Alternative Metrics for Comparing Domestic Climate Change Mitigation Efforts and the Emerging International Climate Policy Architecture

Joseph E. Aldy* and William A. Pizer†

Introduction

The issue of comparability of domestic efforts to mitigate global climate change has long played an important role in notions of successful international climate agreements. Indeed, much of the game-theory literature on international environmental agreements highlights the need for institutions to support broad and typically comparable emission-mitigation efforts by countries in order to avoid the free-riding that can cause global agreements to unravel (e.g., Barrett 2003). Similar efforts among similar countries would likely be consistent with notions of equity and contribute to a "fair" deal (e.g., Cazorla and Toman 2001). Based on various formulas for emission targets, other research has sought to identify exactly what might be a fair deal (e.g., Bosetti and Frankel 2012). Some have argued that a cost-effective and potentially efficient international environmental agreement would be one in which marginal abatement costs among emission sources are comparable across the world (e.g., Aldy et al. 2010).

The international climate architecture that has emerged from the 2014 Lima climate talks focuses on comparability of mitigation effort in a new way. Rather than exploring how, in theory, comparable effort can mitigate free-riding or define a scheme for burden-sharing, the current approach is to identify how comparability can fit into the practical framework of a

* Harvard University, Resources for the Future, the National Bureau of Economic Research, and the Center for Strategic and International Studies; Harvard Kennedy School, 79 JFK Street, Mailbox 57, Cambridge, MA 02138; Telephone: 617-496-7213; e-mail: Joseph_Aldy@hks.harvard.edu.

† Duke University, Resources for the Future, the National Bureau of Economic Research, and the Center for Global Development; 190 Rubenstein Hall, 302 Towerview Drive, Box 90311, Duke University, Durham, NC 27708; Telephone: 919-613-9286; e-mail: billy.pizer@duke.edu.

Napat Jatusripitak, Carlos Paez, and Ryan Powell provided excellent research assistance on this project. We have benefitted from comments at workshops hosted at American Enterprise Institute and Resources for the Future, a presentation at the 2014 Lima UN climate change talks, and from Charlie Kolstad, Warwick McKibbin, Adele Morris, Ian Parry, and a referee.

Review of Environmental Economics and Policy, volume 10, issue 1, Winter 2016, pp. 3–24
doi:10.1093/reep/rev013
Advance Access published on November 20, 2015

pledge-and-review mechanism for countries' contributions to greenhouse gas emission mitigation. The process of evaluating and comparing mitigation efforts would operationalize such a pledge-and-review system (Schelling 1997; Pizer 2007).[1]

How can we measure and compare effort among countries in this emerging framework? We seek to address this question by developing a set of principles for considering a variety of comparability metrics and then, based on these principles, evaluating these metrics. We focus on emission mitigation,[2] with a country's mitigation "effort" reflecting a determined action or policy deviation from its status quo with the intent to reduce greenhouse gas emissions. Here "determined" emphasizes that the actions taken are specifically aimed at reducing emissions. "Deviation from its status quo" emphasizes that the outcome of the effort is due not merely to chance but to the actions—policies, regulations, budget expenditures—undertaken by a country. We examine three broad categories of metrics: emissions (levels, intensity, and reduction from forecast), prices (carbon and energy), and costs (share of national income or consumption). Using published data, we present illustrations of the first two groups of metrics for a set of large developed and developing countries. In the case of cost metrics, we discuss the additional modeling that is required. We find that none of the metrics performs well according to all of the principles and argue that the emerging architecture could rely on a suite of metrics (rather than a single metric) for comparing mitigation effort across countries. Such an approach is akin to using multiple economic statistics to evaluate the health of the macroeconomy.

To frame our analysis, we first review the issue of comparability of effort in international climate negotiations and highlight the role of economics in informing this debate. Next we propose a set of principles for using metrics to compare effort. This is followed by a review of various metrics. We then examine how explicit demonstration of comparability of effort can facilitate international climate policy coordination and ultimately result in stronger international agreements. The final section presents a summary and conclusions.

Comparability of Mitigation Effort in International Climate Negotiations and the Role of Economics

The notion of comparability has become increasingly important in international negotiations over the past several decades. The 1992 UN Framework Convention on Climate Change and the 1997 Kyoto Protocol set explicit emission targets for developed countries, establishing the first and most enduring notion of comparability. Emission targets were tied to a 1990 base year and, under the Kyoto Protocol, negotiated in terms of country-specific percentage reductions from their 1990 emissions. There were no formal targets for developing countries, thus leaving it unclear what, if any, effort was needed for comparability on their part (Aldy and Stavins 2012a).

The 2007 Bali Action Plan, which launched a negotiating process to produce a post-Kyoto agreement, maintained the distinction between developed and developing countries, noting

[1]The emerging architecture calls for countries to make unilateral commitments (the "pledges"), to review others' commitments, to potentially adjust their commitments, and (ultimately) to make new commitments in the future.

[2]Comparability of effort could be important in other international climate policy contexts as well, such as adaptation, climate finance, geo-engineering, and research and development. However, these are beyond the scope of this article.

that "comparability of efforts" should guide consideration of developed countries' emission-mitigation efforts (1(b)(i)). As a result of the 2009 Copenhagen Accord, however, all major developed and developing countries—about eighty countries in total—voluntarily submitted emission-mitigation targets, policies, and/or actions. To promote the transparency of these mitigation pledges, the Copenhagen Accord and the 2010 Cancun Agreements included "international consultations and analysis" and "measurement, reporting, and verification"—review mechanisms comprised of reporting, technical analysis, and a period of consultation with other parties.[3]

The 2011 Durban Platform created a negotiating process for reaching a new international agreement by the end of 2015. The emerging international climate architecture has further advanced the Copenhagen–Cancun concept of "pledge and review." Complementing this multilateral process, in November 2014, the United States and China announced a bilateral mitigation pledge. Several commentators framed their evaluation of this agreement in terms of the comparability of the commitments.[4] Within this new architecture, economics can play several potential roles in comparing and leveraging commitments.

Burden Sharing

First, scholars from several disciplines have attempted to identify a fair or equitable "burden" or "effort" for countries in mitigating greenhouse gas emissions under international environmental agreements (e.g., Michaelowa, Butzengeiger, and Jung 2005; den Elzen et al. 2006; Gupta 2007; Hof and den Elzen 2010; Bosetti and Frankel 2012). These so-called burden-sharing exercises take a normative approach in developing a set of emission targets for countries, reflecting ethical principles and some long-term global objective (e.g., atmospheric concentration stabilization or a temperature goal).

Descriptive Analysis of Pledges

Rather than defining what should be, economic analysis can provide a richer description of countries' mitigation pledges. In this context, economists have assessed the total costs, marginal costs, and emission abatement relative to a no-policy reference case associated with emission targets dating back to the 1990s (e.g., Weyant and Hill 1999; McKibbin, Morris, and Wilcoxen 2011). Emission-mitigation pledges now take many different forms: targets versus 1990 or 2005 base-year emissions, percentage improvements in the CO_2-to-GDP ratio, percentage abatement versus a no-policy reference case, renewable power goals, energy efficiency goals, afforestation goals, and specific mitigation policies. An apples-to-apples comparison of countries' pledges requires economic tools that translate the pledges into emissions, energy, and economic effects.[5]

[3]However, these requirements were defined separately for developed and developing countries (see Decision 2/CP.17)
[4]See http://www.brookings.edu/blogs/planetpolicy/posts/2014/11/25-us-china-climate-deal-sussman (accessed October 14, 2015); http://www.slate.com/blogs/future_tense/2014/11/12/us_china_climate_agreement_china_more_likely_to_meet_its_targets.html (accessed October 14, 2015).
[5]Likewise, detemining the global impact on emissions through an aggregation of countries' pledges would also require the application of economic tools.

