Sustainable Development Law & Policy

Volume 6
Issue 2 *Winter 2006: Climate Law Special Edition*

Article 8

A Statistically-driven Approach to Offset-based GHG Additionality Determinations: What Can We Learn?

Mark C. Trexler

Derik J. Broekhoff

Laura H. Kosloff

Recommended Citation

Trexler, Mark C., Derik J. Broekhoff and Laura H. Kosloff. "A Statistically-driven Approach to Offset-based GHG Additionality Determinations: What Can We Learn?" Sustainable Development Law & Policy, Winter 2006, 30-40.

A Statistically-Driven Approach to Offset-Based GHG Additionality Determinations:

What Can We Learn?

*by Dr. Mark C. Trexler, Derik J. Broekhoff, and Laura H. Kosloff**

Introduction

Seventeen years ago, researchers at the World Resources Institute ("WRI") put together the first additionality and baseline analysis for a greenhouse gas ("GHG") mitigation project. The project was the CARE Agroforestry Project in Guatemala, funded by AES Corporation in 1989 as the first corporate carbon offset project.[1] There was little question that the CARE Agroforestry Project happened only because of AES Corporation's concerns about climate change. AES would not have pursued the project otherwise; in other words, it was clearly "additional." Such clarity of action cannot be seen in all the carbon-offset projects that have followed. Indeed, "additionality" remains the single most contentious issue in the development of today's voluntary and compliance-based carbon offset programs and GHG markets.

> **Additionality: Never has so much been said about a topic by so many, without ever agreeing on a common vocabulary, and the goals of the conversation.**
>
> – Dr. Mark C. Trexler, presentation at Additionality Side Event, COP-10 in Buenos Aires (2004).

What is there about differentiating between "non-additional" and "additional" projects that has vexed offset-based emissions trading efforts for so many years? Why are so many elements of today's additionality debate basically unchanged from the debates of five or ten years ago? Is additionality even that important? Is it key to today's offset-based emissions trading programs, as some observers argue, or should we simply drop it in favor of "getting things done," as others argue? If it is key, is there a viable path through the additionality conundrum, or do we need to scale back expectations for offset-based emissions reduction programs?

To help the search for a viable path, this article takes a step back from the day-to-day additionality debates taking place in today's voluntary and compliance-based GHG markets. Instead, the article looks at additionality through the lens of statistical hypothesis testing – *i.e.* the task of controlling for "phantom reductions" ("false positives") and "lost opportunities" ("false negatives") when testing a hypothesis – and explores how a statistical approach to additionality might be practically applied to the design of environmentally sound markets for offset-based emissions reductions.

This article addresses four key areas. The first section defines additionality and differentiates between the concept of additionality and its application in practice. The second section presents ways in which the elements of statistical hypothesis testing can be meaningfully applied to additionality testing for climate change mitigation purposes. The third section discusses what the statistical basis of additionality testing means for the environmental integrity of the supply of offset credits entering the market. The final section presents policy recommendations related to the design of offset-based GHG credit markets.

Some of the analysis and recommendations presented in this article are directly applicable to the ongoing additionality debate surrounding the Kyoto Protocol's Clean Development Mechanism ("CDM"). The analysis applies equally well to offset-based emissions trading at any level, whether state-specific (*e.g.* design of a California trading system), region-specific (*e.g.* design of a trading system under the Regional Greenhouse Gas Initiative), country-specific (*e.g.* design of Canada's offset system), or internationally (*e.g.* the CDM). The conceptual challenges facing offset-based emissions trading are basically the same, regardless of the specific trading system involved or what kinds of offsets are being considered.[2]

** Dr. Mark C. Trexler is President of Trexler Climate + Energy Services, Inc., in Portland, Oregon. He has worked in the climate change field since 1988 when he joined the World Resources Institute in Washington, DC. He carried out the baseline analysis for the first carbon offset project, the CARE Guatemala Agroforestry Project pursued by AES Corp. in 1989. He has reviewed hundreds of offset projects, worked extensively on offset-based emissions trading mechanisms, and written extensively on the subject of additionality. Dr. Trexler welcomes comments at mtrexler@climateservices.com.*

Derik J. Broekhoff is a Senior Associate with the World Resources Institute in Washington, DC. He is a primary author of the WRI/WBCSD GHG Protocol for Project Accounting, and has worked extensively on offset-based additionality analysis. Mr. Broekhoff welcomes comments at dbroekhoff@wri.org.

Laura H. Kosloff is Vice President and General Counsel of Trexler Climate + Energy Services, Inc. She has worked on climate change legal and policy developments of since 1992. She oversees TC+ES's negotiations and contracting for carbon offset projects, and has developed pathbreaking legal agreements among private parties, governmental agencies, and nongovernmental organizations addressing legal issues associated with the long-term reliability of offset projects.

ADDITIONALITY: CONCEPT VS. APPLICATION

Carbon offsets allow GHG emitters to continue to emit GHGs in one place by procuring GHG "credits" from somewhere else (whether next door or around the world), thus meeting either voluntary or mandatory emissions reduction targets. In strictly conceptual terms, the need for carbon offsets to be "additional" is easy to understand. Emissions trading systems are premised on capping overall emissions from a certain set of sources at an absolute level. An "offset credit" allows emissions from these capped sources to increase with the understanding that this increase is "offset" by a reduction from a source whose emissions are not capped, leaving net emissions unchanged. To accomplish this objective, the reduction from the uncapped source must be a response to the presence of the offset crediting mechanism. If emissions reductions would have happened regardless of any offset credits, then issuing credits for them would allow global emissions to rise beyond what was intended under the cap. Credited reductions must therefore be *additional* to reductions that would have occurred in the absence of the trading system.[3]

For offset-based trading additionality, this boils down to why a given project is being undertaken. There are many potential reasons for implementing emissions-reducing projects. For additionality purposes,[4] the question is whether the availability of offset credits is a decisive reason (although not necessarily the *only* reason) for pursuing the emissions reduction project. The question boils down to a kind of thought experiment: holding everything else constant, would a project have happened in the absence of the offset crediting mechanism (*i.e.* if it and all other projects were not eligible for offset credits)?[5] If yes, then the project is not additional; if no, then the project is additional.[6]

Unfortunately, it is impossible to definitively answer this thought experiment. Even if we could read the minds of project developers, they themselves may not know what they would have done under different circumstances. It is not even a "hypothetical" question, since a hypothesis can be empirically tested. We are forced to seek a second-best solution, namely designing questions that *are* answerable. For additionality, these questions have taken the form of what are generally called "additionality tests." A variety of tests have been developed; including mechanisms designed to

TABLE 1 ILLUSTRATIVE* ADDITIONALITY "TESTS"

