Simulating Tenant Responses to Energy Policy Interventions with Transaction-Cost-Aware LLM Agents

Weijie Xia[1], Stefanie Horian[1], Hanyue Huang[2], Queena K. Qian[1], Jie Yang[1] and Pedro P. Vergara[1]

[1]Delft University of Technology, Delft, The Netherlands

[2]Technical University of Munich, Munich, Germany

w.xia@tudelft.nl, s.horian@tudelft.nl, k.qian@tudelft.nl, J.Yang-3@tudelft.nl, p.p.vergarabarrios@tudelft.nl, hanyue.huang@tum.de

Keywords:
Large Language Models, Persona Modeling, Transaction Cost Theory, Policy Simulation, Fine-Tuning.

Abstract:
Recent studies use Large language models (LLMs) to simulate human opinions and decisions by prompting models with demographic, attitudinal, or persona-based descriptions. Yet such simulations rarely model the practical, cognitive, or social frictions that shape how people respond to policy interventions. Perceived transaction cost (PTC) provides a useful lens for modeling the practical frictions that shape policy responses, such as information burden, administrative effort, coordination demands, and perceived uncertainty. We use this lens to develop a friction-aware persona modeling approach for LLM-based simulation. In the context of energy-efficient renovation (EER), tenants are represented not only by who they are demographically, but by how they perceive the costs, benefits, barriers, and uncertainties associated with proposed renovation plans. Using survey data collected from 1,068 tenants in the Netherlands, comprising approximately 40,548 survey question and answer pairs, we compare prompt-only and fine-tuned settings across GPT-3.5-turbo, Ministral-8B-Instruct, and Llama-3.1-8B-Instruct, and evaluate supervised fine-tuning (SFT) and Group Relative Policy Optimization (GRPO) for local open-weight models. Results show that incorporating PTC-based personas and reasoning consistently improves model performance across both prompt-only and fine-tuned settings, suggesting that PTC-based persona design provides a useful bridge between institutional policy theory and interpretable LLM-based policy simulation.[a]

1 INTRODUCTION

The effectiveness of energy policy interventions depends not only on tenants' preferences, but also on their ability to act on those preferences (Liu et al., 2023). Energy-efficient renovation (EER), for example, may require tenants to understand proposed measures, assess financial consequences, coordinate with landlords or contractors, tolerate disruption, and evaluate uncertain future benefits. These informational, administrative, and coordination burdens can shape policy responses even when tenants support the policy's underlying objective (Li et al., 2025). Policymakers therefore need tools that represent both tenants' attitudes and the practical frictions affecting their decisions.

Large language models (LLMs) offer a possible basis for such tools because they can generate context-sensitive responses from textual descriptions of people and policy settings. Previous studies have used LLMs as simulated survey respondents, economic agents, and interactive social agents (Argyle et al., 2023; Aher et al., 2023; Horton et al., 2023; Park et al., 2023). Although LLMs can reproduce some aggregate patterns in human responses, their outputs may be biased, insufficiently variable, prompt-sensitive,

[a]Code and data are available at: Personal Repo and TU Delft Repo.

and unreliable for subgroup inference (Bisbee et al., 2024; Qu and Wang, 2024). Evidence from climate and energy research further indicates that demographic conditioning alone is often insufficient: simulation fidelity improves when prompts include issue-specific attitudes and covariates (Lee et al., 2024; Fell, 2024). However, the action-related frictions through which citizens evaluate and respond to policy interventions remain underrepresented in LLM persona design.

Transaction cost (TC) theory originally emphasized the costs of searching for information, negotiating, coordinating, monitoring, and implementing exchanges (Coase, 1937; Williamson, 1981; North, 1990), and environmental policy research similarly shows that such costs shape policy design and performance (McCann et al., 2005; McCann, 2013). In household energy decisions, relevant frictions include time and cognitive effort, procedural burden, disruption, distrust, and uncertainty (Mundaca, 2007; Lundmark, 2024). We focus on tenants' perceptions of these frictions alongside perceived policy benefits, such as comfort, health, and energy savings, so that the framework captures both obstacles to action and motivations for acting. To this end, we adopt perceived transaction costs (PTCs) as a theoretically grounded way to represent these frictions.

Motivated by the gap between PTC and current LLM-based persona design, we propose a PTC-aware LLM framework for simulating tenant responses to EER policies. Rather than relying on demographic priors alone, the framework represents tenants through empirically derived profiles of perceived barriers and benefits, and its main contribution is to introduce PTC-aware persona modeling as a bridge between institutional policy theory and LLM agent design. We operationalize this framework with survey data from 1,068 respondents in the Netherlands, evaluating a prompt-only GPT-3.5-turbo baseline alongside two fine-tuned open-weight models, Ministral-8B-Instruct (Mistral AI, 2024) and Llama-3.1-8B-Instruct (Ollama, 2024), adapted with QLoRA under supervised fine-tuning (SFT) and Group Relative Policy Optimization (GRPO) (Shao et al., 2024). Overall, our results show that grounding LLM personas in PTC consistently improves simulation accuracy over demographic-only prompting, providing a concrete bridge between institutional policy theory and interpretable LLM-based policy simulation.

2 LITERATURE REVIEW

2.1 LLM-Based Social and Citizen Simulation

LLMs have recently been proposed as tools for computational social science because they can classify, explain, and generate social data in ways that complement traditional survey and annotation pipelines (Ziems et al., 2024). One line of work uses LLMs to simulate individual or group behavior. Aher, Arriaga, and Kalai introduce "Turing Experiments" to test whether LLMs can replicate human-subject experiments across economic, psycholinguistic, and social-psychological settings (Aher et al., 2023). Horton and colleagues frame LLMs as simulated economic agents that can be assigned endowments, preferences, and information before being placed in experimental scenarios (Horton et al., 2023). In human-computer interaction, Social Simulacra and Generative Agents show how LLM-driven agents can populate social environments, remember experiences, plan, and generate plausible interactions (Park et al., 2022; Park et al., 2023).

For public-opinion research, Argyle et al. introduce the idea of conditioning LLMs on demographic backstories from real survey respondents and evaluating whether the resulting "silicon samples" reproduce human response patterns (Argyle et al., 2023). This work motivates survey-conditioned simulation, but subsequent studies caution against treating synthetic responses as direct substitutes for human survey data. Bisbee et al. find that LLM-generated survey responses can match broad averages while producing too little variation, unstable results across prompt changes, and regression patterns that differ from human surveys (Bisbee et al., 2024). Qu and Wang similarly document performance variation and demographic bias when simulating public opinion across countries and topics (Qu and Wang, 2024). Together, these studies suggest that LLM-based citizen simulation should be empirically benchmarked, context-specific, and explicit about its conditioning variables.

