Research agenda draft for GPI Economics

Table of contents

Introduction	1
1. General issues in global prioritisation	2
1.1 Strategic issues in altruistic decision-making	2
1.2 Forecasting	3
1.3 Conceptual issues in cost-effectiveness analysis	4
1.4 Welfare criteria and decision procedures	5
2. Applications to impactful issues	6
2.1 Growth, population and inequality	6
2.2 Catastrophic risks	6
2.3 Artificial intelligence	7
2.4 Intergenerational governance and policy-making	8
References	9

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 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 crowding-out effects, forecasting and the distribution of cost-effectiveness across interventions, as well as normative questions related to welfare criteria and ethical values. 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 relative lack of prior research on these topics suggests that further research may be particularly likely to yield important and decision-relevant insights.

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 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.

- An altruistic decision-maker that funds a charitable intervention may crowd out funding from other actors (e.g., governments or philanthropists). Empirically, what is the magnitude of such displacement effects for charitable causes that are particularly impactful? (cf. Andreoni & Payne 2003).
- 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 determines 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 are the optimal levels of altruistic 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:

- 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?
- To what extent are forecasting methods informative for assessing the probability of global catastrophic risks and other future events of special importance for social welfare? (Karger *et al.* 2023)
- 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 predictability may be more important than long-run forecasting (Millner and Heyen 2019). In which settings do the conditions of these models apply, and how widely applicable is this conclusion?

1.3 Conceptual issues in cost-effectiveness analysis

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 interventions in global health and beyond have significant externalities on individuals other than the direct recipients of the intervention. 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. Luyten et al. 2022)
- The cost-effectiveness of health interventions is often measured in terms of disabilityadjusted life years (DALYs) per dollar. However, the DALY approach has well-known limitations, and there are other approaches such as using wellbeing-adjusted life years (WELLBYs) (e.g., Layard and Oparina 2021). How do these metrics differ in evaluations of cost-effectiveness of different interventions, and which metrics are most appropriate?

1.4 Welfare criteria 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 many highly decision-relevant issues in these fields that have been largely neglected. We list some of these issues below.

- 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)
- The value of normative parameters such as the social rate of pure time preference, the elasticity of marginal utility to consumption, and the consumption level corresponding to a 'zero' level of utility (i.e. the consumption level at which being alive and being dead are equally valuable) are crucial for philanthropic and governmental priority-setting. How should these parameters be estimated?
- Suppose that a decision maker wishes to maximise the expectation of some social welfare function. Given that the state space is sufficiently complex, explicit maximisation of expected utility may be computationally intractable (cf. Camara 2022). What heuristics should a social decision-maker use to approximate expected social welfare maximisation?
- 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, e.g., the social rate or pure time preference (Millner 2020). Prior literature has used tools from decision theory (Dietrich and Jabarian 2022), social choice (Tarsney 2019) and bargaining theory (Greaves and Cotton-Barratt, forthcoming) to provide methods for handling such normative uncertainty. What are the best procedures for handling normative uncertainty?

2. Applications to impactful issues

2.1 Growth, population and inequality

One of the most important developments of the last hundred years is the massive growth in global income per capita and population levels that has occurred since the industrial revolution. Similarly, the future of economic and population growth may be of crucial importance for long-run welfare, especially if the economy transitions into a new growth mode with significantly higher or lower long-run growth rates. Although economic growth has 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 likely is it that the global economy transitions into a new growth mode, with a significantly higher or lower long-run economic growth rate? (E.g. Gordon 2017; Aghion *et al.* 2018; Roodman 2020; Davidson 2021; Nordhaus 2021; Jones 2022)
- What is the relationship between economic growth, population growth, inequality and long-run social welfare? (E.g. Asheim and Weitzman 2001; Stevenson & Wolfers 2013; Jones & Klenow 2016; Klenow et al. 2023)
- Population growth has both positive externalities (e.g., increased idea production) and negative externalities (e.g., increased resource depletion). In light of such externalities, what is the marginal social value of a larger population size? (E.g. Dasgupta 2019; Eden and Kuruc 2023)
- What is the impact of economic growth, population growth and inequality on global catastrophic risks? (E.g. Baranzini & Bourguignon 1995; Jones 2016; 2023; Aschenbrenner 2019)

2.2 Catastrophic risks

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 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; hence 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. Moreover, there is almost no literature in economics on existential risks, i.e., those that could result in human extinction or civilisational collapse. 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 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 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; Aschenbrenner 2019). Taking considerations such as this into account, how is the annual risk of global and existential catastrophes likely to develop over time?

2.3 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. Moreover, given the possibility that AI at some point may reach or exceed human-level intelligence, a new type of principal-agent problem is raised by the fact that the interests of AIs and humans may not be sufficiently aligned. 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.

• Many have suggested that explosive growth from AI will be bottlenecked by inability for AI to increase economy-wide productivity. How likely is it that these bottle-necks could

preclude the possibility of explosive growth? (Aghion, Jones and Jones, 2018; Bersiroglu and Erdil, 2023)

- Some have called for stronger regulations related to the development of more capable AI systems, including a moratorium on AI capabilities research if necessary. How should governments approach regulating potentially transformative technologies when information may be learned through their roll-out, or when there are political economy considerations? (Acemoglu, 2023)
- 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?
- Humans may afford greater power to AI systems, who may also provide information to humans. Can we design mechanisms to ensure that AI systems exhibit desirable behaviour such as truth-telling or a lack of deception?

2.4 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 incentives or commitment devices could be provided for an impatient government to implement policy consistent with placing a higher valuation on the future (Song et al 2012)? What factors predict a more or less patient government (MacKenzie 2016; Bonfiglioli and Gancia 2013; Healy and Malhorta 2009; Rogoff 1990; Jacobs 2016; Read and Toma, 2023)? Can long-term intergenerational mechanisms help to overcome this principal-agent problem? What might they look like (Rangel 2003; Harstad 2020)?
- Is there a strong case for enfranchising future generations (Kavka and Warren 1983; Goodin 2007; Tännsjö 2007; Beckman 2009: Ch. 7)? If so, how should this be implemented (Gonzalez-Ricoy and Gosseries 2017)? How effectively can contemporary individuals act as representatives for future generations?

- What influences when and how the 'right' evidence reaches the key decision-makers, and updates their beliefs (Vivalt and Coville 2023; Vivalt, Coville, KC 2023)? Can policy tools or interventions be introduced to promote evidence utilisation (Chen et al 2023; Toma and Bell 2023)? 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?
- Do 'broad' approaches to improving effective governance, and ultimately serving the far future, tend to be more or less effective in expectation than 'narrow' approaches (such as working on reducing the risk of bioengineered pandemics)?

References

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.

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.

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.

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

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

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

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

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

Greaves, H., Cotton-Barratt, O. (forthcoming). A bargaining-theoretic approach to moral uncertainty. *Journal of Moral Philosophy*.

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.

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.

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

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.

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.

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.

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.

Shulman, C., & Thornley, E. (2023). 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.

Smith, J. E., & Winkler, R. L. (2006). The optimizer's curse: Skepticism and postdecision surprise in decision analysis. *Management Science*, *52*(3), 311-322.

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.

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. (2023). Ethical Consumerism. Draft.

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.