Additionality Test	General Description of the Test as It Is Commonly Formulated
Legal, Regulatory, or Institutional Test	The offset project must reduce GHG emissions below the level required by any official policies, regulations, guidance, or industry standards. If it does not reduce emissions beyond these levels, the assumption is that the only real reason for pursuing the project is compliance; the project, therefore, is not additional. Under some versions of this test, the converse is true – if the project reduces emissions beyond required levels, it is assumed that the only real reason for pursuing the project is to earn credits, and the project is therefore additional.
Technology Test	The offset project and its associated GHG reductions are considered additional if the offset project involves a technology specified as not being "business as usual." The default assumption is that for these "additional" technologies, GHG reductions are a decisive reason (if not the only reason) for using the technology in a particular project.
Investment Test	The most common version of this test (often termed financial additionality) assumes an offset project to be additional if it can be demonstrated that it would have a lower than acceptable rate of return without revenue from GHG reductions. The underlying assumption is that GHG reductions must be a decisive reason for implementing a project that is not an attractive investment absent revenues associated with those reductions. Under some versions of this test, an offset project with a high or competitive rate of return could still be additional, but must demonstrate additionality through other means.
Barriers Test	Under some versions of this test, an offset project is assumed to be additional if it faces significant implementation barriers (*e.g.* local resistance to new technologies, institutional constraints, etc.). Under other versions of the test, it must further be shown that at least one alternative (*e.g.* the business as usual alternative) to the offset project does not face these barriers. The underlying assumption is that GHG reductions are a decisive reason that a project is able to overcome the identified barriers (particularly if realistic alternatives do not face these barriers).
Common Practice Test	The offset project must reduce GHG emissions below levels produced by "common practice" technologies that produce the same products and services as the offset project. If it does not, the assumption is that GHG reductions are not a decisive reason for pursuing the project (or conversely, that the only real reason is to conform to common practice for the same reasons as other actors in the same market). Therefore, "common practice" technologies are not considered to be additional.
Timing Test	The offset project must have been initiated after a certain date (*e.g.* the date of initiation of a GHG trading program) to be considered additional. The assumption is that any project started before that date must have had motivations other than GHG reductions. Under most versions of this test, offset projects started after the required date must also establish additionality through a second test.
Performance Benchmark Test	The offset project must demonstrate an emissions rate that is lower than a predetermined benchmark emissions rate for the particular technology or practice. This test is premised on the assumption that most, if not all, projects that beat the specified benchmark are ones in which climate change mitigation is a decisive factor in the decision to exceed the benchmark. The benchmark may also be used to calculate baseline emissions.
Project In, Project Out Test	The offset project must have lower GHG emissions than a scenario in which the project had not been implemented. If GHG emissions associated with the project are lower, then it is assumed that reducing emissions was a decisive reason for the project and that the project is additional.

* This table is a summary and an introduction to the variety of additionality tests that have been circulated over the past decade. It is not an exhaustive list of additionality tests, nor is it intended to provide precise definitions of the different tests.

measure environmental additionality, financial additionality, regulatory additionality, technology additionality, and others (see Table 1). At their root, these tests all are trying to answer the same question: would a project have occurred regardless of the existence of drivers created by the trading system, or not?

In practice, it has proven extremely difficult for stakeholders to agree on what tests to apply, as well as the circumstances in which particular tests are appropriate. In part, this is because people disagree about how well different tests perform with respect to the underlying objective of the tests, *i.e.* judging whether the project would have happened in the absence of an offset crediting mechanism (or more generally, without concern for climate change mitigation). As we shall show, however, it is also because they disagree on the practical importance of getting the answer to this question right or wrong when designing a working market for offset-based GHG emissions reductions.

"Phantom Reductions" and "Lost Opportunities": A Statistical Perspective on Additionality Testing

This section presents several concepts from the field of statistical hypothesis testing and explains their relevance to thinking about additionality testing. Our premise is that it is possible – and necessary – to think about additionality tests in the same way one thinks about tests in any other area of science or public policy.

Statistical Concept #1

There is no such thing as a perfect test in statistics. Any test in almost any field – whether home pregnancy kits or eligibility screening for social welfare programs – will, in addition to "correct" results, yield false positive and false negative results.[7] How well a particular test works depends on how frequently it correctly returns a positive result (its "true positive" rate) and how frequently it correctly returns a negative result (its "true negative" rate).[8]

Likewise, additionality tests are not perfect. They sometimes will falsely indicate that a project is additional when the project would have happened regardless of concerns about climate change (*i.e.* a false positive). Emissions reductions from such a project are effectively illusory, or "phantom reductions." Alternatively, an additionality test may indicate that a proposed project is not additional when in fact it is (*i.e.* a false negative). The potential reductions from such rejected projects can be thought of as "lost opportunities." As in statistical hypothesis testing, any given additionality test will produce both types of errors (see Figure 1). Additionality tests can thus be thought of as having "true positive" and "true negative" rates, although empirically and precisely determining these values may be impossible. The likely performance of different tests against key evaluative criteria, however, can be qualitatively evaluated (see Table 2).

Statistical Concept #2

The relative proportions of false positives, false negatives, and true results can vary widely depending on how a test is constructed and the nature of what is being tested. The proportions of false positives and false negatives produced by a test will depend on its rate of identifying "true positives" and "true negatives," as well as the relative proportion of true positives and negatives in the real world. For example, a profiling test used to catch criminals might falsely implicate an innocent person one out of one thousand times; one would say that it has a "true positive" rate of 99.9 percent. Nevertheless, if only one out of every million people profiled is actually a criminal, then on average about one thousand innocent people would be tagged as criminals for every true criminal profiled.[9] In other words, because criminals are so rare, the proportion of false positives to true positives can be enormous, despite the test's high "true positive" rate.

The relative proportions of "phantom reductions" and "lost opportunities" making up the final credit pool will depend not only on which additionality tests are employed (and how they are designed), but also on how many non-additional projects exist relative to additional projects in the underlying population. It is impossible to know empirically what these proportions are. Nevertheless, projects that reduce emissions relative to historical levels occur all the time and for many reasons without regard for climate change mitigation, and the relative number of these projects can be quite high (see Figure 2). Under these circumstances, even additionality tests that almost always correctly identify non-additional projects could still result in a lot of "phantom reductions" in relative terms, since many non-additional projects are slipping through – especially if the tests are not particularly good

Figure 1: The Nature of Additionality Rules: Any Attempt to Define Additionality Creates False Positives and Negatives

