Combined application of cost–benefit and multi-criteria analysis for decision support in air quality management policy: a case study in Lima and El Callao, Peru.

Gerardo Sanchez Martinez

Department of Applied Economics III (Econometrics and Statistics)

Faculty of Economics and Business

2017

eman ta zabal zazu

Universidad del País Vasco

Euskal Herriko Unibertsitatea

Aplicación combinada de análisis coste-beneficio y multicriterio para apoyo de toma de decisiones en políticas de calidad del aire: un caso de estudio en Lima y El Callao, Perú.

1. Introducción

La colección sistemática y utilización de datos y evidencia para la evaluación de impacto potencial de políticas ambientales se ha convertido en norma en la mayoría de países desarrollados. Aunque de manera menos sistemática, estos insumos también se utilizan para la toma de decisiones en países en desarrollo. Existe una gran variedad de herramientas para organizar y procesar esta información, y su tipología y ventajas han sido estudiadas exhaustivamente (por ejemplo, por Crabbé & Leroy, 2008). Los dos tipos de herramientas más utilizados en este tipo de evaluaciones son el análisis coste-beneficio (ACB) y el análisis multicriterio (AMC). Sólidos argumentos justifican el uso generalizado del ACB y AMC en la evaluación de políticas públicas ambientales (OECD, 2008). Ambas técnicas son lo suficientemente flexibles para aplicarse a casi cualquier tipo de evaluación. Además, existe un gran acervo de literatura e información acerca de sus bases conceptuales, aplicación en la práctica y casos de estudio –aunque algo menor en el caso del AMC, más reciente. Por otra parte, sus perfiles de ventajas y desventajas son complementarios, lo que sugiere un cierto valor añadido de su uso combinado. Sin embargo, existe poca información en la literatura científica y técnica acerca de las posibilidades del uso combinado de ACB y AMC. En concreto, apenas hay estudios sobre las limitaciones o el valor conceptual y práctico de las posibles combinaciones de ambas técnicas, particularmente en el campo de políticas ambientales. Es esta escasez la que dio origen al tema de esta tesis doctoral.

2. Objetivos

Los objetivos principales de esta tesis doctoral son:

- Estudiar los ejemplos del uso combinado de ACB y AMC y sus implicaciones, y desarrollar un marco conceptual para la integración de ambos métodos;
- Distinguir y analizar las principales opciones metodológicas para este uso combinado y aplicarlas a un caso de estudio real de priorización de políticas locales de gestión de la calidad del aire; e
- Investigar las ventajas y desventajas, fortalezas y debilidades de cada opción de combinación

En resumen, se trata de explorar la factibilidad y valor añadido de las posibilidades de combinación a través de un caso de estudio, que actúa como hilo conector del análisis. Tras la revisión de la literatura relevante, este proceso de análisis conllevó los siguientes pasos:

- Evaluación de impacto ambiental y en salud de las políticas propuestas;
- ACB de las políticas del caso de estudio;
- AMC de las políticas del caso de estudio;
- Combinaciones posibles de uso combinado de ACB y AMC, y discusión al respecto; y
- Conclusiones.

Esta tesis también aporta producción de conocimiento e innovación. El ACB de las políticas de calidad del aire en Lima se añade a un muy pequeño conjunto de evaluaciones de este tipo en América Latina. En el caso del AMC, además de lo excepcional de su aplicación en este tema y esta región, todo el proceso fue guiado y canalizado por parte de los actores clave y representantes de partes afectadas, un aspecto infrecuente en la literatura. Y finalmente, la aplicación combinada de ACB y AMC es, de acuerdo a la búsqueda de literatura, el único caso en la región; además, entre los enfoques de combinación se propone una técnica novedosa con posible valor de aplicación en el futuro.

3. Caso de estudio: políticas de calidad del aire en Lima y El Callao, Perú

Los efectos en salud de la contaminación del aire urbano han sido estudiados minuciosa y exhaustivamente desde el punto de vista epidemiológico (Dockery & Pope, 2006; Dockery et al., 1993; Katsouyanni et al., 2001; Pope et al., 1995; Pope et al., 2002). Estos efectos incluyen mortalidad prematura, severas enfermedades respiratorias y cardiovasculares agudas y crónicas, y múltiples efectos más leves o subclínicos. La carga de enfermedad de la contaminación afecta de manera desproporcionada a grupos vulnerables de población como niños, ancianos y enfermos crónicos (Brook, 2010; Eftim et al., 2008; Samet & Krewski, 2007).

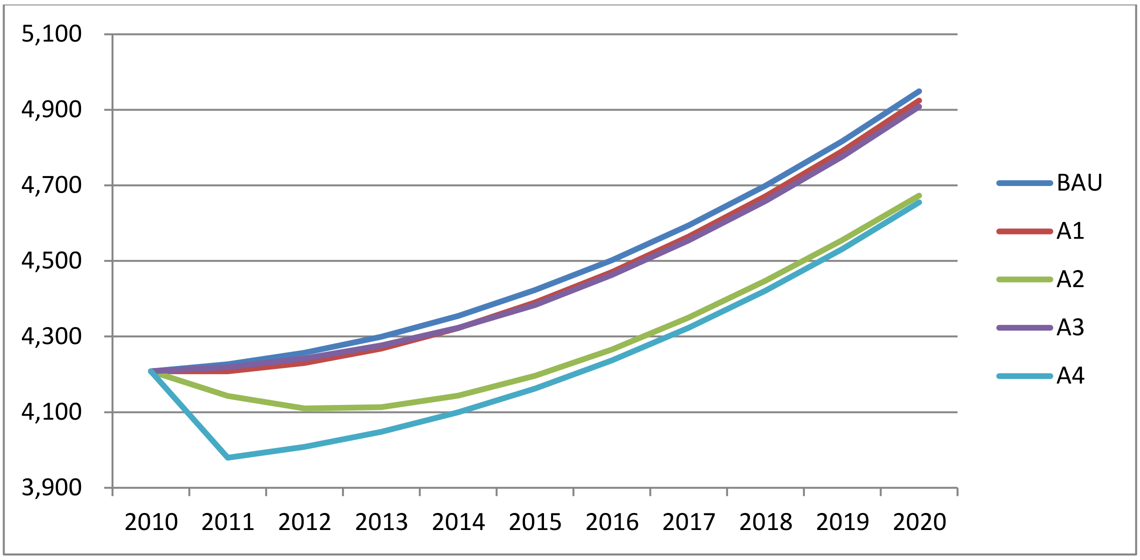
A pesar de la escasa evidencia epidemiológica en Lima, la pésima calidad del aire llevó a los gobiernos nacional y local a tomar medidas para disminuir la contaminación atmosférica en la ciudad. Estos esfuerzos iniciales continuaron a finales de la década de los 90 con la incorporación de Lima a la Iniciativa de Aire Limpio para América Latina del Banco Mundial, que proporcionó incentivos y asistencia técnica. A través de legislación, control de cumplimiento y programas específicos se consiguieron en la década pasada considerables descensos en las concentraciones de importantes contaminantes como los óxidos de nitrógeno y el dióxido de azufre. Sin embargo, las partículas inhalables se han mantenido en promedio muy por encima de los límites permisibles locales, de 50 y 15 microgramos por metro cúbico para PM_{10} y $PM_{2.5}$, respectivamente (DIGESA, 2012).

Esta permanencia en altas concentraciones es preocupante, puesto que la materia particulada inhalable es el contaminante urbano con mayores y más severos efectos en salud pública (WHO, 2006). Dada la situación, los gobiernos locales de la Zona Metropolitana de Lima y el Callao (ZMLC) y el gobierno de Perú diseñaron el "Segundo Plan Integral de Saneamiento Atmosférico para Lima y Callao" (PISA II), estructurado en torno a tres programas básicos: P1: reducción de las emisiones y concentraciones de materia particulada; P2: prevención y reducción de la contaminación en general; y P3: investigación en contaminación del aire. El diseño e implementación del plan se coordinaron desde el Comité Gestor de la Iniciativa de Air Limpio (CGIAL), el cual estableció por su urgencia la reducción de las concentraciones de materia particulada como el área prioritaria de acción. Las autoridades peruanas solicitaron al Banco Mundial asistencia técnica para evaluar y priorizar estas acciones, así como sugerir herramientas de apoyo de toma de decisiones al respecto. El Banco Mundial contrató mi colaboración, y mi propuesta fue la de realizar tanto un análisis coste-beneficio (ACB) como un análisis multicriterio (AMC) de las medidas. Todas las partes aceptaron esta propuesta y el uso de los resultados para esta tesis doctoral. Las actividades clave del programa para la reducción de las partículas son:

- Acción 1 (A1), renovación de la flota vehicular particular, basada en un sistema de incentivos, variaciones impositivas y restricciones a la importación de vehículos de segunda mano;
- Acción 2 (A2), un bono de chatarreo para vehículos obsoletos de transporte público, diseñado como incentive para la renovación de la atomizada flota vehicular de transporte público en la ZMLC;
- Acción 3 (A3), promoción del uso del Gas Natural Vehicular a través de subsidios para la adaptación de vehículos de transporte público;
- Acción 4 (A4), fomentar el mantenimiento e inspección vehicular en todas las flotas.

