THE GLOBAL PRIORITIES INSTITUTE

ECONOMICS RESEARCH AGENDA

Version 1

November 2024

GLOBAL PRIORITIES INSTITUTE

UNIVERSITY OF OXFORD

Table of Contents

Introduction

There are many problems in the world. Because resources are scarce, it is impossible for any given actor to solve them all. A government, philanthropist or individual seeking to improve the world therefore needs to prioritise, both among the problems themselves and among policies and interventions for addressing them. This task of prioritisation requires careful analysis. Some opportunities are likely to be vastly more cost-effective than others. Identifying such opportunities—focus areas, policies, and interventions—requires grappling with a host of complex questions.

The aim of the Global Priorities Institute (GPI) is to conduct foundational academic research that informs the decision-making of individuals and institutions seeking to do the most good in the world. In particular, we focus on research that makes progress towards figuring out what the world's most pressing problems are and how these problems can be solved.

This document outlines some of the core research priorities for the economics team at GPI. Section 1 centres on general issues in global prioritisation. This includes empirical and theoretical questions about e.g. forecasting, optimal philanthropy, and the distribution of cost-effectiveness across interventions, as well as normative questions related to welfare criteria and decision procedures. Section 2 centres on applied issues where further research in economics may be particularly impactful such as the economics of growth, population, inequality, catastrophic risks and artificial intelligence.

Within each research topic, we have aimed to list examples of research questions that are both decision-relevant for altruistic actors *and* that have been relatively unexplored in prior research. The emphasis on topics that are unexplored in prior research is motivated by the belief that there are diminishing returns to research and that further research on these topics may therefore be particularly likely to yield insights that are important and decision-relevant.

The current version of the document is a draft and subject to revision. The topics and questions that have been included only constitute a fraction of research questions that fall within the scope of GPI's mission. We stress that in most cases, what we list here are relatively broad research themes, rather than the more specific research questions that would naturally correspond to individual research papers. Within each such theme, the first step is to do significant further work identifying and articulating the most fruitful and impactful research questions.

1. General issues in global prioritisation

1.1 Strategic issues in altruistic decision-making

Altruistic decision-making (whether by governments, philanthropists or individuals) raises several difficult considerations that are not raised to the same extent by self-interested decision-making. For example, altruistic decision-makers must take into account the possibility of crowding out efforts by other altruists. Although such considerations may be of great importance for the effectiveness of altruistic efforts, they have received relatively little attention in economics compared to considerations that loom large in the analysis of self-interested and profit-maximising behaviour. We list some of the key research questions on optimal philanthropy below.

- Funding a charitable intervention may increase or decrease the funding that the intervention receives from other actors. Empirically, what is the magnitude of such displacement effects for charitable causes that are particularly impactful? (C.f. Andreoni & Payne 2003).

- Altruistic decision-makers may have preferences over the aggregate supply of various goods and the consumption utility of other agents in the economy. Both of these depend on prices, which can depend in complex ways on individual spending behaviour. In light of this, how can we better understand the optimal spending behaviour of decision-makers with such external preferences? (Kaufmann *et al.* 2024; Trammell 2024; Wilkinson 2022).

- Altruistic decision-makers must choose what fraction of their resources they wish to spend immediately and what fraction they wish to invest for the sake of future spending. What is the optimal spending schedule for altruistic decision-makers? (Andreoni 2018; Trammell 2021)

- Strategic interaction between altruistic decision-makers with heterogeneous beliefs or preferences may result in coordination failures (Kalai & Kalai 2001). What mechanisms can be used to coordinate funding decisions? (Brandl *et al.* 2022)

- Small actors (e.g. individuals donating to charity) may maximize their expected altruistic impact by exclusively funding the charitable intervention that they deem to have the highest expected cost-effectiveness. In contrast, larger

actors (e.g. governments or philanthropic foundations) typically diversify their spending to account for diminishing marginal returns and the information value obtained from funding new interventions. What is the optimal degree of funding diversification for different types of altruistic decision-makers? (Snowden 2019)

