|t|	95% Confidence Interval [Lower, Upper]
sexo_num (Reference: femenino)					
masculino	0.1778292	0.2035121	0.87	0.384	[-0.2249153, 0, 5.805738]
1.notperuv (Reference: Peruvian)	0.2098954	0.1580006	1.33	0.186	[-0.1077831, 0.52225740]
Constant	11.873260	0.1805816	[illegible]	0.000	[11.5158900, 12.2306200]

Distribution of Annual Rental Values: Before and After Log Transformation

Log transformation reduces right-skewness and improves normality.



Conclusion

- As the dataset grows, it is possible that more observations of sex and nationality could prove to be significant.
- Additionally, a larger dataset would allow for other variables to be incorporated into the model, such as literacy, profession, and marital status.
- Growing the dataset, given the current exploration of AI and OCR, will take time. Any advancement in this technology would be monumental for this work.

Limitations

- Due to image blur, image distortion, and page deterioration, the archives are very difficult to digitize.
- Due to image quality, the OCR is imperfect, missing, and misinterpreting many observations. So, all of the data thus far has required manual corrections to ensure accuracy.
- The greatest limitation of the model is the size of the dataset.
- 586 observations for the predictor variable nationality are missing.
- The OLS regression assumes a linear relationship between the predictors and the log of annual rental value.
- OLS regression treats each observation (property) as an independent observation.
- The model includes a few covariates, so there may be unobserved influences.
- Annual rental rate is assumed as a reasonable proxy for property value.

Discussion

- The regression analysis did not observe a statistically significant relation between property owner sex or nationality and annual rental value (p-values > 5%).
- Based on the sample, males' and non-Peruvians' properties had slightly higher annual rental rates on average, but these differences were not statistically significant.
- Although there is no significance, there is still food for thought. Women were very oppressed in this period of Peruvian history, possessing limited rights.
- Married women were not allowed to hold property separate from their husbands, yet ⅓ of the observations for sex were female.
- Future research could assess the reasons for this similarity across gender.
- These results are consistent with the previous findings of a diverse economic elite.
- These findings serve as a preliminary analysis exploring what traits relate to property value.
- These findings of insignificance serve as an initial attempt at using archival tax records to understand a pivotal point in Peruvian economics.

References & Acknowledgments

Thanks to the QAC professors for their time and insights during the 2026 Summer apprenticeship program. Thank you especially to Professor Castillo-García and Professor Oleinikov.