Reciprocity

Finally, economic reasoning suggests that comparable effort should leverage reciprocal mitigation actions. For example, under the 2009 Copenhagen Accord, the European Union and Japan both announced a willingness to implement more ambitious domestic emission targets if other developed countries committed to comparable reductions. [6] This concept of reciprocity has also characterized successful international trade negotiations. For example, in a review of the literature on compliance with international agreements, Simmons (1998) notes that compliance is likely better under rules "prescribing reciprocal rather than uni-obligational behavior" (p. 87). Finger, Reincke, and Castro (1999) conclude that "a sense of fairness, of appropriate contribution, was an important concept" (p. 7) in the success of the Uruguay Round of trade talks. Ostrom (1998) focuses on the importance of norms in guiding individual efforts toward collective action, emphasizing that "all reciprocity norms share the common ingredients that individuals tend to react to the positive actions of others with positive responses and the negative actions of others with negative responses" (p. 10).

Our discussion here avoids normative prescription and thus focuses on the latter two roles. We believe that negotiators, stakeholders, and the citizens of countries participating in international climate negotiations would benefit from a richer assessment of the mitigation efforts being proposed by other countries. Economic data and analysis can help to provide this information, which, in turn, can enhance the credibility of countries' pledges (Barrett 2012). Given the iterative nature of the negotiation game, this can leverage reciprocal action, which may take the form of more ambitious execution of current pledges or, following ex post analysis, more ambitious pledging in subsequent rounds of negotiations. With this in mind, we propose a set of principles for establishing "economics-informed" metrics that can be used for comparing mitigation effort across countries.

Principles for Developing Effective Metrics

Effective metrics should characterize mitigation effort in a straightforward manner that can be used to inform an international pledge-and-review process among countries. More specifically, we believe that metrics for comparing mitigation effort across countries should be based on the following four basic principles:

- Comprehensive: An ideal metric should capture the entire effort undertaken by a country to mitigate emissions. This means that the metric should clearly reflect all climate-related policies and measures (and exclude nonpolicy drivers of climate outcomes), resulting in a measure that is sufficient for comparing effort across countries.
- Measurable: An ideal metric for comparing effort should also focus on the observable characteristics of effort, such as emissions, energy and carbon prices, or use of zero-carbon technologies. This creates an incentive for countries to undertake emission-mitigation actions that other countries can easily observe, thus facilitating transparency. Metrics that allow quantitative ex ante estimates—to inform the current round of talks—and

[6] http://unfccc.int/resource/docs/2011/sb/eng/inf01r01.pdf (accessed October 14, 2013).

quantitative ex post assessments—to inform the next round of talks—would serve international climate negotiations well.
- Replicable: Analysts and stakeholders should be able to recreate a metric given (a) the inputs used in the original calculation and (b) available public information. This would ensure the legitimacy of such comparability analysis. This also requires transparent methods that allow for third party review, which would help to increase trust in the assessment and comparison of effort. This suggests that simple, less-complex metrics are preferable.
- Universal: Given the global nature of the climate challenge, the metric should be constructed for—and applicable to—as broad a set of countries as possible.

In practice, there will be trade-offs among these principles when identifying and constructing metrics. For example, emission levels may be measurable and universally available in all countries, but this measure may not comprehensively represent mitigation effort. Carbon prices—explicit or implicit—may be a better measure of effort, but such prices raise measurability issues when making comparisons across currencies and/or facing other energy-market distortions.[7]

We organize our assessment of metrics for measuring mitigation effort into three broad categories: emissions, prices, and costs. This separates physical outcome measures (the first category) from those reflecting market signals, incentivizing private sector behavior, and governing long-term investments (the second category). The third category includes measures of the economic resources diverted away from current consumption or nonmitigation investment/government spending. We discuss each of these categories in more detail next. Given the historical focus on emissions in international climate negotiations, we begin with emissions.

Emissions Metrics

The 1997 Kyoto Protocol (and the recent extension of the second-period emission commitments for some Annex I parties) established economy-wide greenhouse gas emission targets. Under the Copenhagen Accord and Cancun Agreements, nations proposed a wider variety of emission-related commitments, including economy-wide emission goals, economy-wide emission-intensity goals, emission reductions relative to a forecast business-as-usual (BAU) emission level, as well as energy efficiency programs, transportation projects, forestry conservation investments, and wind- and solar-power capacity goals. We review three subcategories of emission metrics: (1) emission levels versus a historical base year, (2) emission abatement versus a forecast of future emissions, and (3) emission-intensity goals.[8]

Emission Levels versus Historical Base Year

Measuring a country's territorial emissions is a relatively straightforward exercise. This is especially true for fossil-fuel carbon emissions, which require only an accounting of the oil, gas, and coal consumed to produce energy.[9] Certain types of greenhouse gas emissions are more

[7]The literature on program evaluation frequently notes this kind of trade-off (Coglianese 2012).
[8]Other physical measures (renewable power, use of public transportation, etc.) can be similarly subcategorized into comparisons against a base year, comparison with a baseline forecast, or relative to a particular output measure.
[9]In contrast to territorial emissions, consumption-based emissions represent an adjustment to territorial emissions based on the embodied carbon in net imports (Peters et al. 2011). As world trade expands, accounting for

Table 1 Percentage change in emissions, 1990–2012 and 1997–2012

Country	1990–2012		1997–2012	
	Fossil CO_2, percentage Change	All GHGs, percentage change	Fossil CO_2, percentage change	All GHGs, percentage change
United States	+5%	+5%	−6%	−5%
United Kingdom	−17%	−25%	−12%	−17%
European Union	−16%	−19%	−11%	−14%
China	+257%	+230%	+178%	+149%
India	+216%	+143%	+114%	+88%
Japan	+20%	+13%	+9%	+4%
Russia	−22%	−32%	+28%	+12%

Note: GHGs refer to greenhouse gas emissions.
Sources: U.S. Energy Information Administration (n.d.); World Resources Institute (n.d.); Russia 1990 data from Le Quéré et al. (2015) and UNFCCC (n.d.).

difficult to measure, such as those related to land-use change. Measurement (and accounting) can also be complicated by market-based mechanisms that allow mitigation commitments and emission offsets (such as those created through the Clean Development Mechanism) to be traded among countries.[10]

The Kyoto Protocol set developed country emission targets as a percentage change from a country's emissions level in 1990, and year-1990 emissions continues to serve as a benchmark for some countries in recent rounds of negotiations. Establishing a base year that predates the climate negotiations eliminates the risk of countries "gaming" the system by increasing emissions during negotiations in order to create a higher benchmark level for evaluation. However, the choice of base year is controversial: should the chosen year favor countries that have already undertaken significant efforts to reduce emissions (e.g., by choosing an early year) or those that have not (e.g., by choosing a later year)?

Moreover, emission levels relative to a base year may not comprehensively represent mitigation effort. Emission trends vary across countries for reasons that go beyond a specific country's policies and efforts to mitigate greenhouse gas emissions. Table 1 presents historical changes in emissions for several large developed and developing countries during the 1990–2012 and 1997–2012 periods.[11] Very few countries had implemented meaningful emission-mitigation policies before the 1997 Kyoto Conference, making 1997 an equally plausible base year.

Consider Russia's 2012 emission levels for all greenhouse gases being 32 percent below 1990 levels. One might conclude from these data that Russia is the world's leader in combating climate change. However, the decline in Russia's emissions does not reflect an extensive climate change mitigation program, as indicated by the 12 percent increase in emissions

the carbon content of traded goods and services could become an important aspect of future assessments of mitigation effort—or the basis for extensive border measures by importers or exporters. We leave this topic for future research.