2.2 Persona Modeling and Survey Conditioning

Persona design is central to LLM simulation because the model's response is shaped by the attributes and context provided in the prompt.

Existing work commonly conditions personas on demographic attributes, political identity, psychographic traits, prior survey answers, or task-specific covariates (Argyle et al., 2023; Lee et al., 2024). In climate-opinion simulation, Lee et al. show that demographic-only prompts can fail to capture global-warming beliefs, while adding issue-relevant covariates such as involvement, interpersonal discussion, and perceived scientific consensus improves fidelity (Lee et al., 2024). Fell's energy-social-survey replications likewise demonstrate the promise of population-representative LLM agents for energy research while emphasizing practical and ethical limitations (Fell, 2024).

These findings motivate persona representations that go beyond demographic identity alone. For energy policy interventions, what matters is not only who a tenant is, but also how difficult it is for that tenant to act. Consider two otherwise comparable respondents evaluating the same home-renovation subsidy. One can quickly find reliable information, compare installers, understand eligibility rules, assemble the required documents, and coordinate the work with contractors. The other struggles to identify trustworthy advice, interpret administrative requirements, estimate the likely benefits, or align decisions with landlords, family members, or service providers. These differences reflect TCs such as information search, administrative burden, coordination demands, and uncertainty, all of which can directly mediate policy response (Mundaca, 2007; Lundmark, 2024). A PTC-based persona therefore provides theoretically grounded context about the mechanisms through which tenants translate policy offers into action. This approach differs from generic role prompting because the persona dimensions are derived from TC theory and energy policy evidence rather than from intuitive or ad hoc descriptions of tenants.

2.3 Transition Cost in Social Science and Policy

TC begins from the observation that economic exchange and institutional coordination are not frictionless. Coase's account of the firm explains organizational boundaries through the costs of using the price mechanism (Coase, 1937). Williamson develops TC economics around the transaction as the unit of analysis, emphasizing uncertainty, asset specificity, bounded rationality, opportunism, and governance structures (Williamson, 1981). North extends this logic to institutions, arguing that formal and informal rules shape human interaction partly by structuring transaction and production costs (North, 1990).

In public and environmental policy, TC affect not only firms but also agencies, intermediaries, and citizens. McCann et al. argue that policy choice and policy design should account for TC and provide guidance for measuring them in environmental and natural-resource policies (McCann et al., 2005). McCann later synthesizes empirical evidence to show that TC interact with policy design, property rights, institutional settings, and abatement costs (McCann, 2013). For energy efficiency, Mundaca shows that tradable white-certificate schemes create costs related to information search, customer persuasion, negotiation, measurement, and verification (Mundaca, 2007). Lundmark's study of Swedish residential energy renovations estimates substantial household TC and connects them to uncertainty, cognitive limitations, social connectedness, and implementation frictions (Lundmark, 2024).

3 METHOD

Figure 1 provides an overview of the proposed framework for simulating tenant responses to energy policy interventions. The framework has two connected components. First, each survey instance is converted into a structured prompt with three layers. A system instruction asks the model to answer as one specific resident and remain consistent with that person's housing situation, attitudes, and priorities. A PTC persona prompt then adds barrier- or benefit-oriented persona descriptions, such as Financial Sensitive and Immediate Utility Seekers, to capture the frictions and motivations most relevant to the intervention. A PTC reasoning prompt further asks the model to reflect briefly on burden, uncertainty, personal gains, and the likely direction of the response before answering the survey question in the required format. Second, we compare different modeling strategies for generating these responses. In addition to a prompt-only GPT-3.5-turbo baseline, we adapt local open-weight models using QLoRA under two training settings: supervised fine-tuning (SFT), which learns directly from survey examples, and Group Relative Policy Optimization (GRPO), which further optimizes outputs with an explicit reward. The following

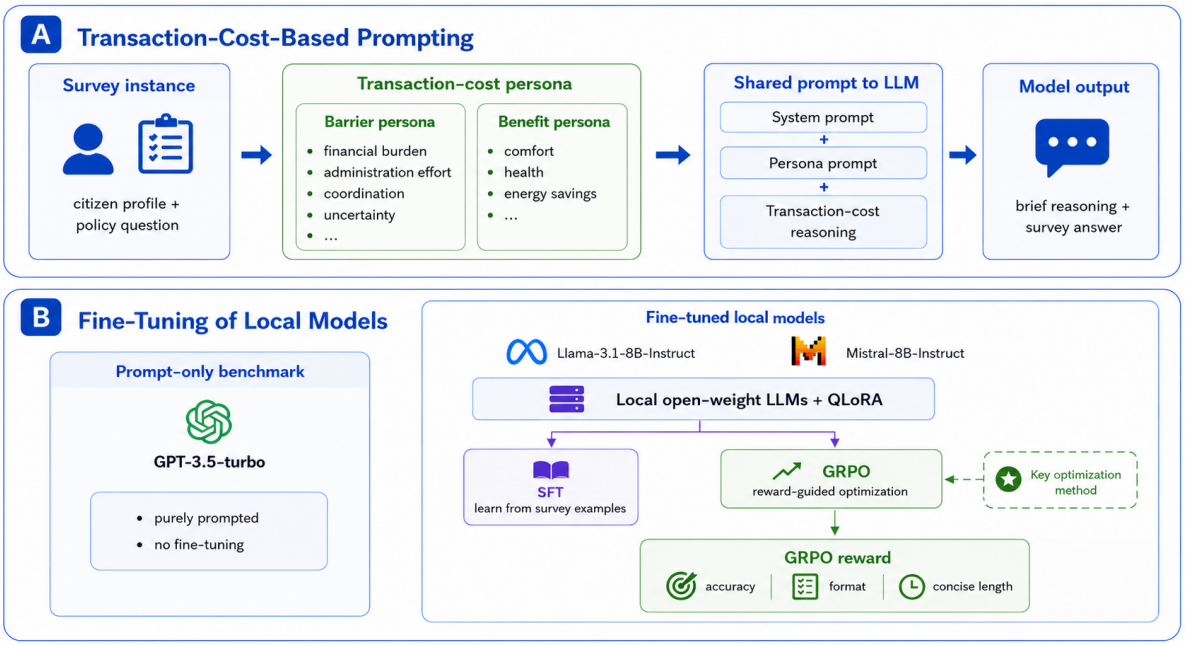

Figure 1: Overview of the proposed framework for simulating tenant responses to energy policy interventions.

subsections explain how PTC personas are designed and assigned, how the shared prompt is structured, and how SFT and GRPO are used to adapt the local models.