En base al modelo de emisiones de fuentes móviles aplicado (ver Gráfico 1), todas y cada una de las acciones planteadas conllevan un descenso relativo en las emisiones de materia particulada en comparación con el escenario base, en el horizonte temporal considerado (2011 a 2020). Sin embargo, en todas ellas hay a la larga un incremento total de emisiones debido a las rápidas tasas de motorización en el país previstas a corto y medio plazo.

Gráfico 1. Emisiones de PM_{10} (Tm) desde 2010 a 2020 bajo distintos escenarios



Fuente: elaborado por el autor, en base a datos del Instituto de Aire Limpio (CAI-LAC)

La acción A4 (mantenimiento e inspección vehicular) aporta las mayor reducción de emisiones, seguida de A2 (chatarreo de vehículos obsoletos de transporte público), y a mayor distancia de A3 (promoción del uso del Gas Natural Vehicular) y de A1 (renovación de la flota vehicular particular). En términos generales, las reducciones correspondientes de concentraciones de contaminantes y los beneficios en salud atribuibles son aproximadamente proporcionales.

4. Análisis coste-beneficio del caso de estudio

El ACB de las acciones se realizó conforme a la metodología estándar del Banco Mundial para evaluaciones ambientales nacionales (ver, por ejemplo, Golub, Klytchnikova, Sanchez Martinez, & Belausteguigoitia, 2014). La principal fuente de beneficios son los efectos en salud (mortalidad prematura y morbilidad) evitados, mediante las metodologías del coste de enfermedad y de valor de una vida estadística. Los costes son calculados a nivel social, incluyendo tanto los costes privados como los gubernamentales de implementación. Las tasas de descuento utilizadas son el 0% y el 3%, y los resultados están listados en la tabla 1.

Tabla 1. Resultados finales del análisis coste-beneficio, sin descuento y con una tasa de descuento del 3%

Indicador	Acción A1	Acción A2	Acción A3	Acción A4
VP Beneficios(M.S/) sin descuento	1,159	8,804	1,325	10,888
VP Beneficios (M.S/) descuento 3%	954	7,109	1,063	8,930
VP Costes (M.S/) sin descuento	8,130	5,122	388	1,280
VP Costes (M.S/) descuento 3%	7,169	4,983	330	1,052
Ratio Beneficio-Coste (RBC) sin descuento	0.14	1.72	3.41	8.50
Ratio Beneficio-Coste (RBC) descuento 3%	0.13	1.43	3.21	8.48

En términos de costes y beneficios a nivel social, la acción A1 no es costo-beneficiosa, con un RBC muy inferior a la unidad (por cada Sol peruano invertido sólo se obtendrían 0.14 Soles), mientras que A2, A3 y A4 sí lo son, con valores crecientes de RBC. En el caso de A3 (promoción del GNV), la discrepancia entre su escaso efecto de reducción de emisiones y su alto RCB se explica por el bajo coste de la medida.

5. Análisis multicriterio del caso de estudio

De entre los posibles modelos AMC, dadas las características y limitaciones del caso de estudio (baja disponibilidad de datos, tiempo de interacción con actores clave limitado, etc.), se optó por modelo del tipo "Outranking", en concreto la metodología PROMETHEE-GAIA (Brans, 1982; Brans, Vincke, & Mareschal, 1986). De este modelo se pueden obtener dos tipos de priorizaciones: una parcial (en que se comparan separadamente la preferencia y la no preferencia) y una completa, que sintetiza todas las preferencias. Para obtener los criterios relevantes, el CGIAL identificó una muestra de 30 actores clave. Posteriormente se diseñó una entrevista semiestructurada (Gregory & Brierley, 2010; Naylor & Appleby, 2013) y se completaron y validaron 24 de las entrevistas. A través de sus respuestas se identificaron 26 criterios clasificables en seis categorías (Ambiental, Salud, Economía, Social, Política e Implementación). Estos criterios se presentaron a los 30 actores clave en un taller al efecto, y se priorizaron de acuerdo a una metodología de tipo "Semáforo". Se entregaron a cada participante tres tarjetas de colores, con el siguiente significado: Verde –criterio clave; Amarillo –criterio relevante, pero no clave; Rojo –criterio no relevante. A cada selección se le asignó un valor numérico (3 al verde, 1 al amarillo, 0 al rojo) y se evaluaron los criterios preliminares. De este ejercicio se obtuvieron los criterios y pesos del AMC, y los criterios se evaluaron posteriormente a través de modelización (criterios cuantitativos) y de cuestionarios (criterios cualitativos). El problema AMC resultante está resumido[1] en la tabla 2.

Tabla 2. Análisis multicriterio del caso de estudio

		Criterios						
		C. Enf.	**Incl.**	**Op. Púb.**	**Cump.**	**Impl.**	**Coste**	**Incentivos**
Parámetro	Unidades	DALYs	N/A	N/A	N/A	Años	M. Soles	N/A
	Optimización	Max.	Max.	Max.	Max.	Min.	Min.	Max.
	Peso	0.1910	0.1348	0.1124	0.1551	0.1865	0.1124	0.1079
Acciones	**A1 renovación**	7263	1.80	1.60	3.80	5.77	15.00	2.40
	A2 chatarreo	55169	2.60	3.80	2.40	6.63	53.96	2.00
	A3 GNV	8308	4.20	4.40	3.80	6.36	80.00	4.80
	A4 mantenim.	68226	2.40	3.00	3.00	6.51	4.31	2.40

Verde: rendimiento máximo; Rojo: rendimiento mínimo

La priorización del AMC da una clara primera posición a la acción A3 (GNV), seguida de A4 (mantenimiento), y a mayor distancia A1, siendo A2 la peor clasificada. Claramente, la influencia de los criterios subjetivos, a su vez evaluados por los actores clave, favorecen a A3. Sin embargo, la primera posición global de A3 enmascara un bajo rendimiento en términos en los criterios cuantitativos (y monetizables), lo cual ejemplifica las limitaciones del AMC puro en considerar costes y beneficios sociales. Aunque no es generalizable, esta discrepancia refuerza la idea de la complementariedad de AMC y ACB.

6. Aplicación combinada de AMC y ACB al caso de estudio

Las metodologías publicadas de combinación del AMC y ACB se pueden clasificar en cuatro categorías:

1) Combinar los resultados del AMC y ACB;
2) Modificar un ACB con los atributos de un AMC;
3) Integrar resultados de un ACB en la estructura de un AMC; o
4) Crear un sistema híbrido.

[1] Abreviaturas, de izquierda a derecha y de arriba a abajo: Carga de enfermedad, Inclusión social, Opinión pública, Cumplimiento, Implementación, Máximización, Minimización, Mantenimiento.

Sin embargo, el único ejemplo encontrado de sistema híbrido no representa una combinación real, sino una presentación alternativa de los resultados de ambas técnicas. Por tanto, no se aplicó al caso de estudio, al que sí se sometió al resto de técnicas.

La combinación de los resultados de AMC y ACB se realizó mediante la metodología COSIMA, propuesta por Barfod, Salling, & Leleur (2011). En esta metodología se obtiene una "Tasa Total de Retorno" que se compone de un término monetario aportado por el ACB y un término procedente del AMC, monetizado mediante una función de valor predeterminada. El peso relativo de uno y otro término es decidido por el analista o por los actores clave. Los resultados de la aplicación al caso de estudio revelan la dificultad conceptual asociada a la función de valor, escasamente replicable y de cuya forma y valoración depende completamente la priorización.

La modificación de un ACB con los atributos de un AMC se realizó mediante la metodología propuesta por Diakoulaki & Grafakos (2004) en el proyecto "External Costs of Energy" (ExternE). En esta metodología se realiza una monetización indirecta de los criterios del AMC basada en el rango de rendimiento de cada criterio y el peso asignado al mismo por los actores clave. Esta evaluación indirecta se suma después a los costes y beneficios del ACB, consiguiendo un ACB ampliado con elementos a priori no monetizables. Los resultados de la aplicación al caso de estudio revelan las restrictivas condiciones teóricas que hacen la metodología poco generalizable en términos de su aplicación; sin embargo, el método de monetización indirecta es transparente y replicable.

Para la integración de los resultados de un ACB en la estructura de un AMC se consideraron varias opciones de entre las publicadas, en concreto las propuestas por Gühnemann, Laird, & Pearman (2012); Nellthorp, Mackie, & Bristow (1998); y Schutte & Brits (2012); las cuales resultaron no aplicables o poco relevantes para el caso de estudio. Por tanto, se propuso una metodología alternativa. Esta metodología es replicable y permite la visualización de contrapartidas entre variables monetizables y no monetizables, aunque esto se consigue a costa de modificar el AMC a posteriori.

Una vez aplicadas al caso de estudio las opciones de combinación AMC-ACB, se hizo una evaluación cualitativa de las características deseables que cada opción presenta. Estas características son múltiples, incluyendo la reproducibilidad, eliminación de doble contabilidad, provisión de solución unívoca, etc. La comparación revela que en todas las opciones de combinación se pierde alguna característica deseable, sugiriendo una elección de tipo de combinación basada en el problema específico a tratar.

