IMF fairness: Calibrating the policies of the International Monetary Fund based on distributive justice

Downloaded from: https://research.chalmers.se, 2026-09-12 22:52 UTC

Citation for the original published paper (version of record):

Daoud, A., Herlitz, A., Subramanian, S. (2022). IMF fairness: Calibrating the policies of the International Monetary Fund based on distributive justice. World Development, 157. http://dx.doi.org/10.1016/j.worlddev.2022.105924

N.B. When citing this work, cite the original published paper.

research.chalmers.se offers the possibility of retrieving research publications produced at Chalmers University of Technology. It covers all kind of research output: articles, dissertations, conference papers, reports etc. since 2004. research.chalmers.se is administrated and maintained by Chalmers Library

(article starts on next page)

World Development 157 (2022) 105924

Contents lists available at ScienceDirect

World Development

journal homepage: www.elsevier.com/locate/worlddev

IMF fairness: Calibrating the policies of the International Monetary Fund based on distributive justice

Adel Daoud [a,b,c,*], Anders Herlitz [d,e], S.V. Subramanian [c,f]

[a] Institute for Analytical Sociology, Linköping University, Kopparhammaren 7, Campus Norrköping, Sweden
[b] The Division of Data Science and Artificial Intelligence, The Department of Computer Science and Engineering, Chalmers University of Technology, Rännvägen 6, 412 56, Gothenburg, Sweden
[c] Center for Population and Development Studies, Harvard T.H. Chan School of Public Health, Harvard University, 9 Bow St, Cambridge, MA 02138, Boston, MA, USA
[d] Department of Global Health and Population, Harvard T.H. Chan School of Public Health, Harvard University, 655 Huntington Ave, Boston, MA 02115, USA
[e] Institute for Futures Studies, Holländargatan 13, Stockholm, Sweden
[f] Department of Social and Behavioral Sciences, Harvard T.H. Chan School of Public Health, Harvard University, 677 Huntington Ave, Boston, MA 02115, USA

ARTICLE INFO

Article history:
Accepted 21 April 2022
Available online 30 April 2022

Keywords:
Causal inference
Children
Health inequalities
International Monetary Fund
Policy studies
Social demography
Distributive justice
Algorithmic fairness
Public policy
Poverty
Governance
International organizations
Ethics

ABSTRACT

The International Monetary Fund (IMF) provides financial assistance to its member countries in economic difficulties but at the same time requires these countries to reform public policies. In several contexts, these reforms have been at odds with population health and material living standards. While researchers have empirically analyzed the consequences of IMF reforms on health, no analysis has yet identified under what conditions tradeoffs between consequences for populations and economic outcomes would be fair and acceptable. Our article analyzes and identifies five principles to govern such tradeoffs and thus define IMF fairness. The article first reviews existing policy-evaluation studies, which on balance show that IMF policies, in their pursuit of macroeconomic improvement, frequently produce adverse effects on children's health and material living standards. Secondly, the article discusses four theories from distributive ethics—maximization, egalitarianism, prioritarianism, and sufficientarianism—to identify which is most compatible with the IMF's core mission of improving macroeconomic conditions, while at the same time balancing the consequences for population outcomes. Using a distributive justice analysis of IMF policies, we argue that sufficientarianism constitutes the most compatible theory. Thirdly, the article formalizes IMF fairness in the language of causal inference. It also supplies a framework for empirically measuring the extent to which IMF policies fulfill the criteria of IMF fairness, using observational data.
© 2022 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

The International Monetary Fund's (IMF) mission is to monitor the global financial system and provide macroeconomic support to its 190 member countries. It conditions this support on a set of policies that recipient governments must implement, the most challenging being structural adjustments and austerity. Given the often invasive nature of IMF policies, their subsequent consequences on population health programs have prompted considerable debate (Babb, 2005; McKee et al., 2012; Stuckler and Basu, 2013; Summers and Pritchett, 1993). Researchers have conducted empirical evaluations of IMF policies to identify those that affect population health (Clements et al., 2013; Daoud et al., 2017; Daoud and Reinsberg, 2018; Dreher, 2009; Kentikelenis et al., 2016; Shandra et al., 2012; Stubbs et al., 2016; Vreeland, 2007). However, these empirical studies lack a crucial component found in distributive justice research: transparency regarding what should be considered fair and desirable outcomes. *Distributive justice* is the study of just allocation of resources (Adler, 2012). While empirical studies favor explanatory or descriptive arguments focused on *is*-issues, justice analyses—also known as ethics-based studies—employ normative arguments that focus on *ought*-issues.[1] While empirical studies evaluate causes and effects, ethical studies critique and propose desirable outcomes. By ignoring ought-

* Corresponding author
E-mail addresses: adel.daoud@liu.se (A. Daoud), Anders.herlitz@iffs.se (A. Herlitz), svsubram@hsph.harvard.edu (S.V. Subramanian).

[1] We will use *justice* and *ethics* interchangeably.

https://doi.org/10.1016/j.worlddev.2022.105924
0305-750X/© 2022 The Authors. Published by Elsevier Ltd.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

issues, empirical studies lack a foundation for making recommendations regarding how the IMF should balance population health and macroeconomic outcomes.

Distributive justice analysis of the IMF's mission entails identifying which principles would reflect fair tradeoffs (Daniels, 1985). In the IMF literature, the tradeoffs of interest are primarily related to two outcomes: *macroeconomic wellness* and *individual wellness*. The former refers to all outcomes that are considered beneficial for a population's economy in the aggregate (e.g., higher economic growth, more trade, stable currency), while the latter refers to all outcomes that are considered beneficial for a person as an atomic unit (e.g., better physical and mental health, reduced stress, access to sufficient material resources). While these two outcomes often reinforce one another, they are sometimes in conflict (Burgard and Kalousova, 2015). Considerable criticism has been directed at the IMF that while its policies aim to improve macroeconomic wellness, they often do so at the expense of individuals wellness. (Babb, 2005; McKee et al., 2012; Stuckler and Basu, 2013; Summers and Pritchett, 1993). However, it remains unclear exactly how the IMF and recipient governments should strike this balance between macroeconomic and individual wellness. Further, given that the IMF has attracted a substantial amount of criticism, it is warranted to take a step back and ask, "what portion of this criticism is warranted and what portion is not?" This and similar questions refer to what we will label *IMF fairness*: that is, what constitutes an acceptable distributive justice outcome for an IMF policy. To evaluate such questions, we need a systematic framework for the analysis of IMF fairness.

As no such framework currently exists, our article aims to fill this knowledge gap by developing a distributive justice framework that identifies the conditions under which IMF policies have fair and acceptable outcomes. To this end, our article (a) reviews the findings of empirical studies focused on the (causal) direction in which IMF policies tend to push individual wellness, focusing on child health as a case study; (b) evaluates which distributive justice theory would be appropriate as a means of governing the IMF Articles of Agreement (the principles that regulate the IMF's mission) while minimizing the potential harm to individual outcomes; and (c) formalizes this distributive justice argument in the language of causal inference for algorithmic fairness (Balgi et al., 2022a; Kusner et al., 2017; Loftus et al., 2018). This formalization will enable applied researchers to study IMF fairness.

We present our arguments in four steps, organized into four sections. Section 2 reviews the empirical findings from previous research, focusing on the impact of IMF policies on children's health and material living standards. While IMF policies have wide-ranging effects on population health and material living standards that extend beyond the experiences of children (Babb, 2005), we have limited our review to the situation of children as a means of making it manageable. The review shows that IMF policies have had an adverse effect on children on several occasions, and provides empirical evidence of the IMF's failure to guarantee individual wellness, thus motivating our IMF-fairness framework. Section 3 discusses the IMF Articles of Agreement and how they direct the focus of the IMF at macroeconomic wellness while downplaying effects on the wellness of individuals. In order to provide a basis for recalibrating IMF policies and increasing the priority given to individual wellness, Section 4 evaluates theories in the field of distributive justice and identifies which theory is most compatible with the goal of balancing the IMF's core macroeconomic mission with the consideration of individual well-being. Having then made our argument in favor of the theory of sufficientarianism, Section 5 formalizes our account of how we can empirically evaluate the degree to which an IMF policy is fair. This section also provides a simulation-based example to exemplify the empirical feasibility of our approach.

The target of our distributive justice framework is not the IMF as an organization but rather the *arrangements*[2] made between the IMF and recipient governments. The IMF does not itself implement austerity policies in recipient countries, as it lacks the mandate to do so. Instead, both parties formulate a contract that contains one or more policies that are intended to boost macroeconomic performance, but which, as a side effect, may adversely affect the wellness of individuals. While the IMF controls sizable financial resources, to which the government in question wishes to obtain access (credit) in exchange for the implementation of a set of policies, the government nonetheless represents a sovereign state. Ultimately, it is governments that choose to implement—for better or worse—policies that may adversely affect the individual wellness of their populations. At the same time, the IMF could be expected to be even more open—despite improvements during the last decade—than it has been in the past to policies that mitigate adverse impacts on individual wellness, sometimes at the expense of macroeconomic performance. As will be shown in Section 3, the IMF will need to be provided with such leeway through a revision of its Articles of Agreement.

2. Reviewing the empirical evidence on the effects of international monetary Fund policies on children's health

Individual wellness covers many dimensions, which include an individual's financial and material living standards and both physical and mental health (Adler, 2012). As IMF policies can affect several of these dimensions (Babb, 2005; Burgard and Kalousova, 2015; Stuckler and Basu, 2013), we have limited our review to a manageable set of dimensions. And because children constitute one of the most vulnerable groups in society, we have focused our review on the effects of IMF policies on health and poverty outcomes among children.

In what directions do IMF policies push children's health and living standards? Thomson et al. (2017) recently reviewed studies on the effect of international financial organization policies on maternal and child health. These organizations included the IMF, the World Bank, and the African Development Bank. Their search identified six studies that focused on the effect of IMF policies on children. We have updated this search, focusing only on IMF programs, to include the period 2017 to 2021, and have found three additional studies.[3] Table 1 outlines these nine studies.

Six of the nine studies focus on the outcome of infant and child mortality. The outcomes examined in the remaining studies are vaccination, child poverty (an index of underlying outcomes), and child health (e.g., malnutrition and access to adequate sanitation and drinking water facilities). While all six of the mortality-focused studies have relied on country-level data, the remainder have employed child-level data. Most of the studies cover the period from the 1990s to the beginning of the 2000s, with the exception of Bird et al. (2020), which included data for the period up to 2015. One study, Oliver et al. (2006), employed a qualitative-comparative analysis, while the rest relied on statistical analyses. Three of the nine studies dealt with the issue of confounding by only adjusting for observed variables. A *confounder* is a variable that affects both the treatment and the outcome, and if a confounder is not adjusted for, the statistical estimate will be biased (Imbens and Rubin, 2015). In IMF policy-evaluation research, one of the most powerful potential confounding variables is a govern-

[2] Henceforth, when we use the terms *an IMF policy* or *IMF program*, we refer to the arrangement between the IMF and the government of the recipient country.

[3] After the literature search period, Balgi et al (2022) was published. This study conducts a counterfactual inference, finding some beneficial effects of IMF programs on children.

Table 1
The effects of IMF programs on children's well-being.

	Dependent variable	Sample	Period	Handles confounding by adjusting for...	Estimation method	IMF impact
(Shandra et al., 2012)	Infant mortality	32 Sub-Saharan countries	1990–2005	Observed & unobserved time-invariant confounding	Two-way fixed effect	Adverse
(Hajro and Joyce, 2009)	Infant mortality	82 low and middle-income countries	1985–2000	Observed & unobserved time-invariant confounding	GEE and fixed effect	Beneficial
(Shandra et al., 2004)	Infant mortality	59 low and middle-income countries	1980–1997	Observed confounding	Lagged-dependent-variables regression	Adverse (in interaction with democracy)
(Oliver, 2006)	Infant mortality and child mortality	Argentina & Uruguay	1980–2000	N/A	Qualitative comparative case study	Adverse
(Pongou et al., 2006)	Child undernutrition	3510 children in Cameroon	1991–1998	Observed confounding	OLS	Adverse
(Daoud et al., 2017)	Child health across five outcomes	1,941,734 children in 67 low and middle-income countries	2000 (±5)	Observed & one unobserved time-variant confounder (political will)	Heckman	Adverse (moderated by education of head of household)
(Daoud and Reinsberg, 2018)	Under-five mortality and child vaccination	128 developing countries	1980–2014	Observed & one unobserved time-variant confounder (political will)	IV	Adverse (for public-sector policies)
(Daoud and Johansson, 2020)	Child poverty	1,941,734 children in 67 low and middle-income countries	2000 (±5)	Observed & one unobserved time-variant confounder (political will)	Heckman	Adverse
(Bird et al., 2020)	Infant mortality	48 low and middle-income countries	1990–2015	Observed confounding	Propensity score matching	Beneficial

Notes: 1. IV refers to instrumental variable methods. 2. Heckman refers to Heckman selection models. 3. GEE refers to Generalized estimating equations.

ment's willingness to implement difficult policy changes, such as government spending cuts or the privatization of state-owned companies (Vreeland, 2007). As social-psychological processes among government officials are difficult to measure, there are no systematic data that directly measure political will, and thus the political will of governments remains unobserved. However, there are statistical methods that enable an analysis to indirectly deal with unobserved confounders of this kind (Imbens and Rubin, 2015; Stubbs et al., 2018), yet only three of the studies have explicitly dealt with the confounding effects of political will.

