Mediation Analysis on Electrical Usage of Apartment Complexes in Seoul

1. Introduction

Given that South Korea has a significant proportion of multi-unit housing, with apartments comprising 81% of this housing type, policymakers have recognized the potential of implementing energy-saving measures in these residential units. In light of this, extensive efforts have been devoted to investigating factors that may influence energy consumption in apartments, aiming to identify effective strategies for reducing energy waste.

To contribute to this area of research, I conducted an analysis to examine the impact of various variables on electricity usage in apartments, with a particular focus on the role of apartment electricity contract methods, the number of elevators, the number of underground parking spaces, and the duration of residents' occupancy. These variables were treated as potential determinants of energy consumption in apartments and/or confounding variables. They were investigated under the frame of mediation analysis, with the apartment's common electrical fee acting as an intermediary in the relationship between the aforementioned variables and electricity usage.

By comprehensively analyzing these factors, I aimed to shed light on the complex dynamics of energy usage in apartments and provide insights that can inform policymakers' decision-making processes. Through this research, the goal is to identify key drivers of energy consumption in apartments and potentially identify areas where energy-reducing measures can be effectively implemented. The findings of this study contribute to the ongoing efforts in South Korea to address energy waste and promote sustainable practices within the apartment sector.

2. Data and Methods

2.1 Data Source:

The study utilizes data from various Open API's, constructed by the Ministry of Land, Infrastructure, and Transport. They were: "공동주택 단지 목록제공 서비스", "공동주택 기본 정보제공 서비스", "공동주택 상세 정보제공 서비스", "공동주택 관리비 정보제공 서비스", "공동주택 에너지 사용 정보". The APIs contain much information about multi-unit housing and apartment complexes, such as a list of all apartment complexes in Korea, the number of CCTV's in the complex, nearby metro station and heating and water usage fees. Among them, I focused on apartment complexes located in Seoul, electricity usage fees and amount, and some variables likely to be related to electricity and/or energy usage: apartment electricity contract methods, the number of elevators, the number of underground parking spaces, and the duration of residents' occupancy.

2.2 Variables:

My goal is to identify the factors that are influential towards the apartment's electricity usage. Thus, a group of candidates were selected for the independent variable, and I conducted analysis on all of them. However, since I suspected all of the variables to be somewhat impactful, all of them were considered to be measured confounder variables when they were not the variable of focus.

- Independent Variable(s):
    - Electricity contract method (Joint contract was 1, individual and any other methods were considered 0)
    - Number of elevators (divided by number of households)
    - Number of underground parking spaces (divided by number of households)
    - Duration of residents' occupancy (measured in days)
- Mediator Variable: Common (shared) electrical fee (measured in Korean Won, divided by number of households)
- Dependent Variable: Electrical usage of the apartment (measured in kwh, divided by number of households)

Unfortunately, usage data on other forms of energy were often missing, thus the focus on electrical usage was inevitable. For electricity contract methods, I assumed that electricity contracts would be effective, because joint contracts involve centralized management of supply, and individual contracts involve more individual decisions. For the number of elevators, and underground parking spaces, since they are facilities requiring electricity, I assumed that they would be effective. For all variables where they are naturally correlated with the size of the apartment complex, I divided by the number of households. In the original dataset, there were many other variables to use as the denominator, but since this is an analysis on people's energy usage, the number of households made more sense than various measures of area.

I gathered the data of all apartment complexes in Seoul, and for electrical fee and usage, I picked one month (October of 2022) to collect data from. If there were any missing data, I dropped the apartment complex in question, and continued as normal.

As I could only have access to public datasets, whose variables often affect the common energy usage than the individual household energy usage. Thus, it only made sense to use the common energy fee data as the mediator variable.

2.3 Mediation Analysis:

Mediation analysis is conducted to assess the extent to which the relationship between the candidate of independent variables and electrical usage, mediated by common electrical fee. The analysis was conducted in these three steps: (1) assessing the relationship between the independent variable of focus, with all other independent variable candidates as measured confounding variables, and common electrical fee, (2) examining the relationship between

common electrical fee, the independent variable of focus, and electrical usage, and (3) evaluating the direct and indirect effects of each independent variable.

3. Results

3.1 Regression Analysis Results:

The mediation analysis is conducted using regression models. The exposure-mediator regression results indicate a significant positive association with all candidates, though the duration of residence seems to be slightly less significant. The mediator-outcome regression results tend to show a positive relation except for model 2, in which the exposure has a negative relationship with the outcome. All of the values are quite significant.

Table 2: Regression Results

Model 1	Coefficient	Standard Error	p-value
Contract Method → Common	3591	657.1	5.41×10^{-8}
Common → Usage	A: 7631 M: 0.7835	A: 1438 M: 0.05545	A: 1.29×10^{-7} M: Less than 2×10^{-16}
Model 2	Coefficient	Standard Error	p-value
Residence Duration → Common	0.3065	0.1133	6.924×10^{-3}
Common → Usage	A: -0.8595 M: 1.217	A: 0.1746 M: 0.06825	A: 9.51×10^{-7} M: Less than 2×10^{-16}
Model 3	Coefficient	Standard Error	p-value
Number of Elevators → Common	79460	21400	2.13×10^{-4}
Common → Usage	A: 3.640e5 M: 0.7936	A: 3.765e4 M: 0.05816	A: Less than 2×10^{-16} M: Less than 2×10^{-16}
Model 4	Coefficient	Standard Error	p-value
Underground Parking → Common	7008	773.1	Less than 2×10^{-16}
Common → Usage	A: 1.468e4 M: 0.4017	A: 8.703e2 M: 0.04669	A: Less than 2×10^{-16} M: Less than 2×10^{-16}

Table 3: Mediation Effect Results

	Controlled Direct Effect (m = 10000)	Natural Direct Effect	Natural Indirect Effect	Proportion Mediated	Total Effect
Model 1	11315	7913.365	4136.6	0.3433	12049.965
Model 2	-1.2092	-0.8863	0.3728	-0.7260	-0.5135
Model 3	411549.8	367664.1	440684.1	0.5452	808348.2
Model 4	17964.4	14944.9	5107.4	0.2547	20052.3

3.3 Mediation Effect:

I feel that I shouldn't directly compare the values, because they aren't in the same range. However, Model 2 has a negative direct effect with a positive indirect effect. In terms of proportion mediated, in absolute proportion, Model 2 has the highest proportion explained indirectly, while Model 4 has the least proportion explained indirectly.

5. Limitations

It is important to acknowledge the limitations of this study. First, I wasn't able to incorporate the longitudinal nature of energy usage data in the analysis. After many attempts, I had to conclude that I didn't have enough time to understand the multi-mediator mediation analysis methods that time-varying mediation analysis was based upon. It truly is a pity. Second, I wasn't able to account for very many confounding variables, when it would have been very beneficial. I had received advice that I should incorporate many variables in the analysis, and I feel that I failed in that aspect. Lastly, I didn't think through the comparison aspect that I wished to do, and did not normalize the data beforehand. I will make sure to do so for future similar analyses.

6. Conclusion

In conclusion, this study demonstrates the effect each independent variable candidate has on the energy usage of apartment complexes. The findings underscore the significance of each variable in common energy usage, which ultimately impact the entire energy usage. These insights can inform policymakers in developing strategies to enhance infrastructural aspects to reduce unnecessary waste of energy.