3.1 Prompt Design and PTC Persona Construction

As shown in Figure 2, each survey prompt combines five components: a system instruction, PTC persona descriptions, a PTC reasoning checklist, the survey question, and an output format template. The system instruction and question format provide structural scaffolding common to all instances, while the persona and reasoning components are the substantive PTC-theoretic elements that vary by respondent. The following subsubsections introduce the two core components in detail, followed by a brief note on the remaining structural elements.

3.1.1 PTC Persona Prompting

We adopt the EER persona setup developed in prior work on Dutch social-housing tenants (Horian et al., 2026). In that study, tenant personas are obtained from survey indicators of perceived renovation benefits and barriers. In our study, we adopt the persona predictor developed in that work, which is built to infer the most likely barrier and benefit persona from non-EER contextual variables, such as age, household composition, tenure length, and other socio-demographic characteristics. Benefit personas (BE1–BE5) measure how strongly tenants value outcomes such as comfort, health, wellbeing, reduced energy consumption, environmental improvement, and neighborhood effects. Barrier personas (BA1–BA7) measure perceived frictions that may prevent agreement with renovation plans, including rent or service-charge increases, temporary relocation, daily-life disruption, lack of time or information, distrust, and uncertainty.

In our LLM framework, each survey instance is assigned one benefit persona and one barrier persona according to the respondent's observed survey pattern and the persona predictors developed in the prior study. The benefit persona captures what kinds of EER gains the respondent is likely to value, while the barrier persona captures the transaction-cost frictions and behavioral constraints that may shape the respondent's answer. This keeps persona construction tied to empirical survey evidence rather than ad hoc role descriptions or demographic stereotypes. Table 1 summarizes the twelve persona types; full definitions appear in the Appendix.

System Prompting:
You answer survey questions for one specific resident. Stay consistent with that person's housing situation, attitudes, and priorities.

PTC Persona Prompting:
Financial Sensitive: This person generally supports energy-efficient renovation but worries about money, especially higher rent or service charges. Disruption matters less than financial risk. They need clear reassurance that renovation will not make their finances worse.

Immediate Utility Seekers: This person mainly cares about direct personal gains, such as better comfort, health, well-being, and lower energy use. Appearance or neighborhood effects matter less. They respond well when short-term benefits are clear.

PTC Reasoning Prompting:
Before the final response, briefly think about:

1. Burden: time, effort, disruption, moving, or financial risk
2. Uncertainty: whether the effects are clear or unclear
3. Personal gains: comfort, health, well-being, and energy savings
4. Likely direction: stay neutral when information is unclear; lean positive when direct benefits are clear; lean negative when burdens or costs are clear

Question:
Please indicate to what extent you personally care about the following benefit of Energy-Efficient Renovation: Increasing thermal comfort for my home (e.g., warmer in winter and cooler in summer).
1–I don't care at all; 2–I don't care; 3–Neutral; 4–I care; 5–I care a lot

Output Format:
<thinking>
brief reasoning in 2–4 sentences
</thinking>
<answer>
single integer only, for example 3
</answer>

Figure 2: Shared system prompt used in both SFT and GRPO. It defines the role, PTC persona, reasoning guide, question, and required output format.

3.1.2 PTC Reasoning Prompting

The reasoning component asks the model to briefly consider four items before producing a final Likert-scale answer: the burden the intervention imposes, sources of uncertainty, expected personal gains, and the likely direction of the response. This design anchors outputs to the explicit decision frictions encoded in the assigned PTC persona, rather than relying on demographic priors alone. Figure 3 illustrates the expected output format, where a short reasoning trace enclosed in <thinking> tags precedes a single integer answer.

Persona type	Index	Personas
Barrier	BA1	Financial Sensitive
	BA2	Practical and Financial Concerned
	BA3	Ambivalent Observer
	BA4	Self-Reliant but Unconvinced
	BA5	Financially Alarmed
	BA6	Confident Acceptors
	BA7	The Uncertain
Benefit	BE1	Immediate Utility Seekers
	BE2	Personal Comfort Seekers
	BE3	Balanced Benefit Idealists
	BE4	Ambivalent Respondents
	BE5	Pessimists

Table 1: PTC-aware personas used for prompt conditioning. Barrier personas (BA1–BA7) encode frictions that may impede renovation agreement; benefit personas (BE1–BE5) encode values that may motivate it. Each respondent is assigned one barrier and one benefit persona. Full definitions are in the Appendix.

<thinking>
Thermal comfort matters to me because it improves daily life right away. I am also careful about money, so I would want to know that renovation will not raise my rent or service charges too much. If the comfort gains are clear and the financial risk is controlled, I would likely care about this benefit.
</thinking>
<answer>
4
</answer>

Figure 3: Formatted answer example corresponding to the shared survey prompt.

3.1.3 System Prompt and Survey Questions

The system instruction briefly positions the model as a survey respondent with the assigned persona characteristics, ensuring consistent role adoption throughout the response. The survey question is drawn verbatim from the original EER questionnaire and appended after the persona and reasoning prompts. The output format template, shown in Figure 2, specifies the required <thinking> and <answer> tag structure so that training targets have a consistent form across both SFT and GRPO.

3.2 Supervised Fine-Tuning

Supervised fine-tuning adapts the base model by maximizing the likelihood of survey-consistent responses conditioned on the prompt. Given a dataset $D = \{(x_i, y_i)\}_{i=1}^{N}$, where x_i is the persona-intervention prompt and y_i is the target response, SFT minimizes the negative log likelihood

$$\mathcal{L}_{\mathrm{SFT}}(\theta) = -\sum_{(x,y)\in D}\sum_{t=1}^{|y|}\log \pi_\theta(y_t \mid x, y_{<t}). \tag{1}$$

This objective is appropriate when the goal is to teach the model the format, domain vocabulary, and empirical mapping between survey-conditioned prompts and observed answers. SFT also provides a transparent baseline because it uses only labeled demonstrations and does not require a separately specified reward function (Ouyang et al., 2022). In our implementation, SFT is QLoRA-based: the base model remains quantized and fixed, while low-rank adapter parameters are trained to reproduce the survey-consistent reasoning trace and final integer answer (Hu et al., 2021; Dettmers et al., 2023).