Seven of the nine articles identified adverse effects. All three studies that dealt with the challenge of political will identified an adverse effect of IMF programs on children—two of these studies, Daoud et al. (2017) and Daoud and Johansson (2020), employed child-level data, providing a higher resolution image of the way in which IMF policies affect children than that obtained by country-level studies. Studies that explicitly deal with unobserved confounding are generally considered more robust than those that only adjust for observed confounders (Imbens and Rubin, 2015; Stubbs et al., 2018). Of the three studies that relied on adjusting only for observed confounders, two reported an adverse IMF effect (Pongou et al., 2006; Shandra et al., 2004), and one a beneficial effect (Bird et al., 2020). Of the nine studies, while four identified adverse direct effects on various child outcomes (i.e., mortality, poverty, vaccination, undernutrition) (Daoud and Johansson, 2020; Oliver, 2006; Pongou et al., 2006; Shandra et al., 2012), three reported adverse indirect effects on children (interactions) via a third variable (e.g., democracy, education, public sector policies) (Daoud et al., 2017; Daoud and Reinsberg, 2018; Shandra et al., 2004). Two studies identified a beneficial direct effect, with Bird et al. (2020) and Hajro and Joyce (2009) reporting that IMF programs had reduced infant mortality. In sum then, irrespective of how the review is filtered—controlling for observed versus unobserved confounding—IMF policies were found to have had an adverse effect on children in the majority of cases.

We will now discuss the content of the nine studies by identifying the direct and indirect pathways via which IMF programs are likely to affect child well-being.[4] The adverse effects noted in seven of the studies described above provide an empirical motivation of the need for IMF programs to explicitly consider potential adverse effects on individual wellness.

2.1. Direct pathways

Governments often subsidize the costs of immunization, food, and health services (Daoud, 2015; Halleröd et al., 2013). Given that IMF policies have the aim of balancing the government-borrowing budget, these subsidies are often the first to be dismantled (Shandra et al., 2011). Another direct pathway through which IMF programs are likely to affect child poverty involves the deregulation or liberalization of health systems (Daoud and Reinsberg, 2018; Oliver, 2006; Stuckler and Basu, 2013). The IMF operates under the assumption that private actors are more efficient providers of healthcare and other public services (Benson, 2001). While the overall quality of health care might improve, privatization tends to favor affluent households (Daoud and Johansson, 2020; Ismi, 2004; Pongou et al., 2006). Those that gain from privatization thus tend to be the economically privileged segments of urban populations, while those that lose out tend to be concentrated in remote rural areas, far removed from the logic of private profit that characterizes modern corporations (Shandra et al., 2012).

IMF programs also serve as a catalyst for foreign investment and aid (Clements et al., 2013). Recent research has found that this effect is more potent in sectors linked to the IMF's core policy areas than it is in those linked to non-core areas. In turn, increased aid may for example provide opportunities for countries to combat

[4] We include further references in addition to the nine studies where helpful.

health deprivation (Stubbs et al., 2016, p. 511), and reduce child poverty (Bird et al., 2020; Hajro and Joyce, 2009).

2.2. Indirect pathways

IMF programs can also affect child poverty through indirect channels, such as via unintended consequences of policy reforms. One such indirect pathway operates via state employees (Daoud and Reinsberg, 2018). The IMF programs implemented within the current framework cap public sector wages (Rickard and Caraway, 2014), while a significant portion of the health spending of low-income countries is allocated to the wages of doctors and other medical staff (Kentikelenis et al., 2015). In these contexts, such wage caps may have a deleterious impact on the provision and quality of health services.

Another indirect pathway runs through international trade. The IMF currently seeks to liberalize countries' trade and investment regimes, for example by recalibrating tariffs, quotas, duties, and taxes (Dreher, 2009). While these policies may benefit countries in the long-term, they are likely to decrease a government's budget in the short-term, as its tax base contracts. A reduced budget will have a negative effect on social spending, which will reinforce the first direct pathway outlined above (Daoud and Johansson, 2020; Pongou et al., 2006). Moreover, IMF policies involve currency devaluations with the aim of fueling increases in exports, but may at the same time inflate the cost of importing critical goods such as medicine, medical equipment, water, and sanitation equipment,[5] and may reduce access to vaccines and other crucial health-related resources (Shandra et al., 2011).

A third indirect pathway involves global politics. The IMF takes different negotiating positions depending on the country with which it is entering into an agreement (Vreeland, 2007), which can affect the level of resources a country will receive. For example, a final agreement tends to be more favorable if the country has some form of substantial influence on the global economy or is compliant with IMF demands (Oliver, 2006; Pongou et al., 2006). Emerging powers, such as Brazil and Turkey, tend to have more influence than Sub-Saharan African countries, which instead have a weaker negotiating position. At the same time, because of their high poverty rates, low-income countries tend to be given differential treatment in forms such as debt forgiveness or favorable policies motivated by the IMF's Poverty Reduction Strategy (Dreher, 2009). Consequently, these geo-historical channels can moderate the strength of the relationship between an IMF program and children's well-being, independently of the factors outlined above.

All of these direct and indirect impacts of IMF programs on child poverty depend on household resilience (Daoud et al., 2017), which refers to the family members' ability to mitigate, adapt to, and recover from economic, natural, or political shocks and stresses (Daoud et al., 2016; Nandy et al., 2016). Households with more resources tend to be more resilient than those with fewer resources. Nevertheless, household resilience can itself be affected by IMF policies, with the strongest pathway going via the labor market. Changes in labor market policies (e.g., adjustments to minimum wage levels, deregulations of employment policies) will affect parents' incomes and their ability to care for their children.

In sum, despite the IMF's efforts to reduce the adverse effects of its programs on population health (Copelovitch, 2010), the majority of the empirical research has found a net adverse effect on children. If we give all studies "equal weight" in an assessment of how IMF policies affect children, seven of nine have reported adverse effects. If we only consider those studies that have used the most sophisticated statistical research designs to deal with unobserved confounding—most notably in the form of a government's political will to implement austerity policies—three out of three have reported adverse effects. Although more research is required to evaluate all possible pathways through which IMF policies can affect outcomes for children, our findings reinforce the argument that the IMF Articles of Agreement need to be transformed to include the element of distributive justice. This will provide IMF programs with a better ethical basis for balancing macroeconomic and individual wellness.

3. The IMF Articles of Agreement

Having reviewed the findings from empirical studies, we now proceed to evaluate how the IMF Articles of Agreement regulate the organization's policies, and in effect hinder it from considering the wellness of individuals. The IMF operates within the framework established by its 190 member countries. The parameters of this framework are defined by the Articles of Agreement, which were adopted at the United Nations Monetary and Financial Conference in Bretton Woods, New Hampshire, on July 22, 1944. These articles have been amended seven times, with the last amendment being adopted on December 15, 2010 (IMF, 2011). In the present context, Article V, section 3(a) of the Articles of Agreement is critical, because it shapes the IMF's focus on macroeconomic wellness (Vreeland, 2007):

> The Fund shall adopt policies on the use of its general resources [...] that will assist members to solve their balance of payments problems in a manner consistent with the provisions of this Agreement and that will establish adequate safeguards for the temporary use of the general resources of the Fund. (IMF, 2011)

This article thus binds the IMF to assist its member countries, but only to resolve balance of payment problems. This and supporting articles obligate the IMF to encourage public policies in borrowing countries that improve their macroeconomic wellness (Clements et al., 2013; IMF, 2016), but which may simultaneously lead to the de-prioritization of other issues associated with the individual wellness of a country's population (health and poverty).

While Article V focuses the IMF on the correction of macroeconomic fundamentals, Article I nonetheless states that one of the purposes of the IMF is "[t]o give confidence to members by making the general resources of the Fund temporarily available to them under adequate safeguards, thus providing them with an opportunity to correct maladjustments in their balance of payments without resorting to measures destructive of national or international prosperity" (IMF, 2011). However, our literature review indicates that IMF programs often fail to adequately protect vulnerable populations (Babb, 2005; Bohoslavsky, 2016; Stiglitz, 2003; Summers and Pritchett, 1993). The criticism for this should in part be directed at the governments that implement IMF policies, but in part also, and more so, at the IMF, since it is often the more powerful party. In response to this criticism, the IMF has sought to incorporate several social policies into its programs (IMF, 2014). By the mid-1990s, the IMF had introduced policies designed to protect poor populations via so-called Poverty Reduction Strategies (Gupta et al., 2000). These strategies ranged from specifying a minimum amount that a country should spend on health, known as priority expenditure floors (de Rato, 2006), to implementing targeted social safety nets (Clegg, 2014). In addition, low-income countries, the most common borrowers, are also given access to concessionary funds (Barro and Lee, 2005).

Despite the IMF's effort to reduce poverty, several empirical studies show that these efforts are likely to be insufficient to counter the adverse effects of its policies on individual well-being.

[5] This is the case in settings where the availability of clean and safe water is a core determinant of infant mortality (Shandra et al., 2011).

Between 1985 and 2014, the IMF operated 1,550 programs in over 130 countries. In the context of these programs, the IMF launched a total of 55,465 policies, but only 822 (1.5 percent) of these were explicitly aimed at reducing poverty. The vast majority (87 percent) were designed to address macroeconomic imbalances (Kentikelenis et al., 2016). On average, an IMF program includes 36 policies. Assuming that the 822 social policies were uniformly distributed, each IMF program—on average—contained only 0.36 social policies.

During the early history of the IMF, its programs were criticized for including policies that were too specific and intrusive (Babb, 2005; Summers and Pritchett, 1993; Vreeland, 2007). Since the onset of the 2007–2008 global financial crisis, however, some have instead raised criticisms that its policies are not specific enough (Ban, 2015; Ban and Gallagher, 2015; Broome, 2014; IMF, 2009). In the 1980s, IMF programs included few policy conditions—between one and six—and most of these targeted aggregate macroeconomic parameters. The number of conditions rose during the 1990s to a range of between one and sixteen (Kentikelenis et al., 2016). Many of these attracted harsh criticism for extending IMF policies into domestic affairs, and for being likely to violate the rights of sovereign states. To take one example, the IMF has in the past often argued that a country must cut its budget deficit, which is consistent with the IMF's macroeconomic orientation. In recent programs, however, the IMF has changed its stance and now eschews explicitly stating what expenditures should be cut, although IMF staff provide guidance—for example suggesting cuts in the government wage bill or reduced military spending. Consequently, these *too-specific* versus *not-specific-enough* criticisms require the IMF to balance two poles of interest: international governance versus domestic affairs.

The IMF Articles of Agreement are closely mirrored in what the organization does empirically: it prioritizes macroeconomic issues over individual well-being (Copelovitch, 2010; Dreher, 2006). This prioritization implies that even if IMF officials are interested in improving the wellness of individuals, the Articles of Agreement tie them to the prioritization of macroeconomic over individual well-being. Thus, IMF programs require a distributive justice theory that will provide a basis for balancing macroeconomic wellness and the wellness of individuals.

4. Distributive justice and the Articles of Agreement

Distributive justice is an area of ethics and political philosophy that addresses different distributions (i.e., allocations) of goods among individuals (Anderson, 1999; Crisp, 2003; Tännsjö, 1998; Temkin, 1993). Distributive theories comprise at least two elements: the *definition* of goods and the *pattern* of distribution. The definition of goods refers to the content of what is desired, and is often thought of in terms of material resources (Dworkin, 1981; Rawls, 1971), welfare (Dorsey, 2012), opportunities (Roemer, 1998), capabilities (Nussbaum, 2000; Sen, 1992), or event outcomes. A *distribution* refers to how goods are allocated among individuals. For the purposes of our argument in the current context, we focus on the health and material outcomes produced by IMF policies and the distributions of these outcomes.

There are four prominent distributive justice theories—*maximization, egalitarianism*, *prioritarianism*, and *sufficientarianism*—and in this section, we discuss which of these is suitable for evaluations of IMF fairness. Foreshadowing our conclusion, we will argue that of the four, sufficientarianism is that which is most compatible with the IMF's core mission—to reinvigorate macroeconomic wellness—and with the simultaneous eradication of harms to individual wellness that may result from IMF policies.

The four distributive justice theories all focus on patterns of distribution. The theory of maximization (of which utilitarianism constitutes one example) argues that whatever the relevant good, what matters is the production of as many goods as possible (Ord, 2013; Tännsjö, 1998). If the relevant good is vaccines for children, then distributions ought to be evaluated in terms of how many children are vaccinated. The theory of egalitarianism argues that whatever the relevant good, what matters is that it is distributed equally (Cohen, 1989; Temkin, 2003, 1993). If the relevant good is material welfare, distributions should be evaluated in terms of how equally welfare is distributed. The theory of prioritarianism maintains that whatever the relevant good, what matters is that benefits are valued higher for the worse off (Adler, 2012; Fleurbaey, 2015; Parfit, 1997). If the relevant good is material welfare, then distribution ought to be evaluated in terms of how much material goods the poor can access. The theory of sufficientarianism holds that whatever the relevant good, what matters is that it is distributed so that all individuals receive a sufficient amount of that good (Casal, 2007; Crisp, 2003; Shields, 2012).

In all these theories, individuals constitute the main unit of analysis. The four theories often yield significantly different evaluations of alternative distributions, but common to all of them is that they focus on the distribution of goods across individuals, not groups, regions, or countries. This commonality implies that, for all four theories, macroeconomic wellness is relevant only if it enhances the wellness of individuals.