[10]The Kyoto Protocol established the Clean Development Mechanism to provide developed countries the option of investing in developing countries' emission-reducing projects. These projects' emission reductions produce "emission offsets" that developed countries are allowed to use for compliance under the Kyoto Protocol.

[11]The online supplementary materials present these data for most countries in the world.

over the 1997–2012 period. Rather, the decline in 1990–2012 emissions resulted from dramatic economic restructuring after the Cold War.

More generally, the significant heterogeneity in countries' economic and population growth suggests that the effective stringency of a common percentage reduction from a common base year could differ dramatically among countries. Using emission levels in 1997 (rather than 1990) as a benchmark, the EU emission-mitigation effort would appear to represent a 14 percent reduction instead of a 19 percent reduction, whereas U.S. emissions would decline 5 percent instead of growing 5 percent. Using either benchmark, China's and India's carbon dioxide emissions have increased significantly, although a meaningful share of this growth occurred between 1990 and 1997.[12] Although emissions relative to a base year is the most commonly used metric for comparing mitigation across countries and is relatively measurable, replicable, and universal, it is far from comprehensive.

Emission Intensity

Emission intensity (tons of carbon dioxide per GDP) is often viewed as a metric that avoids the problem of penalizing economic growth that occurs when using the emission base-year metric. The concept emerged in 1999, when the government of Argentina proposed an emission commitment specified as a convex function of economic output (Aldy 2004). In 2002, the United States proposed an emission goal structured as a percentage reduction in the ratio of emissions to GDP. In the run-up to the 2009 Copenhagen talks, China and India proposed similarly structured emission-intensity goals.

There are several drawbacks to using emission intensity as a measure of mitigation effort: (1) emissions will continue to grow unless the reduction in emission intensity exceeds the economic growth rate, (2) countries that are growing tend to experience an automatic decline in emission intensity owing to technology improvement and the changing structure of the economy, and (3) analysis has shown that emission intensity targets often become more stringent if a country grows slower than expected and less stringent if it grows faster than expected (Aldy 2004). The third drawback suggests using an "indexed" approach, where an emission target is adjusted ("indexed") based on deviations from expected economic activity, rather than the strict proportionality implied by an emission-intensity metric (Newell and Pizer 2008).

As with emissions, it is relatively straightforward to measure and use GDP to compute intensity. However, comparing emission intensities across countries at a point in time requires converting local currencies into a single currency. The first two columns of table 2 indicate that the use of market exchange rate versus purchasing power parity affects multicountry comparisons.[13] For example, under market exchange rates, emission intensity in China is about five times higher than in the United States and about 35 percent higher than in India. Emission intensity in India is more than three times higher than in the United States. If purchasing power parity exchange rates are used, China's intensity is about 60 percent higher than energy intensity in the United States. Alternatively, an emission-intensity metric could focus on changes in

[12]The UNFCCC does not receive timely information on greenhouse gas emissions from developing countries (Aldy 2013). Some of the largest developing countries have reported no more than one emission inventory per decade to the UNFCCC since the 1992 Earth Summit. The question is whether this reflects technical or political obstacles. Thus, we have used the World Resources Institute (WRI) Climate Analysis and Indicators Tool (CAIT) database for greenhouse gas emissions.

[13]See online supplementary materials for an expanded table 2.

Table 2 Emissions intensity (CO_2/GDP ratio, 2012) and CO_2/GDP (percentage change in local currency)

Country	CO_2/GDP (2012$, market exchange rate)	CO_2/GDP (2012$, purchasing power parity)	CO_2/GDP (LCU) annual rate of change (1997–2012)	GDP (LCU) annual rate of change (1997–2012)
United States	0.32	0.33	−2.6%	+2.2%
United Kingdom	0.17	0.21	−2.8%	+2.0%
China	1.54	0.54	−2.5%	+9.3%
India	1.14	0.29	−1.7%	+6.8%
Japan	0.23	0.28	0%	+0.6%

Note: Measures represent one metric ton of CO_2 from fossil-fuel combustion per $1,000 of GDP. LCU denotes local currency unit.
Sources: U.S. Energy Information Administration (2000, n.d.); World Bank (n.d.); World Resources Institute (n.d.).

intensity over time (adjusting for inflation). The last two columns of table 2 illustrate the earlier point that growing economies generally have declining intensities, making it difficult to isolate mitigation from background trends.

Emission Reductions versus a Forecast of Future Emissions

In recent years, some large developing countries have proposed emission abatement goals as percentage reductions from a reference-case forecast of the future level of emissions. Brazil, Indonesia, Korea, and South Africa chose such targets under the Copenhagen Accord. In theory, by capturing the emission reduction directly associated with a climate-mitigation program, this metric has the potential to be a more comprehensive indicator of mitigation effort than emission levels benchmarked to a historical base year or adjusted based on GDP.

In practice, however, calculating the reference-case emission estimate used to define emission reductions requires subjective and uncertain judgments. Unlike the other two emission metrics, where observed outcomes are compared with historical data (or with other countries), this metric compares outcomes with assumptions about what would have happened without the mitigation policy. It is easy to imagine how countries might "game" the calculation, especially developing countries experiencing volatile economic growth. With emissions growth tied to economic growth, assuming faster economic growth would produce a higher reference-case forecast emission level. If the country has slower-than-forecast economic growth, then the goal is that much easier to achieve. Even with an unbiased reference-case forecast, differences between observed and forecast outcomes may be unrelated to mitigation. To illustrate, figure 1 presents actual 2010 emissions versus a year-2000 forecast for 2010 for several countries. By definition, the forecasts published in 2000 do not include policies proposed and implemented after 2000, including the EU Emission Trading Scheme (ETS). Thus, for the European Union, the changes in emissions relative to the year-2000 forecast could be due to explicit climate-change policies. However, given the absence of significant policies in the other countries, the differences between actual and forecast emissions for China, India, Brazil, and the United States reflect the difficulty of predicting future emissions. Given the forecasting errors shown in figure 1, it is easy to see how a country could strategically (and credibly) claim a more

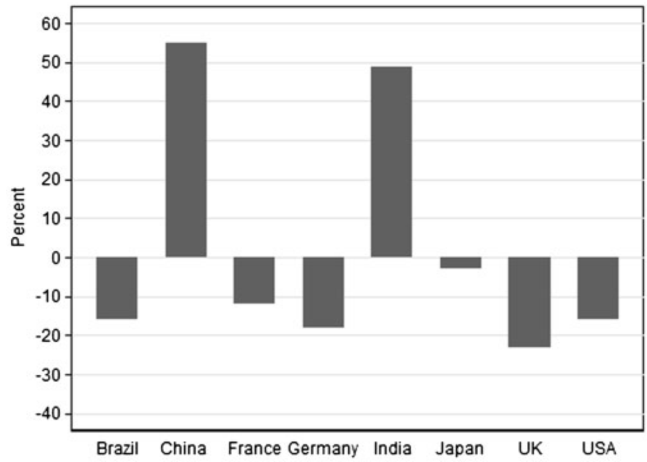

Figure 1 Actual 2010 emissions relative to the "business-as-usual" forecast for 2010
Note: Forecast was published in 2000.
Sources: Le Quere et al. 2015; U.S. Energy Information Administration 2000.

favorable baseline, as well as how countries might easily achieve their emission targets due to events that are unrelated to mitigation efforts. A retrospective analysis that revises the reference-case forecast based on new data but subtracts any effects of mitigation policy could address these problems. However, this approach would still require subjective assumptions and modeling to account for the impacts of mitigation policy.