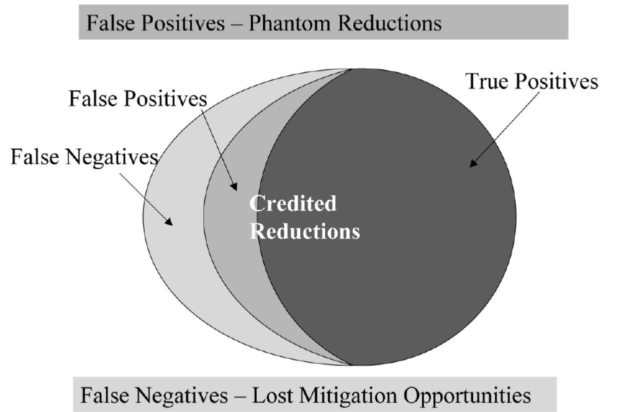


Figure 2: The Context for Additionality: The Real World Includes A Mix of Additional and Non-Additional Reductions

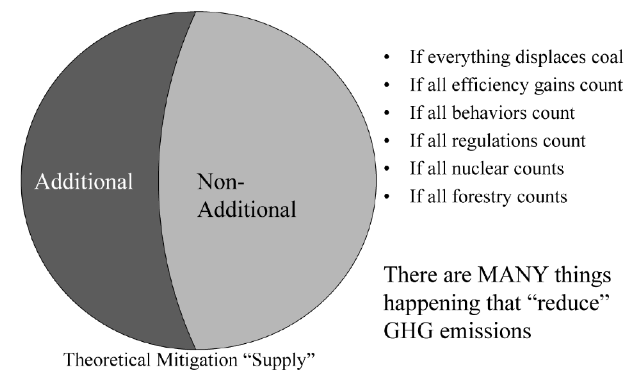


TABLE 2 Qualitative Assessment of Additionality Tests' Characteristics

Additionality Test	Ease of Development	Ease of Application	"True Positive" Rate*	"True Negative" Rate**
Legal, Regulatory, or Institutional Test	Easy. Benchmarks already exist.	Easy/Moderate. Once relevant legal requirements are identified, reviewing projects against them is generally straightforward. Transaction costs are low.	High, though not perfect. Most additional projects will have lower emissions than required by law. However, not all legal requirements are enforced; a project with emissions that are no better than required could be additional. This test would reject such a project.	Moderate/Low. Many non-additional projects may also reduce emissions below legal requirements.
Technology Test	Moderate. Done correctly, requires considerable assessment of what technologies are likely to be additional.	Easy. Usually simply a matter of checking whether a project is on the list of specified technologies. Transaction costs are low.	Low. Depends on which technologies are included, but many additional projects could use technologies not on the recognized list.	Low/Moderate to Very High. Depends on which technologies are included.
Investment Test	Easy. The real work under this test is left to its application.	Moderate to Difficult. Can require detailed financial analyses and possible disclosure of confidential information. Test results are often subjective, and easily manipulated. Transaction costs can be high.	High in theory, Moderate to Unknown in practice. In theory, additional projects will be uneconomical without considering the benefit of an emissions trading system. However, some economically attractive projects can be additional (*e.g.* because they face non-financial barriers). It can also be difficult to objectively define and ascertain a project's economic viability – *e.g.* the definition of "economically attractive" can differ dramatically from company to company – leading to uncertain outcomes.	High to Moderate in theory, Moderate to Unknown in practice. Again, non-additional projects are in theory projects that are economical without carbon credit revenues. However, some uneconomical projects may also be non-additional (*e.g.* because they are required by law). Finally, in practice, determination of non-additionality using investment analysis is fraught with subjectivity and uncertainties. Project developers can establish radically different risk-adjusted hurdle rates for projects in similar contexts.
Barriers Test	Easy. The real work under this test is left to its application.	Moderate to Difficult. Requires project developers to substantiate the existence of barriers and convincingly argue their significance. Test results are often subjective. Transaction costs can be high.	Moderate to High, depending on how test is formulated. Additional projects are likely to face barriers, but not all barriers are easily identified. Barrier tests can have burdensome evidentiary requirements; some additional projects may be excluded automatically.	Moderate to Low, depending on how test is formulated. Many (if not most) non-additional projects will also face some barriers. A project that faces significantly greater barriers than its alternatives may be non-additional if it has a high-expected payoff without GHG credit revenues.
Common Practice Test	Easy/Moderate. Done correctly, requires considerable assessment of how to define and identify "common practice."	Easy/Moderate. Once "common practice" is defined, it is generally straightforward to compare projects to the definition. Transaction costs can vary.	High theoretically, Moderate in practice. If "common practice" is perfectly defined, most additional projects will not be common practice. But specifying common practice across all sectors and regions is almost impossible. Even "common practice" projects might be additional in certain contexts, particularly if common practice is defined at the international, regional, or national levels.	High to Moderate. Many non-additional projects will correspond to common practice. However, depending on the technologies, generally there will be some non-additional projects that are beyond common practice.
Timing Test	Easy. Setting the date is relatively arbitrary.	Easy. Requires knowing only when the project was, or will be, implemented. Transaction costs very low.	High. If a project is truly additional, it is likely to be under development or only recently implemented. This may not be true for all additional projects, however. And many non-additional projects will also have started after any given date.	Low, for the same reasons.
Performance Benchmark Test	Difficult. Done correctly, requires considerable assessment to identify sector and country-specific benchmarks.	Easy. Once a benchmark is defined, it is easy to compare to the project emission rate. Transaction costs low.	Variable, depending on how benchmark is set. Sensitivity will be high for a lenient benchmark, since most additional projects will have lower than average emission rates. A stringent benchmark could effectively exclude many additional projects.	Variable, depending on how benchmark is set. Specificity will be low for lenient benchmarks, and relatively high for stringent ones. Even a stringent benchmark will probably recognize some non-additional projects as additional.
Project In, Project Out Test	Easy.	Easy to Moderate. Usually requires only comparing project emissions to historical emissions. Transaction costs vary.	Very High (Perfect). Practically all additional offset projects will have lower emissions than what would have occurred in their absence.	Very Low (Zero). All non-additional offset projects would pass this test. Only projects that do not ostensibly reduce emissions would be excluded, but these would not be candidates for offset projects.

* This column characterizes how likely it is that a truly additional reduction will generate a "yes" result when the additionality test is applied to it. This rating, however, should not be looked at in isolation. The "true negative" rate is also very important, since a low "true negative" rate means a lot of "false positives" are slipping into the credit pool.

** This column characterizes how likely it is that a truly non-additional reduction will generate a "no" result when the additionality test is applied to it. This rating, however, should not be looked at in isolation. The "true positive" rate is also very important, since a low "true positive" rate means a lot of "false negatives" are being excluded from the credit pool.

at correctly identifying the additional projects. The implication is that we could easily face relatively large proportions of phantom reductions or lost opportunities when applying specific additionality tests, depending on the true proportions of additional and non-additional projects that are being tested.