7. Conclusiones

Tras analizar y aplicar a un caso de estudio las opciones para la combinación de AMC y ACB para la toma de decisiones en políticas ambientales, se puede concluir que:

- La combinación de AMC y ACB tiene valor añadido, permitiendo visualizar simultáneamente una perspectiva integral de aspectos monetizables y no monetizables, así como las contrapartidas entre unos y otros;
- Específicamente, es potencialmente de gran utilidad para los actores clave y tomadores de decisiones la visualización de discrepancias entre criterios subjetivos, variables no monetizables y la representación de las preferencias de la sociedad, ilustradas por el ACB;
- Sin embargo, no existe por el momento una modalidad de combinación que pueda conservar las características deseables de ambas técnicas, lo cual indica una elección del tipo de combinación basada en el problema de decisión planteado;

Bibliografía

Barfod, M. B., Salling, K. B., & Leleur, S. (2011). Composite decision support by combining cost-benefit and multi-criteria decision analysis. *Decision Support Systems*, *51*(1), 167–175. http://doi.org/10.1016/j.dss.2010.12.005

Brans, J. (1982). L'ingénièrie de la décision; Elaboration d'instruments d'aide à la décision. La méthode PROMETHEE. In R. Nadeau & M. Landry (Eds.), *L'aide à la décision: Nature, Instruments et Perspectives d'Avenir* (pp. 183–213). Québec, Canada: Presses de l'Université Laval.

Brans, J., Vincke, P., & Mareschal, B. (1986). How to select and how to rank projects: the PROMETHEE method. *European Journal of Operational Research*, *24*, 228–23.

Brook, R. (2010). Particulate Matter Air Pollution and Cardiovascular Disease : An Update to the Scientific Statement From the American Heart Association. *Circulation*, *121*, 2331–2378.

Crabbé, A., & Leroy, P. (2008). The Handbook of Environmental Policy Evaluation. London: Earthscan.

Diakoulaki, D., & Grafakos, S. (2004). ExternE-Pol Multicriteria Analysis. Final report on Work Package 4. Retrieved from http://www.externe.info/externe_d7/sites/default/files/expolwp4.pdf

DIGESA. (2012). Estudio de saturación de la calidad del aire en Lima y Callao 2011: vigilancia confiable para la protección de la salud de las personas y de su entorno. Dirección General de Sanidad Ambiental, Peru.

Dockery, D. W., & Pope, C. A. (2006). Health effects of fine particulate air pollution: lines that connect (2006 critical review). *Journal of the Air & Waste Management Association*, *56*, 709–742.

Dockery, D. W., Pope, C. A. I., Xu, X., Spengler, J. D., Ware, J. H., Fay, M. E., … Speizer, F. E. (1993). An association between air pollution and mortality in six US cities. *New England Journal of Medicine*, *329*, 1753–1759.

Eftim, S., Samet, J., Janes, H., McDermott, A., & Dominici, F. (2008). Fine particulate matter and mortality: a comparison of the six cities and American Cancer Society cohorts with a medicare cohort. *Epidemiology*, *19*(2), 209–216.

Golub, E., Klytchnikova, I., Sanchez Martinez, G., & Belausteguigoitia, J. C. (2014). Environmental Health Costs in Colombia: The Changes from 2002 to 2010. *World Bank Occasional Paper Series*, (92956).

Gregory, C. E., & Brierley, G. J. (2010). Development and application of vision statements in river rehabilitation: the experience of Project Twin Streams, New Zealand. *Area*, *42*(4), 468–478. http://doi.org/10.2307/40890904

Gühnemann, A., Laird, J. J., & Pearman, A. D. (2012). Combining cost-benefit and multi-criteria analysis to prioritise a national road infrastructure programme. *Transport Policy*, *23*, 15–24. http://doi.org/10.1016/j.tranpol.2012.05.005

Katsouyanni, K., Touloumi, G., Samoli, E., Gryparis, A., Le Tertre, A., Monopolis, Y., … Schwartz, J. (2001). Confounding and effect modification in the short term effects of ambient particles on total mortality: results from 29 European cities within the APHEA2 project. *Epidemiology*, *12*(5), 521–531.

Naylor, C., & Appleby, J. (2013). Environmentally sustainable health and social care: Scoping review and implications for the English NHS. *Journal of Health Services Research & Policy*, *18*(2), 114–121. http://doi.org/10.1177/1355819613485672

Nellthorp, J., Mackie, P., & Bristow, A. (1998). Measurement and Valuation of the Impacts of Transport Initiatives. Retrieved from http://ec.europa.eu/environment/archives/tremove/pdf/measurement_maintext.pdf

OECD. (2008). Introductory Handbook for Undertaking Regulatory Impact Analysis (RIA). Paris, France. Retrieved from https://www.oecd.org/gov/regulatory-policy/44789472.pdf

Pope, C. A. I., Burnett, R. T., & Thun, M. J. (2002). Lung cancer, Cardiopulmonary mortality, and Long-term exposure to fine particulate air pollution. *Journal of the American Medical Association*, *287*, 1132–1141.

Pope, C. A. I., Thun, M. J., Nambudiri, M. M., Dockery, D. W., Evans, J. S., Speizer, F. E., & Heath, C. W. J. (1995). Particulate air pollution as a predictor of mortality in a prospective study of US adults. *American Journal of Respiratory and Critical Care Medicine*, *151*, 669–674.

Samet, J., & Krewski, D. (2007). Health effects associated with exposure to ambient air pollution. *J Toxicol Environ Health A*, *70*(3–4), 227–242.

Schutte, I., & Brits, A. (2012). Prioritising transport infrastructure projects: towards a multi-criterion analysis. *Southern African Business Review*, (3), 97–117.

WHO. (2006). Air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide. Global update 2005. Summary of risk assessment. Geneva, Switzerland: World Health Organization. Retrieved from http://www.who.int/phe/health_topics/outdoorair/outdoorair_aqg/en/

This page was intentionally left blank

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS

AF: attributable fraction
AHP: analytic hierarchy process
AMR-D: subdivision of the WHO region of the Americas
ARAPER: (Spanish) automotive dealers association of Peru
BAU: "business as usual" scenario
BCR: benefit cost ratio
BoD: burden of disease
CAI-LAC: Clean Air Institute for Latin American cities
CBA: cost–benefit analysis
CE: choice experiments
CEA: cost effectiveness analysis
CGIAL: (Spanish) management committee for the Clean Air Initiative
CNG: compressed natural gas
COFIDE: (Spanish) Peru development finance corporation
COI: cost of illness
CONFIEP: (Spanish) national confederation of private companies of Peru
COSIMA: Composite Model for Assessment
CVM: contingent valuation methods
DALYs: Disability-Adjusted Life Years
DEFRA United Kingdom Department of the Environment Food and Rural Affairs:
DETR: United Kingdom Department of Transportation Planning
DIGESA: (Spanish) Peru environmental health directorate of the ministry of health
DKK: Danish Kroner
DTU: Technical University of Denmark
ECAs: (Spanish) Peru environmental air quality standards
ELECTRE: ELimination and Choice Expressing REality
EUAC: equivalent uniform annual cost
ExternE: the External Costs of Energy, a research project funded by the European Union
GAIA: Geometrical Analysis for Interactive Aid
GDP PPP: gross domestic product adjusted by purchasing power parity
ICD-10: International Statistical Classification of Diseases and Related Health Problems 10th Revision
INEI: (Spanish) Peru national institute of statistics
IRR: internal rate of return
LAC: Latin America and the Caribbean
LAM: linear additive model
LMP: (Spanish) Peru air quality maximum permissible thresholds
MALC: metropolitan area of Lima and el Callao
MAUT: Multi attribute utility theory
MCA: multi-criteria analysis
MCCBA: multi-criteria cost–benefit analysis
MCDA: multi-criteria decision analysis
MCDM: Multi-criteria Decision Methods
NBER: National Bureau of Economic Research
NOx: nitrogen oxides
NPV: net present value

NRA: National Roads Authority of Ireland
OECD: Organization for Economic Cooperation and Development
PISA-II: (Spanish) second plan for air quality improvement in Lima and Callao
PM10: particulate matter smaller than 10 microns in diameter
PM2.5: particulate matter smaller than 2.5 microns in diameter
PPI: potential Pareto improvement
PROMETHEE: Preference Ranking Organization METHods for Enrichment Evaluations
PROTRANSPORTE: transportation authority of Lima and Callao
PV: present value
ROD: ranked order distribution
SEB: social economic benefit
SEC: social economic cost
SEU: subjective expected utility
SMART: specificity measurability attainability relevance and time-boundedness
SO_x: Sulphur oxides
SUNAT: (Spanish) Peru national tax and tariffs superintendence
SWF: social welfare function
SWING: standardized weight elicitation model
TOPSIS: Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS)
TRR: total rate of return
USEPA: United States Environmental Protection Agency
VIKOR: multi-criteria optimization and compromise solution
VNG: vehicular natural gas
VOCs: volatile organic compounds
VSL: value of a statistical life
WB: World Bank
WHO: World Health Organization
WSI weight stability intervals:
WTA: willingness to accept
WTP: willingness to pay
YLD: years of life lost due to disability
YLL: years of life lost due to premature mortality

ACKNOWLEDGEMENTS

I gratefully acknowledge the help and collaboration of: Sergio Sánchez, Executive Director of CAI-LAC; Juan Carlos Belausteguigoitia, former lead environmental economist of the LAC region at the World Bank; the World Bank Peru Country Office; all the members of the CGIAL and its coordinator Gladys Macizo; all the stakeholders involved in this policy process, particularly Peter Davis (ARAPER), Cecilia Mendiola (ASPEC), José Estrada (PETROPERU), Francisco Fuentes (DIGESA); Congressman Jaime Delgado Zegarra; Francisco Pimentel (COFIDE); and all the helpful and kind personnel at the Ministry of Housing of Peru.