1.2 Forecasting

When prioritising resources, we often need to judge effectiveness without solid empirical evidence. This problem is especially severe when assessing unprecedented risks (e.g. from new technologies) and evaluating interventions to mitigate them. Forecasting could potentially partially solve this problem, to the extent to which subjective estimates accurately estimate crucial quantities. These subjective estimates might come from the decision-maker or groups of other forecasters and might be supported by formal statistical models. However, there are several open questions about how to optimally select forecasters, how to reward them, how to exchange information between them, and how to aggregate different forecasts into a single decision-guiding judgement. Such questions include:

- What are the most informative forecasting methods for assessing the probability of global catastrophic risks and other future events of special importance for long-term social welfare? (Karger *et al.* 2023)

- Is it possible to incentivise accuracy in long-term forecasts, given that we may not live to see the realisation of the event we wish to forecast? Can we make use of incentives to report 'subjective truths' for this purpose? (Prelec 2004) To what extent does learning from the short-run forecasts for which we can observe accuracy translate to other domains?

- Forecasts of standard econometric time series models will often collapse into simple predictions when applied to the far future (Granger and Jeon 2007). Are these predictions still informative for long-term decision-making? Can econometric tools be adjusted to make more substantive long-run forecasts?

- In at least some modelling frameworks, short-term forecasting may be more important than long-run forecasting due to the possibility of gradually adapting to the future (Millner and Heyen 2019). In which settings do the conditions of these models apply, and how widely applicable is this conclusion?

See also section 2.3 (Forecasting) in GPI's research agenda in psychology.

1.3 Welfare and decision procedures

Assessing and comparing the social value of interventions requires a welfare criterion. When information is incomplete, it also requires a normatively compelling account of decision-making under uncertainty. Although these issues have been studied extensively in welfare economics and normative decision theory, there are still many decision-relevant issues in these fields that have been largely neglected. We list some of these issues below.

- The value of normative parameters are crucial for altruistic priority-setting. Examples of such parameters include the social rate of pure time preference, the elasticity of marginal utility to consumption, the social value of existing and future lives, and the zero (neutral) level of wellbeing. How should these parameters be determined? (Greaves 2017; Groom *et al.* 2022)

- It has been argued that welfare analysis should sometimes take the wellbeing of non-human animals into account (e.g., Espinosa and Treich, 2024). In particular, this may go beyond evaluations of consumer willingness to pay to purchase products like free-range rather than caged eggs. How should non-human animals be factored into welfare evaluations? See also section 4.4.1 (Animal ethics) of GPI's research agenda in philosophy.

- Suppose that a decision maker wishes to maximise the expectation of some social welfare function. When the space of possible actions and possible states is complex, explicit maximisation of expected utility may be intractable (cf. Camara 2022). What heuristics should a social decision-maker use to approximate expected social welfare maximisation? See also section 1.2 (Epistemic challenges) of GPI's research agenda in philosophy.

- Decision-makers may be uncertain about normative issues such as which welfare criterion to use or how to determine the value of important normative parameters (see e.g., Millner 2020). Prior literature has used tools from decision theory (Dietrich and Jabarian 2022), social choice (Tarsney 2019) or bargaining theory (Greaves and Cotton-Barrat 2023) to provide methods for handling such normative uncertainty. What are the best procedures for

handling normative uncertainty? See also section 4.3.3 (Moral uncertainty) of GPI's research agenda in philosophy.

- Most of the economics literature on welfare criteria has focused on environments with a fixed finite population. Many real-world policies and charitable interventions are, however, likely to affect the size of the population. Moreover, if the universe is spatially or temporally infinite, there might be an infinite number of individuals that should be assigned positive ethical weight (Pivato 2023). How should welfare criteria be extended to settings with variable or infinite populations? (Blackorby *et al.* 1995; 2005; Asheim 2010). See also section 4.3.1 (Population Ethics and Aggregation) and section 4.1.2 (Infinities) of GPI's research agenda in philosophy.

See also section 1.1 (Understanding people's values) and section 8 (Future wellbeing) of GPI's research agenda on psychology.