Statistical Concept #3

False positives and false negatives can never be fully eliminated. Generally, as one tries to eliminate one error by modifying a test or testing procedures, one will increase the magnitude of the other error (see Figures 3(a) and 3(b)). In other words, if you're most concerned about minimizing the number of false positives coming out of a test, be prepared for more false negatives.

With GHG additionality, we should expect efforts to squeeze down "phantom reductions" to lead to more "lost opportunities." Using an additionality test (or a combination of such tests) to rule out all non-additional projects would lead to many truly additional projects being excluded from the credit pool.[10] An extreme example would be a "technology test" that allows only projects involving practices that have no conceivable purpose besides climate change mitigation (*e.g.* flaring coalmine methane at an abandoned mine). Such a test would ensure that recognized projects are additional, but would exclude a whole universe of truly additional projects in other technology sectors.

Figure 3a: The Relationship Between False Positives and False Negatives in Additionality Testing

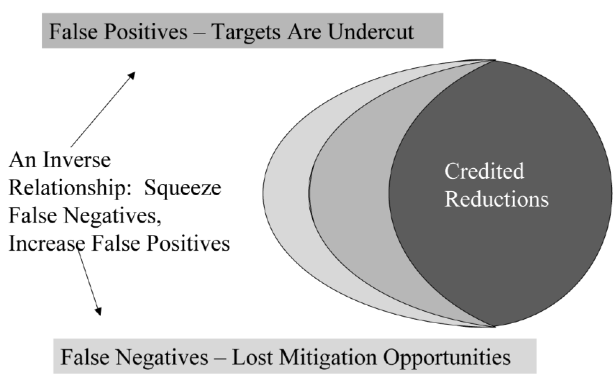
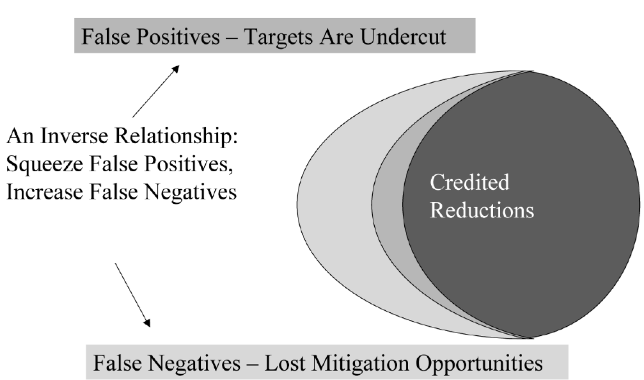


Figure 3b: The Relationship Between False Positives and False Negatives in Additionality Testing



Statistical Concept #4

Arriving at an acceptable balance of false positives and negatives is a key part of designing and choosing a particular test. There is no "one size fits all," partially because the relative importance of false positives or false negatives can vary widely for different testing situations. In the criminal profile test described earlier, for example, whether one thousand false positives for every true criminal is acceptable or not could depend on whether the "criminals" are petty thieves or terrorists carrying nuclear bombs. Defining the appropriate balance can be thought of as a policy decision more than a technical determination, although technical data should obviously inform the process.

GHG additionality testing is no different. Defining the acceptable balance between "phantom reductions" and "lost opportunities" in the context of additionality testing is ultimately a policy rather than a technical decision. The appropriate balance will depend on weighing competing objectives, including cost-effectiveness and the priorities of getting a trading system into operation vs. near-term environmental integrity, among others. Nevertheless, the four statistical testing concepts presented here are not commonly discussed when debating additionality standards. There is a common but misplaced notion that there is a technical solution to the additionality conundrum. There is almost no discussion of false positives and false negatives, much less of their inevitability and the need to balance them. There is even less discussion of the fact that this balancing must ultimately fall to policymakers in determining the objectives of the additionality testing in the first place.

Critical to our arguments in this article is the assumption that one can reliably project the market outcomes of particular additionality standards and choices. We believe that one can make such projections with the right data and analysis.[11] Although many GHG market observers have expressed surprise at how the Kyoto Protocol's CDM has been evolving, in particular with respect to small-scale and sustainable development projects, this outcome could have been predicted – and, indeed, was predicted – years ago as the framework of the CDM began to firm.

Before worrying too much about how to address the challenge of additionality in designing offset-based emissions trading systems, it is useful to answer the fundamental question of whether additionality testing really matters in advancing climate change mitigation objectives.[12] The following section takes a quantitative approach to looking at this question.

Does Additionality Matter in Today's International GHG Mitigation Market?

Many business and even some environmental observers have argued that additionality simply is not the most important factor at this stage in the development of offset-based emissions trading mechanisms. They point to the importance of getting emissions trading frameworks into place for the future, noting that near-term emissions reduction targets and trading are only a first step toward long-term climate policy. These observers don't want to see the near-term gridlock on additionality impede establishment of an emissions trading frameworks and are willing to trade off near-term environmental integrity in favor of getting a trading system into place.

Evaluating the relative importance of different objectives in the design of an offset-based trading system is a policy judgment. Technical analysis of GHG market fundamentals, however, can identify the implications of different design decisions. This section will review some of these GHG market fundamentals with this goal in mind.

GHG Market Fundamental #1

The universe of potential offset-based emissions reductions is enormous. It is not difficult to identify many gigatons of potential offset-based reductions; the volume of these reductions rises over time (see Figures 4 and 5 for project-based supply curves). These projects include many non-additional reductions that are occurring as fuel sources change for power generation, energy supply, and demand technologies become more efficient, and fossil fuel prices rise.

Figure 4: Project Based 2010 Marginal Abatement Cost Curve

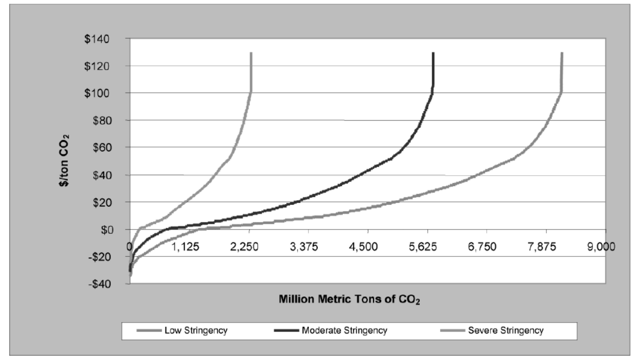
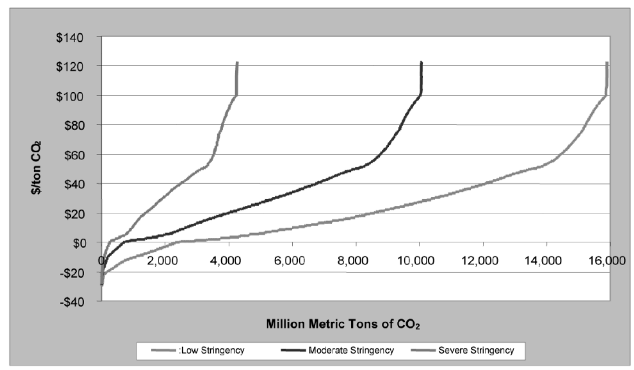


Figure 5: Project Based 2020 Marginal Abatement Cost Curve



Note: "Severe" stringency would equate with minimizing non-additional "phantom reductions" while allowing a large proportion of additional reductions to go unrecognized ("lost opportunities"). "Low" stringency would equate with recognizing many ostensible emissions reductions without regard to additionality.

