Reply on RC1

(Note: The symbol "➢" denotes the beginning of each response.)

Major comments

The biggest comment I have is that the paper needs to be more precise in how it discusses the LLM agents. I appreciate that anthropomorphizing them to some extent can be helpful, but the paper needs a better grounding in what is actually happening in the LLM. It's not until the Discussion that the actual mode of operation is actually discussed—these models generate a series of words based on a corpus of training data as well as the "conversation" that has occurred between the operator and model. This is an important enough concept that it should be described when LLMs are first mentioned in the Introduction and used to frame the results and discussion.

As a result, these models are not capable of "reasoning," a term which appears throughout the paper; they simply generate "the most likely next word." (See Bender et al., 2021, "On the dangers of stochastic parrots.") However, the paper frequently seems to buy in to the illusion. For example, at lines 314-316:

> While LLM models are often perceived as opaque, LLM-powered agents can offer the compelling ability to articulate human-comprehensible reasoning for their actions, providing a window into the decision-making processes that drive their behaviour.

It certainly is compelling and articulate, but the LLM is neither reasoning nor making decisions. I'm not just trying to be pedantic: This over-anthropomorphization of the model seems in some places to bleed into the authors' interpretation of their results. For example, at lines 280-281: "S1.2 seems to provide reasoning in hindsight to justify a decision made in the absence of such reasoning." It's not "reasoning" or "justifying" anything. It's simply producing the most plausible string of text given that it's already said "+2 tax increase." This is why the results can differ so much based on the order of "decision" vs. "explanation" (Agents S1.1 vs. S1.2) and should not have been surprising.

The authors also must add discussion of the potential biases that are possible with this kind of setup. One is that it'd be possible for the human in the loop to introduce their own bias when interacting with the LLM; I hardly consider this a dealbreaker, but it should be at least mentioned. More importantly, the outputs of an LLM will reflect any biases in its training dataset. It might thus be difficult to spur an LLM to enact policies that defy political-economic orthodoxy. It's going to be biased toward "conventional wisdom," making me skeptical of statements like "[the use of LLMs] provides an opportunity to search for novel insights into human behaviour" (lines 71-72) and "modellers can get useful inspiration from this communication" (line 498).

➢ **Response to Major Comments:**

1. Clarification on the Functioning of LLMs

We agree with your observation and will revise the Introduction and Results sections to clarify the mechanics of LLMs, emphasizing that they generate the "most likely next word" based on their training data and input prompts. We will introduce the concept of LLMs earlier in the Introduction and explicitly avoid language and interpretation that over-anthropomorphizes these models. For instance, we will replace terms like "reasoning" with "pattern generation", "simulated rationale", or other words that fit in the context more precisely.

2. Rephrasing language to avoid over-anthropomorphization

Your understanding of the stochastic text generation essence of the large language models is accurate. Although there exist a number of papers using the words “reasoning”, “thinking”, and “planning” loosely in the LLM research field, these models cannot conduct real reasoning like humans. They are often trained and fine-tuned on extensive high-quality text and can mimic language patterns, which can produce contextually coherent text that gives us an illusion of reasoning. However, contextual coherence does not equal logic consistency. We appreciate your suggestion and will handle this issue by carefully phrasing the language in the revised manuscript.

3. Discussing bias

It is true that biases can sneak into the model. You correctly highlighted two important sources of biases, including prompt design and training dataset. We will add a new subsection in the Discussion to explore potential biases. This will include biases introduced by the training data and the potential for human operators to influence results inadvertently. We will highlight how these biases might constrain the ability of LLMs to produce innovative or unconventional policy solutions.

Minor comments

- Fig. 1:
  - Does not seem to be referenced at all in text.
  - “Use LLM to assist with refinement” does not seem to be mentioned anywhere in the text.

➢ Thank you for pointing out this issue. Fig. 1 should be referenced in section 2.1. We will ensure that Fig. 1 is explicitly referenced in the revised manuscript.

➢ “Use LLM to assist with refinement” is related to “Utilizing ChatGPT as a drafting tool”. We will change this to “Utilizing tools powered by LLMs for drafting, such as ChatGPT” to explicitly link the figure to the text.

- Lines 141-143: I thought the agent was supposed to represent an institution, but here it says there’s an “institutional environment within which the agent operates.” Are those other institutions? Are there non-institutional agents as well?

➢ We agree that the word “institutional environment” may cause confusion. The “environment” actually means the programmed model that the institutional agents interact with. In this paper, the programmed model refers to the CRAFTY land use model, which cotains numerous rule-based agents to mimic land users. We will revise this sentence to improve its clarity.

- Lines 209-210: Add text explaining that the policy actions represent change in tax level.

➢ We will explicitly state in the Methodology section that policy actions are defined as changes in tax levels.

- Table 1:
  - What “experience” is being referred to by “experiential learning”? Is the agent looking at what its policy was for previous years and considering what changes to make?
  - Second sentence of Description box for Agent Q is unclear. I don’t know what sequencing the roles” or “conversational endpoint setting” are.

➢ The experience used by the agent refers to its previous output, including the policy actions and the simulated rationale behind these actions. We will improve the clarity of this part by explicitly explaining what means experience here.

➢ “sequencing the roles” means arranging the order of the agents involved in a conversation. As you correctly pointed out that the order of LLM outputs would influence the outcomes, who talks after whom might make differences in a multi-agent system.

“conversation endpoint settings” simply means how to end a conversation between the LLM agents. Because the agents are text generators, without carefully setting how a conversation should end, the agents may get trapped in an endless response loop. A simple example is that two agents say “goodbye” to each other forever. We will add explanations to address these words in the revised manuscript.