1.4 Evidence and cost-effectiveness

Cost-effectiveness analysis is crucial for prioritising resources aimed at promoting social welfare. However, it also raises important empirical, statistical and normative questions. Resolving these questions is likely to have important practical implications for global priority-setting. While some of the key questions concerning cost-effectiveness have been studied in detail (e.g., discounting), others have been relatively neglected in the economics literature. We list some of the key conceptual questions in cost-effectiveness analysis below.

- The distribution of cost-effectiveness within and between charitable causes may have important implications for priority setting. For instance, it could provide a Bayesian prior for adjusting noisy cost-effectiveness estimates of interventions that have not yet been thoroughly evaluated. Empirically, what is the distribution of cost-effectiveness within and between the most impactful charitable causes, and what does this imply for priority setting? (Jamison et al. 2006; Vivalt 2015; 2020)

- Many charitable interventions may have significant externalities on individuals other than the direct recipients of the intervention (e.g., Egger *et al.* 2022). How should such indirect effects be incorporated in cost-effectiveness analysis?

- How do social welfare gains from improvements in people's quality of life (e.g., improvements in health or consumption) compare to social welfare gains from extending people's lives or bringing new people into existence? (Cf. Cookson *et al.* 2021; Luyten *et al.* 2022) Relatedly, what are the primary constitutive elements of wellbeing?

- How much do findings from impact evaluations generalise and how should policymakers update their beliefs about optimal policy when obtaining evidence from impact evaluations? (Vivalt 2020)

2. Applications to some important issues

2.1 Economics of catastrophes

Global catastrophic risks are often defined as threats that could lead to the deaths of at least a tenth of the global population or have a comparable social impact. Potential examples of such risks include severe pandemics, asteroids, nuclear war, climate change and risks from emerging technologies. A global catastrophe would not only be disastrous for the current generation, but may also have negative effects on the welfare of future generations. Catastrophic risk mitigation is therefore not only a global public good, but also an intergenerational public good, suggesting that it is likely to be severely underprovided by markets and governments. Nonetheless, the economics literature on catastrophes is small and almost exclusively focused on climate change. We list some of the key research questions on the economics of global catastrophic risks below.

- Large disasters stemming from pandemics, asteroids and military conflicts have occurred throughout history. Based on the historical record of such events, what is the tail distribution of harmful impacts (e.g., fatalities) from pandemics, asteroids, wars, and other potential disasters? (E.g., Marani *et al.* 2021; González-Val 2016).

- Catastrophic events may have persistent effects on the welfare of future generations. What are the long-term effects of catastrophes on population levels, economic output and social welfare? (E.g., Jordà *et al.* 2022; Alexandrie and Eden, forthcoming).

- There is a wide range of proposed interventions that may reduce global catastrophic risks. How much, if anything, should society be willing to pay to reduce the probability of global catastrophes and how cost-effective are the proposed interventions? (E.g., Pindyck & Wang 2013; Shulman & Thornley, forthcoming).

- It has been suggested that the risk of catastrophic events might decline over time as economic growth reduces people's marginal valuation of consumption and increases their marginal valuation of life (Cf. Jones 2016; Trammell and Aschenbrenner 2024). Taking considerations such as this into account, how is

the annual risk of global and existential catastrophes likely to develop over time?

- Managing an economy during or after a catastrophic event is likely to raise unprecedented policy challenges. What can be done now to ensure that societal institutions respond appropriately to such challenges?

See also section 1.3.1 (Extinction and Other Catastrophic Risks) of GPI's research agenda in philosophy and section 9 (Cause-Specific Issues) in GPI's research agenda in psychology.

2.2 Economics of artificial intelligence

Advances in artificial intelligence (AI) may be transformative for the global economy and political institutions. This raises important questions about the future of technological development and governance of AI. While economists have conducted important research on the implications of AI for the future of labour markets, these broader topics have been studied in much less detail. We list some key research questions at the intersection of economics and AI below.

- While some have suggested that AI may lead to explosive economic growth, others have argued that growth will be bottlenecked by the failure to automate some human tasks. How likely is it that such bottlenecks could prevent explosive economic growth? (Aghion *et al.* 2017; Erdil and Besiroglu 2023).