One reason additionality continues to be so hotly debated is that the Kyoto Protocol's flexibility mechanisms were originally designed for a market that included U.S. demand.

GHG Market Fundamental #2

Alternative additionality standards could dramatically affect the supply curve available to the GHG market. Figures 4 and 5 show that the supply curve for the global GHG market looks very different depending on the strictness of the additionality standard.[13] Supply curves this different would have significant impacts on the market-clearing price for reductions as market demand rises.

GHG Market Fundamental #3

Sources of offset credits differ radically in their additionality "profiles" and in how they advance particular policy objectives for the ultimate credit pool. For example, some potential mitigation sectors are likely to prove almost entirely additional in today's market context (*e.g.* coal-mine methane flaring at abandoned mines); others will be almost entirely non-additional (*e.g.* existing nuclear power installations). Many sectors will be characterized by a more diverse range of additionality outcomes, making it difficult to differentiate between "true positive" and "false positive," as well as "true negative" and "false negative" reductions. In addition, sectors vary widely in costs, contribution to sustainable development objectives, and other characteristics.

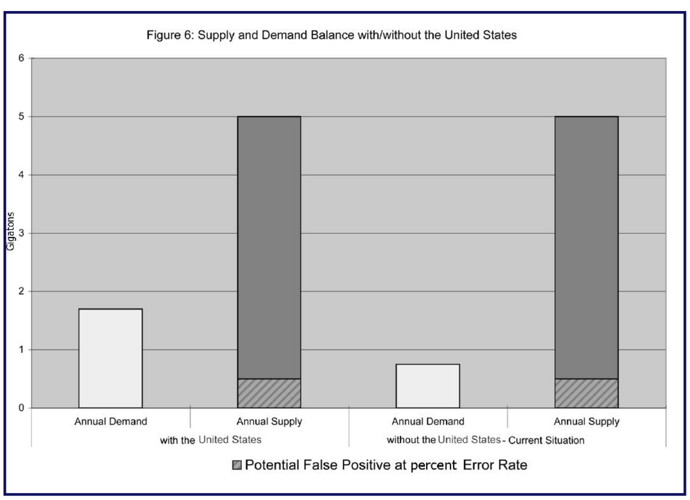
GHG Market Fundamental #4

In an overall supply curve for GHG reductions, the distribution of additional and non-additional reductions will not be random. All else being equal, non-additional reductions will tend to cluster towards the low end of the potential supply curve; *i.e.* they will tend to be low cost and the easiest to get to market quickly. If one thinks of non-additional projects as those that are already going to happen without a trading system, then there is no incremental economic cost for them to "produce" GHG reductions beyond the transaction costs associated with documenting and selling the reductions.

GHG emissions reduction projects face significant uncertainties in the market. Uncertainty about post-2012 targets, and about different countries' reliance on domestic policy and measures, is creating uncertainty about overall offset-based demand. Uncertainty about the disposition of "hot air" and functioning of the CDM is creating uncertainty about overall offset-based supply. The combination of these factors creates considerable uncertainty about credit prices, and makes it more difficult for project developers to evaluate the business case for investing in truly additional emissions reduction projects. This makes it abundantly clear that non-additional projects – if they are able to get credited – offer both the lowest risk and lowest cost sources of credits for the market. Moreover, even where non-additional projects have GHG reduction costs (including transaction costs) comparable to additional projects,[14] the non-additional projects are likely to be brought to market faster because the GHG returns represent an upside potential, rather than being key to a project's underlying viability.

GHG Market Fundamental #5

Demand in today's international GHG market is not what we had anticipated in 1997 when the Kyoto Protocol was negotiated. Without the United States, the primary source of anticipated market demand is absent, yet most of the supply is still available to the market. Figure 6 illustrates a representative balance between supply and demand anticipated in 1997 and the situation today. The graph shows that absence of the United States has resulted in a much smaller demand today than the two gigatons of GHG reductions per year that were expected for the first commitment period. It has not, however, resulted in a smaller potential supply (except for supply that would have come directly from the United States), conservatively estimated here at five gigatons of GHG reductions per year. One reason additionality continues to be so hotly debated is that the Kyoto Protocol's flexibility mechanisms were originally designed for a market that included U.S. demand. The reality, however, is that the current GHG market does not include U.S. demand. In effect we are developing additionality standards for a multi-gigaton GHG market, while current demand is a fraction of that amount.



GHG Market Fundamental #6

When you combine the potentially radical imbalance of supply and demand in today's market, with an understanding of the distribution of non-additional reductions in the supply curve, it becomes clear that a large fraction of the projects whose reductions are offered to the market could be non-additional. In fact, the total number of available reductions from such projects could conceivably swamp demand.

Whether they do swamp demand, of course, depends on how effectively additionality tests are used to keep "phantom reductions" out of the market. The proportion of "phantom reductions" in implementing additionality tests becomes paramount. Even a ten percent false positive rate in a market where five gigatons of non-additional reductions are available could result in 500 million tons of "phantom reductions." With demand on the order of 700 million tons, this level of "phantom reductions" could severely undermine the market's effectiveness in keeping total emissions within the cap agreed to by industrialized countries.

At their root, these [additionality] tests all are trying to answer the same question: would a project have occurred regardless of the existence of drivers created by the trading system, or not?

GHG Market Fundamental #7

A different problem would occur if demand were much higher. If annual market demand exceeded five gigatons, 500 million tons of "phantom reductions" might be acceptable, and the focus of policy makers might switch to reducing "lost opportunities." Compliance costs are a major issue when emissions targets are ambitious and demand is high. Every additional project that is erroneously rejected by an additionality test means higher costs as buyers have to move further up the supply curve. Thus, while the rate of "phantom reductions" for additionality tests is paramount when demand is low, the rate of "lost opportunities" becomes important when demand is high.