3.3 Group Relative Policy Optimization

GRPO is a reinforcement-learning method introduced to improve LLM reasoning while reducing the memory overhead of critic-based PPO (Shao et al., 2024). As with SFT, the GRPO run uses QLoRA adapters rather than full-parameter updates. For each prompt q, the current policy samples a group of G candidate responses $\{o_1, \ldots, o_G\}$. The reward function assigns each response a scalar reward r_i, and the reward is normalized within the sampled group to obtain a token-level advantage:

$$\begin{aligned} \hat{A}_{i,t} &= \frac{r_i - \mu}{\sigma}, \\ \mu &= \frac{1}{G}\sum_{i=1}^{G} r_i, \quad \sigma = \sqrt{\frac{1}{G}\sum_{i=1}^{G}(r_i-\mu)^2 + \varepsilon}. \end{aligned} \tag{2}$$

The policy is then updated with a GRPO objective and a KL penalty that keeps the adapted model close to a reference policy:

$$\begin{aligned} p_{i,t}^{\theta} &= \pi_\theta(o_{i,t} \mid q, o_{i,<t}), \\ p_{i,t}^{\mathrm{ref}} &= \pi_{\mathrm{ref}}(o_{i,t} \mid q, o_{i,<t}). \end{aligned} \tag{3}$$

$$\begin{aligned} \mathcal{L}_{\mathrm{GRPO}}(\theta) = -\frac{1}{\sum_{i=1}^{G}|o_i|}\sum_{i=1}^{G}\sum_{t=1}^{|o_i|}\Bigg[& \frac{p_{i,t}^{\theta}}{[p_{i,t}^{\theta}]_{\mathrm{sg}}}\hat{A}_{i,t} \\ & - \beta_{\mathrm{KL}} D_{\mathrm{KL}}(\pi_\theta \| \pi_{\mathrm{ref}})\Bigg], \end{aligned} \tag{4}$$

where $[\cdot]_{\mathrm{sg}}$ denotes stop-gradient. The KL term is estimated at each generated token as

$$D_{\mathrm{KL}}(\pi_\theta \| \pi_{\mathrm{ref}}) = \frac{p_{i,t}^{\mathrm{ref}}}{p_{i,t}^{\theta}} - \log\frac{p_{i,t}^{\mathrm{ref}}}{p_{i,t}^{\theta}} - 1. \tag{5}$$

Unlike actor-critic PPO, GRPO does not require a learned value model; the group baseline provides the relative advantage estimate. This makes it attractive for resource-constrained fine-tuning of open models while still allowing explicit optimization of response quality.

In this paper, GRPO is used to compare whether reward-guided adaptation improves policy-response simulation over SFT alone. The reward combines survey-label accuracy, output-format validity, and length control:

$$r_i = w_1 r_{\mathrm{acc},i} + w_2 r_{\mathrm{format},i} + w_3 r_{\mathrm{len},i}. \tag{6}$$

The accuracy reward decreases smoothly as the predicted integer answer $\hat{y}_i$ moves away from the observed survey label y_i:

$$r_{\mathrm{acc},i} = r_{\min} + \frac{r_{\max} - r_{\min}}{1 + e_i^2}, \quad e_i = |\hat{y}_i - y_i|. \tag{7}$$

The format reward checks whether the model returns a parsable reasoning field and a final single-integer answer field:

$$r_{\mathrm{format},i} = \begin{cases} 1, & \text{if the required format is correct,} \\ 0, & \text{otherwise.} \end{cases} \tag{8}$$

In implementation, this check can be applied to XML-like tags such as $\langle \mathrm{thinking} \rangle \cdots \langle /\mathrm{thinking} \rangle$ and $\langle \mathrm{answer} \rangle \cdots \langle /\mathrm{answer} \rangle$, or to the equivalent Thinking/Answer fields in the shared prompt. Finally, the length reward penalizes reasoning traces and answer fields that deviate from the target lengths:

$$\begin{aligned} r_{\mathrm{len},i} = -\Big(& \alpha_\ell |L_{r,i} - L_r^{\mathrm{target}}| \\ & + \beta_\ell |L_{a,i} - L_a^{\mathrm{target}}| \Big), \end{aligned} \tag{9}$$

where $L_{r,i}$ and $L_{a,i}$ are the reasoning and answer lengths. This reward design favors outputs that are accurate with respect to the survey label, short enough to remain inspectable, and strict enough to be parsed automatically.

4 EXPERIMENTS AND RESULTS

4.1 Evaluation Metrics

We evaluate the models at the individual-response level using two complementary metrics.

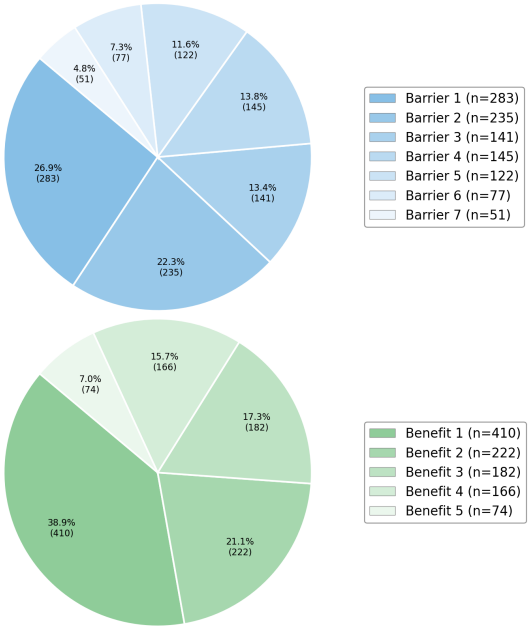

Figure 4: Distribution of the twelve PTC personas in the dataset.

For ordinal prediction error, we use mean absolute error (MAE) between the model's integer response $\hat{y}_i$ and the observed survey label y_i:

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N} |\hat{y}_i - y_i|. \tag{10}$$

Lower MAE indicates closer agreement with the observed survey labels.

For classification performance, we use two accuracy metrics. Exact-match accuracy (Acc) counts only predictions that equal the true label:

$$\text{Acc} = \frac{1}{N}\sum_{i=1}^{N} \mathbf{1}[\hat{y}_i = y_i]. \tag{11}$$

Relaxed accuracy (Acc±1) counts predictions within one scale point of the true label, capturing near-miss agreement on the ordinal response scale:

$$\text{Acc}\pm1 = \frac{1}{N}\sum_{i=1}^{N} \mathbf{1}[|\hat{y}_i - y_i| \le 1]. \tag{12}$$

Together, MAE, Acc, and Acc±1 distinguish exact classification, near-miss agreement, and overall ordinal distance from the observed survey responses.

