Those hoping that AI-related job loss will help AI Safety wrongly assume that people will mourn the death of high paid knowledge work.

When the UK imposed an insignificant inheritance tax increase on farmers, a large convoy of tractors, led by car-show presenter turned hobbyist farmer Jeremy Clarkson, descended upon London to protest. This is not a unique phenomenon. From the boulevards of Paris to the highways of Haryana, farmers have consistently mobilised in the face of hostile legislation, often with great success. Each of the three protests mentioned here resulted in the government either rolling back or pausing future legislation.

Farmer's protests are a visceral example of the power of worker mobilisation in bringing about legislative change. Many within AI Safety think it is similarly worth harnessing workers whose jobs have been automated to push for AI Safety related legislation. Some have already written about potential harms of this approach, arguing that mass movements focused on job loss might not push for measures that improve AI Safety.

This piece takes a different approach. I argue that proponents of AI job-loss based movements often wrongly assume that those who initially lose their jobs because of AI are groups with high amounts of political cachet. This will likely not be the case. Even if a mass movement focused on AI job loss adopted the right policies, it would by no means be likely to succeed.

Which jobs are most likely to be threatened by AI initially? We should expect routine knowledge work to be most prone to AI disruption. These jobs can be done fully remotely, and involve repetitive tasks that the AI can be trained to excel at. These jobs include customer service, routine HR work and so on.

Importantly, much entry-level high-paying white-collar work also takes this form. Software engineering, management consulting and radiologists are examples of high paid jobs which are likely to be disrupted by AI.[1] OpenAI is employing more than a 100 former bankers to help train a model to automate junior banking work. Unlike previous waves of automation, AI related disruption will likely affect high paid white collar work. It's not hard to see why; the promise of laying off high earning workers is an alluring one. Some people are already changing their behaviour as a result.

If the core of the AI labour movement is to be formerly high earning white collar workers, we should be much more sceptical of the movement's success. These sorts of High Earners who are Not Rich Yet are some of the most maligned figures in politics. They both face the highest marginal tax rates and receive little benefits from the state in return. They are also often not unionised, because the benefits of these unions seemed low previously. Yet even in researching this piece it has been hard to feel much in return. Mourning the loss of childcare benefits for someone who has started earning £100000 a year is a difficult thing to do.

This group of people, most exposed to AI automation, is poorly placed to bring about policy change. I suspect many will only feel a sense of schadenfreude when newly minted university graduates can no longer glide into their high paying jobs. Even if we also included workers in lower paying knowledge work, this will not lead to an AI labour movement that can bring about effective policy. Much call centre labour, for instance, has been offshored to countries like India and the Philippines. These countries, and thus their workers, will likely have limited leverage over frontier AI capabilities.

---

[1] Interestingly, though radiology seems amenable to automation, mass displacement has not yet occurred. For For interesting discussion see here.

The harsh reality is simply that some industries carry more salience than others. Sectors like farming, heavy manufacturing and steel production have plenty of political cachet. Images of helping ordinary blue collar works increasingly motivate both sides of the political divide, from JD Vance to Bernie Sanders. Whilst it's true that industries like banking also have political influence via lobbying, this influence mostly extends to protecting the profit margins of banks, rather than their workers. In other words, insofar as industries like banking have political capital, this will mostly be exerted on enabling further automation. Nobody will mourn the death of entry level investment banking.

One objection to this view is that even if AI begins by automating jobs like software engineering, rapid increases in capabilities will mean that even those jobs with greater political credit will be on the chopping block rapidly. Against this, I would contend that in worlds where AI related unemployment spreads rapidly through the economy, workers will be disempowered so rapidly that their objections become moot. If AI has automated away most forms of labour already, then it becomes hard to see how an anti-AI job movement could achieve much. Slowing the pace of AI at that point would have too high a cost. For an anti-AI job movement to effectively move the dial on AI Safety, slowing down AI development has to be a cost that governments are willing to bear. At the point that much of the economy has been automated, this seems unlikely.

In conclusion, the politics of AI job loss are far more complex than they might seem at first glance. Whilst proponents of this sort of movement for AI Safety assume that it would have a high amount of political leverage, this may not be the case. There are two relevant scenarios I have discussed so far:

1. AI-caused unemployment spreads slowly by sector. In this case, the first groups to lose their jobs (those in highly paid entry level knowledge work, or low paid service sector jobs) lack the political capital to form an effective labour movement.
2. AI-caused unemployment spreads rapidly by sector. In this case anti-AI measures have a high economic cost. Governments would be less likely to act in this case.

Either way, the AI job movement faces significant political headwinds, and we should be sceptical of this approach as a way of bringing about AI Safety. Management consultants will not be an effective vanguard for the AI Safety revolution.