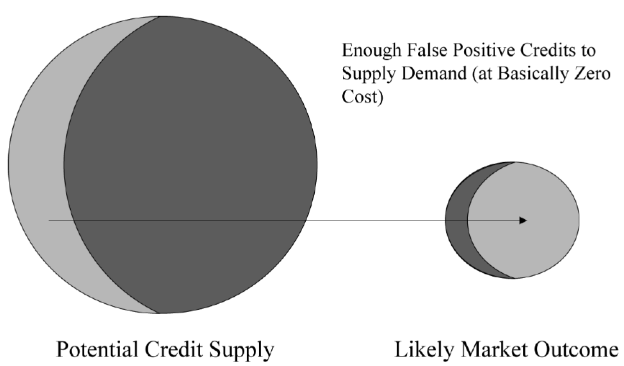
In today's market, potential credit suppliers overwhelmingly focus on ensuring that the rules allow their project into the market. There is little understanding that this ultimately is likely to magnify the supply and reduce market-clearing prices. The result could be that participating in the market would be of relatively little interest to anyone. Ironically, the only people benefiting from the market under these circumstances are those supplying "phantom reductions." With a low market-clearing price, project proponents will not be able to justify pursuing projects that are additional.

In fact, potential credit suppliers have significantly varying market interests. Developers with truly additional project opportunities should be advocating stricter additionality standards, which will tend to raise market-clearing prices sufficiently for them to participate. Instead, even project developers with clearly additional projects argue for weaker additionality standards or for abandoning additionality standards altogether because they want to "get going." They don't understand that this is a self-defeating outcome.

Additionality is pivotal to the environmental outcomes associated with GHG emissions trading programs that incorporate offsets from uncapped sectors or countries. Given market fundamentals, flooding the GHG market with non-additional reductions would be virtually inevitable if additionality were simply ignored in the design of the emissions trading system. It may be that getting an emissions trading system into place is more important to policymakers than ensuring the market's near-term environmental integrity. But if that is the approach policymakers wish to take, it should be made transparently clear, so that in-depth debates over additionality rules can be largely avoided.

More significantly, any suggestion that additionality should be ignored must grapple with at least two key questions: (1) is environmental integrity a necessary tradeoff in getting a trading system up and running; and (2) is it realistic to expect, once a trading system is up and running, that it will be possible to retroactively fix the system to reinstate environmental integrity? We do not believe that environmental integrity is a necessary tradeoff, as discussed in the next section of this article. Also, we question policymakers' political and institutional ability to successfully implement such a retroactive change.

Figure 7: The Implications of A Dramatic Imbalance Between Supply and Demand



From Additionality Facts to Additionality Policy

This analysis shows that there is no "correct" additionality objective or additionality test. Additionality testing, however, will play a crucial role in determining every aspect of market supply (*i.e.* the supply of recognized reductions or GHG credits), including:

- The magnitude of the supply pool available to the market (by specifying what can count);
- The cost curve associated with that supply pool (based on the costs of potentially qualifying offsets);
- The magnitude and proportion of "phantom reductions" in the final supply pool;
- The magnitude of "lost opportunities" being kept out of the supply pool and the opportunity costs associated with their exclusion from the market.

If participants in the GHG market and in the development of offset-based emissions trading strategies accept this analysis, it could radically change the nature of today's additionality debate. Instead of talking past each other about what the perfect additionality test is, we could move to the concrete step of asking how to prioritize different policy objectives inherent in establishing emissions trading programs. We could move to the action step of thinking about how to pursue an effective additionality policy that accomplishes these objectives. This section profiles how a better understanding of the underlying principles of additionality testing can guide the development of additionality policy.

Additionality Policy Design Principle #1

The objectives of an additionality policy need to be identified in order to guide the development and use of additionality testing. Policymakers cannot bypass this obligation by simply directing a technical body to design additionality standards. Key design questions include:

- How important are the physical reductions associated with the offset-based trading system as opposed to other objectives, *e.g.* simply establishing a trading system and/or creating incentives for climate-friendly technologies and practices without concern for the integrity of an emissions cap?
- What is politically acceptable, particularly with respect to the cost of credits under the trading system? Additionality tests, by affecting the available supply, can radically affect the resulting credit market-clearing price.
- How big a pool of offset-based credits is needed? The size of the pool has implications for additionality test design. Too strict an additionality screen against a given level of demand can lead to a constrained supply curve and much higher prices. Too weak an additionality screen against a given level of demand can lead to an unacceptably high pool of "phantom reductions" available to meet that demand.
- Is the goal to develop a standard for the near-term GHG markets with limited demand or to develop an additionality standard for a future, larger GHG market? What tradeoffs are we willing to accept during the transitional phase?
- Should the emissions trading system favor certain sectors and projects?

Once these questions are answered, it becomes possible to develop additionality standards that will advance the chosen policy decisions. Not everyone will be happy with those standards at any given time; there will be winners and losers in terms of who can play in the market on the supply side, and what it will cost participants to satisfy their demand for reductions. We believe that most market players would be happy with a functioning GHG market that delivers cost-effective reductions and satisfies environmental integrity objectives, without agonizing over each project developer's motivations.

Additionality Policy Design Principle #2

In order to understand whether offset-based emissions trading programs are supporting the integrity of the overall emissions reduction targets, it is necessary to understand the relative proportions of "phantom reductions" and "lost opportunities" to real reductions making it into the credit supply pool. As such, the analysis necessary for such an understanding has to be planned and budgeted as part of the development of the emissions trading system.

Additionality Policy Design Principle #3

Additionality rules need to be adapted to market circumstances on an ongoing basis. In the face of significant changes in market supply or demand, static additionality tests cannot effectively balance the policy objectives of acceptably small magnitudes of "phantom reductions" and "lost opportunities," and the cost-effectiveness of the overall credit pool.

Delivering a cost-effective pool of truly additional reductions in a 300 million-ton trading market is an entirely different challenge than delivering a cost-effective pool of additional reductions in a five gigaton market. Additionality rules designed for a large market (or designed without attention to market size at all) can result in little or no environmental integrity during the "small market stage." Strict additionality tests could limit "phantom reductions" in a small market, but the price of credits could rapidly rise as demand rises (due to a shortage of supply). A single static additionality standard will not be appropriate for every level of demand.

Additionality Policy Design Principle #4

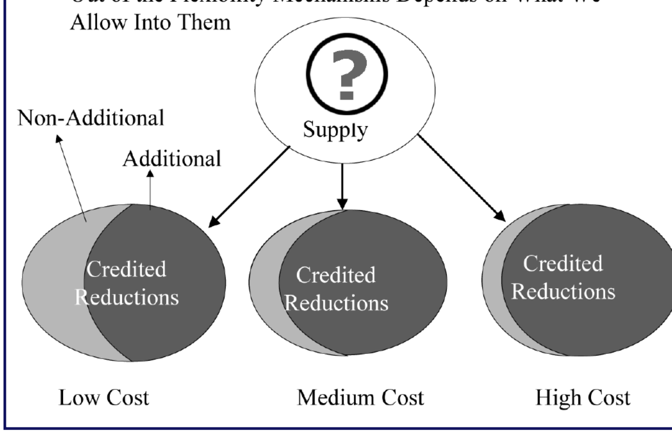
The ideal goal is to pursue an additionality standard that gets us as close as possible to both a low "phantom reduction" risk and low "lost opportunity" risk. When offset-based credit demand is relatively low, this is politically easier to do given the existence of "low-hanging fruit" opportunities. As demand increases and these opportunities disappear, it becomes much harder to keep both sources of error low because pressure will increase to ease rules that minimize "phantom reductions" and promote those that minimize "lost opportunities."