- Lines 217-220: For reproducibility, please give more details of the genetic algorithm setup in an Appendix or Supplement.

➢ We will add an appendix with a detailed description of the genetic algorithm, including its parameters and operational logic, to enhance reproducibility.

- Fig. 4:
  - Please avoid using red and green on the same figure, as these are hard to distinguish for people with the most common form of color blindness.
  - What are the Y-axis units?
  - What is “demand force”?

➢ We will revise the color scheme of all figures to be more accessible, ensuring they are distinguishable for readers with color blindness.

➢ In the current settings, the unit of production is omitted by normalization across different ecosystem services. Here, we only need the relative magnitude of production to see the influence of LLM agents. We will explain this explicitly in the figure caption.

➢ “Demand force” means the driving force of the demand that steers the change of meat production in the experiments. We will change it to “Demand” to avoid causing confusion.

- Lines 242-246: Is the average error in Fig. 5a (Agent B1) just a re-presentation of the data from Fig. 4? If so, please mention that in the text here. If not, please explain.

➢ Yes. Your observation is correct. Fig. 5a (Agent B1) is just a re-presentation of Fig.4. We will mention this in the revised manuscript.

- Sect. 3.2.1 (performance of Agent S1.1:
    - You're right that the policy actions are "generally understandable," but one weird thing is the drop to baseline taxation levels around years 35-50. Any idea why that happened? … I see now that this is discussed in Sect. 3.3, Action III. Please add a reference to that in Sect. 3.2.1.
    - Why are non-negative changes in taxation "plausible"?
    - Line 257: "Significant"—was there a statistical test? If not, please rephrase for clarity. If so, please explain.

➢ As per your suggestion, we will add a reference to Sect. 3.2.1 to mention the sudden drop is explained in Sect. 3.3.

➢ Non-negative changes are plausible because we set a challenge for the agents to apply taxes to maintain the current level of meat supply, which is driven to grow by the increasing demand. Ideally, the tax should increase as the gap becomes large but decrease to zero as the gap shrinks to be minimal. We will explain this explicitly in Sect. 3.2.1 in the revised manuscript.

➢ "Significant" will be changed to "noticeable" to improve clarity.

- Fig. 5: Having errors be negative when there's too much meat production feels counterintuitive. This is obviously just a personal preference, but in any case, the chosen convention should be mentioned in the figure caption.

➢ Thank you for your suggestion. The errors are negative because they are calculated as "policy goal minus meat supply". When the policy goal is lower than the meat supply, the error becomes negative. The target for the agents is to maintain the meat supply at the initial level to overcome the driving force of the increasing demand. As presented in Fig. 5, without any intervention, the land use model tends to produce meat to meet the increasing demand. When the policy goal is prominently below the meat supply, the error is negative, indicating the meat supply is more than expected, and it needs to be reduced. On the contrary, if the policy goal is higher than the meat supply, the error is positive, indicating the meat supply needs to be bolstered. We will ensure the chosen convention will be explained clearly in the figure caption.

- Lines 348-349: The way this is phrased makes it sound like the agent was given two goals: maintaining supply levels and matching supply to demand. In reality, it seems like it just made up the latter. Right?

➢ Yes, your understanding is accurate. The agent has only one goal, which is to maintain the initial meat supply level. The agent may sometimes misunderstand its target and generate erroneous outcomes. We will explicitly discuss this in the manuscript to ensure clarity.

- Lines 450-451: What "conventional methods" do you mean? Hard-coded agent behavior can't produce (the appearance of) reasoning, but in that case it doesn't need to—you already know the rules governing agent behavior.

➢ We agree that the comparison between rule-based and LLM-based agents is not fair enough. Instead of replacing one another, these two methods are much more complementary than competitive. We will modify the sentence here to clarify.

- Lines 460-461: "As the policy objective nears realisation, Agent S1.1 judiciously reduces tax levels to mitigate potential over-adjustment." Does it? I thought policy actions represented changes in tax levels. Because those aren't negative after the first few years, this means that taxes never go down.

➢ Your understanding of the policy actions representing tax changes is precise. The agents actually need to find an appropriate tax level to counterbalance the driving force of the growing demand in order to maintain the meat supply at the target level. The agents incrementally adjust the taxes until the taxes can offset the meat producers' benefit derived from the high market demand. For instance, if the demand for meat leads to 100 units of profit, then the tax should cause 100 units of loss. Hence, the tax levels should be increased to a proper level and maintained at that level (because the demand stops changing eventually in the simulations), which means the tax increase should ideally drop to zero if the policy goal is precisely met.

- Lines 490-492: Please explain why conventional modeling techniques can't represent these interactions.

➢ It might be difficult for conventional modelling techniques, such as rule-based agents, to model the interactions between institutional agents, such as lobbyists, law consultants, and research suppliers, because their interactions involve extensive unstructured information. For example, land user associations and environmental NGOs may have conflicting advocacies expressed in words, which are challenging to formalize using mathematical equations or code. Although we can simplify their interactions to fit conventional methods, it often involves oversimplification and abstraction. LLM agents provide a more straightforward way to simulate these unstructured interactions that are challenging to formalize. We will elaborate on this point in the revised manuscript.

Technical corrections / typos

- Line 199: "plausibility" is probably not the right word. "Desire for"?
- Line 327: Word missing (of?) in "investigation the large".
- Line 444: "conversions" should be "conversations".
- Line 472: Missing period.

➢ Thank you for pointing out these issues. We will correct these typos and rephrase them as suggested.