

On AI for Quantitative Social Science*

Carlisle Rainey†

October 27, 2026

The emergence of agentic AI¹ has radically changed how we can conduct quantitative social science.² It seems clear that we *should* use agentic AI in our work—critics can no longer dispute the benefits.

But how should we think about how AI models can help? How can we make sure that we do not *underuse* or *overuse* the models? I cannot answer these questions confidently. However, I can describe how I am *thinking about* these questions.

What We Do (or Perhaps “What We Did”)

In social science, we write papers to help other social scientists understand the world. For *quantitative* social science, I think of projects as having three parts:

1. A *dataset*.
2. A *paper*.
3. A *process* that takes the researcher from the dataset to the paper.

I’ll focus on this *process* in this essay.

At our best, the process consists of well-documented computer code that implements our principled decisions (that we carefully defend in the paper).

The process is *long*. A quick look across several replication archives in political science suggests that a typical project uses one to six scripts and 500 to 2,500 lines of code. Some use many more scripts and many more lines of code.

The process is *important*. We might even argue that the *process* is the primary research product. Buckheit and Donoho (1995, 5) write:

* *Initial draft; still very early*. This semester, I challenged my students to think about the role of AI in their research. This is my sketch of the promise and pitfalls.

† Carlisle Rainey is Professor of Political Science, Florida State University, 540 Bellamy, Tallahassee, FL 32306 (crainey@fsu.edu).

¹ Such as Anthropic’s Claude Code interface and Opus 5.5 model, as of October 2026.

² Perhaps I should write “computational” social science for this essay. That’s what I mean—social science that combines quantitative datasets, data wrangling, and statistical modeling.

Buckheit, Jonathan B., and David L. Donoho. 1995. “WaveLab and Reproducible Research.” In *Wavelets and Statistics*, eds. Anestis Antoniadis and Georges Oppenheim, 55–81. New York: Springer. Page numbers cited refer to the Stanford version, <https://web.stanford.edu/dept/statistics/cgi-bin/donoho/wp-content/uploads/2018/08/wavelab.pdf>

An article about computational science in a scientific publication is *not* the scholarship itself, it is merely *advertising* of the scholarship. The actual scholarship is the complete software development environment and the complete set of instructions which generated the figures.

So what should this process look like—what features should it have? I suggest three:

1. *Principled*: the researcher made reasonable decisions in analyzing their data.
2. *Implemented*: the researcher actually implemented those decisions.
3. *Documented*: the researcher made it easy to check that they actually implemented those decisions.

We could imagine, for example, that a researcher comes up with a clever research design but fails to code an important variable correctly. Or perhaps they did everything correctly, but they pointed-and-clicked their way to their final estimates and left no way to verify they did anything at all.

I’ve long advocated for what Healy (2020) calls “plain-text social science” because tools like CSV datasets, R scripts, and LaTeX files make it easier to document your process (and thus easier to implement). As Healy (2020, 8) notes:

[S]oftware applications are not all created equal, and some make it easier than others to do the Right Thing.

This leaves us with two uncontroversial claims:

1. We should make smart choices, implement those choices, and document those choices.
2. We should use tools that make it easy to do those things.

Enter AI

AI models enter almost seamlessly into this framework, especially when used agentially.

It is indisputable that AI models are tools that can help *implement* our choices and *document* those choices. They can write clear and correct R scripts. They can refactor and debug human-written R scripts. They can create careful records. They can document revisions.³

Healy, Kieran. 2020. “The Plain Person’s Guide to Plain-Text Social Science.” Version 2020-10-24. <https://plain-text.co>

³They can also help make principled decisions, but work in this dimension is more complicated, I think.

This is clear: *at a minimum*, we should use AI models *extensively* for error-checking (Aronow 2026).

Mistakes We Can Make

We can make two types of mistakes:

1. We can use AI too little.
2. We can use AI too much (or for the wrong things).

And for PhD students and junior faculty, it is important that they thread this needle.

Too Little

The risks of too little seem clear: we fail to discover and correct errors that AI models would have detected.⁴

Real examples of errors.

- *Example 1:* A paper describes a prior distribution as “ $N(0, 2)$.” From context, “2” refers to a variance. The Stan model uses `normal(0, 2)`, and Stan treats 2 as a standard deviation.
- *Example 2:* A paper describes a complex algorithm implemented in C++. A subtle but important difference exists between the algorithm described in the paper and the algorithm implemented by the C++ code.

While these errors are nearly impossible for a human to detect, an AI model will almost always discover them.

Too Much

The risks of too much seem less clear. But here is my thinking.

First, we might turn over tasks to AI models in which the human should remain deeply involved. I suspect that current models are not great at things like justifying the substantive credibility of assumptions used for causal inference. I suspect they are not great at finding open and important questions.⁵

Second, we might use AI models to quickly complete tasks that ought not be completed at all. Thus, rather than making less work, they make *more* work. It is clear that we can use agents to do lots of tasks very

Aronow, P. M. 2026. “A Paper without an LLM Disclosure Is Cause for Suspicion.” Unpublished manuscript, October 4, 2026. <https://pmaronow.github.io/llm-disclosure.pdf>

⁴Others seem to think that AI allows us to work faster, incorporate more data, or push further theoretically. Perhaps there is lots of opportunity there, as well.

⁵I don’t worry about researchers overusing models for these things, though. I’m actually optimistic that researchers care about the quality of the papers they put out under their name. I suspect we have (or will quickly develop) good ways to filter out “slop.”

quickly and with many fewer errors. What matters now is that we point those agents towards *worthwhile* tasks.⁶

An Attitude for the Future

Some researchers seem optimistic that agents will allow us to work faster, collect more data, or push theories further. Other researchers seem worried about being overwhelmed by slop. My feelings are different.

I'm excited about this: AI agents will allow us to work *more carefully*.⁷

And I'm worried about this: AI agents will allow us to pursue all sorts of computational projects that are not worth doing at all.

So here's an attitude for the future: let's continue our work. Let's continue to do work that's worth doing. Let's continue to write papers worth reading. And let's continue to use tools that help us do that work well (even if it slows us down a little).

⁶ Take the example of data collection. Prior to the emergence of agents, extensive data collection was difficult, expensive, and time-consuming. The high costs made us think carefully about what we wanted and why before we went ahead. The difficulty encouraged and rewarded discernment and judgment. With AI agents, extensive data collection is easy and cheap—you can do it while you sleep. The low costs make it easy to “just go ahead” before we've put any thought into the matter.

⁷ When I've used them best, I feel like agents have *slowed me down*, not sped me up, without a huge jump in theoretical or empirical ambition.