Focusing on distributions across individuals does not imply that the four theories are insensitive to the distribution of goods among groups of individuals. Distributive justice scholars often recognize the importance of groups, but primarily as an explanatory or justifying factor for the existence of specific distribution patterns (Anderson 1999). For example, the unequal distribution of opportunity (a good) among individuals in a society such as the United States may be explained in terms of groups (e.g., race and ethnicity). In this example, the distribution of opportunities (*explanandum*) is often explained by race and ethnicity (*explanans*). All four distributive justice theories argue that it is misguided to evaluate only between-group distributions, as they mask within-group distributions.

4.1. Sufficientarianism

Although the core mission of IMF policies involves improving macroeconomic wellness, from an ethics perspective, this mission must also involve taking responsibility for any deterioration in the wellness of individuals. Because the literature review presented in Section 2 testifies to the existence of an adverse link between IMF policies and individual wellness outcomes, the IMF and its partnering governments would benefit from a fairness framework that provides a basis for making ethical tradeoffs.

Sufficientarianism can provide a framework of this kind, that will provide for the consideration of individual wellness while maintaining the IMF's focus on macroeconomics. The other three distributive justice theories add additional principles than are not compatible with the IMF Articles of Agreement. As noted previously, sufficientarianism holds that a distribution is fair if all individuals in a population have a sufficient amount of goods, and is unfair if one or more individuals fall short of the given threshold.

There are at least three reasons for favoring sufficientarianism. First, regardless of which of the four distributive justice theories one subscribes to, there are at least instrumental reasons for preferring sufficientarianism. From a maximizer's perspective, although the focus of maximization is directed at increasing the total amount of goods in a population, this perspective would nonetheless endorse providing a sufficient amount of resources

to each member of the population, if this were to lead to even more goods being produced. If individuals can remain productive and continue contributing to their community, this is likely to lead to greater utility than if some individuals fall below the threshold. Egalitarians, who believe that increasingly equal outcomes are desirable in general, would also agree with the sufficientarians' specific focus on equal outcomes based on the threshold determined. The larger the number of people who reach beyond the baseline threshold, the lower is the level of inequality, if viewed in terms of the number of individuals who have sufficient resources. Prioritarians would agree with sufficientarians if the outcome is considered to be of high priority to the maintenance of well-being (e.g., human needs).

Second, we would argue that sufficientarianism fulfills a principle of human dignity that is embraced by most international organizations, including the IMF (Abouharb and Cingranelli, 2008; IMF, 2001). Upholding such dignity involves ensuring that all individuals having secured the minimum amount of goods required to live a decent life (Sen, 1992). In addition, having sufficient resources is a fundamental prerequisite for participation in economic, social, and political life (Anderson, 1999).

Third, adopting sufficientarianism means focusing primarily on those who fall below a given threshold, who are often the poor. This focus on the poor resonates with the poverty goals that the IMF has already committed to (Clements et al., 2013; IMF, 2016). At the same time, sufficientarianism should also appeal to those who have criticized the IMF on the basis of the "do no harm" principle (Beauchamp and Childress, 2012; Feldstein and Feldstein, 1998). Doctors must avoid actions that cause suffering, and by analogy, the IMF, as a doctor of macroeconomics, should likewise make sure its policies do not cause suffering for individuals (Feldstein and Feldstein, 1998).

Accepting sufficientarianism as the optimal fairness framework involves introducing one fundamental principle into any IMF policy: the threshold principle. This principle states that for each IMF policy, or program (set of policies), the IMF and the recipient government must define a threshold against which a policy will be tested against—Section 5 elaborates on two versions of sufficientarianism and how they can be measured. A policy is *unfair* if it for some individuals leads to a shortfall below the defined threshold for satisfying their requirements; a policy is *fair* if it is invariant to individuals' chances of satisfying their requirements. For example, in the case of child nutrition, an established threshold for when a child is to be regarded as severely malnourished is three standard deviations below the median international reference population (Nandy and Svedberg, 2012). If an IMF policy increases the number of malnourished children based on this threshold, then the IMF policy is producing an unfair outcome for the affected children.

4.2. Thresholds

Sufficientarianism does not embrace a specific threshold or deprivation approach, making it flexible in relation to varying historical contexts. It is compatible with a variety of approaches—such as the World Bank's monetary approach (Ravallion, 2015), the Alkire-Foster framework (Alkire and Foster, 2011), and the Bristol Method (Gordon et al., 2003) for multidimensional adult and child poverty. Nonetheless, the threshold principle requires that both the IMF and the recipient government use a threshold that makes sense for the relevant society and economy (Daoud, 2010, 2007; Nussbaum, 2011; Reddy and Daoud, 2020; Sen, 1999).

Although thresholds will vary depending on the context, as a general recommendation we suggest the use of a deprivation approach to evaluate the level of fairness of IMF policies. While a monetary approach would measure the shortfall in an individual's wellness on one dimension (disposable income), a deprivation approach captures a potential shortfall on several dimensions. For example, the World Bank has defined a threshold for international (extreme) poverty at US$1.9 per day in 2011 purchasing-power parity. Country-specific thresholds exist, which take local context into account and which range from less than one dollar per day for rural populations in India to more than US$ 30 per day in the United States (Atkinson, 2016). One advantage of the monetary approach is that it facilities measuring the wellness of individuals and conducting international comparisons. A disadvantage is the loss of precision in measuring the wellness of individuals. Disposable income thresholds only relate to what an individual possesses financially and say nothing about how this income is actually consumed. Nor do they provide any information about the resources required to satisfy the basic needs associated with promoting individual wellness (Reddy and Pogge, 2010).

A deprivation approach achieves precisely this—measuring what is consumed—by focusing on basic needs. Although a deprivation approach requires an internationally harmonized means of measuring the wellness of individuals, and also requires more data than a monetary approach, it provides a granular method for identifying which individuals are falling short in relation to the satisfaction of their basic needs, and by how much. Two well-established deprivation approaches can be found in the Alkire-Foster framework (Alkire and Foster, 2011) and the Bristol Method (Gordon et al., 2003). While the Alkire-Foster framework is inspired by Amartya Sen's capability approach (Alkire and Foster, 2011; Nussbaum, 2000; Reddy and Daoud, 2020), the Bristol method emerged from the practical realities of child well-being, as formulated by UNICEF and the research of Peter Townsend (Gordon et al., 2003; Gordon and Nandy, 2012). Nevertheless, both methods measure the wellness of individuals in terms of their basic needs.

For the purpose of evaluating IMF fairness, we suggest using the Bristol method to measure the wellness of children and the Alkire-Foster framework for the wellness of adults and households. Table 2 presents the basic-needs dimensions and their thresholds using the Bristol method. This method operationalizes children's well-being on seven dimensions: child nutrition, vaccine access, housing quality, water access, education provision, information sources, and sanitation quality (Gordon et al., 2003). The thresholds presented are for *severe* deprivation. The wellness of individuals exists on a continuum from complete lack (extreme deprivation) to complete fulfillment. Table 2 shows the thresholds for severe deprivation, but there are several other internationally accepted thresholds, which we recommend using when appropriate (Minujin and Nandy, 2012). For example, as previously mentioned, if a child's weight-for-height for its age is below the severe threshold of three standard deviations below the median international reference point, then the child is starving and will require immediate medical attention (Nandy and Svedberg, 2012). The threshold for moderate malnutrition is two standard deviations below the median and is likely to be more appropriate for most middle-income countries.

The Alkire-Foster framework operationalizes adult (or household) deprivation on three dimensions: health, education, and material living standards (Alkire et al., 2017). As Table 3 shows, each category has at least two subdimensions. For example, the category *living standards* has six subdimensions covering whether the individual's home is electrified, and including improved sanitation and access to drinking water. As with the Bristol Method, the Alkire-Foster framework has different severity thresholds that can be used for different contexts. Table 3 presents definitions for *moderate* threshold (Alkire and Santos, 2014; Alkire and Seth, 2015).

Although the Bristol method and the Alkire-Foster framework provide empirically measurable and established definitions for measuring the wellness of individuals, they have their limitations.

Table 2
Defining the wellness of children: dimensions and their thresholds (Bristol method).

Water: Children who only have access to surface water (e.g., rivers) for drinking or who live in households where the nearest source of water is more than 15 min away. Children < 18 years old.
Malnutrition: Children whose heights and weights for their age are more than 3 standard deviations below the median of the international reference, that is, severe anthropometric failure. Children < 5 years old.
Shelter: Children in dwellings with more than five people per room and/or with no flooring material. Children < 18 years old.
Sanitation: Children who have no access to a toilet of any kind in the vicinity of their dwelling, that is, no private or communal toilets or latrines. Children < 18 years old.
Health: Children who have not been immunized against diseases or young children who have had a recent illness involving diarrhea and not received any medical advice or treatment [polio, measles, DPT (diphtheria, pertussis, and tetanus), tuberculosis (Bacillus Calmette-Guérin), recommended by the WHO (cf. 2013)]. Children < 5 years old.
Information: Children with no access to radio, television, telephone, or newspaper at home. Children 3–17 years old.
Education: Children who have never been to school and are not currently attending school, in other words, no professional education of any kind. Children 7–17 years old.

Notes: "Proportion of *n* deprived" signifies the proportion of deprived children relative to the age-filtered sample, not relative to the total sample.

Table 3
Defining the wellness of adults and children: dimensions and their thresholds (Alkire-Foster framework).

Dimensions	Sub-dimension (indicator)	Threshold definitions
Education	Years of Schooling	No household member has completed five years of schooling.
	Child school attendance	Any school-aged child is not attending school in years 1 to 8.
Health	Child mortality	Any child in the family has died.
	Nutrition	Any adult to a child for whom there is nutritional information is malnourished.
Living standard	Electricity	The household has no electricity.
	Improved sanitation	The household's sanitation facilities are not improved (according to MDG guidelines), or are improved but shared with other households.
	Improved drinking water	The household does not have access to safe drinking water (according to MDG guidelines) or safe drinking water is more than 30 min walk from home, roundtrip.
	Flooring	The household has dirt, sand, or dung floor.
	Cooking fuel	The household cooks using dung, wood, or carbon.
	Asset ownership	The household does not own more than one of the following assets: radio, television, telephone, bicycle, scooter or refrigerator, and does not own a car or a truck.

For example, they do not cover other wellness dimensions (e.g., mental health) and they are suboptimal for use in high-income countries. IMF operations are rarely focused on these countries, but when they are—as was the case in Greece in 2010—scholars will need refined thresholds that are suitable for these societies. Moreover, since the two approaches are tailored to measuring poverty and ill-health in low- and middle-income countries, more research will be required to refine the threshold definitions to cover high-income countries (Guio et al., 2018; Guio et al., 2016). Nonetheless, regardless of the selected threshold and dimensions of interest, what ultimately matters to sufficientarians is that as many individuals as possible are located above the relevant thresholds (Daoud, 2018; Shields, 2012).

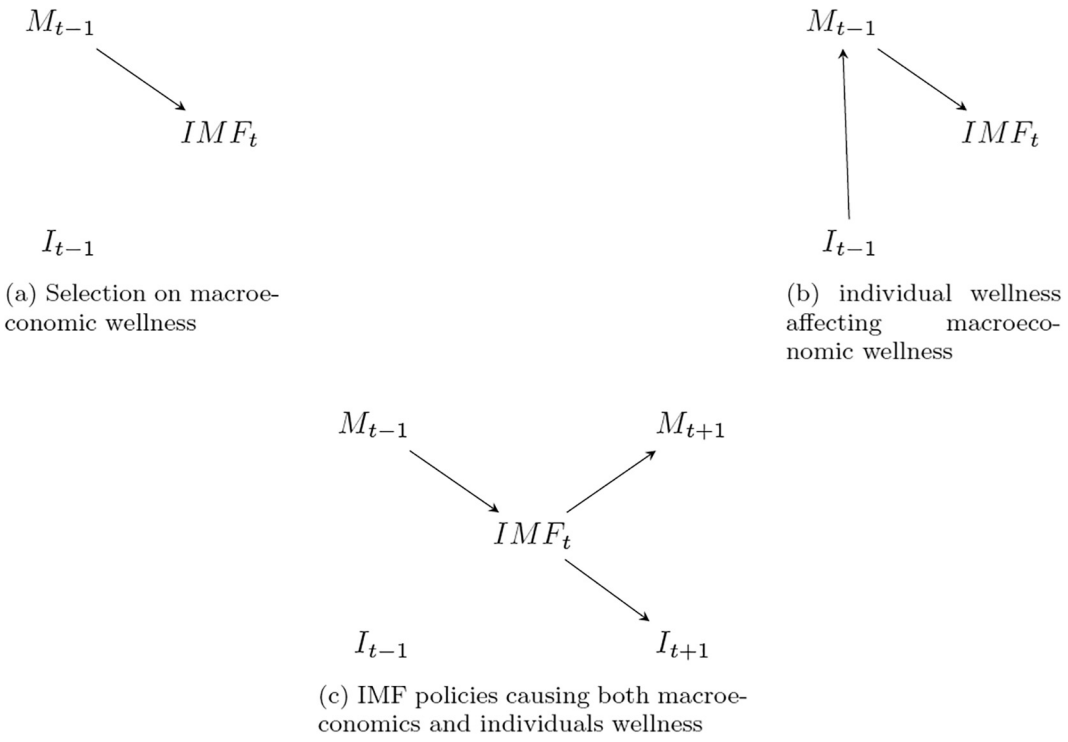
5. Formalizing and measuring IMF fairness

Having identified sufficientarianism as the most suitable distributive justice theory for the purpose of revising the IMF Articles of Agreement, and having discussed what thresholds can be used to measure individual wellness, in this section, we move on to formalize our IMF-fairness framework. This formalization aims to define quantities of interest for empirically evaluating when an IMF program complies with the principles of sufficientarianism: IMF fairness. Our formalization combines elements from causal inference (Imbens and Rubin, 2015; Pearl, 2009) and algorithmic fairness (Loftus et al., 2018). This combination enables us to define what IMF fairness means empirically and to evaluate it statistically. Our proposal for the evaluation of IMF fairness is divided into two parts: the first part relates to evaluating the *causes* of a country selecting into an IMF program (i.e., the reasons a country does so), and the second part accounts for the *effects* of an IMF program on the country's macroeconomic and individual wellness. This section concludes with a worked example that demonstrates how our approach can be deployed empirically.