Special thanks to my PhD supervisors, Dr Maria Paz Moral Zuazo of the University of the Basque Country (UPV) and Dr Aline Chiabai of the Basque Centre for Climate Change (BC3).

I would like to dedicate this dissertation to my growing family, especially Nina, who unlike me never doubted that I would get it finished.

“Systems were made for men, and not men for systems”

C.H. Douglas, Economic Democracy (1920)

Chapter 1. Introduction

1.1. Evidence in the general and environmental policy process

While the use of empirical data and rational evidence to inform policy is far from new, the notion that rational decision-making should systematically prevail before ideologically-driven politics is relatively recent. The term "evidence-based policy-making" was popularized during the mandate of A.C.L. ("Tony") Blair, Premier of the United Kingdom from 1997 to 2007. The UK government itself openly committed to using such approach (UKCO, 1999), stating that "[The government must] produce policies […] shaped by the evidence rather than a response to short-term pressures".

The arguments in favor of evidence-based policy-making have been discussed extensively elsewhere (Banks, 2009; Bullock et al., 2001; Sutcliffe & Court, 2005). Some commonly used ones are:

- Transparency: by relying on and making use of objective data in a formalized framework, decision-making is less prone to be affected by personal and/or sectorial interests.
- Accountability: the availability of information may help hold decision-makers and administrations accountable to institutions and the public.
- Efficiency: evidence can steer public resources away from costly and ineffectual policies and into interventions that provide the best financial and/or social returns.
- Relevance: evidence can highlight the urgency of an issue and how it fits the real needs of the communities, which in turn may result in improved outcomes for the selected interventions.

However, the approach also draws criticism. A few frequently cited by researchers in the field are:

- Resources: the collection and analysis of solid evidence frequently entails a financial cost and capacities that may not be readily available or affordable by the decision-making body.
- Scope: evidence is just one element of decision-making, and neither can nor should displace other important ones such as values, political considerations, etc.
- Speed: decision-making in the policy process often needs to happen in shorter timeframes than those required to gather adequate data and produce sound scientific evidence.
- Uncertainty: very frequently, data and science cannot provide conclusive evidence on one or more aspects of the actions analyzed. Moreover, this sometimes creates a communication gap between analysts and decision-makers regarding the interpretation of uncertainty.

There are many more arguments in favor and against the reliance on evidence for policy-making proposed by researchers, policy-makers, analysts and other relevant actors. This dissertation does not enter that debate, and simply assumes significant value in the availability and use of relevant evidence for decision-making.

Researchers in the field commonly disaggregate the policy process in different stages. These stepwise models can be used to identify "entry points" for policy inputs, including evidence. There are numerous conceptual models describing these stages in a linear (or rather, circular) way, the scope of which is besides the intent of this dissertation.

A simplified conceptual model by Young & Quinn (2002) based on Anderson (1994) among others, can be summarized in the following steps:

1. Problem definition/agenda setting: at some point in time or periodically, one or more groups within society identify a particular issue that is deemed to require government action. Once acknowledged the issue and need for action by the government, it is included into the political agenda.
2. Policy alternatives/formulation: once the problem is defined and in the government's agenda, and upon gathering enough detail information, policy alternatives are defined. Ideally those should be mutually exclusive, so that adequate prioritization can occur.
3. Selection of preferred policy option/s: once policy alternatives are defined, one or more policy options are selected based on a set of evaluation criteria. These criteria can vary widely (e.g. effectiveness, cost, feasibility, etc.) but once agreed upon, they can serve both to prioritize the most suitable alternative/s and to defend the legitimacy of selected policy options.
4. Policy design: once the policy option/s are selected, government authorities need to decide how to implement them most effectively, including choices as to which instruments to use, mechanisms of enforcement if necessary, etc.
5. Implementation and monitoring: the policy option/s is/are implemented according to the policy design. Simultaneously, information on process and intermediate outcome/s is collected in order to enable an appropriate evaluation.
6. Evaluation: in order to accurately measure the effectiveness of any policy it is necessary to evaluate its outcomes. Such evaluation is particularly important as inputs for further policy-making processes.

A common approach to gathering, analyzing and providing evidence in a workable format for the different stages of the policy cycle is the use of analytical tools imported from several disciplines. Depending on the specific needs of each policy process, analysts can resort to one or more of the vast pool of available analytical techniques that each discipline provides. The amount and diversity of those techniques is large and well beyond the objectives of this document, which will focus only on a subset of techniques commonly used to evaluate environmental policy.

In the case of environmental policy as in several other areas, public policy-makers require information on the foreseeable consequences of diverse policy interventions in order to make informed decisions regarding their implementation or revision. The range of approaches to gathering and organizing that information is large and growing, and although none of the existing tools are exclusive to the field of environmental policy, certain types of tools are systematically favored in this area (Table 1 lists some of the most widely used).

This dissertation deals only with tools which can include a central economic component that can be applied to environmental policy. The analysis is applied in the context of air quality management. The tools analyzed are cost–benefit analysis[1] (CBA) and multi-criteria analysis (MCA), for several reasons. Firstly, they are comprehensive enough to be applied to almost any evaluation, which helps explaining why they are the most widely used tools regulatory impact assessment (OECD, 2008). Secondly, there is a very large body of literature on them (more so on CBA, but MCA is growing rapidly) spanning theoretical foundations, practical rules for application and several case studies to draw from. And thirdly, in combination they offer possibilities to consider economic impacts in broad terms, integrating monetized impacts with those that do not admit monetization. Moreover, there is little research on the possibilities of this combination, and how to draw the positive aspects of both tools.

[1] Cost-Effectiveness Analysis is treated in this report as a specific case of CBA in which expected outcomes are not converted into benefits.

Table 1. Commonly used tools for the evaluation of environmental policies

Evaluation tool	Summarized description
Needs analysis	Evaluation of the relevance of a policy by confronting expected outcomes with policy needs
Program theory evaluation	Analysis of the assumptions and rationales upon which the policy is based, and its contribution to outcomes
Case study evaluation	Analysis of real government intervention case studies that can serve as examples
Experiment and quasi-experiment	Measures effects of policies in selected groups, before and/or after the intervention
Formative/developmental evaluation	Evaluates effectiveness and the quality and performance of policy implementation
Goal-free evaluation	The evaluator is not informed of the goals of the policy, rather focusing on needs of the impacted
Impact assessment	Process of evaluating different policy impacts before their implementation
Cost–benefit and Cost-Effectiveness Analysis	Comprehensive and systematic evaluation of costs versus either economic benefits or impact of policies
Log frame method	Provides and analyzes a logical summary of a project through a preset matrix of events vs. available information
Multi-criteria analysis	Explicitly considers criteria and their relative importance in the evaluation of one or more policies
Realistic evaluation	Includes the specific context in the evaluation to maximize impact in a specific situation

Source: Adapted from Crabbé & Leroy (2008).

MCA and CBA are used in difficult policy problems, including those concerning environmental risks. In urban settings, poor air quality is frequently among the most severe such environmental risks. Increasingly aware of the large burden that air pollution imposes, governments around the world are trying to tackle its causes and effects. However, adopting policies that effectively reduce urban air pollution is challenging, since such policies typically affect sensitive pillars of urban life, such as people's mobility (Molina & Molina, 2005). Often times, actions required to improve air quality (e.g. fuel quality improvement) cannot be readily acted upon by local governments. The types of actions included in air quality management program frequently involve long timeframes and large investments. Importantly, considerable uncertainty as to the real effects of each measure plagues the decision-making in the process of improving urban air quality. These uncertainties have been frequently used as an argument in favor of the use of multi-criteria analysis (MCA) in urban environmental policy-making. MCA processes fit the characteristics of local urban decision-making regarding air quality, namely: 1) Hard budget constraints; 2) Limited capacity and data availability; 3) Conflicting views and criteria; 4) Need for inter-sectorial and multi-disciplinary action; and 5) Consensus is needed for an effective implementation.

Notwithstanding this alignment, the role of MCA in air quality policy evaluation remains minor in comparison with CBA. In this dissertation, the two tools (CBA and MCA) are applied to the same case study, separately and in combination, in order to obtain insights as to their relative and joint value for decision support. The case study, based on an actual policy process in a Latin American megacity, supports the conceptual exploration of the opportunities and pitfalls of the combination of CBA and MCA. The background and specifics of this case study are laid out in the next section.

1.2. Case study: air quality policies in Lima and El Callao, Peru

1.2.1. Background: health effects of air pollution

Evidence on the health effects of urban air pollution has been substantial for decades, with extensive studies showing the association between certain air pollutants and respiratory and cardiovascular mortality, chronic bronchitis, respiratory infections, and several other related disorders, both acute and chronic (Golub et al., 2014). These health outcomes affect disproportionately some vulnerable groups, including children, the elderly and those with preexisting cardiovascular and respiratory conditions (Dockery & Pope, 2006; Dockery et al., 1993; Katsouyanni et al., 2001; Pope et al., 1995; Pope et al., 2002). Although health effects are not the only impacts[2] of air pollution in cities, they are the most important and best studied. Epidemiological studies show an association between certain air pollutants and respiratory and cardiovascular mortality, chronic bronchitis, respiratory infections, and several other disorders. Research shows the strongest effects for inhalable particulate matter, particularly $PM_{2.5}$ (smaller than 2.5 microns in diameter). After comprehensive reviews on the effects of particulate matter on health in the late 1990s and early 2000s (e.g. Dockery & Pope, 2006; Pope et al., 2002) large epidemiological studies, including meta-analyses, have followed. These included multiple locations, mostly in Europe and North America (e.g. Samet et al., 2000). Large cohort studies (Brook, 2010; Eftim et al., 2008; Samet & Krewski, 2007) have since confirmed significant effects of inhalable particles on mortality and morbidity in different age groups.