4.2 Experimental Setup

The dataset contains approximately 1,068 tenant consultation records and around 40,548 question–answer pairs collected from real respondents. Of these pairs, 24,242 are attitude-related and the remainder are demographic. For fine-tuning, we use 4,600 attitude-related question–answer pairs, corresponding to around 20% of the attitude-related subset, and evaluate the models on the remaining attitude-related pairs. Each attitude-related pair captures a tenant's response to an energy policy item, including benefit- and barrier-related questions answered on a 5-point Likert scale. Figure 4 shows the distribution of the twelve PTC personas in the dataset. In addition to survey responses, each record includes basic socio-demographic information, such as age, income or salary level, and household composition (e.g., family size and presence of dependants). These attributes support the construction of the PTC persona used during prompting and fine-tuning.

We use GPT-3.5-turbo as a prompt-only baseline under three prompting conditions that progressively add PTC information. The first condition uses no PTC persona or reasoning prompt; the second adds the PTC persona prompt; and the third adds both the PTC persona prompt and the PTC reasoning prompt. For the local open-weight models, we use Ministral-8B-Instruct and Llama-3.1-8B-Instruct, both served via Ollama. These models are first evaluated without fine-tuning under the fully enriched prompt, which includes both persona and reasoning components, and are then adapted with QLoRA-based SFT and QLoRA-based GRPO. Table 2 summarizes the nine experimental conditions. The GPT-3.5-turbo conditions (a–c) isolate the contribution of PTC prompting, while the open-weight model conditions (-1 through -3) keep the fully enriched prompt fixed and vary only the adaptation method.

Both open-weight models are adapted using QLoRA with 4-bit NF4 quantization, double quantization, and fp16 compute precision. For SFT, we attach LoRA adapters (r=16, α=32, dropout 0.05) to all linear layers, use a learning rate of 1×10^{-6} and an effective batch size of 8, and supervise only the final response token. For GRPO, we use the same QLoRA configuration and learning rate, with a per-device batch size of 4, 4 completions sampled per prompt, gradient clipping at 0.1, and bf16 mixed precision.

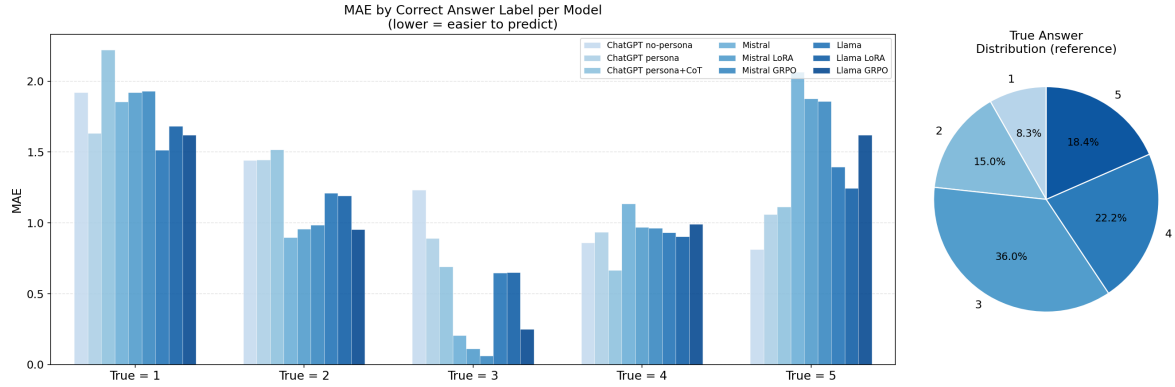

Figure 5: MAE by true answer label for each model condition. Lower values indicate that the corresponding response category is easier to predict.

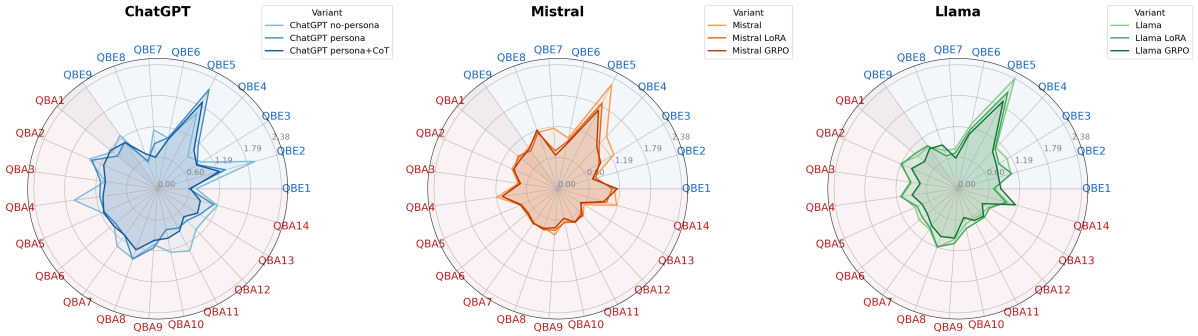

Figure 6: Question-wise MAE across all model conditions. QBE and QBA index the benefit- and barrier-related survey questions. Note: these question indices are distinct from the PTC persona labels (BE1–BE5, BA1–BA7) in Table 1.

Model	PTC Per.	PTC Reas.	Fine-tune
GPT-3.5-turbo-a	No	No	—
GPT-3.5-turbo-b	Yes	No	—
GPT-3.5-turbo-c	Yes	Yes	—
Ministral-8B-Instruct-1	Yes	Yes	None
Ministral-8B-Instruct-2	Yes	Yes	SFT
Ministral-8B-Instruct-3	Yes	Yes	GRPO
Llama-3.1-8B-Instruct-1	Yes	Yes	None
Llama-3.1-8B-Instruct-2	Yes	Yes	SFT
Llama-3.1-8B-Instruct-3	Yes	Yes	GRPO

Table 2: Experimental conditions. PTC Per./Reas. = PTC persona/reasoning prompting; open-weight models served via Ollama, fine-tuned with QLoRA.