Our two-part argument is depicted in a directed acyclic graph (DAG) in Figure 1. A *directed acyclic graph* captures a postulated or hypothesized causal system; a node indicates a causal factor, and an arrow from A to B indicates that A causes B. Mathematically, this arrow implies that the conditional distribution of B depends on A and can be described in some (parametric or nonparametric) functional form, $f(\cdot)$. This dependence statistically implies, $f(B|A)$. For simplicity, we primarily focus on probabilities denoted $P(\cdot)$. The graph is acyclic, meaning that a factor cannot cause itself. For example, A_t cannot affect itself at a time point t but it can affect future states of itself A_{t+1}. In the DAGs in Figure 1, we assume no common causes (confounding). In the following subsection, we will account for the case in which common causes exist and discuss how we propose capturing IMF fairness on the basis of observational data using Pearl's *do*-calculus (Pearl, 2009).

We formulate five principles that define IMF fairness. Grounded in the spirit of the IMF Articles of Agreement, we argue that the reasons for the IMF and a recipient government to agree on a program should be solely based on the issue of macroeconomic wellness (M): this is Principle 1. This principle is largely consistent with the way in which the IMF currently operates. Our argument also entails that the IMF and the recipient government must evaluate whether enrolment in a program would be agonistic to individual wellness (I): this is Principle 2. Thus, our argument leads to the first two principles of IMF fairness, both of which refer to the situation prior to the IMF and the government in question initiating a program (pre-program conditions),

Principle 1. The "based only on macroeconomic wellness" criterion (pre-program). The IMF and the recipient government should agree to implement an IMF program solely based on the issue of



(a) Selection on macroeconomic wellness

(b) individual wellness affecting macroeconomic wellness

(c) IMF policies causing both macroeconomics and individuals wellness

Figure 1. Directed acyclic graphs depicting stylized causal systems for the relationships between IMF policies, macroeconomic wellness, and the wellness of individuals.

macroeconomic wellness. Formally, this macroeconomic focus is represented by panel a of our DAG (Figure 1), and in the language of probability, this selection criterion is expressed as:

$$P(IMF_t = 1|M_{t-1} = 0) \geq P(IMF_t = 1|M_{t-1} = 1) \quad (1)$$

Principle 1 states that in order to determine whether self-selection into IMF programs is fair in a sufficientarian sense, we must evaluate the probabilistic inequality in equation (1). Let the binary random variable IMF_t be whether a government chooses to enter an IMF program at time t; the binary random variables M_{t-1} and I_{t-1} denote the country's macroeconomic and individual wellness, respectively, in the year prior to choosing whether to enter an IMF program, $t-1$. When these variables equal "1", they indicate that wellness is high; when they equal "0", wellness is low. Equation (1) states that the probability of a government choosing to enter an IMF program, $IMF_t = 1$, should be greater when the level of macroeconomic wellness indicates "low performance" (i.e., $M_{t-1} = 0$) than when it indicates "high performance" (i.e., $M_{t-1} = 1$). To concretize and simplify our thesis, we use binary-valued indicators of all our random variables with no loss of generality. We can consider any macroeconomic indicator of interest separately or together with others (e.g., economic growth, inflation, balance of payments). Following the temporal flow of cause and effect, M_{t-1} captures all the relevant macroeconomic parameters at $t-1$, i.e., prior to the point at which a government decides whether to enter an IMF program (at time point t).

While Principle 1 refers to the desired presence of macroeconomic wellness (M_{t-1}), Principle 2 stipulates an agnostic position in relation to individual wellness (I_{t-1}).

Principle 2. The "agnostic to individual wellness" criterion (pre-program): The probability of the IMF and the recipient government agreeing to a program should be independent of the individual wellness. This independence is indicated in panel a of Figure 1 by the absence of a direct arrow from M_{t-1} to IMF. The probability statement for this absence of a direct arrow is expressed in the following equation:

$$P(IMF_t = 1|M_{t-1}, I_{t-1}) = P(IMF_t = 1|M_{t-1}) \quad (2)$$

This probability statement implies that individual wellness, denoted I_{t-1}, does not provide any causal information on the probability of a government and the IMF agreeing to a program. In other words, their decision to agree to a program should be independent of the health and material living standards associated with individual wellness.

In many cases, however, low levels of individual wellness (e.g., poor health) have an adverse effect on macroeconomic wellness (Banerjee and Duflo, 2012; Deaton, 2015). A situation of this kind is depicted in panel *b* in Figure 1. Because this is a common situation, it is valid to ask whether IMF programs should consider second-order effects of individual wellness when determining which countries should enter a program (*second-order effects* are indirect effects, *first-order effects* are direct effects (VanderWeele, 2016)). In Principle 1, we have already stated that a determination of whether to enter a program should solely be based on macroeconomic considerations. In line with the spirit of this principle, we suggest that second-order effects (the effects of individual wellness on macroeconomic wellbeing) should be viewed as secondary, and that these should thus not play a role in determining whether or not to enter an IMF program. While second-order effects and long-term human development fall within the mission statements of the World Bank, the World Health Organization (WHO), Food and Agricultural Organization (FAO), UNICEF, and other global organizations, the IMF's primary mission—unlike these other UN organizations—is to oversee global macroeconomic stability, not to engage itself in issues of population health or poverty reduction. To handle second-order effects effectively, the IMF needs to coordinate its policies with these organizations. Although the IMF does

already coordinate many of its policies with the World Bank, more coordination is required with UNICEF, WHO, and FAO Bank to manage adverse policy effects on children and other vulnerable populations.

While the division of labor among these UN organizations exists for historical reasons (Ziring et al., 2005), blurring this division of labor would lead to incompatibilities in relation to the organizations' mission statements—but we would argue that inter-organizational cooperation, as stated in the Paris Declaration on aid effectiveness, is necessary when evaluating the causal effects of IMF programs. In cases where there is a mixture of first- and second-order effects, which is likely to constitute the default condition, the IMF and recipient governments would benefit from such cooperation, and more cooperation between the IMF and other international organizations is therefore needed. Nevertheless, the critical aim of Principles 1 and 2 is to stipulate that the primary motivation for IMF program selection ought to be based on M_{t-1}, not I_{t-1}.

Once a government has agreed to enter an IMF program, a sufficientarian approach would argue that the IMF and the government should design the IMF program in a way that takes the program's causal effects on both macroeconomic and individual wellness into consideration. While Principles 1 and 2 refer to pre-program conditions, Principles 3, 4, and 5 below regulate post-program outcomes. Post-program outcomes, depicted in panel *c* in Figure 1, are described in terms of causal arrows pointing from IMF_t to M_{t+1} and I_{t+1}. We define one principle for the arrow pointing from IMF_t to M_{t+1}, and two for the arrow connecting IMF_t and I_{t+1}.

Principle 3. *Improving macroeconomic wellness* (*post-program*): *With regard to macroeconomic wellness, we stipulate that the following probability inequality has to hold to for a program to be considered successful*:

$$P(M_{t+1} = 1|IMF_t = 1) > P(M_{t+1} = 1|IMF_t = 0) \quad (3)$$

On the basis of the IMF Articles of Agreement, an IMF program is expected to have a positive effect on macroeconomic parameters (IMF, 2011). In statistical terms, the probability of an acceptable level of macroeconomic wellness, $M_{t+1} = 1$, should be higher after a program (i.e., $IMF_t = 1$) than when no program has yet been introduced (i.e., $IMF_t = 0$).

As has been shown by our literature review, what is often not given sufficient consideration is the way in which the IMF's macroeconomic policies affect the wellness of individuals. The review shows that IMF programs tend to have adverse effects on children, suggesting that IMF_t may also have adverse effects on I_{t+1} more generally (Babb, 2005). While the question of how IMF fairness should be evaluated is essential in relation to any IMF program, in the context of our own review, the critical question is "how many children (or individuals) have fallen below a threshold of interest.".

As will be discussed in the section on empirical identification below, it will be challenging to evaluate whether sufficientarian ethics have been respected if individual-level data are lacking. To deal with this challenge, we propose two sufficientarian approaches: one lax, the other stringent. When individual data are lacking and only population (country) data are available, an analyst will only be able to evaluate IMF fairness on the basis of the lax-sufficientarian approach. In this case, the focus is directed at the average effect of IMF policies on a population. One limitation of this lax approach is that although an analyst might find that an IMF program does not have an adverse effect on the average level of population health, the program may still have a negative effect on the health of specific individuals or subpopulations. Thus, the lax approach does not evaluate the potentially considerable variation (e.g., heterogeneity) that IMF policies may produce (Daoud and Johansson, 2020). Nevertheless, in the absence of individual-level data, the lax-sufficientarian approach is the only option. In this approach, $I_{t+1} = 1$ is a measure of sufficient population health in the aggregate. For example, a measure of I_{t+1} could be based on the mean values of all the dimension in the Alkire-Foster or Bristol method.

Principle 4. Satisfying IMF fairness based on lax sufficientarianism (post-program). We define an ethically fair lax-sufficientarian IMF effect as fulfilling the following probability condition:

$$P(I_{t+1} = 1|IMF_t = 1) \geq P(I_{t+1} = 1|IMF_t = 0) \quad (4)$$

This equation states that populations on which IMF programs are focused are either better off after the program (i.e., the "greater than" part of $\geq$) or are unaffected (the "equal to" part of $\geq$), compared to populations with no such program.

For stringent sufficientarianism, Principle 4 is insufficient, because it only considers the mean population effect of IMF_t on I_{t+1}. The stringent approach follows logically from sufficientarianism, and is thus disaggregated.

Principle 5. Satisfying IMF fairness based on stringent sufficientarianism (post-program). A stringent sufficientarian approach accepts no adverse heterogeneity in the impact of an IMF policy: all impact heterogeneity has to be either invariant to the IMF program or beneficial to all individuals *i* in a country. In this case, $I_{i,t+1} = 1$ represents individual-level health and poverty outcomes:

$$P(I_{i,t+1} = 1|IMF_t = 1) \geq P(I_{i,t+1} = 1|IMF_t = 0) \quad (5)$$

This probability implies that an individual should be either better off in an IMF program or unaffected by it, compared to a situation in which the same individuals were not in the same program. To identify individual-level effects—*counterfactual* inference—the analyst must stipulate the entire structure causal model in a DAG, including the noise terms for each random variable (Balgi et al., 2022a; Balgi et al., 2022b; Kusner et al., 2017; Pearl, 2009). In other words, assumptions must be made about the entire data-generating process.

We now refine our definition of *I* based on Principles 4 and 5. The variable *I* represents any measure of wellness (e.g., malnutrition, education, sanitation), indicating whether a person has enough resources to satisfy a minimum level of individual wellness that is sufficient to avoid deprivation and poverty. While for maximizers, more of *I* is always better, for sufficientarians, any level beyond the defined minimum *I*-threshold matters little. This distinction between maximizers and sufficientarians is essential. Both lax and stringent sufficientarian definitions (Principle 4 and 5) argue that IMF programs should not be held accountable for a lack of improvement if there would have been no improvement even without the intervention of the IMF. The IMF and the recipient government should be held to account for their policies only when they deteriorate the situation of people who lie close to the wellness threshold, pushing them below this threshold. A lax sufficientarian focuses on the average IMF effect and will have to tolerate situations in which some may fall below the threshold, while stringent sufficientarianism instead argues that an IMF policy has unfair effects if the policy pushes at least one individual, *i*, below the threshold. This logic also implies that if members of a population were above this well-being threshold prior to an IMF program, but some or all of them lost a substantial portion of this well-being as a result of the program, but nonetheless stayed just above

the threshold, then stringent sufficientarians would argue that this is a fair outcome. It is fair because all individuals still have sufficient well-being. Even though a significant loss of well-being is undesirable from a public-policy perspective, this outcome would be tolerated on the basis of a sufficientarian perspective. This tolerance follows from the threshold principle discussed in Section 4. In addition, as was also discussed in the section on Thresholds, the level and quality of each threshold will vary across different national contexts. Thus, we emphasize that any thresholds have to be defined on the basis of each country's social and cultural situation (Sen, 1992).