- Potential risks from advanced AI systems have led some to call for stronger regulations of AI, including a moratorium on some types of AI research. How should governments approach regulating potentially transformative technologies when information may be learned through their roll-out, or given political economy considerations? (Acemoglu and Lensman 2024, Koh and Sanguanmoo 2024)

- Advances in AI systems may result in unprecedented power for non-state entities as well as cement or exacerbate inequalities. How do economic models predict the impact of advanced AI systems on political institutions and inequalities, and how should political institutions govern powerful non-state entities? (Trammell and Korinek 2023).

See also section 3 (Artificial Intelligence: Risks and Opportunities) of GPI's research agenda in philosophy.

2.3 Modelling artificial agents

The possibility that AI may at some point reach or exceed human-level intelligence raises a new type of principal-agent problem stemming from the fact that the interests of AIs and humans may not be aligned. Given that the problem of incentive provision is at the heart of modern microeconomic theory, tools from this literature could have insightful applications to AI.

- In the future, humans may afford ever greater power to AI systems, who may in turn provide information to humans. Can we design mechanisms to ensure that AI systems exhibit desirable behaviours such as truth-telling and a lack of deception? (Che *et al.* 2024; Oesterheld et al. 2023; Chambers *et al.* 2019).

- If we are to gradually delegate more tasks to increasingly capable AI, how should we design the dynamic monitoring of such agents? How should we adapt the "contract" agreed with the AI? (Hadfield-Menell & Hadfield 2019; Ekmekci *et al.* 2022).

- Modelling AI agents raises decision-theoretic questions. Some have proposed to use causal models à la Pearl to model optimising AI agents (Everitt *et al.* 2021; Hammond *et al.* 2023; Spiegler 2016). What should be the rational benchmark, for instance, when facing uncertainty or Newcomb-like situations (Cerreia-Vioglio *et al.* 2011; Oesterheld 2021)? How can we model "corrigible" agents that have no incentives to manipulate, resist, or deceive their operators? These agents may be induced to change their preferences over time (Soares *et al.* 2015; Dietrich & List 2013; Boissonnet *et al.* 2023, Thornley 2024).

- Interactions between AI systems pose novel multi-agent safety problems beyond aligning a single agent to our goals. Moreover, AI interactions have specific features (e.g., much greater mutual transparency, ability to simulate each other, greater commitment power) which raises new questions: will they cooperate/collude better? Will unusual equilibria emerge? (Calvano *et al.* 2020; Banchio & Mantegazza 2022; Conitzer & Oesterheld 2023; Kovarik *et al.* 2024)

See also section 3.1 (Catastrophic Risk from AI) of GPI's research agenda in philosophy.

2.4 Population, inequality and long-term welfare

One of the most important developments of the last hundred years is the massive growth in global income per capita, both at the median and at the frontier, and in population levels that has occurred since the industrial revolution. Although these topics have been studied extensively in economics, there remain questions that have been relatively unexplored. We list some key research questions regarding economic growth and population growth below.

- How does long-run social welfare relate to key macroeconomic variables such as economic growth, population growth and inequality? (E.g. Asheim and Weitzman 2001; Stevenson and Wolfers 2013; Jones and Klenow 2016; Adhami *et al.* 2024) How can welfare improvements be sustained over time and expanded broadly?

- How likely is it that the economy will transition to a new growth regime with a significantly higher or lower rate of economic growth? (E.g. Gordon 2017; Aghion *et al.* 2018; Roodman 2020; Davidson 2021; Nordhaus 2021; Jones 2022)

- Population growth may have both positive externalities (e.g. increased idea production, economies of scale) and negative externalities (e.g. increased resource depletion, animal welfare concerns). In light of such externalities, what is the marginal social value of a larger population size? (E.g. Dasgupta 2019; Eden and Kuruc 2019; 2023; Kuruc 2024)

- What is the impact of economic growth, population growth and/or inequality on global catastrophic risks? (E.g. Baranzini and Bourguignon 1995; Jones 2016; 2023; Trammell and Aschenbrenner 2024)

2.5 Intergenerational governance and policy-making

Policymakers and institutions might fail to take the best actions accounting for future generations' welfare, either because of limited information about these actions' effects, or because of limited concern for future welfare. Moreover, current and future generations might face conflicting incentives. In evaluating the long-term consequences of policies and actions, it is therefore important to reckon with questions of how to influence the behaviour of future policymakers, and how to 'coordinate' optimally in the face of constraints on that influence.