Price Metrics

In the context of climate policy, most industry stakeholders and economic policymakers focus on the second category of metrics—price, specifically the prices for fuels and electricity. These market signals drive private-sector behavior and determine long-term investments and also influence trade. Ultimately, the delivered price of fossil energy reflects a combination of global and local resource costs, other tax and subsidy policies, and any explicit carbon price. Thus we consider both explicit carbon prices as well as energy prices and taxes.

Carbon Prices

The observed carbon price is directly connected to climate mitigation effort because it measures the marginal cost of those emissions from sources covered by a country's climate-change program. Households and firms should undertake all covered mitigation opportunities that are less expensive per ton than the carbon price. A comparison of carbon prices across countries indicates the extent to which a country is creating incentives for more or less expensive (per ton) mitigation efforts in the sectors it regulates. However, because countries implement domestic carbon taxes in their local currencies, it is unclear how these different currencies should be valued. Although market exchange rates are most appropriate in the context of competitiveness concerns and traded goods, purchasing power parity exchange rates allow us to compare domestic carbon prices in relation to a common basket of domestic goods and wages. This is relevant for examining notions of burden—that is, what households are giving up in order to mitigate emissions.[14] Thus, a comprehensive evaluation would likely need to compare carbon prices using both market and purchasing power parity exchange rates.

In addition to the challenge of comparing different currencies, explicit carbon prices may not accurately reflect mitigation effort. First, an explicit carbon price may be an excessively narrow measure of a country's efforts to mitigate greenhouse gas emissions. For example, it may cover only a subset of a country's emissions (e.g., only power plants), and it may fail to account for the effect of other, nonprice policies that reduce greenhouse gas emissions. For example, efficiency standards or regulations supporting renewable energy can have significant emission consequences and represent significant effort but are not reflected in carbon prices (or energy prices). There is also the risk that a country may undermine the effectiveness of the carbon price by adjusting taxes downward (or increasing subsidies) for firms covered by the carbon price.[15] Finally, the effort represented by a carbon price—in terms of resources expended—depends on both the price and the amount of emissions reduced. For a country with particularly inelastic demand—and relatively few opportunities to reduce emissions—a high carbon price may do little to change emissions.[16]

An alternative to using explicit carbon prices is to consider implicit (or "effective") carbon prices, which estimate the average cost of abatement associated with a specific climate policy or collection of policies (see Organisation of Economic Co-operation and Development 2013). Such implicit prices have the advantage of potentially being applied more broadly but the disadvantage of not being directly observed (i.e., they are derived from model simulations). Because estimated costs are divided by estimated abatement, implicit carbon prices will be high for expensive policies and/or those that produce little net abatement. Moreover, in contrast with high explicit carbon prices, high implicit prices are not market signals, do not suggest that all less expensive mitigation options are being incentivized or taken, and are not as relevant for investment or trade decisions.

Table 3 illustrates some of these issues by presenting the explicit carbon prices observed under specific carbon tax and cap-and-trade policies and two implicit carbon prices estimated from renewable subsidies in Germany. The relatively high implicit carbon price associated with Germany's solar feed-in tariff could reflect significant effort, less effectiveness, or a combination of both.[17] There is also considerable within-country variation. For example, Norway's carbon tax in 2009 set prices of \$58/tCO_2 for gasoline, \$34/tCO_2 for diesel, \$31–33/tCO_2 for natural gas, and exempted coal (Aldy and Stavins 2012b); some sectors in Norway were subject to the EU ETS with a market price in the range of €5–10/tCO_2 in 2012. In Germany, half of the economy faced the EU ETS market price, but wind and solar feed-in tariffs were associated with an implicit price of € 62/tCO_2 and €537/tCO_2, respectively (Marcantonini and Ellerman 2013).

[14]Market exchange rates also face the challenge of currency devaluation. For example, during the Mexican peso crisis, Mexico's currency devalued by one-third; during the Asian financial crisis, Korea's currency devalued by one-half.
[15]This is typically referred to as fiscal cushioning (see Wiener 1999).
[16]In their analysis of the Copenhagen mitigation commitments, McKibbin, Morris, and Wilcoxen (2011) find that the ranking of effort changes depending on whether the analysis is based on carbon prices or economic costs (foregone consumption). We address this issue in our discussion of cost metrics.
[17]In the underlying study, Marcantonini and Ellerman (2013) show that it is a bit of both.

Table 3 Carbon prices under various energy and climate policies

Country	Program	Cost ($)/ tCO_2	Source/date
EU ETS	Cap-and-trade	€8	Average daily price, 2012 (Datastream International BlueNext Series)
Germany	Wind feed-in-tariff	€62	Average abatement cost, 2010 (Marcantonini and Ellerman 2013) (2011€)
Germany	Solar feed-in-tariff	€547	Average abatement cost, 2010, (Marcantonini and Ellerman 2013) (2011€)
Australia	Cap-and-trade/tax hybrid	A$23	2012 fixed price (Environmental Defense Fund and the International Emissions Trading Association 2013)
New Zealand	Cap-and-trade	NZ$3.15	January 3, 2014 closing spot price (Point Carbon 2014)
Regional Greenhouse Gas Initiative	Cap-and-trade	$2.67	September 2013 auction clearing price (Regional Greenhouse Gas Initiative 2013)
California	Cap-and-trade	$11.75	December 2013 futures price (November 22, 2013) (CME Group 2013)
Shenzen, China	Cap-and-trade	US$7	August 2013 (Climate Group 2013)
Quebec	Cap-and-trade	C$10.75	Auction price floor 2013, (Government of Quebec n.d.)
British Columbia	Tax	C$30	2013 (Government of British Columbia n.d.)
Alberta	C performance standard	C$15	2012 (Government of Alberta 2012)
Denmark	Tax	€3–90	2009, Industry, varies by type of industry and voluntary agreement (Aldy and Stavins 2012b)
Finland	Tax	€20	2009 (Aldy and Stavins 2012b)
Norway	Tax	NOK92–363	2009 (Aldy and Stavins 2012b)
Sweden	Tax	€114	Households and services, 2012 (Aldy and Stavins 2012b)

This within-country variation suggests that, for comparison purposes, it would be helpful to construct a single economy-wide average carbon price that weights implicit and explicit carbon prices throughout the economy.[18]

Energy Prices and Taxes

While carbon prices reflect explicit efforts to reduce emissions, it is energy prices that matter for both the supply and demand for energy, investment in energy technologies by businesses and households, and the trade flows associated with particular energy-intensive activities. Thus, energy prices are likely to provide a more comprehensive picture of the long-term incentives to decarbonize the economy. Higher overall energy prices will drive more investment in energy efficiency, and higher relative prices for more carbon-intensive energy sources will spur investment in low- and zero-carbon technologies. Energy prices are transparent and measurable with high frequency. Energy prices allow for a net assessment of all price-based policies, including carbon pricing, and thus can mitigate concerns that a country engages in fiscal cushioning by

[18]See Organisation of Economic Co-operation and Development (2013) for a discussion of such an aggregation.

Table 4 Average energy prices and energy taxes (in 2012) and change since 1997 for select OECD countries

Country	Average energy price (2012US$/MMBTU)	1997–2012 change in energy price (percentage)	Average energy tax (2012US$/MMBTU)	1997–2012 change in energy tax (percentage)
United States	9.2	+17%	4.1	−6%
United Kingdom	19.4	+94%	2.4	+16%
France	26.6	+66%	5.3	+10%
Germany	27.3	+98%	7.4	+87%
Japan	42.6	+44%	3.2	0%
Canada	9.3	+50%	1.0	−80%
Australia	45.4	+294%	15.3	+85%
Mexico	9.1	+33%	0.9	+65%

Notes: The average energy price (tax) reflects a consumption-weighted measure of end-user prices (based on a market exchange rate in 2012 dollars) per million British thermal units (MMBTU) basis.
Source: International Energy Agency (2013).

simultaneously imposing a carbon tax and targeted tax relief. However, for countries pursuing significant nonprice regulation, energy prices may be a poor measure of effort.[19] Although energy-price metrics could capture some of the effects of nonprice regulations, most nonprice regulations mitigate emissions largely without influencing energy prices (which explains why they often attract political support).