In summary, of the available distributive justice principles, stringent sufficientarianism is the approach that lies closest in spirit to an ideal-world sufficientarianism. At the empirical level, however, the stringent version of sufficientarianism requires much more data and stronger statistical assumptions than the lax version (Balgi et al., 2022a). In situations where data is scarce or an analyst is unwilling to commit to such strong assumptions, it will be impossible to evaluate whether IMF policies are fair in terms of stringent sufficientarianism. A stringent approach requires a causal estimate (counterfactual inference) for each individual (e.g., child), whereas lax sufficientarianism can be routinely captured under the milder assumptions associated with commonly used causal-inference methods based on observational data.

Before considering how to empirically measure whether IMF policies fulfill lax or stringent sufficientarianism, we conclude this section by proposing an amendment to Article V of the IMF Articles of Agreement to include sufficientarian principles. Revising the governing articles of the IMF would empower the organization, and recipient governments, to promote macroeconomic wellness without producing detrimental effects on health and poverty. A revised Article V, section 3(a) could be formulated along the following lines (our amendments are in brackets):

> The Fund shall adopt policies on the use of its general resources ... that will assist members to (i) solve their balance of payments problems, [and (ii) will ensure that any adverse effects of these policies do not push individuals or populations below a minimum well-being threshold] ...that will establish adequate safeguards for the temporary use of the general resources of the Fund.

5.1. Estimands for lax and stringent sufficientarianism

To empirically evaluate whether IMF policies comply with sufficientarianism, we have to identify the causal effects of these policies. Because it is not possible to use randomized controlled experiments to evaluate the effects of IMF policies, scholars use observational data. Estimating causal effects from observational data is challenging (Imbens and Rubin, 2015; Pearl, 2009). As was discussed in Section 2, the challenges faced by the literature on IMF policy evaluations are primarily associated with the need to account for a government's political will: the government's willingness to enter an IMF program and implement often complex policies (Vreeland, 2007). Once the influence of this and other confounders have been accounted for in observational data, we will have secured *empirical identification*, which then allows us to use observational data to estimate causal effects.

Once the requirement of empirical identification has been satisfied, there are a number of estimators (i.e., methods) that can be used to calculate causal effects. If all the necessary data are observed, matching or regression would be appropriate. Conversely, when political will remains unobserved, scholars can use Heckman selection models, instrumental variables (IV), or the generalized method of moments (GMM) to circumvent the influence of political will, under certain assumptions (Stubbs et al., 2018). These are important estimators, with different strengths and weaknesses.

The remainder of our discussion focuses on defining the relevant estimands, i.e., defining IMF fairness. *Estimands* are the quantity of interest. In our case, all estimands define what IMF fairness means statistically, and in the remainder of this section these estimands are defined on the basis of the five principles discussed above. Estimands can be formulated independently of data or models. When an estimand has been defined, empirical identification specifies the conditions required to calculate an estimand using observational data. Once empirical identification is complete, we apply an estimator to a sample and thereby calculate (estimate) an approximation of an estimand. In other words, *estimators* are the methods (e.g., IV, matching, GMM) used to produce an estimate of an estimand (Lundberg et al., 2021). The IMF-fairness framework outlined in this article is first and foremost focused on sharpening the definitions of the quantities of interest: the estimands used to measure the extent to which an IMF policy meets the criteria of sufficientarianism.

The DAGs in Figure 1 contain all of our estimands. In order to simplify our argument, these DAGs are based on the assumption that there are no confounders, and confounding is thus dealt with subsequently. Principle 1 stipulates that governments should self-select into IMF programs based only on macroeconomic parameters. We define the estimand that captures this principle as a difference α, shown in Equation (6). This difference assesses the probability (propensity) of a government selecting into an IMF program, conditioned on the country's macroeconomic wellness. For a government to select into a program, we would expect the first probability term (the one containing $M_{t-1} = 0$) to be larger than the second term (the one containing $M_{t-1} = 1$). It should be larger because poor macroeconomic performance (indicated by $M_{t-1} = 0$) should be the primary driver for entering an IMF program.

Estimand for Principle 1 (*The "based only on macroeconomic wellness" criterion***):.** For Principle 1 to be fulfilled, the difference α must be $\alpha > 0$,

$$\alpha = P(IMF_t = 1|do(M_{t-1} = 0)) - P(IMF_t = 1|do(M_{t-1} = 1)) \quad (6)$$

We rely on Pearl's (2009) *do*-calculus, and use the $do(\cdot)$ operator to signify the fixing of the causal factor of interest. *Fixing* means intervening, and is analogous to an intervention in a randomized controlled trial (Pearl, 2009). Generally, $do(X = x)$ implies that any causal factor of interest, X, can be fixed and set to x. For all our estimands, we fix the causal variable of interest to mimic an intervention in the causal system. For causal estimation using any binary outcome Y of interest, our estimand is then:

$$P(Y|do(X = 1)) - P(Y|do(X = 0)) \quad (7)$$

For example, calculating $do(M_{t-1} = 1)$ implies that an experimenter intervenes in the data by changing M_{t-1} to 1. This causal logic is helpful, because it allows us to evaluate IMF fairness in terms of interventional logic. When we later identify and estimate α, and it turns out to have a positive value, we conclude that selecting into an IMF program is fair from a sufficientarianism perspective.

When Principle 1 is fulfilled, we evaluate Principle 2. We can check whether the individual wellness is independent of IMF program selection by evaluating that I_{t-1} provides no information that can explain a government selection into a program.

Estimand for Principle 2 (*The "agnostic to individual wellness" criterion***):.** For Principle 2 to be fulfilled, the difference β must be $\beta = 0$ in the following equation:

$$\beta = P(IMF_t = 1|do(M_{t-1}), I_{t-1}) - P(IMF_t = 1|do(M_{t-1})) \quad (8)$$

If $\beta = 0$ in Equation (8), this means that IMF_t and I_{t-1} are statistically independent and that the terms $P(IMF_t = 1|do(M_{t-1}), I_{t-1})$ and $P(IMF_t = 1|do(M_{t-1}))$ have the same value.

Principle 3 is defined by the estimand γ. This is the difference that captures the effect of an IMF program on macroeconomic performance. Because an IMF program $IMF_t = 1$ is expected to improve a country's economy, we expect γ to have a positive value. If this value is indeed positive, we conclude that a program has had the desired outcome for the country's macroeconomic wellness.

Estimand for Principle 3 (*Improving macroeconomic wellness***):.** For Principle 3 to be fulfilled, the difference γ must be $\gamma > 0$ in the following equation:

$$\gamma = P(M_{t+1} = 1|do(IMF_t = 1)) - P(M_{t+1} = 1|do(IMF_t = 0)) \quad (9)$$

Principle 4 evaluates the effect of IMF policies on the wellness of individuals from a lax-sufficientarianism perspective. To do so, we define the causal effect of IMF_t on I_{t+1} as the following difference δ :

Estimand for Principle 4 (*Satisfying IMF fairness based on lax sufficientarianism***):.** For Principle 4 to be fulfilled, the following difference δ must be $\delta \geq 0$.

$$\delta = P(I_{t+1} = 1|do(IMF_t = 1)) - P(I_{t+1} = 1|do(IMF_t = 0)) \quad (10)$$

If δ has a positive value (i.e., IMF policies have produced a beneficial effect) or zero (i.e., there is no effect), then we conclude that the IMF program has impacted individuals fairly; otherwise, an IMF program will have produced unfair outcomes in terms of lax sufficientarianism.

The estimand for Principle 5 is defined at the individual level. For this estimand, which captures stringent sufficientarianism, counterfactual inference has to be conducted at this level (Balgi et al., 2022b). In do-calculus, random variables are indexed to denote average effects. To facilitate the discussion of individual-level treatment effects (counterfactuals), we use the language of potential outcomes (Imbens and Rubin, 2015). The potential outcome for individual *i* in a population who is subject to an IMF program is $I_{i,t+1}(IMF_t = 1)$, while without the program the potential outcome is $I_{i,t+1}(IMF_t = 0)$. The difference δ_i between these outcomes constitutes the causal effect of an IMF program on individual *i*. Thus, the estimand for Principle 5 is as follows:

Estimand of Principle 5 (*Satisfying IMF fairness based on stringent sufficientarianism***):** For Principle 5 to be fulfilled, the following difference δ_i must be $\delta_i \geq 0$.

$$\delta_i = I_{i,t+1}(1) - I_{i,t+1}(0) \quad (11)$$

Estimating stringent sufficientarian fairness builds on the same assumptions as the lax-sufficientarian approach, but also on two additional assumptions (Pearl, 2009). The first is that besides the two potential outcomes for each individual *i*, an individual's characteristics (covariates X_i) are also observed. The second assumption is that there is a DAG that defines the causal structural conditions that also define the nuisance variables U_i for each individual (Balgi et al., 2022a). A nuisance variable represents all the exogenous factors associated with a DAG. Based on these two assumptions, an estimate of stringent IMF fairness requires data $\{I_{i,t+1}(1), I_{i,t+1}(0), X_i, U_i\}^n$ for all *n* individuals—or at least assumptions about the distribution of these data.

The requirement of having information on U_i is one of the most onerous (Balgi et al., 2022a; Balgi et al., 2022b). When U_i is measurable, stringent-sufficientarian fairness can be estimated using Pearl's do-calculus for counterfactuals. Computing counterfactuals requires three steps (Kusner et al., 2017; Lillie et al., 2011; Loftus et al., 2018; van and Petersen, 2007). First, existing evidence is used to measure U_i. This evidence is based on previous literature and assumptions about the relationship between IMF programs, individual wellness, and macroeconomic wellness. Second, the assumed DAG is modified twice by intervening on IMF_t by first fixing it to $IMF_t = 1$ and then fixing it to $IMF_t = 0$. Third, the modified DAGs under $IMF_t = 1$ and U_i are used to compute $I_{i,t+1}(1)$; and the modified DAGs under $IMF_t = 0$ and U_i are used to compute $I_{i,t+1}(0)$. The difference between the two is then taken to obtains δ_i.

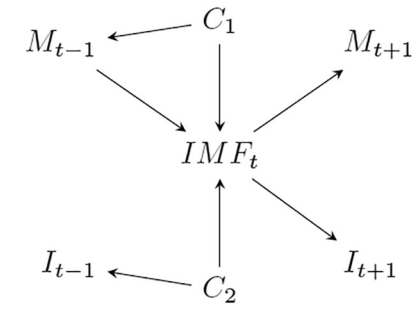
When U_i remains unobserved, δ_i can only be approximated (Kusner et al., 2017; Lillie et al., 2011; Loftus et al., 2018; van and Petersen, 2007). The synthetic control method is one of the few estimators that can be used to approximate δ_i (Abadie et al., 2015), but this method requires high-quality panel data for each individual, which are rarely available. A technique that does not require panel data involves imputing the Conditional Average Treatment Effect (CATE) (Künzel et al., 2018), where the expectation $E[\cdot]$ is equivalent to the probability $P[\cdot]$ for binary variables:

$$\delta(x_i) = E[I_{t+1}|do(IMF_t = 1), X = x_i] - E[I_{t+1}|do(IMF_t = 0), X = x_i] \quad (12)$$

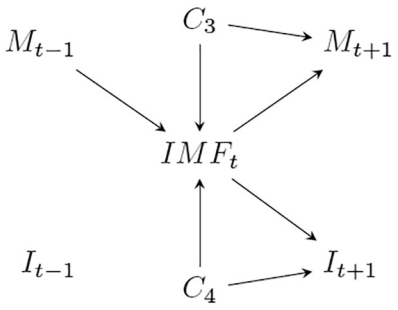
Here, X is a set of covariates, including confounders, that characterize each individual. This CATE-quantity $\delta(x_i)$ allows for an approximation of $\delta_i \approx \delta(x_i)$, which is the best estimate of individual-level treatment effects when U_i is unobserved (Künzel et al., 2018). This equation implies that each individual's causal effect is estimated based on the sub-population to which the individual belongs, as defined by the covariates x_i (Daoud and Dubhashi, 2021). Daoud and Johansson (2020) have used this method to estimate the impact of IMF programs on child poverty in cross-sectional data. Empirically, without additional assumptions about U_i, CATE is the most granular effect that can be estimated for an individual based on the individual's (granular) subpopulation membership. Even though a stringent-sufficientarian approach is measurable only on the basis of strong causal assumptions, it remains a critical perspective for theorizing about IMF fairness. In the next section, we will deal with how confounding can be incorporated into our estimands, and we will illustrate how to estimate $\delta(x_i)$.

5.2. Adjusting for confounding to identify IMF fairness

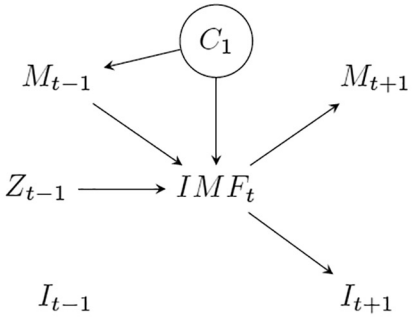
Because the evaluation of IMF policy relies on observational data, all the estimands that we have defined $(\alpha, \beta, \gamma, \delta, \delta_i)$ assume an empirical analysis in which an "as-if random" situation has been secured. However, the IMF and recipient governments agree to collaborate for several reasons besides macroeconomic considerations, violating the "as-if random" assumption. These other reasons are key confounders and must be accounted for in the empirical identification of causal effects. Figure 2 adds confounding to the DAGs, with the variable *C* representing one or potentially several confounders. In panel *a*, C_1 represents the relevant confounders that affect the process of selecting into an IMF program. As previously discussed, a government's political will to implement demanding and unpopular policies is often cited as one of the most critical confounders (Daoud and Johansson, 2020; Dreher, 2009; Stubbs et al., 2018; Vreeland, 2007). Conditioning (adjusting or controlling) for these confounders is critical to the production of unbiased estimates of $(\alpha, \beta, \gamma, \delta, \delta_i)$. *Conditioning* means producing a measure of *C* and including it when estimating any of the equa-



(a) Common causes in the selection process



(b) Common causes in IMF-policy impact



(c) Instrumental variable scenario.