- What influences when and how evidence reaches key decision-makers and how does evidence update policymakers beliefs (Vivalt and Coville 2023; Vivalt, Coville, KC 2024)? Can policy tools or interventions be introduced to promote evidence utilisation (Mehmood et al 2023; Toma and Bell 2023)?

- What incentives or commitment devices could be provided for an impatient government to implement policy consistent with placing a higher valuation on the future? What factors predict a more or less patient government (Bonfiglioli and Gancia 2013; Healy and Malhorta 2009; Rogoff 1990; Jacobs 2016)? Can long-term intergenerational mechanisms help to overcome this principal-agent problem? What might they look like (Rangel 2003; Harstad 2020)?

- How could future generations be represented in political decision-making (Gonzalez-Ricoy & Gosseries 2016)? How effectively can contemporary individuals act as representatives for future generations?

- Under what conditions should a social planner preserve 'option value' by delaying an important, irreversible decision to acquire more information, thereby delegating decision-making authority to future agents with potentially different values and preferences (cf. Bishop 1982; Dixit and Pindyck 1994)? How can evidence be disseminated most effectively?

See also section 1 (The Long-Term Future) of GPI's research agenda in philosophy and section 7.1 (Psychological obstacles to effectively improving the future in policy contexts) of GPI's research agenda in psychology.

References

Acemoglu, D., & Lensman, T. (2024). Regulating transformative technologies. *American Economic Review: Insights, 6(3),* 359-376.

Adhami, M., Bils, M., Jones, C. I., & Klenow, P. J. (2024). Population and Welfare: The Greatest Good for the Greatest Number (No. w31999). National Bureau of Economic Research.

Aghion, P., Jones, B. F., & Jones, C. I. (2017). Artificial intelligence and economic growth (Vol. 23928). Cambridge, MA: National Bureau of Economic Research.

Alexandrie, G., & Eden, M. (forthcoming). Is extinction risk mitigation uniquely cost-effective? Not in standard population models. In Barratt, Greaves, Thorstad (eds.) *Essays on Longtermism*.

Andreoni, J. (2018). The benefits and costs of donor-advised funds. *Tax Policy and the Economy*, *32*(1), 1-44.

Andreoni, J., & Payne, A. A. (2003). Do government grants to private charities crowd out giving or fund-raising?. *American Economic Review*, *93*(3), 792-812.

Asheim, G. B. (2010). Intergenerational equity. *Annual Review of Economics*, *2*(1), 197-222.

Asheim, G. B., & Weitzman, M. L. (2001). Does NNP growth indicate welfare improvement? Economics Letters, 73(2), 233-239.

Baranzini, A., & Bourguignon, F. (1995). Is sustainable growth optimal?. International Tax and Public Finance, 2, 341-356.

Banchio, M., & Mantegazza, G., (2022). Artificial Intelligence and Spontaneous Collusion, Papers 2202.05946, arXiv.org.

Bishop, R. C. (1982). Option Value: An Exposition and Extension. Land Economics, 58(1), 1. doi: 10.2307/3146073

Blackorby, C., Bossert, W., & Donaldson, D. (1995). Intertemporal population ethics: critical-level utilitarian principles. *Econometrica: journal of the Econometric Society*, 1303-1320.

Blackorby, C., Bossert, W., & Donaldson, D. J. (2005). *Population issues in social choice theory, welfare economics, and ethics* (No. 39). Cambridge University Press.

Boissonnet, N., Ghersengorin, A., Gleyze, S. (2023). Revealed deliberate preference change. *Games and Economic Behavior 142, 357-367.*

Bonfiglioli, A., & Gancia, G. (2013). Uncertainty, electoral incentives and political myopia. The Economic Journal, 123(568), 373-400.