Of course, not all energy-price differences across countries or over time represent policy choices. Different resource endowments coupled with transportation constraints have led to significant regional disparities in coal and natural-gas prices.[20] Moreover, price changes over time may reflect fundamental supply-and-demand shifts that are unrelated to policy changes. The transparency of taxes and other price-based policies means that the policy component of a given energy price can be conveniently presented alongside the net price to consumers. This suggests a matrix of energy-price metrics that includes levels and changes over time, for both the net price and the policy component.

Table 4, which presents such a matrix for a set of Organisation for Economic Co-operation and Development (OECD) countries over the 1997–2012 period, indicates that average energy prices vary by a factor of four, while average energy taxes vary by a factor of more than ten. Many countries experienced significant growth in energy prices, owing to higher crude-oil prices and, in some regions of the world, higher natural-gas prices (the United States is a notable exception in the case of natural-gas prices). The average energy tax grew significantly over 1997–2012 in some countries, such as Australia, Germany, and Mexico, while falling significantly in Canada.[21]

[19]The energy price metric is also less relevant for some countries in which a large fraction of their greenhouse gas emissions occur beyond the energy sector (e.g., Brazil and Indonesia due to land-use change or New Zealand due to agriculture). In these cases, additional metrics to compare action will be necessary.
[20]Coal and natural gas are characterized by regional markets with contemporaneous prices varying by 5X (natural gas) to 10X (coal) across these markets in recent years. While oil prices may differ depending on crude-oil characteristics and location, these differences are typically on the order of no more than 10 percent to 20 percent; this is why most analysts treat oil as having a world price.
[21]See online supplementary materials for an extension of table 4 to all OECD countries and for data on gasoline prices for most countries in the world.

Fossil-energy prices could also be useful in identifying both any initial energy-price distortions and the impact of a carbon tax on the prices of the underlying fuels delivered to end-users. Perhaps the most comprehensive metric would be an “effective carbon tax” for end-users of each fossil fuel (as well as electricity) that takes into account all price-based policies affecting that fuel. This would reflect a net energy “tax” on carbon-dioxide emissions. In contrast to the implicit carbon price discussed earlier and the more general cost metrics discussed in the next section, both of which focus on the cost of abated carbon dioxide, this carbon-tax metric would capture the price charged for unabated carbon dioxide from fossil-fuel use.

Cost Metrics

The third category of metrics—the mitigation costs of domestic climate policy—is the one most consistent with economists’ notion of mitigation effort. This metric is comprehensive and, if expressed as a share of national or per capita income, could be scaled to be comparable across countries of vastly different sizes. Moreover, concern about the costs of combating climate change is one of, if not the most, significant impediment to meaningful global action. Thus, a metric that compares mitigation effort based on costs could promote confidence in the fairness of global international action by ensuring that comparable countries bear comparable costs. In addition, when combined with information about emission reductions, it could be used to identify the potential advantages of certain policies (i.e., those that reduce more emissions with lower mitigation costs) over others.

In the same way that emissions are easily observed but reductions are not, explicit prices are easily observed but costs are not. Thus, emission reductions and costs both require models to construct counterfactual outcomes. But estimating costs typically requires even more assumptions and modeling.

Partial-Equilibrium Analysis of Costs

One option would be to use simple partial-equilibrium analyses of the mitigation costs associated with different policies. For sectors with explicit carbon prices, estimates of emission reductions could be combined with these explicit prices to produce estimates of total costs.[22] For other sectors, this approach would be comparable with the construction of implicit carbon prices, which are typically defined as total costs divided by reductions in a given sector. A key weakness of this approach is that it fails to consider how policies in different sectors may interact.

Integrated Modeling of Mitigation Costs

Alternatively, one could pursue a more integrated modeling approach. For more than two decades, researchers in the energy-economic modeling literature have published mitigation cost estimates of climate change policy.[23] These models provide important insights (e.g., the

[22]For example, assuming linear marginal costs rising from zero to the observed carbon price level, one could estimate costs as ½ × (marginal costs) × (emission reductions).

[23]See Gaskins and Weyant (1993) for a comparison of U.S. mitigation costs, Weyant and Hill (1999) for a comparison of Annex I mitigation costs under the Kyoto Protocol, and McKibbin, Morris, and Wilcoxen (2011) for a recent assessment of Copenhagen mitigation commitments

economic gains to trading among countries and across time) that have informed the debate on climate policy design and implementation. There are two limitations to extending such integrated modeling to analysis of comparability of effort. First, most models focus on only a small set of large countries and regions. For example, the McKibbin, Morris, and Wilcoxen (2011) assessment of the Copenhagen commitments assumes a world economy with six countries and five regions. Second, these models are most appropriate for evaluating economy-wide carbon price policies. However, in practice, countries implement myriad overlapping sectoral policies. The United Nations Framework Convention on Climate Change (UNFCCC 2011) reports that Annex I countries have implemented more than one thousand mitigation policies. It may be beyond the capacity of these models to incorporate all or many of these instruments, although recent efforts through the Stanford Energy Modeling Forum have attempted to include small combinations of policy instruments (such as carbon pricing, fuel economy standards, and power-sector renewables mandates). To the extent that a broader range of policies can be modeled, the assessment of costs could combine an evaluation of the observed mix of possibly cost-ineffective policies with an assessment of a cost-minimizing effort to achieve the same reductions. This could inform countries about the efficacy of their mitigation programs.

The Issue of Negative Costs

Any cost metric would need to recognize the potential for policies to have negative costs. Consider two examples. First, suppose a country implements a tax swap (i.e., imposing a tax on carbon while reducing labor and/or capital tax rates). Depending on the distortions of the preexisting factor taxes and the nature of the tax swap, a country could experience greater economic growth while reducing emissions. Second, suppose a country eliminates fossil-fuel consumption subsidies. For some countries, this could significantly reduce emissions while contributing to faster economic growth (International Energy Agency et al. 2010). Even with negative overall costs, both policy options could involve significant costs among some, if not many, stakeholders, thus requiring substantial political, as opposed to economic, effort.

Using Metrics of Comparability in the Design and Implementation of International Climate Policy: Issues and Challenges

Combinations of these metrics could play a variety of roles in the design and implementation of the emerging international climate architecture. Before considering how metrics could inform the future international climate policy architecture, we summarize in table 5 how the various metrics perform against our four design principles. The table indicates that no single metric does well against all of the principles. Emission levels and intensity perform particularly poorly in terms of comprehensiveness. This is because many factors that are unrelated to emission-mitigation policies can influence emissions. Carbon prices do well on effort per abated ton, but actual effort depends on the total tons abated. Although energy prices capture the market signals for low carbon investment and behavioral change, they fail to reflect nonmarket policies. For many countries, including the United States, nonmarket policies dominate the landscape.