Because of the severity of their health effects and their preponderance in the burden of disease caused by urban air pollution, fine particles are commonly used in studies on the cost of environmental degradation as the exposure indicator of choice to assess attributable mortality and morbidity (Golub et al., 2014). Notwithstanding, several types of anthropogenic emissions are associated with adverse health outcomes, including (but not limited to) sulfur oxides, nitrogen oxides, volatile organic compounds (VOCs), carbon monoxide, lead, and especially ozone (WHO, 2006). Tropospheric (i.e. ground-level) ozone can trigger a large number of respiratory effects and aggravate certain chronic diseases, thus increasing outcomes such as health care usage or absenteeism, with high costs to society (USEPA, 2011, 2012). Overall, however, both the population-wide observed effect and the evidence on impacts of particles is greater than those of any other major air pollutant (WHO, 2006).

Besides the solid body of literature from developed countries, a growing body of evidence is being established from cities in Latin American countries. There are currently several studies from large and medium cities in Mexico, Brazil and Chile (for instance Bell et al., 2011; Bell et al., 2006; O'Neill et al., 2008). However, the amount of locally-generated information available regarding the health effects of air pollution in Peru is scarce, in particular with regard to peer-reviewed epidemiological studies.

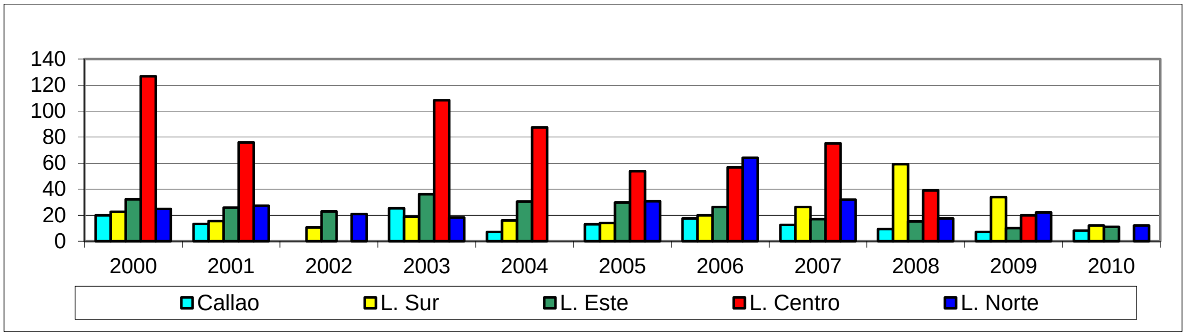
1.2.2. Air pollution and air quality policy in Lima and El Callao

Air quality monitoring in the Metropolitan Area of Lima and El Callao (hereafter MALC) is carried out through five monitoring stations placed in different parts of the city: Callao, South, East, Center, and North. According to governmental data (DIGESA, 2012), the concentrations of Sulfur dioxide (SO_2) and Nitrogen Dioxide (NO_2) have dropped sharply in the capital area since the year 2000, when systematic measurements started, to well below the established national regulatory thresholds (80 and 100 micrograms per cubic meter, respectively) by 2010.

[2] Other typical impacts of urban air pollution include loss of visibility, damages to buildings and cultural heritage, corrosion of materials and various impacts on vegetation.

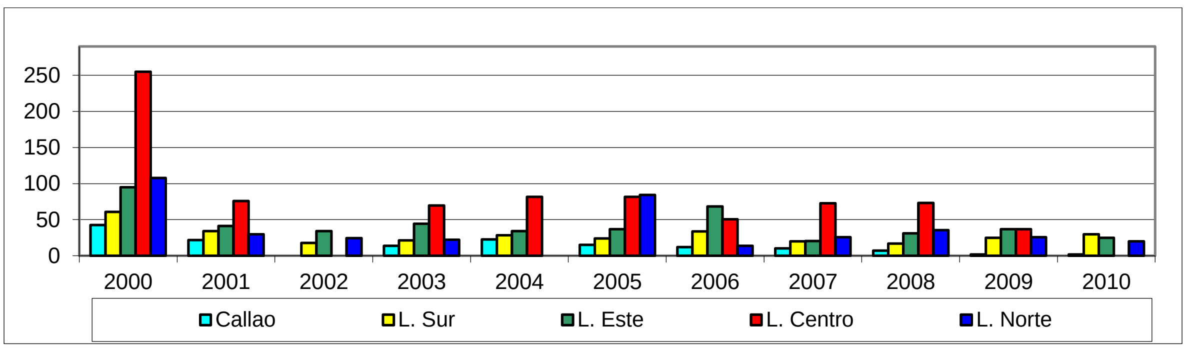
The same, however, cannot be said about inhalable particulate matter (PM) which on average has stayed over the regulatory thresholds (50 and 15 micrograms per cubic meter for PM_{10} and $PM_{2.5}$, respectively). These trends are visible in Figure 1, Figure 2, and Figure 3 below[3].

Figure 1. Yearly average of daily concentration of SO_2 in Lima-Callao 2000-2010



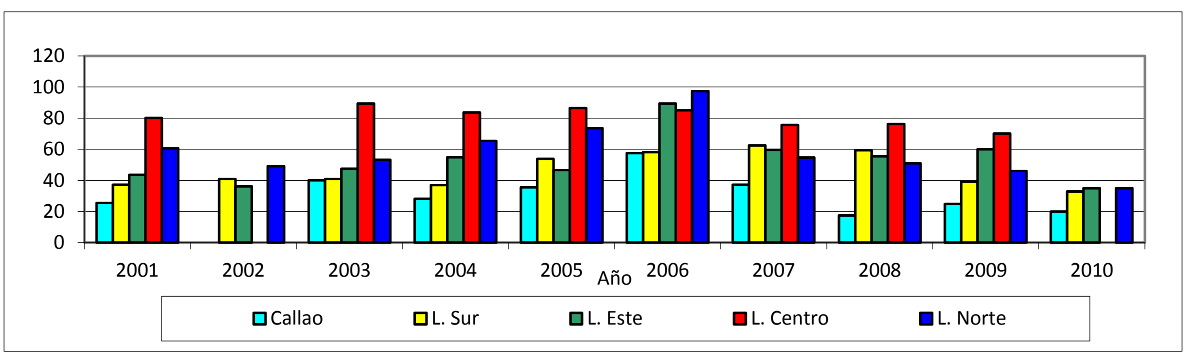
Source: DIGESA, 2012. Years in X axis. Concentration units (Y axis): μg/m³

Figure 2. Yearly average concentration of NO_2 in Lima-Callao 2000-2010



Source: DIGESA, 2012. Years in X axis. Concentration units (Y axis): μg/m³

Figure 3. Yearly average concentration of PM_{10} in Lima-Callao 2000-2010



Source: DIGESA 2012. Years in X axis. Concentration units (Y axis): μg/m³

[3] The monitoring stations are denoted by their coded names in Spanish: Callao (for El Callao); L. Sur (for South); L. Este (for East); L. Centro (for Center) and L. Norte (for North). Additional detail on the situation is provided in Chapter 3.

In summary, and according to the available data, the MALC has a severe air pollution problem, largely due particles emitted by mobile sources (i.e. vehicles), which supposes an environmental health risk causing premature mortality and disease. These health outcomes are largely preventable through coordinated policies that could effectively reduce population exposure to air pollutants, and specifically inhalable particles.

Efforts towards a comprehensive management of air quality in Peru started in earnest in the late 1990s with the incorporation of Lima to the World Bank-sponsored Clean Air Initiative for Latin America, followed by regulations establishing maximum thresholds for key pollutants. The 2001 National Air Quality Standards decree (Reglamento D.S. Nº 074-2001-PCM) established basic thresholds (locally known as Estándares de Calidad Ambiental del Aire) that were subsequently developed and sometimes modified by other regulatory tools. In the late 2000s the Maximum Permissible Thresholds (Spanish acronym: LMP) were progressively approved in order to limit emissions by mobile and stationary sources in all urban areas, but with special focus on the MALC. Whereas the LMPs are unlikely to change in the short term, ECAs have been progressively lowered in a continuous policy process.

Currently, most competencies regarding environmental issues are carried out by the Peruvian Ministry of the Environment. However, air quality remains an exception, with competencies shared among the Ministry of Housing, the environmental health directorate (Spanish Acronym: DIGESA) of the Ministry of Health and the Ministry of the Environment. Of note is the role of the "Committee for the management of the clean air initiative for Lima and Callao" (CGIAL), a multi-stakeholder organ maintained by the Ministry of Housing which has in effect designed and developed most air quality regulations and policies.