Brandl, F., Brandt, F., Greger, M., Peters, D., Stricker, C., & Suksompong, W. (2022). Funding public projects: A case for the Nash product rule. *Journal of Mathematical Economics*, *99*, 102585.

Calvano, E., Calzolari, G., Denicolò, V., & Pastorello, S. (2020). Artificial Intelligence, Algorithmic Pricing, and Collusion. *The American Economic Review*, *110*(10), 3267–3297.

Camara, M. K. (2022, July). Computationally Tractable Choice. In Proceedings of the 23rd ACM Conference on Economics and Computation (pp. 28-28).

Cerreia-Vioglio, S., Ghirardato, P., Maccheroni, F., Marinacci, M., Siniscalchi, M. (2011). Rational preferences under ambiguity. *Econ Theory* 48, 341–375.

Chambers, C.P., Healy, P.J., Lambert, N.S., (2019). Proper scoring rules with general preferences: A dual characterization of optimal reports. *Games and Economic Behavior, 117, 322-341.*

Che, Y. K., Kim, J., & Mierendorff, K. (2024). Predictive enforcement. arXiv preprint arXiv:2405.04764.

Conitzer, V., & Caspar Oesterheld, C. (2023). Foundations of cooperative AI. In Proceedings of the Thirty-Seventh AAAI Conference on Artificial Intelligence and Thirty-Fifth Conference on Innovative Applications of Artificial Intelligence and Thirteenth Symposium on Educational Advances in Artificial Intelligence (AAAI'23/IAAI'23/EAAI'23), Vol. 37. AAAI Press, Article 1722, 15359–15367.

Cookson, R., Skarda, I., Cotton-Barratt, O., Adler, M., Asaria, M., & Ord, T. (2021). Quality adjusted life years based on health and consumption: A summary wellbeing measure for cross-sectoral economic evaluation. *Health economics, 30*(1), 70-85.

Davidson T. 2021. *Could advanced AI drive explosive economic growth?* Open Philanthropy, June 25. https://www.openphilanthropy.org/could-advanced-ai-drive-explosive-economic-growth

Dasgupta, P. (2019). *Time and the generations: population ethics for a diminishing planet*. Columbia University Press.

Dietrich, F., & List, C. (2013). Where do preferences come from?. International Journal of Game Theory, 42, 613-637.

Dietrich, F., & Jabarian, B. (2022). Decision under normative uncertainty. *Economics & Philosophy*, *38*(3), 372-394.

Dixit, A. K., & Pindyck, R. S. (1994). Investment under uncertainty. Princeton, NJ: Princeton University Press.

Eden, M., & Kuruc, K. (2019) *The long-run relationship between per capita incomes and population sizes*, CEPR Discussion Paper DP19353.

Eden, M., & Kuruc, K. (2023). Marginal Benefits of Population: Evidence from a Malthusian Semi-Endogenous Growth Model. PWI Working Paper.

Egger, D., Haushofer, J., Miguel, E., Niehaus, P., & Walker, M. (2022). General equilibrium effects of cash transfers: experimental evidence from Kenya. *Econometrica, 90*(6), 2603-2643.

Ekmekci,M., Gorno, L., Maestri, L., Sun, J., and Wei, D. (2022). "Learning from Manipulable Signals." *American Economic Review*, 112 (12): 3995–4040.

Erdil, E., & Besiroglu, T. (2023). *Explosive growth from AI automation: A review of the arguments*. arXiv preprint arXiv:2309.11690.

Espinosa, R. & Treich, N. (2024). Animal welfare as a public good. *Ecological Economics, 216,* 108025.

Everitt, T., Carey, C., Langlois, E.D., Ortega, P.A., Legg1, S. (2021). Agent Incentives: A Causal Perspective. *The Thirty-Fifth AAAI Conference on Artificial Intelligence (AAAI-21)*.

González-Val, R. (2016). War size distribution: Empirical regularities behind conflicts. *Defence and Peace Economics*, *27*(6), 838-853.

González-Ricoy, I., & Gosseries, A. (Eds.). (2016). Institutions for future generations. Oxford University Press.