Table 5 Synthesis of metrics and principles

Metric	Principle			
	Comprehensive	**Measurable**	**Replicable**	**Universal**
Emission levels	A poor estimate of effort because it conflates natural trends	Yes	Yes; public domain data for energy and fossil CO_2 available	Fossil CO_2 data exist for all countries; additional work needed for all GHGs
Emission intensities	Better than emission levels as it controls for economic trends, but a noisy signal	Yes	Yes; public domain data for energy and fossil CO_2 available	Yes for fossil CO_2/GDP; additional work needed for GHG/GDP
Emission abatement	Most comprehensive among emission-related metrics	Challenging—requires modeling tools / subjective choices to determine counterfactuals	Different model structures with different assumptions could yield different outcomes	No, few modeling platforms evaluate more than ~10 countries
Carbon prices	Captures effort per ton, but says little about tons	Explicit, yes; implicit requires detailed analyses	Yes for explicit prices; implicit prices may depend on analytic assumptions	No, given few explicit C pricing policies; modeling tools necessary for implicit C prices
Energy prices and taxes	Inadequate for nonenergy emissions; fails to account for nonmarket regulatory instruments	Yes, but unclear how to aggregate	Yes	Yes, but requires more detailed data collection than currently in public domain
Abatement costs	Best measure of effort, still requires benchmarking	Challenging—requires modeling tools / subjective choices to determine counterfactuals and model costs	Different model structures with different assumptions could yield different outcomes	No, few modeling platforms to comprehensively evaluate more than ~10 countries

Note: GHGs refer to greenhouse gas emissions.

Emission abatement and abatement costs would appear to be the metrics that best indicate mitigation effort. However, they are also the most difficult to measure, requiring assumptions and modeling to implement them. Credible differences in opinion on assumptions or modeling approach could produce different results for abatement and costs, suggesting that estimated measures may not be replicable. Moreover, thus far, most modeling tools have only been applied to the largest developed and developing countries. This suggests that a portfolio of metrics is necessary and that considerable work remains to be done to construct more comprehensive metrics for estimating emission reductions and costs in other countries.

What roles might such a portfolio play in the emerging international architecture? First, the assessment and comparison of effort could influence the legal structure of mitigation commitments and enable a two-part or hybrid approach in which emission-mitigation pledges are voluntary while the review of effort is legally binding. Second, estimating metrics will require a rigorous system of policy surveillance that can draw from experiences in other multilateral policy contexts, such as economic and trade policy. The implementation of surveillance could aim to produce comparability metrics for ex ante review of proposed commitments and ex post assessment of the performance of those pledged mitigation actions. Finally, the application of metrics will naturally lead to consideration of how to benchmark effort for the countries participating in the international climate agreement.

The Legal Structure of Emission-Mitigation Commitments

National governments are more likely to take stronger action within an international agreement to combat climate change if they have clear assurances that all countries will be making a fair contribution to the global effort. Given the evolution in Copenhagen and Cancun toward a pledge-and-review approach to emission-mitigation commitments, one way to operationalize the comparability of effort would be to modify the notion of what constitutes a legally binding commitment.

For example, an international climate policy architecture could require all countries to submit two-part emission commitments.[24] In the first part, each country would pledge emission-mitigation goals, policies, and/or actions (e.g., an economy-wide emission goal versus a base year, elimination of fossil fuel subsidies, a carbon tax). In the second part, each country would regularly produce data and analysis to indicate the impacts of its pledged commitment, thereby facilitating the comparison with other countries' pledged commitments.[25] Because some countries and some stakeholders continue to call for "legally binding" commitments, the second part of this approach—the provision of data and analysis to promote transparency and facilitate comparisons—could be deemed "legally binding." In addition to providing a comparison of effort, this two-part commitment could provide performance measures that permit a more general evaluation of various policy approaches.

Importance of Policy Surveillance

Distinct from the ultimate use of metrics for comparing effort across countries is the process of producing and validating those metrics. Many countries have neither the resources nor the capacity to evaluate other countries' commitments and performance and may be suspicious of assurances provided by those countries that have such capacity. This suggests the need for professional, regular, and independent assessment of countries' policies, actions, and emissions, ideally through an official, agreed-upon process, which would help to inform the periodic rounds of international climate negotiations.

Models of policy surveillance

There are several models for such official policy surveillance (see Aldy 2013, 2014 for details). For example, the International Monetary Fund undertakes so-called Article IV consultations of member governments' economic, fiscal, and monetary policies. Under the World Trade Organization, the Trade Policy Review Mechanism evaluates the trade policies of members,

[24]New Zealand introduced such a proposal in the international climate negotiations in March 2014. See https://unfccc.int/files/documentation/submissions_from_parties/adp/application/pdf/adp2-4_submission_by_new_zealand_submission_20140312.pdf (accessed October 14, 2015).

[25]Such analysis could be subject to agreed-upon guidelines to facilitate comparability.

with greater frequency for the largest trading nations. These treaty organizations have created teams of experts that conduct policy surveillance by evaluating economic, fiscal, monetary, and trade policies in member states. In contrast, the G-20 tasked international organizations (the World Bank, OECD, International Energy Agency, and Organization of the Petroleum Exporting Countries [OPEC]) with identifying fossil-fuel subsidies and evaluating the performance of G-20 nations in reducing their fossil-fuel subsidies pursuant to the 2009 Pittsburgh G-20 Leaders' Agreement (Aldy 2015). Under the UNFCCC, ad hoc groups of experts evaluate the emission inventories submitted by developed countries, as well as these countries' national communications.[26]

Suite of metrics

Because there is no single metric that fully satisfies our comprehensive principle, we recommend a policy-surveillance mechanism that includes a suite of metrics, which would provide a richer characterization of effort. Just as an economic analyst would review a suite of data to understand the strength of a national economy (e.g., GDP, unemployment rate, interest rates, business investment), so too would a climate policy analyst benefit from a more comprehensive assessment of emission-mitigation effort that includes detailed information on carbon and energy prices, abatement costs, and emission reductions.

Ex ante and ex post review

Such a surveillance mechanism could play an important role in both ex ante and ex post assessments of mitigation commitments (or contributions). In the former, countries could propose mitigation commitments that facilitate the comparability of effort exercise. Countries could be required to submit their own data and analyses to demonstrate the effort they plan to undertake to meet their commitments. A cadre of experts could process, compile, analyze, and construct comparisons based on submitted data and analyses as well as third-party data and analyses. This could inform the negotiations as commitments are finalized. A similar process could be followed for ex post assessments of countries' efforts to satisfy their pledged commitments. Past experience suggests there are likely to be political stumbling blocks to securing multilateral agreement on such a policy-surveillance regime.

Thus, scholars and nongovernmental experts may need to fill this analytic gap. Unofficial but independent expert analysis could further synthesize these data to estimate metrics that require forecasts and modeling (e.g., Aldy, Pizer, and Akimoto 2015). In turn, stakeholders and other users could provide feedback on the feasibility, integrity, and precision of available metrics and estimates. This would enable further refinement and improvement of estimates going forward. In addition, the work on developing metrics for ex ante comparisons of effort could be used to inform the data collection and analysis needs of ex post reviews. Finally, the retrospective review of pledges will be more informative and more effective if countries plan in advance for such reviews by implementing data collection and dissemination protocols. Given that the 2015 UN climate talks in Paris are just the beginning of an ongoing process of policy commitments, these

[26]Thompson (2006) noted that the UNFCCC assessments of national communications do not permit a comparison of effort by a given country from one round of national communications to the next, let alone a meaningful comparison of effort across countries.

types of refinements and improvements would ultimately promote greater confidence and stronger levels of effort among all countries.

Benchmarking for Comparability

After constructing a metric, a logical next step is benchmarking. Each metric lends itself to ranking countries based on their raw metric values, but negotiators, policymakers, and stakeholders may want to go a step further. That step may involve setting an absolute threshold to distinguish "good" and "bad" values, which in turn might be differentiated by country.[27] Alternatively, it may involve grouping countries in ways that make rankings and relative comparisons more appropriate.