4.3 Main Results

Table 3 presents the full benchmark results. The three GPT-3.5-turbo conditions show that PTC-based prompting improves performance step by step. The baseline without persona or reasoning prompts (GPT-3.5-turbo-a) performs worst overall (MAE 1.4783, Acc 0.2476, Acc±1 0.6780). Adding a PTC-based persona prompt (GPT-3.5-turbo-b) substantially improves performance, reducing MAE to 1.0859 and raising Acc to 0.3131 and Acc±1 to 0.7168. Adding PTC-based reasoning (GPT-3.5-turbo-c) yields a further gain in ordinal fit, with the lowest MAE (1.0335) and highest Acc±1 (0.7379) among the three GPT-3.5-turbo settings, although exact-match accuracy remains slightly below GPT-3.5-turbo-b (0.3086 versus 0.3131). Taken together, these results show that PTC-based persona and reasoning prompts consistently improve simulation quality, supporting the effectiveness of the proposed PTC-based modeling strategy.

Table 3 also shows that fine-tuning local open-weight models further outperforms the prompt-only baseline. Compared with the best GPT-3.5-turbo prompt-only setting (GPT-3.5-turbo-c: MAE 1.0335, Acc 0.3086, Acc±1 0.7379), both SFT and GRPO achieve lower MAE and higher accuracy on Ministral-8B-Instruct and Llama-3.1-8B-Instruct. The strongest overall result comes from Llama-3.1-8B-Instruct with GRPO, which achieves the lowest MAE (0.8694) and highest Acc±1 (0.7843), while Ministral-8B-Instruct with SFT attains the highest exact-

Experiment	MAE	Acc	Acc±1
GPT-3.5-turbo-a	1.4783	0.2476	0.6780
GPT-3.5-turbo-b	1.0859	0.3131	0.7168
GPT-3.5-turbo-c	1.0335	0.3086	0.7379
Ministral-8B-Instruct-1	0.9855	0.3466	0.7338
Ministral-8B-Instruct-2	0.8962	0.3702	0.7608
Ministral-8B-Instruct-3	0.8802	0.3683	0.7665
Llama-3.1-8B-Instruct-1	1.0012	0.3320	0.7482
Llama-3.1-8B-Instruct-2	0.9780	0.3383	0.7530
Llama-3.1-8B-Instruct-3	0.8694	0.3643	0.7843

Table 3: Benchmark results across prompt-only and fine-tuned settings. Lower MAE and higher Acc/Acc±1 indicate better agreement with observed survey responses.

match accuracy in the table (0.3702). For both model families, GRPO yields the best MAE and Acc±1, suggesting that fine-tuning improves not only exact matches but also the ordinal closeness of predictions to the observed survey responses.

4.4 Question-Wise Analysis

Figure 5 shows the MAE by true answers for each model. Middle response categories are generally easier to predict than extreme ones. Across model conditions, MAE is lowest for label 3 and remains relatively low for labels 2 and 4, whereas labels 1 and 5 are consistently harder. The pie chart suggests one reason for this pattern: the extreme categories are also less common in the data, with label 1 accounting for only 8.3% of responses and label 5 for 18.4%, compared with 36.0% for label 3 and 22.2% for label 4. In other words, the rarest response categories are also the most difficult to model accurately. Figure 5 also helps explain the differences in Table 3: PTC-based prompting and model adaptation often reduce ordinal error by moving predictions closer to the correct label, even when exact-match accuracy improves only modestly.

Figure 6 plots question-wise MAE across all model conditions, allowing comparison of how prompting and fine-tuning affect performance on individual benefit and barrier items. The figure shows that these gains are not uniform across questions, but two broad patterns are clear. For GPT-3.5-turbo, the improvement comes from prompt design: adding a PTC-based persona generally improves the baseline across many items, and adding PTC-based reasoning often yields further but smaller gains. For the open-weight models, the main improvement comes from model adaptation. Both Ministral-8B-Instruct and Llama-3.1-8B-Instruct show lower question-wise error after SFT and GRPO than in their untuned settings, indicating that fine-tuning provides a more consistent benefit across the full question set. One notable exception is BE5, which remains difficult across all model families and settings.

4.5 Persona-Wise Analysis

Tables 4 and 5 report MAE disaggregated by barrier persona and benefit persona type, where each cell shows the average prediction error for survey instances assigned to that persona. Three patterns emerge that complement the aggregate results in Table 3.

First, persona difficulty is uneven and follows a recognizable structure. Among barrier personas, Financially Alarmed (avg. MAE 1.20) and Confident Acceptors (avg. 1.18) are consistently the hardest across all models, while The Uncertain (avg. 0.71) and Self-Reliant but Unconvinced (avg. 0.84) are the easiest. Among benefit personas, Personal Comfort Seekers (avg. 1.18) and Balanced Benefit Idealists (avg. 1.08) are hardest, while Ambivalent Respondents (avg. 0.73) is easiest. The easy personas in both groups represent moderate or undecided tenants who tend to cluster near the neutral middle of the response scale, making them easier to predict.

Second, fine-tuning with GRPO generally dominates but not universally. Ministral-8B-Instruct-3 wins 4 of 7 barrier columns and 3 of 5 benefit columns. However, two notable exceptions arise: Confident Acceptors (barrier) and Balanced Benefit Idealists (benefit) are both best predicted by GPT-3.5-turbo-b, the persona-only prompt without reasoning or fine-tuning. This suggests that for polarized-positive tenant profiles, PTC persona prompting alone captures sufficient context, while reward-guided adaptation does not provide additional benefit.

Third, Ministral and Llama models exhibit different persona-level weaknesses. Ministral models consistently struggle on Confident Acceptors and Financially Alarmed (MAE above 1.2 even after fine-tuning), whereas Llama models are more moderate on those columns. Conversely, Llama-3.1-8B-Instruct-3 outperforms Ministral-8B-Instruct-3 on Personal Comfort Seekers (1.04 versus 1.08). Finally, all Ministral fine-tuned models perform worse than GPT-3.5-turbo-a on Balanced Benefit Idealists, suggesting that fine-tuning may introduce bias for this holistic persona type.