Figure 2. Directed acyclic graphs depicting stylized causal systems with confounding. Notes: (1) a circle around *C* denotes the presence of unobserved confounding, while the lack of a circle means it is an observed variable. (2) *Z* represents an instrumental variable.

tions defined above, as in $P(\cdot|C)$ where the dot "·" represents any variable of interest. We can condition on any number of variables that are required to empirically identify IMF fairness.

Thus all of our five estimands can be empirically identified by conditioning on the appropriate adjustment set *C*. For example, when evaluating Principle 1, we add C_1 to $P(IMF_t|do(M_{t-1}), C_1)$, enabling us to calculate an unbiased estimate of α. As shown in Figure 2, conditioning on C_1 blocks the influence of C_1 as a common cause. Likewise, we use C_2 to identify the conditional independence between I_{t-1} and IMF_t, and can thus identify β. Similarly, as shown in panel *b* of Figure 2, we calculate γ and δ by conditioning on C_3 and C_4 in their respective formulas. To identify the estimands δ for lax and stringent δ_i sufficientarianism, we must include C_4 in the adjustment set.

As previously mentioned, when key confounders remain unobservable, alternative identification strategies must be used. In such cases, scholars frequently use an instrumental variable approach (Angrist et al., 1996; Dreher, 2009; Lang, 2016) or a Heckman selection model (Heckman, 1979; Vreeland, 2007). These approaches rely on identifying one or several instruments. An *instrument* is a variable that affects the exposure but not the outcome. This causal situation is depicted in panel *c* in Figure 2, where *Z* is the instrument. One critical assumption associated with this method is that an instrument can be found that is uncorrelated with both the confounder (i.e., the absence of a causal arrow pointing to the confounder) and the outcome. While not entirely unproblematic, countries' "voting patterns at the UN" and "the number of countries already enrolled in IMF programs" have proven to be valuable instruments (Lang, 2016; Stubbs et al., 2018). When evaluating the sufficientarianism fairness of IMF policies in the presence of unobserved confounding, scholars can rely on instruments of this kind.

The identification of causal effects depends on the hypothesized causal system captured in a DAG. The DAGs presented in Figures 1 and 2 are stylized and need to be adapted to the domain of interest (Daoud and Dubhashi, 2020). A *domain* defines a causal system, which primarily contains an outcome, an IMF policy, and confounders. As discussed in the review of empirical studies, the causal pathways connecting IMF programs to individual or macroeconomic wellness depend on substantive theories that will vary depending on the outcomes of interest (e.g., child health, poverty, inequality). Thus, a DAG needs to be adapted to the domain of interest *d*, and we denote the DAG of a domain with G_d (as in *graph*). The subscript *d* clarifies that a DAG encoded in a *G* refers to a specific domain. For each domain, the DAG, $G_1, G_2, \cdots G_d$ in relation to which IMF fairness is to be evaluated must be specified. One way to clarify this reference system is to index $P(\cdot)$ with the domain of interest (i.e., $G_1, G_2, \cdots G_d$). For example, we would denote the probability of an outcome in *WoI*, conditional on an IMF policy intervention, $do(IMF)$, in G_d as follows:

$$P_{G_d}(I|do(IMF), C) \tag{13}$$

Because DAGs encode different joint probability distributions, they are likely to produce different IMF-fairness metrics for each of the five principles. Our IMF fairness approach does not favor any particular DAG and the appropriate DAG should be chosen on the basis of substantive theory. If there are several competing causal systems, such as G_1, G_2, G_3, then a reasonable procedure is to sum the estimated effects across the proposed causal systems and weigh them by their evidence base (our belief about how likely each *G* is).This weighting of DAGs can be accomplished via ensembles, which are used in causal approaches in computer science (van der Laan et al., 2007) and can be found in the algorithmic-fairness literature (Kusner et al., 2017; Loftus et al., 2018).

5.3. A simulation: Illustrating the measurement of IMF fairness

To demonstrate the empirical feasibility of our approach, we have produced a worked example based on a simulation. The Supplement specifies the code for this simulation. The data were generated using the DAG presented in panel b of Figure 2, with confounding. We generated M_{t-1} to have a positive effect on IMF selection IMF_t; M_{t-1} has no effect on IMF_t; IMF_t has a positive effect on both future macroeconomic performance M_{t+1} and individuals' future outcomes I_{t+1}. To estimate the quantities of interest $(\alpha, \beta, \gamma, \delta, \delta_i)$, we use an estimator named the *generalized random forest* (GRF) (Athey et al., 2019), which is a machine-learning estimator tailored for causal inference and is especially suitable for approximating δ_i. [6] GRF was used to estimate all quantities of interest (estimands).

As Table 4 shows, under this data-generated model, IMF fairness is empirically measurable, with the model being designed to satisfy all five principles. Values above zero mean beneficial causal effects, while values below zero mean adverse effects. Zero values mean no effect. Principle 1 evaluates whether a government is entering an IMF program due to poor macroeconomic wellness, and here we expect to find $\alpha > 0$. Our model produces $\widehat{\alpha} = 0.18$ with a standard deviation of 0.02. The hat in $\widehat{\alpha}$ means that we are looking at an estimate of the estimand α. An estimate could vary if a different sample or model were employed. A positive and statistically significant value of $\widehat{\alpha}$ provides evidence that Principle 1 is fulfilled. To evaluate Principle 2—that a government is not entering an IMF program based on considerations of individual wellness—we estimate two models, one that includes I_{t-1} and one that does not. The model that includes I_{t-1} produces an estimate of 0.184 with a standard deviation of 0.023; the model without I_{t-1} produces an estimate of 0.180 with a standard deviation of 0.024. We then take the difference of these two models, and obtain the result $\widehat{\beta} = 0$. We can thus conclude that Principle 2 is fulfilled. The logic for evaluating Principles 3 and 4 is similar.

For Principle 5, we produce Figure 3 and check that for all individuals, across their covariate values $X = x_i$, the IMF effect is either zero or beneficial (above zero). The covariates X denote the country and household context in which individuals (e.g., children) are living. We generated our data in such a way that the effect is around $\delta(x_i) = 0$ when the x_i covariates range between -2 to 0, and $\delta(x_i) > 0$ when the covariates x_i range between 0 to 2. Even though that for most instances $\delta(x_i) > 0$, because is $\delta(x_i) < 0$ when x_i is close to -1, we conclude that IMF fairness is *not* fulfilled for stringent sufficientarianism.

6. Conclusions

IMF policies have attracted a considerable amount of criticism for not assuming sufficient responsibility for the effects of austerity measures, and thereby increasing levels of hardship in vulnerable countries and populations (Babb, 2005; McKee et al., 2012; Stuckler and Basu, 2013). Focusing on children's living conditions, our literature review shows that IMF policies tend to affect children adversely. Our literature review identified an adverse effect of IMF policies in seven of the nine articles we examined. Considering only those studies that had adjusted for a key unobserved confounder, i.e., a government's political will to implement IMF policies, we found that three studies out of three reported an adverse effect. Our review thus suggests that attempts to redesign IMF policies so that they do not harm vulnerable populations have to date been insufficient.

Despite the findings of our literature review, we have argued that there is reason to assess the ethical principles under which the IMF and its partnering governments should be held responsible for a deterioration in individual wellness while pursuing improvements to macroeconomic wellness. Such an assessment provides a basis for assigning responsibility for policy outcomes in terms of what we have denoted *IMF fairness*. As no IMF-fairness framework currently exists, our article fills this knowledge gap. The proposed framework combines distributive justice theory (Adler, 2012) with the use of causal inference and algorithmic fairness in the field of public policy (Balgi et al., 2022a; Kusner et al., 2017; Loftus et al., 2018). We have formulated five principles that are empirically measurable using statistical causal inference tools.

Our article shows how the IMF Articles of Agreement could be qualitatively reformulated based on our five principles. Of the central distributive justice theories, sufficientarianism is that which best fits the mission of the IMF. In the context of IMF policies, sufficientarianism holds that a fair outcome is achieved when individuals still have sufficient resources subsequent to policy implementation. Infusing IMF programs with sufficientarianism would enable the IMF to continue targeting macroeconomic wellness while at the same time assuming an ethical responsibility for the wellness of individuals. Assuming such an ethical responsibility would mean that the two parties who formulate and endorse an IMF program—the IMF and the partnering government—must account for the causal effects that an IMF policy is likely to have on the health and material living standards of individuals.

One objection to our IMF-fairness framework is that the IMF's current focus on macroeconomic wellness already reflects the importance of individual wellness. However, as we have argued in Section 4, and as has been noted by many before us, macroeconomic wellness is not of value in and of itself (Adler, 2012; Daoud, 2018; Nussbaum, 2000; Reddy and Daoud, 2020; Sen, 1992). While it is true that without a functioning economy, population health would be expected to suffer, the economy is only a means to an end, in the same way as other macro (e.g., health) systems. It is individuals, not systems, that constitute the focus of any distributive justice theory.

A second objection is that our IMF-fairness framework may appear to be a roundabout way of saying that "the IMF should do a better job protecting the poor."[7] As has been discussed in Sections II and III, the aim of IMF policies is to improve macroeconomic wellness (Principle 1) and should be agnostic to protecting the poor (Principle 2). In fact, what sufficientarianism implies is that in the pursuit of improvements to macroeconomic performance (Principle 3), policymakers may have to accept some decline in individual wellness, and that this may be reasonable as long as the level of individual wellness remains above the well-being threshold. The IMF should be held ethically accountable only when its programs push individuals below this well-being threshold (Principle 4 or 5). The doctor analogy is helpful to clarify what we mean. In pursuing improvements to a patient's health (e.g., removing a cancerous tumor), a doctor may subject the patient to treatment that causes a significant short-term deterioration in health (e.g., surgery or radiation therapy). As long as this deterioration stands in proportion to the illness and the patient's long-term improvement, the doctor should not be held ethically accountable for causing this short-term health deterioration. By analogy, as the IMF and a government reform an economy, some shortfall for individuals may be accepted, as long as it reflects the principles of IMF-fairness.

A third objection to our IMF-fairness framework is that it requires a level of knowledge and precision beyond that which

[6] See Künzel et al. (2019) for an overview of similar algorithms.

[7] We thank our reviewers for asking us to clarify this objection.

Table 4
Simulation results.

	Definition	Estimand	Estimate	IMF fairness principle satisfied
Principle 1:	*"Based only on macroeconomic wellness" criterion*	$\alpha > 0$	$\widehat{\alpha} = 0.18(0.02)$	Yes
Principle 2	*"Agnostic to individual wellness" criterion*	$\beta = 0$	$\widehat{\beta} = 0(0.00)$	Yes
Principle 3	*Improving macroeconomic wellness*	$\gamma > 0$	$\widehat{\gamma} = 1.62(0.07)$	Yes
Principle 4	*Satisfying IMF fairness based on lax sufficientarianism*	$\delta \geq 0$	$\widehat{\delta} = 0.33(0.05)$	Yes
Principle 5	*Satisfying IMF fairness based on stringent sufficientarianism*	$\delta_i \approx \delta(x_i) \geq 0$	$\widehat{\delta}(x_i) =$ see Figure 3	No

Notes.
1. "^" implies an estimate of the quantity of interest (estimand).
2. The quantities in parentheses are the standard deviation of the estimate.

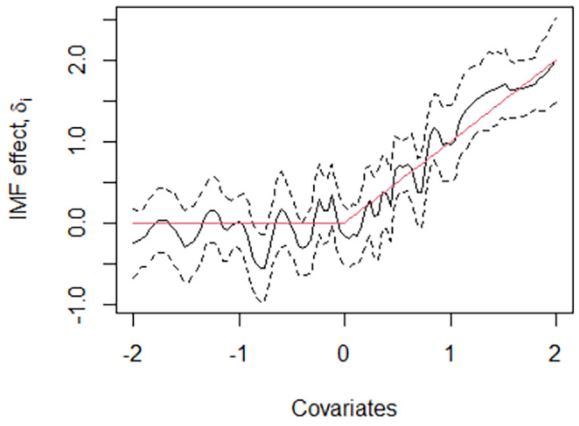


Figure 3. Evaluating Principle 5 using the Conditional Average Treatment Effect (CATE).

the IMF and partnering governments could produce in terms of empirical evidence. The absence of conclusive evidence about IMF fairness would then be likely to cripple the IMF and potential recipient countries, thus harming macroeconomic performance. Indeed, the estimation of causal effects on the basis of observational data does constitute a practical difficulty in relation to the implementation of our IMF-fairness framework (Imbens and Rubin, 2015). As with most social science research, producing high-confidence and precise empirical evidence to underpin IMF-fairness metrics does constitute a challenge. Moreover, knowledge is fallible and changing (Bhaskar, 2008). Adherence to the proposed standard of IMF fairness should therefore be interpreted in terms of probabilistic statements. The larger the number of studies that focus on measuring IMF fairness, the better the evidence base on the ethical implications of IMF policies will become for scholars and policymakers. This evidence will also help the IMF to improve its policies.