In the MALC, the capital area and largest metropolitan concentration of Peru, the first comprehensive air quality management program was the "Plan integral de saneamiento atmosférico para Lima y Callao 2005-2010" (General air quality improvement plan for Lima and Callao 2005-2010), a heterogeneous pool of measures affecting mobile and stationary sources, fuel quality, economic incentives, urban management and traffic calming. Such measures were unequally implemented but marked a strong improvement in general air quality in the MALC, even in a context of rapid motorization rates. In summary, Nitrogen and Sulfur Oxides have decreased significantly in the last decade, but inhalable particles (PM_{10} and $PM_{2.5}$) remain on average above national standards and continue to represent a significant public health hazard.

Acknowledging this situation, the "Second plan for air quality improvement in Lima and Callao" (hereon PISA-II) set out to curb air pollution trends, and specifically to reduce PM levels. PISA-II is structured into three work programs:

- P1: Reduction of the emissions and concentrations of particulate matter;
- P2: Mid-term pollution prevention and abatement; and
- P3: Research in air pollution.

According to the CGIAL (and sensibly, in terms of public health) the most urgent program is P1, aimed at reducing particulate matter, so this dissertation focuses on said program. The proposed timeframe of the plan was from 2011 to 2015. However, most of the policies (including those pertaining mobile sources) were as of the end of 2014 neither evaluated nor implemented.

The key activities within the PISA-II plan with potential for reduction of particulate matter can be integrated into four actions, formulated as follows:

- Action 1 (A1), private vehicle fleet renewal: this action is based on the implementation of a vehicle fleet renewal scheme, including incentives for the purchase of new vehicles and disincentives for the use of older vehicles, which in this case are proposed as variations of the existing tax regime as applied to light vehicles. This would be complemented by a sweeping ban on imports of used light vehicles and their incorporation into the vehicle fleet: Peru's private vehicle fleet is one of the oldest in Latin America, vehicles are on average more than 12 years old (ARAPER, 2011).
- Action 2 (A2), bonus for scrapping older public transportation vehicles: public transportation in the MALC is atomized into myriads of providers operating obsolete and polluting vehicles. This policy is meant to incentivize the replacement of these older vehicles by newer, cleaner ones through a bonus for scrapping, a measure used in multiple settings across the world.
- Action 3 (A3), promotion of vehicular natural gas (VNG) use in public transportation vehicles, specifically through the allocation of grants for the retrofitting of public transportation vehicles in order to be able to use VNG, as well as a line of loans with no interest to facilitate the purchase of new, native VNG user vehicles.
- Action 4 (A4), enforcement of better vehicle maintenance: this entails an overall effort to enforce better vehicle maintenance, thus reducing net pollutant emissions through better functioning engines and exhausts. In this direction, a renewed effort of compliance monitoring can ensure that inspections are carried in accordance with national standards. It will be further enforced through random checks of vehicle emissions in streets and roads: this measure intends to increase the proportion of vehicles that actually undergo annual inspection from the pool that is mandated by law to have such inspections. Compliance levels are currently rather low, inducing poor maintenance which in turn results in higher pollutant emissions.

These actions and their implications are further detailed in Chapter 4 (CBA), and they constitute the basis of the CBA and MCA problems, and of the subsequent integration which is the focus of this dissertation. The purpose of the evaluation in the case study is not to discard alternatives in favor of an "optimal" one, but rather to have an idea of the trade-offs and potential benefits of each, so that their implementation can be prioritized or postponed in a context of challenging governance, competing public demands and constrained resources.

This case study resulted from a request of technical assistance from the competent authorities in air quality management in Peru (the Ministry of Housing) to the Clean Air Institute in Washington, DC. My help (in collaboration with the Clean Air Institute) was requested in exploring decision-support tools for the prioritization of sectorial policies aimed at improving air quality in the capital city area, and specifically to consider cost–benefit analysis and multi-criteria Analysis. The PISA-II air quality management plan was established as an illustrative case study, specifically its actions to reduce particulate matter. Despite the timeframe of PISA-II, the analyzed actions are as yet neither evaluated nor implemented. These actions are also not mutually exclusive, and in principle all will eventually be implemented. Relevant information was obtained from direct sources of the Ministries of the environment, housing, energy, health and finance; the national institute of statistics, the national geographic institute, the municipalities of Lima and El Callao, COFIDE (Corporación Financiera de Desarrollo, a public-private development bank), and two industry associations (ARAPER and CONFIEP, Asociación de Representantes Automotrices del Perú and Confederación Nacional de Instituciones Empresariales Privadas, respectively).

1.3. Objectives and structure of this dissertation

Both MCA and CBA present numerous advantages (as well as some disadvantages, explored in the literature review chapter) for the problem at hand. Moreover, advantages of each approach roughly complement each other, raising the question of whether a combined approach could preserve the main qualities of each analysis while simplifying the output for decision support. This question has received little attention in the relevant literature, and thus is relatively open. Therefore, the main objectives of this dissertation are:

- To provide a background to understand the implications of the combined use of CBA and MCA, and to develop a conceptual framework for the integration of the two methods.
- To lay out and analyze the main methodological options for CBA-MCA combination and apply them to a real case study for the prioritization of air quality policies in the local governance context (an application not yet featured in the published literature, to my knowledge).
- To explore advantages and disadvantages, strengths and conceptual pitfalls of each combination option.

In summary, I intend to explore the feasibility and added value of the possibilities of combination through a case study of air quality management options in a large metropolitan area. The case study serves as a connecting thread for the conceptual discussion. The different steps of the core analysis of the case study are illustrated in the next page (Figure 4). The structure of the dissertation from here onwards is as follows:

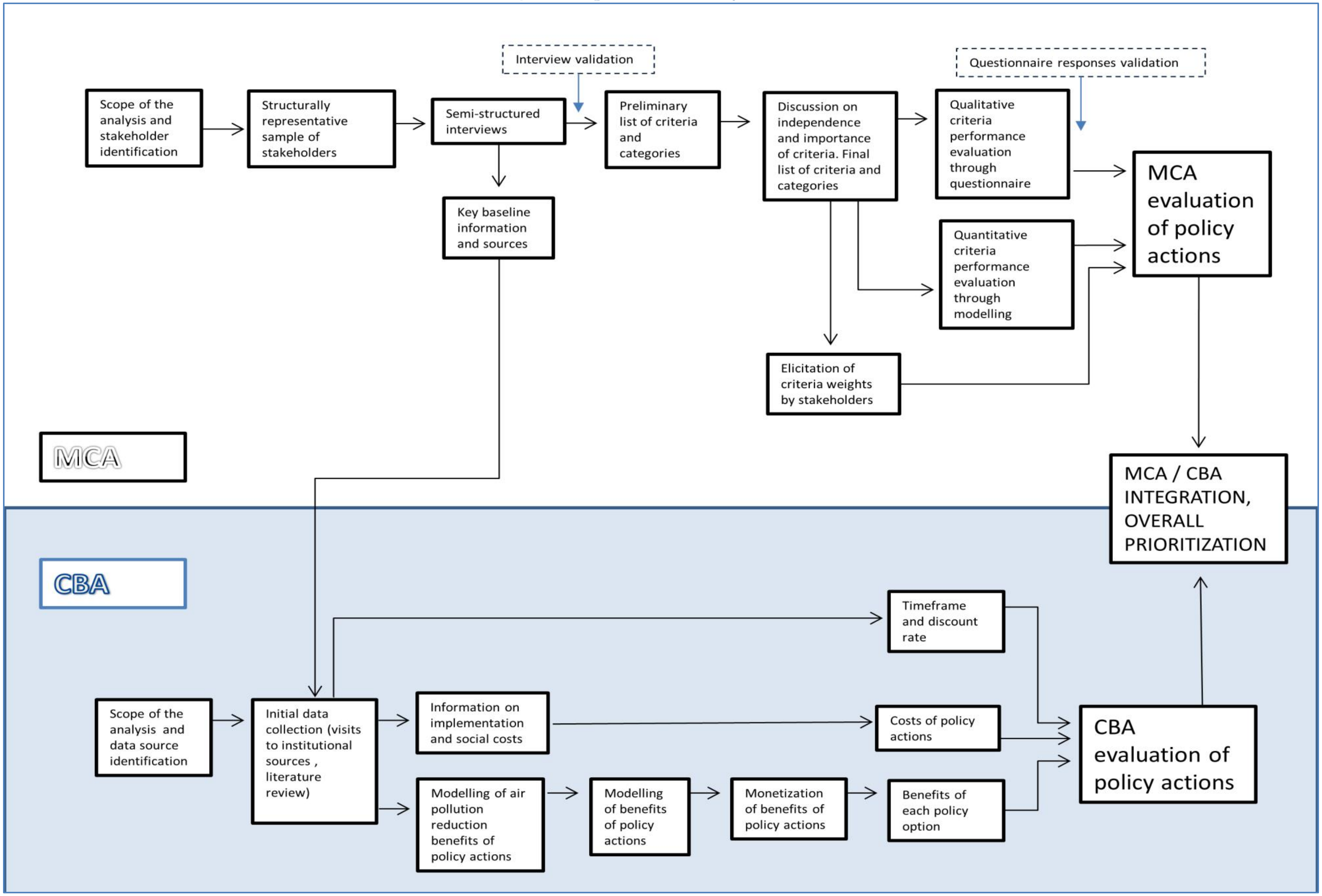
- A review of the relevant literature on CBA and MCA in the context of air quality (Chapter 2);
- An impact assessment of the actions proposed in the case study (Chapter 3);
- A CBA of the actions proposed in the case study (Chapter 4);
- A MCA of the actions proposed in the case study (Chapter 5);
- A set of possible combinations of the CBA and the MCA applied to the case study, with a discussion on the value of each combination approach (Chapter 6);
- Conclusions of the process (Chapter 7)

In addition, this work is bound to add some value for research and a degree of novelty:

- In the case of the CBA, it adds to a very small pool of examples of an application of the technique at the local level for air quality management in the region of Latin America and the Caribbean;
- In the case of the MCA, the added value is two-fold: firstly, its application in this context is rare for cities in LAC; secondly, the entire process is driven by the stakeholders, including the scoping, criteria ascertainment, and weighing; this is also rather uncommon in the existing literature;
- The application of a CBA-MCA combination to local air quality policy in LAC is, to my knowledge, not yet featured in the published literature. In addition, after applying published CBA-MCA integration approaches to the case study, a novel approach for the combination of both techniques is proposed.