Gordon, R. (2017). The rise and fall of American growth: The US standard of living since the civil war. Princeton University Press.

Granger, C., & Jeong, Y. (2007). Long-term forecasting and evaluation. *International Journal of Forecasting, 23*(4), 539-551.

Greaves, H. (2017). Discounting for public policy: A survey. *Economics & Philosophy, 33*(3), 391-439.

Greaves, H., & Cotton-Barratt, O. (2023). A bargaining-theoretic approach to moral uncertainty. *Journal of Moral Philosophy, 1*(aop), 1-43.

Groom, B., Drupp, M. A., Freeman, M. C., & Nesje, F. (2022). The future, now: a review of social discounting. *Annual Review of Resource Economics, 14*(1), 467-491.

Hadfield-Menell, D., & Hadfield, G.K., (2019). Incomplete Contracting and AI Alignment. In *AAAI/ACM Conference on AI, Ethics, and Society (AIES '19), January 27–28, 2019, Honolulu, HI, USA.*

Hammond, L., Fox, J., Everitt, T., Carey, R., Abate, A., Wooldridge, M. (2023). Reasoning about causality in games. Artificial Intelligence Volume 320, 103919.

Harstad, B. (2020). Technology and time inconsistency. Journal of Political Economy, 128(7), 2653-2689.

Healy, A., & Malhotra, N. (2009). Myopic voters and natural disaster policy. American Political Science Review, 103(3), 387-406.

Jacobs, A. M. (2016). Policy making for the long term in advanced democracies. Annual Review of Political Science, 19(1), 433-454.

Jamison, D. T., Breman, J. G., Measham, A. R., Alleyne, G., Claeson, M., Evans, D. B., ... & Musgrove, P. (Eds.). (2006). *Disease control priorities in developing countries (volume 2).*

Jones, C. I. (2016). Life and growth. *Journal of Political Economy*, *124*(2), 539-578.

Jones, C. I. (2022). The past and future of economic growth: A semi-endogenous perspective. *Annual Review of Economics*, *14*, 125-152.

Jones, C. I. (2023). The AI dilemma: Growth versus existential risk. Preliminary draft.

Jones, C. I., & Klenow, P. J. (2016). Beyond GDP? Welfare across countries and time. *American Economic Review*, *106*(9), 2426-2457.

Jordà, Ò., Singh, S. R., & Taylor, A. M. (2022). Longer-run economic consequences of pandemics. *Review of Economics and Statistics*, *104*(1), 166-175.

Kalai, A., & Kalai, E. (2001). Strategic Polarization. *Journal of Mathematical Psychology, 45*(4), 656–663

Karger, E., Rosenberg, J., Jacobs, Z., Hickman, M., Hadshar, R., Gamin, K., ... & Tetlock, P. E. (2023). *Forecasting Existential Risks Evidence from a Long-Run Forecasting Tournament*. FRI Working Paper No. 1.

Kaufmann, M., Andre, P., & Kőszegi, B. (2024). Understanding markets with socially responsible consumers. *The Quarterly Journal of Economics,* qjae009.

Koh, A. & Sanguanmoo, S. (2024), Robust regulation of technology. Working paper, arXiv:2408.17398.

Kovarik, V., Caspar Oesterheld, C., & Conitzer, V. (2024). Recursive Joint Simulation in Games. Working paper, arXiv:2402.08128v2.

Kuruc, K. (2024). Population Growth, Ideas and the Speed of History.

Luyten, J., Verbeke, E., & Schokkaert, E. (2022). To be or not to be: future lives in economic evaluation. *Health economics*, *31*(1), 258-265.

Mehmood, S., Naseer, S., Chen, D., (2023). Training Policymakers in Econometrics. *Working Paper*.

Marani, M., Katul, G. G., Pan, W. K., & Parolari, A. J. (2021). Intensity and frequency of extreme novel epidemics. *Proceedings of the National Academy of Sciences*, *118*(35), e2105482118.

Millner, A. (2020). Nondogmatic social discounting. *American Economic Review*, 110(3), 760-775.