The hurdle for valid inference is higher for stringent sufficientarianism than for lax sufficientarianism. Estimating individual-level treatment effects requires more data, knowledge, and statistical assumptions than the estimation of average treatment effects (Pearl, 2009), since the data must be collected at the individual level and the causal system must be well defined. The collection of such data may appear daunting, but several recent studies have assembled data of this kind, demonstrating that it can be achieved (Balgi et al., 2022b; Daoud, 2021, 2020; Daoud et al., 2017; Daoud and Johansson, 2020). The evaluation of lax sufficientarianism is achievable with existing causal methods because it does not necessarily require individual-level data. Thus, estimating lax IMF fairness is achievable and with a level of methodological difficulty that matches that of most observational studies (Imbens and Rubin, 2015). The field of IMF policy-evaluation research has an established methodological tradition for dealing with the difficulties that exist (Stubbs et al., 2018). What has been lacking in this research is the ability to infer ethical conclusions on the basis of observational studies. Our IMF-fairness framework fills this gap. The simulated example presented in Section 5 shows how to empirically translate causal findings into drawing ethical conclusions.

As we argued in Section 5, stringent sufficientarianism constitutes the approach that is best aligned with the principles of sufficientarianism, and the lax approach should therefore only be considered when the necessary individual-level data are not available. The more fine-grained the data, the greater our ability to assess IMF fairness. We recommend two things that would facilitate systematically evaluating stringent sufficientarianism in relation to IMF programs. First, both recipient governments and the IMF need to enlarge their surveying and data-measuring capabilities at the individual level. In a way similar to that used by the United States to monitor the effects of its foreign aid on the basis of the Demographic and Health Surveys (Corsi et al., 2012), the IMF and recipient governments should develop similar surveys to measure the impact of IMF policies on individuals' health and material living standards. Second, based on these data, IMF policy-evaluation research would benefit from starting to research the methodological links between public policy and personalized medicine (Balgi et al., 2022a; Lillie et al., 2011). Like a doctor tailoring treatment to a patient to optimize the treatment effect, IMF programs and recipient governments could then start to incorporate tailored policies to benefit specific sub-populations.

Our IMF-fairness framework has at least three implications. First, to implement our approach, the IMF and recipient governments have to establish mechanisms to ensure that the combined effect of an IMF program (a set of policies) will safeguard sufficient health and material living standards for each member of the targeted population. This should be defined as a critical condition in the agreement between the IMF and the recipient government. As governments represent sovereign states, the IMF lacks the mandate to intervene directly in domestic affairs. However, it is in the best interest of both parties to ensure that the wellness of individuals is safeguarded. Agreements should therefore stipulate that a government will assemble a committee to monitor and evaluate the effects of IMF policies on crucial dimensions of health and material living standards in the affected population. This committee should preferably consist of experts who are able to conduct policy evaluations adapted to the political, social, and cultural context of the relevant population (Woolcock, 2009).

Second, we recommend that applied studies develop their analyses to include IMF-fairness metrics ($\alpha, \beta, \gamma, \delta, \delta_i$). In Sections III and IV, we have discussed all the tools—thresholds, methods, and concepts—required to start evaluating IMF fairness. While empirical studies currently focus on *is*-questions, they can now also incorporate *ought*-questions. These metrics will provide a disciplined and transparent approach to constructively improving IMF policies, rather than relying on politically motivated criticism that is intended to cripple IMF operations.

Third, our IMF-fairness framework will provide support for a substantive shift from economics first to health first (Daniels, 1985). Our framework reinforces the mandate of IMF programs to improve macroeconomic wellness while balancing the cost to the wellness of individuals. When these concerns conflict with one another, tradeoffs have to be based on empirically grounded outcomes of interest (Deaton, 2015; Marmot and Wilkinson, 2005; McKee et al., 2012). Although more research is needed to provide a basis for making trade-offs between current and future economic and health outcomes, our IMF-fairness framework provides the tools needed to start evaluating these complex interactions.

Author contributions

Led the research: AD. Conceived and designed the research topic: AD, AH. Interpreted the results: AD. Wrote the ethics section: AH. Wrote the first draft: AD. Revised the manuscript: AD, AH, and SVS.

Acknowledgments

We thank the editors and the reviewers for their valuable comments. We want to also thank Dave Shannon for enhancing the readability of our manuscript. Adel Daoud thanks *The Royal Swedish Academy of Letters, History and Antiquities* for providing generous supporting his research.

References

Abadie, A., Diamond, A., & Hainmueller, J. (2015). Comparative Politics and the Synthetic Control Method. *American Journal of Political Science, 59*, 495–510. https://doi.org/10.1111/ajps.12116.

Abouharb, M. R., & Cingranelli, D. L. (2008). IMF programs and human rights, 1981–2003. *Review of International Organizations, 4*, 47–72. https://doi.org/10.1007/s11558-008-9050-5.

Adler, M. D. (2012). *Well-being and fair distribution: Beyond cost-benefit analysis*. New York: Oxford University Press.

Alkire, S., & Foster, J. (2011). Counting and multidimensional poverty measurement. *Journal of Public Economics, 95*, 476–487. https://doi.org/10.1016/j.jpubeco.2010.11.006.

Alkire, S., Roche, J. M., & Vaz, A. (2017). Changes Over Time in Multidimensional Poverty: Methodology and Results for 34 Countries. *World Development, 94*, 232–249. https://doi.org/10.1016/j.worlddev.2017.01.011.

Alkire, S., & Santos, M. E. (2014). Measuring Acute Poverty in the Developing World: Robustness and Scope of the Multidimensional Poverty Index. *World Development, 59*, 251–274. https://doi.org/10.1016/j.worlddev.2014.01.026.

Alkire, S., Seth, S., 2015. Identifying Destitution Through Linked Subsets of Multidimensionally Poor: An Ordinal Approach (SSRN Scholarly Paper No. ID 2755165). Social Science Research Network, Rochester, NY. https://doi.org/10.2139/ssrn.2755165.

Anderson, E. S. (1999). What Is the Point of Equality? *Ethics, 109*, 287–337. https://doi.org/10.1086/233897.

Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). Identification of Causal Effects Using Instrumental Variables. *Journal of American Statistical Association, 91*, 444–455. https://doi.org/10.2307/2291629.

Athey, S., Tibshirani, J., & Wager, S. (2019). Generalized random forests. *Annals of Statistics, 47*, 1148–1178. https://doi.org/10.1214/18-AOS1709.

Atkinson, T. (2016). Monitoring Global Poverty: Report of the Commission on Global Poverty. *The World Bank*. https://doi.org/10.1596/978-1-4648-0961-3.

Babb, S. (2005). The Social Consequences of Structural Adjustment: Recent Evidence and Current Debates. *Annual Review of Sociology, 31*, 199–222.

Balgi, S., Pena, J.M., Daoud, A., 2022a. Personalized Public Policy Analysis in Social Sciences using Causal-Graphical Normalizing Flows. Assoc. Adv. Artif. Intell. AI Soc. Impact Track.

Balgi, S., Peña, J.M., Daoud, A., 2022b. Counterfactual Analysis of the Impact of the IMF Program on Child Poverty in the Global-South Region using Causal-Graphical Normalizing Flows.

Ban, C. (2015). Austerity versus Stimulus? Understanding Fiscal Policy Change at the International Monetary Fund Since the Great Recession. *Governance, 28*, 167–183. https://doi.org/10.1111/gove.12099.

Ban, C., & Gallagher, K. P. (2015). Recalibrating Policy Orthodoxy: The IMF Since the Great Recession. *Governance, 28*, 131–146. https://doi.org/10.1111/gove.12103.

Banerjee, A., & Duflo, E. (2012). *Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty* (Reprint edition). New York: PublicAffairs.

Barro, R. J., & Lee, J.-W. (2005). IMF programs: Who is chosen and what are the effects? *J. Monet. Econ., 52*, 1245–1269. https://doi.org/10.1016/j.jmoneco.2005.04.003.

Beauchamp, T. L., & Childress, J. F. (2012). *Principles of Biomedical Ethics* (7 ed.). New York: Oxford University Press.

Benson, J. S. (2001). The impact of privatization on access in Tanzania. *Social Science and Medicine, 52*, 1903–1915. https://doi.org/10.1016/S0277-9536(00)00308-7.

Bhaskar, R. (2008). *A realist theory of science* (Rev. ed.). London and New York: Routledge.

Bird, G., Qayum, F., & Rowlands, D. (2020). The effects of IMF programs on poverty, income inequality and social expenditure in low income countries: An empirical analysis. *J. Econ. Policy Reform, 1–19*. https://doi.org/10.1080/17487870.2019.1689360.

Bohoslavsky, J. P. (2016). Economic Inequality. *Debt Crises and Human Rights, 41*, 23.

Broome, A. (2014). Back to Basics: The Great Recession and the Narrowing of IMF Policy Advice. *Governance, 28*, 147–165. https://doi.org/10.1111/gove.12098.

Burgard, S. A., & Kalousova, L. (2015). Effects of the Great Recession: Health and Well-Being. *Annual Review of Sociology, 41*, 1–21. https://doi.org/10.1146/annurev-soc-073014-112204.

Casal, P. (2007). Why Sufficiency Is Not Enough. *Ethics, 117*, 296–326. https://doi.org/10.1086/510692.

Clegg, L. (2014). Social spending targets in IMF concessional lending: US domestic politics and the institutional foundations of rapid operational change. *Rev. Int. Polit. Econ., 21*, 735–763. https://doi.org/10.1080/09692290.2013.833958.

Clements, B., Gupta, S., & Nozaki, M. (2013). What happens to social spending in IMF-supported programmes? *Applied Economics, 45*, 4022–4033. https://doi.org/10.1080/00036846.2012.744136.

Cohen, G. A. (1989). On the Currency of Egalitarian Justice. *Ethics, 99*, 906–944. https://doi.org/10.1086/293126.

Copelovitch, M. S. (2010). *The International Monetary Fund in the Global Economy: Banks, Bonds, and Bailouts*. Cambridge: Cambridge University Press.

Corsi, D. J., Neuman, M., Finlay, J. E., & Subramanian, S. V. (2012). Demographic and health surveys: A profile. *International Journal of Epidemiology, 41*, 1602–1613. https://doi.org/10.1093/ije/dys184.

Crisp, R. (2003). Equality, Priority, and Compassion. *Ethics, 113*, 745–763. https://doi.org/10.1086/373954.

Daniels, N. (1985). *Just Health Care*. New York: Cambridge University Press, Cambridge Cambridgeshire.

Daoud, A. (2021). *The International Monetary Fund's intervention in education systems and its impact on childrens chances of completing school*. ArXiv220100013 Econ Q-Fin.

Daoud, A. (2020). The wealth of nations and the health of populations: A quasi-experimental design of the impact of sovereign debt crises on child mortality. *ArXiv Prepr.*. ArXiv201214941.

Daoud, A. (2018). Unifying Studies of Scarcity, Abundance, and Sufficiency. *Ecological Economics, 147*, 208–217. https://doi.org/10.1016/j.ecolecon.2018.01.019.

Daoud, A. (2015). Quality of Governance, Corruption, and Absolute Child Poverty in India. *J. South Asian Dev., 10*, 1–20.

Daoud, A. (2010). Robbins and Malthus on Scarcity, Abundance, and Sufficiency. *Am. J. Econ. Sociol., 69*, 1206–1229. https://doi.org/10.1111/j.1536-7150.2010.00741.x.

Daoud, A. (2007). (Quasi)Scarcity and Global Hunger. *J. Crit. Realism, 6*, 199–225. https://doi.org/10.1558/jocr.v6i2.199.

Daoud, A., & Dubhashi, D. (2021). Melting together prediction and inference. *Obs. Stud., 7*, 1–7.

Daoud, A., Dubhashi, D., 2020. Statistical modeling: the three cultures. ArXiv201204570 Cs Stat.

Daoud, A., Halleröd, B., & Guha-Sapir, D. (2016). What Is the Association between Absolute Child Poverty, Poor Governance, and Natural Disasters? A Global Comparison of Some of the Realities of Climate Change. *PLoS ONE, 11*. https://doi.org/10.1371/journal.pone.0153296 e0153296.

Daoud, A., Johansson, F., 2020. Estimating Treatment Heterogeneity of International Monetary Fund Programs on Child Poverty with Generalized Random Forest. SocArXiv. https://doi.org/osf.io/preprints/socarxiv/awfjt.

Daoud, A., Nosrati, E., Reinsberg, B., Kentikelenis, A. E., Stubbs, T. H., & King, L. P. (2017). Impact of International Monetary Fund programs on child health. *Proceedings of the National Academy of Sciences, 114*, 6492–6497. https://doi.org/10.1073/pnas.1617353114.

Daoud, A., & Reinsberg, B. (2018). Structural adjustment, state capacity and child health: Evidence from IMF programmes. *International Journal of Epidemiology, dyy251*, 1–10. https://doi.org/10.1093/ije/dyy251.

de Rato, R. (2006). Renewing the IMF's Commitment to Low-Income Countries [WWW Document]. *IMF*. http://www.imf.org/external/np/speeches/2006/073106.htm.