What follows in Chapter 2 is an overview of the scientific literature regarding the concept and framework of both CBA and MCA, as well as their application to air quality management. The chapter also explores the published approaches for the combination or integration of CBA and MCA for decision-support purposes, including their theoretical basis, general application and specific use for environmental policy with a focus on air quality management. The main limitations of each technique are also listed and elaborated on.

Figure 4. Steps in the analysis of the case study



Chapter 2. Literature review

2.1. Methods

In a well-developed area like the application of decision-support tools to environmental policy, any overview of the current state of the science requires a comprehensive literature review. Given the potentially large number of references, I developed a basic search strategy and basic selection criteria for inclusion in the review. These restrictions were particularly important in the review of evidence in the application of each of the analysis (CBA and MCA) where an indiscriminate search would have provided an unmanageable number of references. However, such restrictions were deliberately loosened in the main topic of concern (the combination of both tools) in order to capture the bulk of relevant literature in that specific area. The principles used in the literature review were thus:

- Search strategy: relevant research concerning the topics covered in the structure above was identified through a search in scientific and economic databases (RePEc, ScienceDirect, JSTOR, EBSCO, NBER) as well as academic search engines (Google scholar and Mendeley) and electronic libraries (Source OECD, Worldbank iLibrary and the Royal Library at the University of Copenhagen). The search terms were deliberately kept broad to capture most relevant studies. These included the single terms and combinations of: multi-criteria analysis, multi-criteria decision analysis, public policy cycle, evidence-based policy-making, decision-support systems, environmental policymaking, cost–benefit analysis, air quality management, air quality policy prioritization, environmental policy evaluation tools, combination multi-criteria cost benefit. After a first review of articles found using these terms, further key literature and authors were identified through referenced citations (backward snowballing), and accessed mostly through direct e-journal consultations. In addition, to capture governmental reports, international organization publications and grey literature, an internet search was conducting using the same search terms in two freely available internet search engines (Google and Duckduckgo).
- Selection criteria: As far as possible, academic journal articles were used. All materials for which proper authorship, date of publication or attribution could not be established were discarded. For large research area topics such as public policy, decision-support systems, or the application of CBA or MCA, only those studies and publications directly related to environmental policies and/or air quality management policies were kept. However, for areas where literature is scarce, such as the combination of MCA and CBA, all relevant material was kept regardless of its relationship with environmental policies.

The results of the literature review are presented in phases: first, cost–benefit analysis in environmental policymaking (framework, application, examples and limitations); second, multi-criteria analysis in environmental policymaking (framework, application, examples and limitations); and last, the combinations of CBA and MCA (strategies and examples). The rest of this chapter summarizes the most relevant information derived from the reviewed literature, in order to establish the theoretical basis for the analysis and to ensure the originality of the work presented.

2.2. Cost–benefit analysis and environmental policy

2.2.1. Fundamentals of cost–benefit analysis

Measuring costs against benefits is a common way to make rational decisions in our daily lives. Moreover, when options are multiple and consequences uncertain, there is value in the systematization and categorization of the information so that decisions can be made in a clear and rational way. The theoretical development, the economic foundations and the framework adaptation and application of CBA are lengthy and complex matters, and have been discussed extensively elsewhere (EU, 2008; Pearce et al., 2006). Thus, such matters will be explored in the present document only insofar as they provide a better understanding of the value of the technique for decision-making in environmental policy. The theoretical foundations for CBA are summarized by Pearce et al. (2006):

a. The preferences of individuals are the source of value. If an individual prefers state A to state B, then his/her well-being, welfare or utility is higher in A than in B.
b. Preferences are measured by a willingness to pay (WTP) for a benefit and a willingness to accept compensation (WTA compensation) for giving up a good or service.
c. Individuals' preferences can be aggregated: social benefit is the sum of all society's individuals' benefits and social cost is the sum of their costs.
d. The basic decision rule to undertake a change is that benefits for all of society outweigh costs. However, accounting for the difficult application of that rule in the practice, a slight but significant variation is introduced: if those who benefit from a change can hypothetically compensate the losers and still have some net gains left over, then the change is desirable from a welfare perspective – though not necessarily Pareto-optimal. This principle is known as the "Kaldor-Hicks compensation test", discussed later in this subsection.

In addition, it is generally assumed that a unit of benefit or cost in the future has a lower weight than the same unit or benefit cost occurring now. This temporal weight is known as the discount factor and its application implies the determination of a discount rate. This notion is frequently challenged partially or totally in areas where high discounting means harm to future generations, such as environmental and public health policy, an issue also discussed later in this subsection.

Costs and benefits can be considered from a private and social perspective. Private costs are typically the market-based ones faced by individuals and firms. However, those private costs may not take full account of the cost of any given action to society, since there might be external costs unaccounted for in the markets. The same applies to the comparison of private and social benefits. By definition, social CBA is more complete that private CBA, since it includes both market-based and other relevant costs. This dissertation will be dealing with CBA exclusively from a social perspective, as it is the common practice in the context of environmental and public health goods and services.

While the concept of CBA was coined around the mid of the XIX century, its systematic use for policy evaluation was mainstreamed by Federal agencies in the United States prospectively evaluating investments in large hydraulic infrastructures in the second half of the twentieth century. Indeed, water works were the field where most practical guidance in the definition and calculation of costs and benefits of policies, programs and projects was issued. Central to this practical guidance was the work of Eckstein (1958), Maass (1962) and McKean (1958), most of it centered on the issue of investment in and development of water resources.

By the 1960s, the theory supporting the application of CBA had been largely established and the technique widely used for assessment and evaluation in government. In the field of environmental policy, CBA started to be used regularly later, in order to evaluate policies that entailed large investments and considerable uncertainties. That, in turn, took some time, namely the years that went by between environmental issues being a minority concern and the establishment of full governmental environmental agendas. In the 1970s, the US Environmental Protection Agency initiated large research programs into the economic valuation of environmental damage (Pearce, 1998). Within that decade and the next, several methods were developed to ascertain revealed and stated preferences regarding the environment, and by the 1990s the systematic application of CBA to environmental policies was commonplace both in the US and in the EU.

Currently, CBA models are hugely numerous, but most follow a similar framework, translated into practice as a standard set of steps. Below is one example of such a set as featured in a British governmental guidebook for appraisal (HMT, 2011):

1. Provide a justification for action.
    - 1.1. Establish the rationale for intervention.
    - 1.2. Determine whether intervention is warranted.
    - 1.3. Identify the scope of the issues involved and the basis for government action.
2. Set objectives.
    - 2.1. Establish what the proposals are intended to achieve
    - 2.2. Create a hierarchy of outcomes, outputs and targets.
3. Appraise the options to establish a Base Case.
    - 3.1. Prepare a list of the range of possible actions to achieve the objectives.
    - 3.2. Value the costs and benefits of the options.
    - 3.3. Adjust, where necessary, the costs and benefits to take account of distributional aspects and relative price changes.
    - 3.4. Discount all costs and benefits to 'present values'.
    - 3.5. Adjust for differences in tax between options.
    - 3.6. Adjust for risk and optimism, and consider the impact on the Base Case of alternative scenarios and changes in key variables.
    - 3.7. Calculate the net benefits or costs for each scenario and option.
    - 3.8. Consider unvalued costs and benefits using 'weighting and scoring'.
4. Develop and implement the solution.
    - 4.1. Identify the 'best' option.
    - 4.2. Determine the affordability of options.
    - 4.3. Through consultation, refine the 'best' option or options into a solution.
    - 4.4. Devise implementation plans.
    - 4.5. Present the results.
    - 4.6. Implement the solution.
    - 4.7. Track the success of the policy, program or project in achieving its objectives.

While these steps might seem straightforward, their implementation frequently runs into significant practical difficulties. These include, but are not limited to, the proper determination of boundaries and time horizons, issues of data scarcity or reliability, monetization of costs and benefits, selecting a discount rate, dealing with risk and uncertainty and accounting for distributional impacts and equity, to name a few (Atkinson & Mourato, 2008; Pearce, 2003).

Countering its practical difficulties there are numerous advantages to CBA as an evaluation tool, mostly stemming from its objectivity and, arguably, its fairness. Based on Pearce et al. (2006) those advantages can be summarized as:

- the fact that the preferences of the individuals are what counts;
- its comprehensive character, attained through the mandate to include all measurable gains and losses in wellbeing and thus staying clear from particular goals or biases;
- it is inherently and exclusively useful as a comparison tool which uses monetary values as a common metric, so the presence of alternatives is a prerequisite;
- CBA allows for a distributional analysis of gains and losses, although that is rarely done in the practice;
- it has the possibility of assessing a point of net maximum benefit, something other evaluation techniques such as Cost-Effectiveness Analysis or MCA cannot do; and
- time variations in value are included through discounting.