Millner, A., & Heyen, D. (2021). Prediction: The Long and the Short of It. *American Economic Journal: Microeconomics, 13*(1), 374-98.

Nordhaus, W. D. (2021). Are we approaching an economic singularity? information technology and the future of economic growth. *American Economic Journal: Macroeconomics*, *13*(1), 299-332.

Oesterheld, C. (2021). Approval-directed agency and the decision theory of Newcomb-like problems. *Synthese* 198 (Suppl 27), 6491–6504.

Oesterheld, C., Treutlein, J., Cooper, E., & Hudson, R. (2023) *Proceedings of the Thirty-Ninth Conference on Uncertainty in Artificial Intelligence*, PMLR 216:1564-1574.

Pindyck, R. S., & Wang, N. (2013). The economic and policy consequences of catastrophes. *American Economic Journal: Economic Policy*, *5*(4), 306-339.

Pivato, M. (2023). Cesàro average utilitarianism in relativistic spacetime. Social Choice and Welfare, 1-29.

Prelec, D. (2004). A Bayesian truth serum for subjective data. science, 306(5695), 462-466.

Rangel, A. (2003). Forward and backward intergenerational goods: Why is social security good for the environment?. American Economic Review, 93(3), 813-834.

Roodman, D. (2020). *On the probability distribution of long-term changes in the growth rate of the global economy: An outside view*. Open Philanthropy Project.

Rogoff, K. (1990). Equilibrium Political Budget Cycles. The American Economic Review, 80(1), 21–36.

Shulman, C., & Thornley, E. (forthcoming). How Much Should Governments Pay to Prevent Catastrophes? Longtermism's Limited Role. In Barratt, Greaves, Thorstad (eds.) *Essays on Longtermism*.

Snowden, J. (2019). 'Should we give to more than one charity?' In *Effective Altruism: Philosophical Issues*, edited by Theron Pummer and Hilary Greaves. Oxford: Oxford University Press.

Spiegler, R. (2016). Bayesian Networks and Boundedly Rational Expectations. *The Quarterly Journal of Economics*, Volume 131, Issue 3, Pages 1243–1290.

Soares, N., Fallenstein, B., Yudkowsky, E., Armstrong, S. (2015). Corrigibility. *In AAAI Workshops: Workshops at the Twenty-Ninth AAAI Conference on Artificial Intelligence, Austin, TX.*

Stevenson, B., & Wolfers, J. (2013). Subjective well-being and income: Is there any evidence of satiation?. *American Economic Review*, *103*(3), 598-604.

Tarsney, C. (2019). Normative uncertainty and social choice. *Mind, 128*(512), 1285-1308.

Thornley, E. (2024). The shutdown problem: an AI engineering puzzle for decision theorists. *Philos Stud*.

Toma, M., & Bell, E. (2024). Understanding and increasing policymakers' sensitivity to program impact. Journal of Public Economics, 234, 105096.

Trammell, P. (2021). *Dynamic public good provision under time preference heterogeneity: theory and applications to philanthropy*. Global Priorities Institute Working Paper 9-2021

Trammell, P. (2024). Ethical Consumerism. Global Priorities Institute WP, 22-2024.

Trammell, P., & Aschenbrenner, L. (2024). Existential risk and growth. Global Priorities Institute WP, 13-2024.

Trammell, P., & Korinek, A. (2023). Economic growth under transformative AI *(No. w31815). National Bureau of Economic Research.*

Vivalt, E. (2015). Heterogeneous treatment effects in impact evaluation. *American Economic Review*, *105*(5), 467-470.

Vivalt, E. (2020). How much can we generalize from impact evaluations?. *Journal of the European Economic Association*, *18*(6), 3045-3089.

Vivalt, E., & Coville, A. (2023). How do policymakers update their beliefs?. *Journal of Development Economics*, 165, 103121.

Vivalt, E., Coville, A., & KC, S. (2024). Local knowledge, formal evidence, and policy decisions. Working Paper.

Wilkinson, H. (2022). Market harms and market benefits. *Philosophy & Public Affairs, 50(2),* 202-238.