Deaton, A., 2015. The Great Escape: Health, Wealth, and the Origins of Inequality, Reprint edition. ed. Princeton University Press, Princeton, NJ.

Dorsey, D., 2012. The Basic Minimum: A Welfarist Approach [WWW Document]. Camb. Core. https://doi.org/10.1017/CBO9781139062084.

Dreher, A. (2009). IMF conditionality: Theory and evidence. *Public Choice, 141*, 233–267. https://doi.org/10.1007/s11127-009-9486-z.

Dreher, A. (2006). IMF and economic growth: The effects of programs, loans, and compliance with conditionality. *World Development, 34*, 769–788. https://doi.org/10.1016/j.worlddev.2005.11.002.

Dworkin, R. (1981). What is Equality? Part 2: Equality of Resources. *Philosophy & Public Affairs, 10*, 283–345.

Feldstein, M., Feldstein, K., 1998. To IMF: First do no harm.

Fleurbaey, M. (2015). Equality versus priority: How relevant is the distinction? *Econ. Amp Philos., 31*, 203–217. https://doi.org/10.1017/S0266267115000085.
Gordon, D., Nandy, S., Pantazis, C., Pemberton, S., & Townsend, P. (2003). *Child poverty in the developing world, Studies in poverty, inequality, and social exclusion*. Bristol, UK: Polity Press.
Gordon, D., Nandy, Shailen., 2012. Measuring child poverty and deprivation, in: Minujin Z., Alberto., Nandy, Shailen. (Eds.), Global Child Poverty and Well-Being: Measurement, Concepts, Policy and Action. Policy Press, Bristol, UK; Chicago, IL, pp. 57–101.
Guio, A.-C., Gordon, D., Marlier, E., Najera, H., & Pomati, M. (2018). Towards an EU measure of child deprivation. *Child Indic. Res., 11*, 835–860. https://doi.org/10.1007/s12187-017-9491-6.
Guio, A.-C., Marlier, E., Gordon, D., Fahmy, E., Nandy, S., & Pomati, M. (2016). Improving the measurement of material deprivation at the European Union level. *J. Eur. Soc. Policy, 26*, 219–333. https://doi.org/10.1177/0958928716642947.
Gupta, S., Dicks-Mireaux, L., Khemani, R., McDonald, C., & Verhoeven, M. (2000). *Social Issues in IMF-Supported Programs* (p. 191). Pap: IMF Occas.
Hajro, Z., & Joyce, J. P. (2009). A true test: Do IMF programs hurt the poor? *Applied Economics, 41*, 295–306. https://doi.org/10.1080/00036840601007229.
Halleröd, B., Rothstein, B., Daoud, A., & Nandy, S. (2013). Bad governance and poor children: A comparative analysis of government efficiency and severe child deprivation in 68 low-and middle-income countries. *World Development, 48*, 19–31.
Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica, 47*, 153–161.
Imbens, G. W., & Rubin, D. B. (2015). *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction, 1* (edition. ed.). New York: Cambridge University Press.
IMF, 2016. Factsheet – Poverty Reduction Strategy in IMF-supported Programs [WWW Document]. URL www.imf.org (accessed 8.8.16).
Imf (2014). *Transcript of the International Monetary and Financial Committee (IMFC) Press Briefing [WWW Document]*. Fund: Int. Monet. https://www.imf.org/external/np/tr/2014/tr041214b.htm.
Imf (2011). *Articles of Agreement of the International Monetary Fund*. Washington, DC: International Monetary Fund.
Imf (2009). *Creating Policy Space - Responsive Design and Streamlined Conditionality in Recent Low-Income Country Programs*. Washington, DC: International Monetary Fund.
Imf (2001). *Human Rights and the IMF* (p. 38). FD: Finance Dev.
Ismi, A. (2004). *Impoverishing a continent: The World Bank and the IMF in Africa*. Ottawa: Canadian Centre for Policy Alternatives.
Kentikelenis, A., Stubbs, T. H., & King, L. P. (2016). IMF conditionality and development policy space, 1985–2014. *Review of International Political Economy, 23*, 543–582. https://doi.org/10.1080/09692290.2016.1174953.
Kentikelenis, A., Stubbs, T. H., & King, L. P. (2015). Structural adjustment and public spending on health: Evidence from IMF programs in low-income countries. *Social Science and Medicine, 126*, 169–176. https://doi.org/10.1016/j.socscimed.2014.12.027.
Künzel, S. R., Sekhon, J. S., Bickel, P. J., & Yu, B. (2019). Metalearners for estimating heterogeneous treatment effects using machine learning. *Proceedings of the National Academy of Sciences, 116*, 4156–4165. https://doi.org/10.1073/pnas.1804597116.
Künzel, S.R., Sekhon, J.S., Bickel, P.J., Yu, B., 2018. Meta-learners for Estimating Heterogeneous Treatment Effects using Machine Learning. ArXiv170603461 Math Stat.
Kusner, M.J., Loftus, J., Russell, C., Silva, R., 2017. Counterfactual Fairness, in: Guyon, I., Luxburg, U.V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., Garnett, R. (Eds.), Advances in Neural Information Processing Systems 30. Curran Associates, Inc., pp. 4066–4076.
Lang, V., 2016. The Economics of the Democratic Deficit: The Effect of IMF Programs on Inequality (Working paper No. 10.11588/heidok.00021875). Heidelberg. https://doi.org/DOI:10.11588/heidok.00021875.
Lillie, E. O., Patay, B., Diamant, J., Issell, B., Topol, E. J., & Schork, N. J. (2011). The n-of-1 clinical trial: The ultimate strategy for individualizing medicine? *Journal of Personalized Medicine, 8*, 161–173. https://doi.org/10.2217/pme.11.7.
Loftus, J.R., Russell, C., Kusner, M.J., Silva, R., 2018. Causal Reasoning for Algorithmic Fairness. ArXiv180505859 Cs.
Lundberg, I., Johnson, R., & Stewart, B. M. (2021). What Is Your Estimand? Defining the Target Quantity Connects Statistical Evidence to Theory 00031224211004187. *American Sociological Review*. https://doi.org/10.1177/00031224211004187.
Marmot, M., Wilkinson, R., 2005. Social Determinants of Health. OUP Oxford.
McKee, M., Karanikolos, M., Belcher, P., & Stuckler, D. (2012). Austerity: A failed experiment on the people of Europe. *Clinical Medicine, 12*, 346–350. https://doi.org/10.7861/clinmedicine.12-4-346.
Minujin, A., & Nandy, S. (2012). *Global child poverty and well-being: Measurement, concepts, policy and action*. Chicago, IL: Policy Press, Bristol, UK.
Nandy, S., Daoud, A., & Gordon, D. (2016). Examining the changing profile of undernutrition in the context of food price rises and greater inequality. *Social Science and Medicine, 149*, 153–163. https://doi.org/10.1016/j.socscimed.2015.11.036.
Nandy, S., & Svedberg, P. (2012). The Composite Index of Anthropometric Failure (CIAF): An Alternative Indicator for Malnutrition in Young Children. In V. R. Preedy (Ed.), *Handbook of Anthropometry* (pp. 127–137). New York: Springer.
Nussbaum, M. C. (2011). Capabilities, Entitlements, Rights: Supplementation and Critique. *Journal of Human Development and Capabilities, 12*, 23–37. https://doi.org/10.1080/19452829.2011.541731.
Nussbaum, M. C. (2000). *Women and human development: The capabilities approach*. New York: Cambridge University Press, Cambridge.
Oliver, H. C. (2006). In the Wake of Structural Adjustment Programs: Exploring the Relationship Between Domestic Policies and Health Outcomes in Argentina and Uruguay. *Canadian Journal of Public Health Santee Publique, 97*, 217–221.
Ord, T., 2013. The Moral Imperative toward Cost-Effectiveness in Global Health 12.
Parfit, D. (1997). Equality and Priority. *Ratio, 10*, 202–221. https://doi.org/10.1111/1467-9329.00041.
Pearl, J. (2009). *Causality: Models, Reasoning and Inference* (2nd ed.). Cambridge, U.K.; New York: Cambridge University Press.
Pongou, R., Salomon, J. A., & Ezzati, M. (2006). Health impacts of macroeconomic crises and policies: Determinants of variation in childhood malnutrition trends in Cameroon. *International Journal of Epidemiology, 35*, 648–656. https://doi.org/10.1093/ije/dyl016.
Ravallion, M. (2015). *The Economics of Poverty: History, Measurement, and Policy* (1 edition. ed.). Oxford University Press.
Rawls, J. (1971). *A theory of justice*. Cambridge, Mass: Belknap Press of Harvard University Press.
Reddy, S. G., & Daoud, A. (2020). Entitlements and Capabilities. In E. C. Martinetti, S. Osmani, & M. Qizilbash (Eds.), *The Cambridge Handbook of the Capability Approach*. Cambridge: Cambridge University Press.
Reddy, S. G., & Pogge, T. (2010). How Not to Count the Poor. In S. Anand, P. Segal, & J. E. Stiglitz (Eds.), *Debates on the Measurement of Global Poverty* (pp. 42–85). New York: Oxford University Press.
Rickard, S. J., & Caraway, T. L. (2014). International Negotiations in the Shadow of National Elections. *Int. Organ., 68*, 701–720. https://doi.org/10.1017/S0020818314000058.
Roemer, J. E. (1998). *Equality of opportunity*. Cambridge, Mass: Harvard University Press.
Sen, A. (1999). *Development as freedom* (1st. ed.). New York: Knopf.
Sen, A. K. (1992). *Inequality reexamined*. Oxford: Clarendon Press.
Shandra, C. L., Shandra, J. M., & London, B. (2012). The International Monetary Fund, Structural Adjustment, and Infant Mortality: A Cross-National Analysis of Sub-Saharan Africa. *Journal of Poverty, 16*, 194–219. https://doi.org/10.1080/10875549.2012.667059.
Shandra, C. L., Shandra, J. M., & London, B. (2011). World Bank Structural Adjustment, Water, and Sanitation: A Cross-National Analysis of Child Mortality in Sub-Saharan Africa. *Organization & Environment, 24*, 107–129. https://doi.org/10.1177/1086026611413931.
Shandra, J. M., Nobles, J., London, B., & Williamson, J. B. (2004). Dependency, democracy, and infant mortality: A quantitative, cross-national analysis of less developed countries. *Social Science and Medicine, 59*, 321–333. https://doi.org/10.1016/j.socscimed.2003.10.022.
Shields, L. (2012). The Prospects for Sufficientarianism. *Utilitas, 24*, 101–117. https://doi.org/10.1017/S0953820811000392.
Stiglitz, J.E., 2003. Globalization and Its Discontents, 1 edition. ed. W. W. Norton & Company.
Stubbs, T., Reinsberg, B., Kentikelenis, A., King, L., 2018. How to evaluate the effects of IMF conditionality. Rev. Int. Organ. https://doi.org/10.1007/s11558-018-9332-5.
Stubbs, T. H., Kentikelenis, A. E., & King, L. P. (2016). Catalyzing Aid? The IMF and Donor Behavior in Aid Allocation. *World Development, 78*, 511–528. https://doi.org/10.1016/j.worlddev.2015.10.010.
Stuckler, D., & Basu, S. (2013). *The body economic: Why austerity kills: Recessions, budget battles, and the politics of life and death*. London: Allen Lane.
Summers, L. H., & Pritchett, L. H. (1993). The Structural-Adjustment Debate. *The American Economic Review, 83*, 383–389.
Tännsjö, T. (1998). *Hedonistic utilitarianism*. Edinburgh: Edinburgh University Press.
Temkin, L. S. (2003). Egalitarianism Defended. *Ethics, 113*, 764–782. https://doi.org/10.1086/373955.
Temkin, L. S. (1993). *Inequality*. Oxford ethics series: Oxford University Press, New York.
Thomson, M., Kentikelenis, A., & Stubbs, T. (2017). Structural adjustment programmes adversely affect vulnerable populations: A systematic-narrative review of their effect on child and maternal health. *Public Health Reviews, 38*, 13. https://doi.org/10.1186/s40985-017-0059-2.
van, der L. M. J., & Petersen, M. L. (2007). Causal Effect Models for Realistic Individualized Treatment and Intention to Treat Rules. *The International Journal of Biostatistics, 3*. https://doi.org/10.2202/1557-4679.1022.
van der Laan, M. J., Polley, E. C., & Hubbard, A. E. (2007). Super Learner. *Statistical Applications in Genetics and Molecular Biology, 6*. https://doi.org/10.2202/1544-6115.1309.
VanderWeele, T. J. (2016). Explanation in causal inference: Developments in mediation and interaction. *International Journal of Epidemiology, dyw277*. https://doi.org/10.1093/ije/dyw277.
Vreeland, J.R., 2007. The International Monetary Fund: politics of conditional lending. Routledge, Taylor & Francis Group, New York, NY.
Who (2013). *Global Vaccine Action Plan: 2011–2020*. Geneva, Switzerland: World Health Organization Press.
Woolcock, M. (2009). Toward a plurality of methods in project evaluation: A contextualised approach to understanding impact trajectories and efficacy. *J. Dev. Eff., 1*, 1–14. https://doi.org/10.1080/19439340902727719.
Ziring, L., Riggs, R. E., & Plano, J. C. (2005). *The United Nations: International organization and world politics*. Cengage Learning.