A good first step in this direction is to identify specific sectors that can form the backbone of an offset-based credit supply. A relatively simple way to "evolve" additionality standards with the size of the market is through a "technology test" that initially focuses on a limited number of carefully selected sectors. As the market expands, the number of sectors allowed into the market can be increased. This could allow creation of a cost-effective credit pool that grows as demand grows, while effectively balancing false positives and false negatives. This approach would avoid the complexity, if not the impossibility, of attempting to develop additionality standards that are intended to apply to all sectors. It is simply not possible to design such a system in a way that satisfies the objectives being pursued through additionality testing; the lower the demand, the harder it gets. The potential supply becomes too large very quickly, and non-additional credits threaten to swamp the market.

Additionality Policy Design Principle #5

There is no true commodity market for GHG credits in an offset-based trading system. GHG credits are not a conventional commodity; the supply and demand for GHG credits depends more on policy decisions than on physical fundamentals. Policymakers have three choices: they need to either choose the additionality standards they want and accept the resulting prices; determine the prices they are willing to accept[15] and design the additionality standards accordingly; or choose a combination of these two approaches. This is illustrated in Figure 8, which shows that what comes out of flexibility mechanisms depends largely on what is allowed into the supply pool. Figure 8 portrays the easiest potential portfolio outcomes, namely creation of a low-cost, low-additionality pool, or a high-cost, high-additionality pool. Generating these two kinds of portfolios is easy. It will be harder, however, to generate credit pools with more complicated (and desirable) characteristics. For example, a low-cost, high-additionality credit pool requires careful preselection of sectors that will tend to deliver low-cost and highly additional reductions. As one adds objectives, including for example the promotion of sustainable development objectives, the complexity of the process increases.

Figure 8: Going Back to First Principles: What We Get Out of the Flexibility Mechanisms Depends on What We Allow Into Them



Additionality Policy Design Principle #6

Well-designed additionality standards can avoid the misplaced urge to address additionality concerns through proxy measures. Several such proxies are already in widespread use:

Proxy A

Making additionality tests more and more complex: With gigatons of potential market supply, and relatively limited near-term demand, even complicated rules are unlikely to prove effective in balancing competing objectives. An excessive number of "phantom reductions" are still likely to make it into the market, in proportion to the limited demand, while the magnitude of "lost opportunities" grows rapidly as the complexity of additionality rules increases.

Table 2 suggests that it is possible to choose additionality tests that are relatively easy to apply at the project evaluation stage. Such tests can also effectively control for "phantom reductions" or "lost opportunities." When applied to individual projects, most of these tests require simple "yes or no" answers to clearly defined questions. A technology test can be restrictive, for example, but it is relatively easy to determine whether a project is "on or off" the list. Within specific technology sectors, a well-defined "common practice" test can eliminate many remaining "phantom reductions." Where "lost opportunities" are a concern, different combinations, perhaps involving legal or timing tests, might effectively balance this concern with environmental integrity. The most versatile kind of test is a performance benchmark. In principle, a benchmark can easily be adjusted to meet different policy objectives; it has the added benefit of automatically setting baseline emissions for multiple projects. Alone or in combination with other "yes or no" tests, benchmarks could radically reduce the transaction costs associated with current approaches to additionality testing.

Designing these tests in the first place, however, would require substantially more effort than the current system of project-by-project additionality testing, where the burden falls on the project developer. They also could be politically problematic, since they automatically exclude certain sectors or segments of the market. Over time, however, these tests would prove preferable to the requirements of barriers and investments tests for extended weighing and interpretation of evidence for each and every project.

Proxy B

Making the credit quantification process more and more conservative: Such conservatisms can reduce the number of "phantom reductions" from individual projects entering the credit pool, but if a lot of false positive projects are making it through the additionality tests, "phantom reductions" may still dominate the pool. This is partially because those same conservative assumptions, applied to the truly additional projects, can significantly increase the proportion of "lost opportunities" and prevent many truly additional projects from making it to the market.

Proxy C

Too aggressively shortening the crediting period for projects and sectors: What we call "baseline creep" can be illustrated by a situation in which a few projects are approved under a given additionality test, but then the additionality hurdle is significantly (and prematurely) raised for future projects of the same type. Policymakers will face pressure to reduce the number of credits generated at a later date if the market remains small. Current CDM rules limit project crediting to three renewable seven-year periods or a single ten-year period; these rules will encourage baseline creep by giving decision-makers the ability to stop crediting a project when its baseline comes up for renewal. Most project developers recognize this problem; review of CDM project applications shows that developers are more likely to choose a single ten-year crediting period, rather than trying for renewable seven-year periods. Baseline creep is one way to address "phantom reductions," but like other solutions to the problem of "phantom reductions," it could easily have the unintended and adverse effect of substantially amplifying "lost opportunities."

Last in terms of our discussion of how to apply additionality facts to additionality policy, it is worthwhile to briefly address the question of policy and sectoral baselines, since they are being actively discussed in the context of the Kyoto Protocol's CDM. The authors do not see the adoption of policy or sector baselines as being "good" or "bad." Much of the discussion around "sectoral" and "policy" baselines for the CDM, however, either fails to address additionality at all, or suggests that there is something fundamentally easier about addressing additionality at the "sectoral" or "policy" levels than at the "project" level. The statistical testing concepts we have described, however, apply equally to "project," "sectoral," and "policy" additionality. The same issues of "phantom reductions" and "lost opportunities" need balancing and resolution, particularly given that the magnitude of "phantom reductions" entering the supply pool could be much larger under "sectoral" and "policy" additionality standards unless those standards are very carefully developed.

Conclusion

In today's polarized debate over additionality, it is difficult to map a path forward. Most participants (observers, analysts, and agencies involved in development of additionality policy) often do not seem to acknowledge the basic statistical principles we present in this article. The CDM Executive Board is not charged with making the kinds of policy decisions called for here; moreover, it does not have the resources to implement the analysis that would allow "phantom reductions" and "lost opportunities" to be estimated in order to appropriately balance policy objectives.

We are not alone in calling for a reevaluation of our approach to additionality. Many participants in today's CDM debates are in effect calling for the same kind of reevaluation, including aggressive calls for streamlining the CDM's additionality and project-review processes. Unfortunately, simply streamlining these processes, without consideration of the implications of changes for the environmental integrity of the trading system, could put such market mechanisms at even more risk of losing their political credibility.