Model	Fin. Sensitive	Prac. & Fin. Concerned	Ambivalent Obs.	Self-Reliant	Fin. Alarmed	Confident Accept.	The Uncert
GPT-3.5-turbo-a	1.1587	1.0640	1.2909	1.1806	1.1410	1.5000	1.1
GPT-3.5-turbo-b	1.2315	1.1014	0.8071	1.1330	1.1801	0.8348	0.9
GPT-3.5-turbo-c	1.2029	0.8750	0.9476	0.8624	1.0651	1.2007	0.7
Ministral-8B-Instruct-1	1.1543	1.0110	0.7468	0.6672	1.4686	1.4222	0.5
Ministral-8B-Instruct-2	1.0358	0.9088	0.7117	0.6029	1.2942	1.3607	0.4
Ministral-8B-Instruct-3	1.0007	0.9156	0.7210	0.5978	1.2255	1.3779	0.4
Llama-3.1-8B-Instruct-1	1.1248	0.9540	0.8819	0.9550	1.1601	0.9530	0.8
Llama-3.1-8B-Instruct-2	1.0930	0.9634	0.8698	0.8896	1.1505	0.8949	0.8
Llama-3.1-8B-Instruct-3	0.9970	0.9326	0.7233	0.6442	1.1397	1.1186	0.4
Avg (all)	1.1110	0.9695	0.8556	0.8370	1.2028	1.1847	0.7

Table 4: MAE [-] by barrier persona type (see Table 1 for persona definitions).

Model	Immed. Utility Seek.	Personal Comfort Seek.	Balanced Ideal.	Ambivalent Resp.	Pessimists
GPT-3.5-turbo-a	1.1223	1.2352	1.1031	1.2551	1.4786
GPT-3.5-turbo-b	1.1876	1.3731	0.8910	0.7449	1.2857
GPT-3.5-turbo-c	1.0125	1.1507	1.0828	0.7868	1.1411
Ministral-8B-Instruct-1	0.9500	1.2332	1.2671	0.6183	0.9169
Ministral-8B-Instruct-2	0.8234	1.1083	1.2353	0.5827	0.8400
Ministral-8B-Instruct-3	0.7838	1.0787	1.2791	0.5743	0.8340
Llama-3.1-8B-Instruct-1	1.0355	1.2267	0.9234	0.7028	1.2984
Llama-3.1-8B-Instruct-2	1.0171	1.1845	0.9109	0.7052	1.1858
Llama-3.1-8B-Instruct-3	0.8239	1.0364	1.0713	0.6072	0.9894
Avg (all)	0.9729	1.1807	1.0849	0.7308	1.1078

Table 5: MAE [-] by benefit persona type (see Table 1 for persona definitions). Model labels follow Table 2. Bold = lowest MAE per column.

5 CONCLUSION

This paper presented a framework for LLM-based simulation of tenant responses to EER interventions, grounded in PTC theory. Rather than using generic role descriptions or demographic profiles, the framework conditions each simulation instance on empirically derived PTC personas from prior survey research on Dutch social-housing tenants. These personas encode the specific frictions and motivations shaping a respondent's evaluation of a renovation policy, and are paired with a structured PTC reasoning prompt that guides the model to deliberate over burden, uncertainty, and personal gain before answering. This design makes transaction costs an explicit part of the model's deliberative context, rather than a post-hoc interpretation of its outputs.

Across the experiments, the results support three main conclusions. First, PTC-aware prompting improves simulation quality even in a prompt-only setting: for GPT-3.5-turbo, adding persona information and reasoning substantially reduces MAE and improves accuracy relative to a no-persona baseline. Second, adapting local open-weight models yields stronger overall performance than prompt-only baselines, with the best overall results achieved by Llama-3.1-8B-Instruct under GRPO. Third, the choice of metric matters: GRPO most consistently improves ordinal closeness, while exact-match accuracy can show smaller or more model-specific gains.

The finer-grained analyses also show that model performance is uneven across response types. Moderate responses are easier to predict than extreme responses, and improvements are distributed differently across questions and model families. These findings suggest that LLM-based policy simulation should be evaluated not only by aggregate averages, but also by response distribution and question-level behavior.

Overall, the results suggest that PTC-based persona design is a useful bridge between institutional policy theory and LLM agent modeling. It improves interpretability, supports inspectable adaptation of open-weight models, and provides a practical direction for building more policy-relevant simulations of tenant behavior.

ACKNOWLEDGEMENTS

- Conflicts of Interest: The authors have no conflicts of interest or competing interests to declare in the context of this paper.
- Financial and other Support: This research was supported by the Align4Energy Project (NWA.1389.20.251) and utilized the Dutch

National e-Infrastructure with the support from the SURF Cooperative (grant number: EINN-5398).

- Author Contributions: Weijie Xia contributed to conceptualization, implementation, experiments, and writing of the manuscript. Stefanie Horian contributed to conceptualization and writing. Hanyue Huang contributed to implementation. Queena K. Qian and Jie Yang contributed to the revision. Pedro P. Vergara contributed to funding acquisition and revision.
- Data Sharing: Code and data are available at Personal Repo and TU Delft Repo.
- AI Tools: The paper used AI tools for text correction purposes only.

REFERENCES

Aher, G. V., Arriaga, R. I., and Kalai, A. T. (2023). Using large language models to simulate multiple humans and replicate human subject studies. In Proceedings of the 40th International Conference on Machine Learning, volume 202 of Proceedings of Machine Learning Research, pages 337–371. PMLR.

Argyle, L. P., Busby, E. C., Fulda, N., Gubler, J. R., Rytting, C., and Wingate, D. (2023). Out of one, many: Using language models to simulate human samples. Political Analysis, 31(3):337–351.

Bisbee, J., Clinton, J. D., Dorff, C., Kenkel, B., and Larson, J. M. (2024). Synthetic replacements for human survey data? the perils of large language models. Political Analysis, 32(4):401–416.

Coase, R. H. (1937). The nature of the firm. Economica, 4(16):386–405.

Dettmers, T., Pagnoni, A., Holtzman, A., and Zettlemoyer, L. (2023). Qlora: Efficient finetuning of quantized llms. In Advances in Neural Information Processing Systems, volume 36.

Fell, M. J. (2024). Energy social surveys replicated with large language model agents. SSRN working paper.

Horian, S., Qian, Q. K., and Visscher, H. (2026). Beyond one-size-fits-all: Data-driven tenant personas for targeted intervention strategies in social housing renovation. Energy Research & Social Science, 138.

Horton, J. J., Filippas, A., and Manning, B. S. (2023). Large language models as simulated economic agents: What can we learn from homo silicus? Working Paper 31122, National Bureau of Economic Research.

Hu, E. J., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., Wang, L., and Chen, W. (2021). Lora: Low-rank adaptation of large language models.

Lee, S., Peng, T.-Q., Goldberg, M. H., Rosenthal, S. A., Kotcher, J. E., Maibach, E. W., and Leiserowitz, A. (2024). Can large language models estimate public opinion about global warming? an empirical assessment of algorithmic fidelity and bias. PLOS Climate, 3(8):e0000429.