While we have illustrated the design and application of metrics as a positive exercise, we believe that the choice of benchmarks—whether it is by emphasizing one metric over others or defining what level is "enough"—is normative. Thus, we do not make the case for specific benchmarks here. Nonetheless, we highlight several lines of reasoning used by others to motivate particular benchmarks.

Adequacy of global ambition

First, an absolute benchmark could be chosen that reflects an assessment of the adequacy of collective effort in realizing the objectives of the UNFCCC or those suggested by scientists. For example, the Copenhagen Accord and Cancun Agreements identify a long-term objective to limit warming to no more than 2°C, which could be linked to emission levels. Alternatively, various estimates of the social cost of carbon could be compared with collective or individual carbon prices or tax metrics. It should be noted, however, that the social cost of carbon and a 2°C warming limit may yield different sets of benchmarks. For example, Nordhaus (2008) shows that a 2°C warming limit is inconsistent with setting a globally harmonized carbon tax equal to the social cost of carbon.

"Common but Differentiated Responsibilities and Respective Capabilities"

Second, absolute and relative comparisons could be designed to reflect the UNFCCC principle of "common but differentiated responsibilities and respective capabilities." For example, it may appear odd to use any metric to compare the mitigation efforts of China and Chad because the former has emissions that are four orders of magnitude greater than the emissions of the latter. Or, it may appear odd to use a metric that compares the mitigation programs of Singapore and Ethiopia because the former's per capita emissions are three orders of magnitude greater than those of the latter. Or, it may appear odd to compare effort metrics for Qatar and Bangladesh because per capita income in the former is two orders of magnitude greater than in the latter. This suggests creating peer groups along various dimensions to provide an opportunity for more appropriate relative comparisons or to set absolute benchmarks.

To illustrate, if the international community agreed on several official metrics for the largest economies, then the comparability-of-effort exercise could be designed in a way that would

[27]For an example, refer to the Climate Action Tracker rating of individual countries' pledged emission contributions under the Durban Platform negotiations: http://climateactiontracker.org/countries.html (accessed October 14, 2015).

allow for comparisons among countries with the greatest mutual mitigation interests (e.g., Major Economies Forum on Energy and Climate [MEF] membership).[28] Indeed, to demonstrate the feasibility and applicability of a comparability of effort process, the MEF countries could voluntarily adopt metrics and present data and analysis regarding their future emission commitments.

National circumstances

The idea of official peer groups, metrics, and benchmarks raises the issue of exactly how to tailor benchmarks to reflect national circumstances. For example, although the United States, Canada, and Australia have had annual population growth on the order of 1 percent per year, Europe and Japan have not. On a per capita basis, China, India, and other emerging economies are not as wealthy as the United States. Some countries have plentiful fossil resources, whereas others do not. How should we recalibrate their benchmarks to reflect these differences? Some countries have taken significant climate mitigation actions, whereas others have lagged behind. How should benchmark differentiation (or the metrics themselves) be balanced to reward the first movers versus providing an incentive to the laggards to increase their action? These questions—and the possibility that official, negotiated solutions may not be forthcoming—suggest a need for continued independent research in this area.

Summary and Conclusions

Metrics aimed at comparing countries' climate change actions are becoming increasingly relevant as we transition to a system of unilateral pledges of domestic action within international negotiations. The emerging architecture asks countries to state what they intend to do, review other countries' commitments and (along with various stakeholders) determine their adequacy, and then react accordingly. This reaction may be in the form of adjustments to their original commitments, follow-through on those commitments, or new commitments in the future. Countries may also react through bilateral or "minilateral" cooperation. For example, countries may assess others' efforts to determine if it would be appropriate for them to link domestic cap-and-trade programs or establish harmonized carbon-tax policies. Moreover, some countries may decide to unilaterally act (through border taxes) against those failing to undertake comparable effort.[29] Over time, these outside reviews may also help countries improve the design of their own policies.

Our examination of metrics for comparing mitigation efforts reveals some deep differences across metrics. First, some metrics—total emissions and explicit emission prices—are relatively easy to observe and measure, but they may be removed from the key concepts of effort and underlying policy implementation. Second, the metrics that more closely reflect effort—emission reductions, implicit prices, and costs—are harder to observe and measure directly, leading to more subjective and possibly inconsistent estimates. Finally, metrics can be

[28]The MEF facilitates dialogue on climate change policy among the largest developed and developing countries, including Australia, Brazil, Canada, China, the European Union, France, Germany, India, Indonesia, Italy, Japan, Korea, Mexico, Russia, South Africa, the United Kingdom, and the United States.
[29]See Aldy and Pizer (2014) for more details on how comparability of effort assessments could inform cap-and-trade linkage and border tax imposition.

constructed or benchmarked in a variety of ways that may or may not adjust for resource endowments, historical behavior, or future growth. More appropriate relative comparisons could be achieved by grouping countries with peers, and more specific absolute benchmarks could be developed based on experience from past climate negotiations.

A serious, professional, transparent, and legitimate process is required to develop metrics and benchmarks for assessing comparability of effort, compiling data and related analysis in light of these metrics, and reporting the results of the assessments, particularly for policy-surveillance purposes. However, reaching an agreement on specific metrics, benchmarks, and a comprehensive surveillance mechanism is a tall order given the UN climate negotiations' poor track record on transparency (Aldy 2014).

In the meantime, further research is needed to improve methods for assessing and comparing mitigation effort. More specifically, it would be helpful to develop an array of easily available metrics, such as those we have presented here, using data collected by existing international organizations. This would facilitate comparisons of mitigation effort in the near term—in advance of any official policy surveillance or benchmarking. Moreover, independent researchers could further analyze and synthesize these data to create some of the more challenging but informative metrics (e.g., emission abatement costs and emission reductions). This would allow stakeholders and other users to provide feedback on the feasibility, integrity, and precision of these metrics, which could be used to further refine metrics and to inform ongoing deliberations concerning metrics, benchmarks, and mitigation contributions.

References

Aldy, Joseph E. 2004. Saving the planet cost-effectively: The role of economic analysis in climate change mitigation policy. In *Painting the White House green: Rationalizing environmental policy inside the executive office of the president*, ed. R. Lutter and J. F. Shogren. Washington, DC: Resources for the Future Press. 89–118.

Aldy, Joseph E 2013. Designing a Bretton Woods institution to address global climate change. In *Handbook of energy and climate change*, ed. Roger Fouquet. Cheltenham, UK: Edward Elgar. 352–73.

———. 2014. The crucial role of policy surveillance in international climate policy. *Climatic Change* 126 (3–4): 279–92.

———. 2015. Policy surveillance in the G-20 fossil fuel subsidies agreement: Lessons for climate policy. *Climatic Change*, forthcoming. DOI 10.1007/s10584-015-1505-0.

Aldy, Joseph E., Alan J. Krupnick, Richard G. Newell, Ian W.H. Parry, and William A. Pizer. 2010. Designing climate mitigation policy. *Journal of Economic Literature* 48 (4): 903–34.

Aldy, Joseph E., and William A. Pizer. 2014. Comparability of effort in international climate policy. Harvard Project on Climate Agreements Discussion Paper 14-62, Harvard University, Cambridge, MA.

Aldy, Joseph E., William A. Pizer, and Keigo Akimoto. 2015. Comparing emissions mitigation efforts across countries. RFF Discussion Paper 15-32, Resources for the Future, Washington, DC.

Aldy, Joseph E., and Robert N. Stavins. 2012a. Climate negotiators create an opportunity for scholars. *Science* 337: 1043–44.

———. 2012b. The promise and problems of pricing carbon: Theory and experience. *Journal of Economic Development* 21 (2): 152–80.