Without a dramatic change in direction, the additionality wars will continue, as they have for the last decade, with relatively predictable implications:

- Continuing efforts to make additionality tests for particular sectors more and more complex to trim down supply sources that key interest groups dislike;
- Continuing efforts to streamline additionality tests to allow more projects into the credit pool, particularly from sectors that key interest groups do like;
- Continuing industry efforts to abandon additionality standards altogether;
- Continuing environmental concerns that the environ-

mental integrity of trading systems is not being maintained; and

- Aggressive and inappropriate use of proxies as an alternative to effective additionality standards.

These outcomes undercut the objective of credible offset-based trading mechanisms. A successful GHG market should advance environmental integrity and cost-effectiveness. It is vital that stakeholders with an interest in this objective step back and revisit the additionality issue in light of the considerations we have reviewed. Appropriate additionality standards *can* be designed for the GHG market – but only if participants properly understand the market's fundamental challenges and act accordingly.

This analysis is structured around the international GHG market; however, the principles are applicable to any offset-based emissions trading system, whether sub-national, national, or international. Additionality standards can reflect and accomplish key policy objectives for the offset-based GHG market, but doing so will always represent a balancing act. This can only be done effectively if the objectives are specified upfront. Additionality *per se* is not the objective; it is a means to an objective.

We do recognize that the GHG market faces significant challenges beyond those posed by debates over additionality. Given the uncertainties over post-2012 emissions reduction targets, project developers do not know what value the CDM's Certified Emissions Reductions ("CERs") will have post-2012. There is also considerable uncertainty over the role "hot air" will play in GHG markets,[16] contributing to uncertainty for project developers of what value their CERs will have pre-2012 either. These factors make it difficult to build a solid GHG market and create strong incentives to get non-additional and low-risk reductions into the market. Nevertheless, these larger market uncertainties do not alter our fundamental conclusion that additionality policy, if it is to advance the cause of credible GHG trading markets, needs to be based on a solid analytical foundation.

We believe the approach in this article yields insights and implications for future offset-based additionality policy. We are not taking a stand in this article as to whether such policy can be successfully implemented in the face of political and institutional challenges. We are simply trying to make a statistically well-founded point that additionality can be operationalized if one recognizes and accepts that no test will ever be perfect, and then adapts the process accordingly.

ENDNOTES: Offset-Based GHG Additionality Determinations

[1] M.C. TREXLER, P. FAETH, AND J. KRAMER, WORLD RESOURCES INSTITUTE, FORESTRY AS A GLOBAL WARMING MITIGATION STRATEGY: AN ANALYSIS OF THE GUATEMALA CARBON SEQUESTRATION FORESTRY PROJECT (1989).

[2] Most additionality thinking takes place within the context of offset-based emissions trading systems. However, some experts increasingly advocate sector- and policy-based emissions reductions crediting as a means to expand credit supply in today's GHG emissions trading programs. The statistical principles presented in this article apply just as much to sector- and policy-based crediting as to offset-based crediting. Indeed, the potential for large-scale approval of "false positive" credits may well be greater under the former crediting mechanisms.

[3] Additionality is not a problem in all environmental commodity markets. In the U.S. sulfur dioxide ("SO_2") market, for example, all emissions are included within the national emissions cap. In setting a national emissions reduction target, environmental regulators don't need to care about the "additionality" of individual reductions. They need only assess whether the national target is achieved at the end of the year, since that should accomplish the intended environmental gain. In the GHG arena, however, we are crediting emissions reductions from non-capped emissions sources and countries, and allowing them to count against regulatory mandates somewhere else.

[4] People often speak of additionality with respect to emission reductions, as opposed to projects. This may be a distinction without a difference. Conceptually speaking, for emissions reductions to be additional they must result from projects that are additional.

[5] In a voluntary context, where no offset credits are generated, the question can be reformulated in terms of considerations about climate change mitigation: would a project have happened in the absence of voluntary concerns about climate change?

[6] Non-additional projects are often referred to as "business as usual" ("BAU"). This term is useful if it is properly understood to mean projects that would have happened in the absence of an offset crediting mechanism. However, "business as usual" can be interpreted in many different ways, some of which do not comport well with the basic concept of additionality. We use the term "non-additional" instead of "business as usual" to avoid confusion.

[7] False positives commonly are referred to as Type One error; false negatives are referred to as Type Two error.

[8] In statistics, the "true positive" rate is referred to as the test's *sensitivity*. The "true negative" rate is referred to as the test's *specificity.*

[9] That is, if the test has perfect sensitivity. With lower sensitivity, even more innocent people might be falsely implicated for every true criminal caught!

[10] This fact could make many market participants unhappy, particularly those wishing to pursue projects that may fall into an expanding "lost opportunities" pool as efforts to control "phantom reductions" become stricter.

[11] The literature evaluating the implications of different approaches to additionality and baselines is extensive. *See, e.g.* Carolyn Fischer, *Project-Based Mechanisms for Emissions Reductions: Balancing Tradeoffs with Baselines,* 33 ENERGY POL'Y 1807-23 (2005).

[12] Some observers argue that it's enough to simply use a "with project" and "without project" approach to quantifying GHG emissions reductions. If a factory replaces a boiler, give project proponents credit for any resulting emissions reductions. If a company builds a gas-fired power plant in a coal-dominated region, give it credit for displacing coal-fired CO_2 emissions. As Figure 2 demonstrated, however, many things are happening that tend to reduce emissions based on a simple "before and after" analysis.

[13] Where "strictness" corresponds to minimizing phantom reductions and allowing more lost opportunities. In Figures 4 and 5, strictness is modeled approximately using the TC+ES Supply Tool, without explicit reference to specific tests.

[14] This could happen, for example, if a non-additional project were smaller and therefore had higher per-ton transaction costs.

[15] Given today's European Union Emissions Trading System ("EU ETS") prices, it might seem that additionality standards are the least of the problems facing the GHG market. However, today's EU ETS prices are a result of supply constraints applied to the first phase of the EU ETS (*e.g.* limitations on the sectors able to generate reductions within the EU), and market variables (*e.g.* the difficulty of getting Certified Emissions Reductions into the market within the first phase of the ETS). Demand has changed significantly from original expectations, based on the results of the National Allocation Plan process. The first phase of the EU ETS has few implications for how markets will behave during the first Kyoto Protocol commitment period.

[16] Hot air, representing excess allowances awarded primarily to Russia and Ukraine as part of the political compromise leading to the Kyoto Protocol, is basically a zero-cost resource; given the absence of U.S. demand, theoretically "hot air" credits could supply the Kyoto Protocol market out to 2012.