Li, W., Qian, Q., Mlecnik, E., He, S., and Song, K. (2025). Forging enhanced collaboration: Investigating transaction costs in pre-design phase of market-oriented community renovation in china. Land, 14(7):1403.

Liu, X., Yang, D., Arentze, T., and Wielders, T. (2023). The willingness of social housing tenants to participate in natural gas-free heating systems project: Insights from a stated choice experiment in the netherlands. Applied Energy, 350:121706.

Lundmark, R. (2024). Understanding transaction costs of energy efficiency renovations in the swedish residential sector. Energy Efficiency, 17(20).

McCann, L. (2013). Transaction costs and environmental policy design. Ecological Economics, 88:253–262.

McCann, L., Colby, B., Easter, K. W., Kasterine, A., and Kuperan, K. V. (2005). Transaction cost measurement for evaluating environmental policies. Ecological Economics, 52(4):527–542.

Mistral AI (2024). Ministral-8B-Instruct-2410. Hugging Face model repository. Instruction-tuned large language model; accessed 13 July 2026.

Mundaca, L. (2007). Transaction costs of tradable white certificate schemes: The energy efficiency commitment as case study. Energy Policy, 35(8):4340–4354.

North, D. C. (1990). Institutions, Institutional Change and Economic Performance. Cambridge University Press, Cambridge.

Ollama (2024). Llama 3.1. https://ollama.com/library/llama3.1. Accessed: 2026-05-05.

Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C. L., Mishkin, P., Zhang, C., Agarwal, S., Slama, K., Ray, A., Schulman, J., Hilton, J., Kelton, F., Miller, L., Simens, M., Askell, A., Welinder, P., Christiano, P., Leike, J., and Lowe, R. (2022). Training language models to follow instructions with human feedback. In Advances in Neural Information Processing Systems, volume 35.

Park, J. S., O'Brien, J. C., Cai, C. J., Morris, M. R., Liang, P., and Bernstein, M. S. (2023). Generative agents: Interactive simulacra of human behavior. In Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology, UIST '23. Association for Computing Machinery.

Park, J. S., Popowski, L., Cai, C. J., Morris, M. R., Liang, P., and Bernstein, M. S. (2022). Social simulacra: Creating populated prototypes for social computing systems. In Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology, UIST '22. Association for Computing Machinery.

Qu, Y. and Wang, J. (2024). Performance and biases of large language models in public opinion simulation. Humanities and Social Sciences Communications, 11(1095).

Shao, Z., Wang, P., Zhu, Q., Xu, R., Song, J., Bi, X., Zhang, H., Zhang, M., Li, Y. K., Wu, Y., and Guo, D. (2024). Deepseekmath: Pushing the limits of mathematical reasoning in open language models.

Williamson, O. E. (1981). The economics of organization: The transaction cost approach. American Journal of Sociology, 87(3):548–577.

Ziems, C., Held, W., Shaikh, O., Chen, J., Zhang, Z., and Yang, D. (2024). Can large language models transform computational social science? Computational Linguistics, 50(1):237–291.

APPENDIX

5.1 PTC Persona Definitions

The prompt uses two persona fields: a barrier persona that captures the respondent's perceived obstacles to renovation, and a benefit persona that captures the respondent's perceived value of renovation. These personas are used as compact PTC profiles in the shared survey prompt.

Financial Sensitive. This persona is generally supportive of renovation but highly sensitive to financial consequences, especially potential increases in rent or service charges. Practical disruptions are less concerning, but the respondent needs strong reassurance that renovation will not worsen their financial situation.

Practical and Financial Concerned. This persona perceives both financial and practical barriers as major obstacles. The respondent worries about temporary relocation, daily-life disruption, nuisance, and long renovation timelines, and may also feel overloaded or lack time and clear information. Step-by-step planning and logistical support are essential.

Ambivalent Observer. This persona feels mostly neutral about renovation barriers and tends to observe rather than actively engage. The respondent is not strongly opposed, but is not proactive either. They prefer minimal involvement and respond better to simple visuals or relatable messages than to detailed explanations.

Self-Reliant but Unconvinced. This persona often feels uncertain about both financial and practical aspects of renovation and may frequently think, "I don't know." The respondent trusts their own judgment but lacks a clear understanding of what renovation would mean for them. Personalized explanations, such as home visits, are more effective than generic campaigns.

Financially Alarmed. This persona is extremely worried about the financial impact of renovation. The respondent strongly believes it will increase rent, energy bills, and service charges. Practical disruptions such as relocation or delays intensify these concerns. Without firm financial guarantees, they are very likely to reject renovation.

Confident Acceptors. This persona perceives very few barriers to renovation. The respondent feels financially secure, trusts the process, and is willing to tolerate temporary inconvenience. They are often already familiar with alternative heating systems and could serve as a positive reference point for others.

The Uncertain. This persona frequently feels unsure when thinking about renovation and often responds with "I don't know." This reflects limited information or difficulty processing the topic rather than active opposition. The respondent benefits most from guided, personal support through trusted intermediaries or in-home assistance.

Immediate Utility Seekers. This persona cares strongly about immediate and tangible benefits from renovation, such as better indoor comfort, improved health, personal wellbeing, and lower energy consumption. Aesthetic upgrades or neighborhood-level improvements are less important. The respondent is generally positive about energy-efficiency renovation when short-term personal gains are explained clearly and concretely.

Personal Comfort Seekers. This persona mainly cares about how renovation affects personal comfort and health. Environmental benefits, building appearance, and neighborhood improvements matter little. The respondent tends to rely on trust-based reassurance and prefers discussion events or face-to-face explanations over purely digital communication.

Balanced Benefit Idealists. This persona values a broad range of renovation benefits. Personal comfort is important, but so are environmental responsibility and positive impacts on the commu-

nity. The respondent responds well to holistic narratives that connect personal wellbeing with social and environmental goals, and appreciates participatory or interactive engagement formats.

Ambivalent Respondents. This persona feels mostly neutral about the benefits of renovation. The topic does not feel particularly relevant or urgent, and the attitude reflects uncertainty rather than clear support or opposition. The respondent prefers simple, low-effort communication that clarifies why renovation should matter personally.

Pessimists. This persona generally does not care much about the proposed benefits of renovation. Most advantages feel irrelevant, although reducing energy consumption has some limited appeal. The respondent tends to avoid engagement, including emails or events, and benefit-focused messaging alone is unlikely to motivate them.