Barrett, Scott. 2003. *Environment and statecraft: The strategy of environmental treaty-making*. Oxford: Oxford University Press.

———. 2012. Credible commitments, focal points, and tipping: The strategy of climate treaty design. *Climate change and common sense: Essays in honour of Tom Schelling*, ed. R. W. Hahn and A. Ulph. New York: Oxford University Press. 29–49.

Bosetti, Valentine, and Jeffrey Frankel. 2012. Politically feasible emission targets to attain 460ppm CO_2 concentrations. *Review of Environmental Economics and Policy* 6 (1): 86–109.

Cazorla, Marina, and Michael Toman. 2001. International equity and climate change policy. In *Climate Change Economics and Policy: An RFF Anthology*, ed. Michael A. Toman. Washington, DC: RFF Press. 235–47

Climate Group. 2013. Carbon price in China's first trading market overtakes EU ETS. August 22. http://www.theclimategroup.org/what-we-do/news-and-blogs/carbon-price-in-chinas-first-trading-market-overtakes-eu/ (accessed January 2, 2014).

CME Group. 2013. California carbon allowance futures. http://www.cmegroup.com/trading/energy/emissions/california-carbon-allowance-cca-futures_quotes_globex.html (accessed November 22, 2013).

Coglianese, Cary. 2012. Measuring regulatory performance: Evaluating the impact of regulation and regulatory policy. http://www.oecd.org/gov/regulatory-policy/1_coglianese%20web.pdf (accessed October 14, 2015).

den Elzen, Michel G. J., Marcel Berk, Paul Lucas, Patrick Criqui, and Alban Kitous. 2006. Multi-stage: A rule-based evolution of future commitments under the Climate Change Convention. *International Environmental Agreements: Politics, Law and Economics* 6 (1): 1–28.

Environmental Defense Fund and the International Emissions Trading Association. 2013. The world's carbon markets: A case study guide to emissions trading—Australia. June. http://www.ieta.org/assets/Reports/EmissionsTradingAroundTheWorld/edf_ieta_australia_case_study_september_2013.pdf (accessed January 2, 2014).

Finger, J. Michael, Ulrich Reincke, and Adriana Castro. 1999. Market access bargaining in the Uruguay Round: Rigid or relaxed reciprocity? World Bank Policy Research Working Paper 2258, The World Bank, Washingon, DC.

Gaskins, Darius W., and John P. Weyant. 1993. Model comparisons of the costs of reducing CO_2 emissions. *American Economic Review* 83 (2): 318–23.

Government of Alberta. 2012. Greenhouse gas emission reduction program results. http://environment.alberta.ca/04220.html (accessed January 2, 2014).

Government of British Columbia. n.d. Carbon tax report and plan. http://www.fin.gov.bc.ca/tbs/tp/climate/Carbon_Tax_Report_and_Plan_Topic_Box.pdf (accessed December 5, 2013).

Government of Quebec. n.d. The carbon market. http://www.mddefp.gouv.qc.ca/changements/carbone/Ventes-encheres-en.htm (accessed January 2, 2014).

Gupta, Joyeeta. 2007. Beyond graduation and deepening: Toward cosmopolitan scholarship. In *Architectures for agreement: Addressing global climate change in the post-Kyoto world*, ed. J. E. Aldy and R. N. Stavins. Cambridge, UK: Cambridge University Press. 116–30.

Hof, Andries F., and Michel G. J. den Elzen. 2010. The effect of different historical emissions datasets on emission targets of the sectoral mitigation approach triptych. *Climate Policy* 10 (6): 684–704.

International Energy Agency. 2013. *Energy statistics of OECD Countries 2013*. Paris: IEA.

International Energy Agency, OPEC, OECD, and the World Bank. 2010. Analysis of the scope of energy subsidies and suggestions for the G-20 initiative. Joint report prepared for submission to the G-20 Summit meeting Toronto, Canada.

Le Quéré, C. et al. 2015. Global carbon budget 2014. *Earth System Science Data* 7: 47–85. http://www.earth-syst-sci-data-discuss.net/7/521/2014/essdd-7-521-2014.html.

Marcantonini, Claudio, and A. Denny Ellerman. 2013. The cost of abating CO_2 emissions by renewable energy incentives in Germany. MIT CEEPR Working Paper 2013-05, MIT, Cambridge, MA.

McKibbin, Warwick J., Adele C. Morris, and Peter J. Wilcoxen. 2011. Comparing climate commitments: A model-based analysis of the Copenhagen Accord. *Climate Change Economics* 2 (2): 79–103.

Michaelowa, Axel, Sonja Butzengeiger, and Martina Jung. 2005. Graduation and deepening: An ambitious post-2012 climate policy scenario. *International Environmental Agreements: Politics, Law and Economics* 5 (1): 25–46.

Newell, Richard G., and William A. Pizer. 2008. Indexed regulation. *Journal of Environmental Economics and Management* 56 (3): 221–33.

Nordhaus, William D. 2008. *A Question of balance: Weighing the options on global warming policies*. New Haven, CT: Yale University Press.

Organisation of Economic Co-operation and Development. 2013. *Effective carbon prices*. Paris: OECD.

Ostrom, Elinor. 1998. A behavioral approach to the rational choice theory of collective action: Presidential address. *American Political Science Review* 92 (1): 1–22.

Peters, Glen P., Jan C. Minx, Christopher L. Weber, and Ottmar Edenhofer. 2011. Growth in emission transfers via international trade from 1990 to 2008. *Proceedings of the National Academy of Sciences* 108 (21): 8903–8.

Pizer, William A. 2007. Practical global climate policy. In *Architectures for agreement: Addressing global climate change in the post-Kyoto world*, ed. J. E. Aldy and R. N. Stavins. Cambridge, UK: Cambridge University Press. 280–314.

Point Carbon 2014. NZ carbon unchanged, untraded in holiday period. http://www.pointcarbon.com/news/reutersnews/1.3531458?date=20140103&sdtc=1 (accessed January 5, 2014).

Regional Greenhouse Gas Initiative. 2013. Auction 21 market monitor report. http://www.rggi.org/docs/Auctions/21/Auction_21_Market_Monitor_Report.pdf (accessed December 5, 2013).

Schelling, Thomas C. 1997. The cost of combating global warming. *Foreign Affairs* 76 (6): 8–14.

Simmons, Beth A. 1998. Compliance with international agreements. *Annual Review of Political Science* 1: 75–93.

Thompson, Alexander. 2006. Management under anarchy: the international politics of climate change. *Climatic Change* 78: 7–29.

UNFCCC. n.d. Greenhouse gas inventory data—Detailed data by party. http://unfccc.int/di/DetailedByParty.do (accessed July 27, 2015).

———. 2011. Compilation and synthesis of fifth national communications: Executive summary. Report FCCC/SBI/2011/INF.1.

U.S. Energy Information Administration. n.d. International energy statistics. http://www.eia.gov/cfapps/ipdbproject/IEDIndex3.cfm (accessed July 27, 2015).

———. 2000. *International energy outlook 2000*. Washington, DC: Department of Energy.

Weyant, John P., and Jennifer Hill. 1999. The costs of the Kyoto Protocol: A multi-model evaluation, introduction and overview. *Energy Journal,* Special issue on the costs of the Kyoto Protocol: a multi-model evaluation 20: vii–xiiv.

Wiener, Jonathan Baert. 1999. Global environmental regulation: Instrument choice in legal context. *Yale Law Journal* 108: 677–800.

World Bank. n.d. World development indicators online database. http://data.worldbank.org/data-catalog/world-development-indicators (accessed July 27, 2015).

World Resources Institute. n.d. Climate analysis indicators tool. http://cait2.wri.org (accessed July 27, 2015).