Further to these general features, CBA application to environmental issues has specificities that warrant caution, regarding: a) individual preferences, b) gains and losses, c) a common metric and d) discounting.

a) *Individual preferences:* in simple terms, economic theory measures one individual's preference for a bundle of goods according to the economic value that such individual place on that bundle, expressed in a willingness to sacrifice resources (i.e. to pay). It is assumed that an individual behaves rationally in her/his choices as if s/he is to maximize her/his own welfare, independently of other individuals or society. Since many of the goods and services that would be accounted for in a CBA as applied to environmental policies (e.g. cleaner air, reduced premature mortality, etc.) do not have a market, their value must be inferred indirectly. The two main methodological approaches for doing so can be categorized into "revealed preferences" and "stated preferences" methods. Revealed preferences methods assume that the preferences of individuals can be inferred from their purchase habits, or more generally, from their market transactions and behavior regarding resources and money (Varian, 2006). The use of revealed preferences in environmental practice is limited by the absence of markets for several of the items evaluated and because "non-use" values cannot be captured by this approach. Stated preferences methods in theory cover that limitation through surveys and questionnaires that elicit preferences by establishing hypothetical markets (Atkinson & Mourato, 2008). The most widely used of these methods are the contingent valuation methods (CVM), which create hypothetical (i.e. contingent) markets where non-marketed goods (e.g. clean air) can be traded. Respondents then express the value they assign to the good by their maximum willingness to pay (WTP) or minimum willingness to accept (WTA) compensation for a change in the amount of that good or service. CVM have been extensively used for environmental valuation (Carson, 2000) and empirical findings seem to support the validity of its estimates. However, skepticism is unlikely to fade anytime soon, since CVM are basically stretching the results of a survey to how we make decisions for a market that does not exist. One main argument in this direction is the hypothetical bias, whereby hypothetical markets are unlikely to elicit real WTP and valuation of a good is commonly overstated in a hypothetical context (Murphy et al., 2005). Also within the stated preferences methods are choice experiments (CE) where pools of characteristics of goods are used to rank or otherwise categorize alternatives according to the respondents' preferences. Each alternative is characterized by a certain set of values assigned to pre-defined attributes. The clear advantage of CE is its ability to reflect situations where the perspectives coming into play for a given decision are multidimensional (Bateman et al., 2002), but it also entails clear psychological complexities (multidimensional, multilevel choices) as well as analytical ones, for instance a much heavier statistical analysis. The extrapolation of individual preferences in the valuation of environmental goods and services requires special caution. For instance, CVM used in environmental valuation often encounter "protest bids" in the form of extreme values of zero or infinity. By refusing to make trade-offs between environmental factors and money, respondents explicitly refuse to treat environmental factors as market goods. People make different choices as citizens and as consumers (Sagoff, 1994). Indeed, some authors argue that neoclassical welfare economics treats human preferences as they ought to be to fit its axioms, rather than trying to fully capture them as they are. A solid body of research supports the thesis of "endogenous preferences" dependent on the social context. Relevant examples are the "endowment effect", "process-regarding" preferences, and "other-regarding" social preferences. The "endowment effect" refers to the observed psychological response regarding the value of losses and or wins. Basically, people require more money to be paid to them to lose something they consider theirs than what they would pay to acquire the same item; in other words, their willingness to accept is larger than their willingness to pay, or losses are systematically valued more than gains, which is counterintuitive to the rational choice model (Kahneman, 2003). Also contrary to the rational choice preference model is the fact that the process through which an outcome is achieved may be as important for the acceptance of the actors involved as the outcome itself ("process-regarding" preferences). Similarly, experiments in game theory show that people act to affect others' welfare positively or negatively, even at their own expense, i.e. "other-regarding" preferences (Fehr & Gachter, 2000). These socially conditioned preferences are directly at odds with the individual preferences that are the foundation of neoclassical welfare economics (Brekke & Howarth, 2000).

b) *Gains and losses (societal welfare and the Kaldor-Hicks principle):* in a social CBA, options that provide the largest net gain in societal welfare are preferred. The optimal situation for achieving that gain would be what economists call a Pareto optimality, whereby a given option would be desirable if no one suffers a welfare loss as a result and at least one individual benefits from it. In a change that is optimal in the sense of Pareto, there are no losers and at least one winner. However, in reality that net gain applies to society as a whole, thus concealing a pool of winners and losers. Who determines whether that specific situation is optimal? In the context of CBA this issue is addressed through the main theorems of welfare economics (on which CBA is supported). The first theorem states that if all individuals are selfish price takers, then a competitive equilibrium is Pareto-optimal. The second theorem states that if the former is true, then almost every Pareto-optimal equilibrium can be achieved through a competitive mechanism, provided the adequate transfers between individuals (and firms) (Gowdy, 2004). One fundamental implication of the second theorem is that, if a particular state of the economy is deemed desirable, it can be achieved through those "transfers". If the benefits achieved through the considered change are larger than the costs incurred, then this change implies a net welfare gain. This type of change is a "Potential Pareto Improvement" (PPI) and it is, according to some authors (Stavins et al., 2003) "the fundamental foundation […] for employing cost–benefit analysis". It is important to note that this PPI is radically different from a pure Pareto optimality. In a PPI there are winners and losers. In the context of our interest (the evaluation of environmental policies) the purpose of the application of CBA is to decide whether a policy or a set thereof improves societal welfare overall or not. That inherently implies that there needs to be a decision rule whereby, in light of available information, such hypothesis can be tested. Whereas the Pareto principle is an obvious choice, there are few real situations where a substantial change is likely to yield only benefits and no harm to anyone. The Kaldor-Hicks compensation principle implies that if the gains from moving to one state of the economy to another are greater than the losses, then the move improves social welfare, even if the real compensation from winner to losers does not take place. Because such compensation is theoretical, a Kaldor-Hicks efficient move can (and usually does) leave some people worse off, something impossible under a theoretical Pareto efficiency. Furthermore, since a real compensation is not happening, there is no need to incur in interpersonal comparisons of utility/welfare, a much debated and controversial possibility in welfare economics (Pearce et al., 2006).

c) *Common metric (monetization):* one of the main advantages of CBA is that it allows comparing hugely different alternatives, since it translates all costs and benefits into a common metric, namely monetary value. Therefore, any given intervention can be compared with other in their Benefit-to-Cost Ratio, for instance. However, in practical terms, for any given intervention there might be a number of impacts that cannot be readily quantified in monetary terms unless unreasonable assumptions are made. For instance, the diffuse effects of a specific policy on social cohesion, public opinion, or equity or social justice can be difficult to quantify, let alone set against a monetary scale. Often times, monetization implies adding a compounding layer of uncertainty on top of indicators that are already indirect estimates. Moreover, for those non-marketed impacts for which money values can be established through various techniques (e.g. hedonic prices, stated preferences, etc.) there might not be available data, or those could be too difficult or expensive to collect. In those cases, the alternatives would often imply extrapolations from different settings or countries (such as value-transfer), which would add an inherent error to the valuation of the impacts.

d) *Discounting:* it refers to the assignation of a lower importance to a cost or benefit unit in the future than now: the more distant in the future, the lower the value. However, whereas this concept might be almost universally accepted for private purposes and finance, its usefulness for social investments is much debated. The issue is of particular importance for environmental and resource economics; when dealing with environmental goods and services, typically moral dilemmas arise regarding the fairness of the use of discounting for future generations. Applying high discount rates would directly be incompatible with any notion of sustainable development (Broome, 1992) while not discounting would by comparison impoverish current generations (Olson & Bailey, 1981). A number of alternatives have been proposed to deal with this dichotomy, which essentially amount to replacing a constant discount rate with some sort of time-declining one (Portney & Weyant, 1999). Far from a mere technicality, the choice of a time-declining discount rate may have deep effects in CBA results as applied to long term environmental issues such as climate change (Stern, 2007). Also related to preferences and with great relevance for environmental goods is the notion that discounting is not necessarily linear (that is, benefits in the future being discounted at a fixed rate). People tend to discount the near future at higher rates than the distant future (hyperbolic discounting) and sometimes even sometimes a higher value is given to something in the future than in the present (Gowdy, 2004). That has profound implications for environmental matters and particularly for the intergenerational equity issues that environmental policy-making frequently involves. The issue of discounting in environmental policy and proposed strategies has been summarized by Chiabai et al. (2012), who suggest the equivalency principle as a rule to identify the discount rate to use in the CBA. Given its conceptual framework and characteristics, CBA has advantages and disadvantages in its application for the evaluation of environmental policies. Its limitations and criticism (explored later in this chapter, subsection 2.2.3) suggest a balanced, rather than exclusive, role of CBA as a policy evaluation tool.

2.2.2. Application of CBA in air quality management

CBA is the most common evaluation method applied prospectively for environmental policy, including air quality management, where it has been applied extensively to appraise interventions. So much so that a comprehensive review is far beyond the scope and purpose of this dissertation. Voorhees et al. (2001) provided a short overview of the history and state of knowledge in this specific field of application of CBA. The focus of most studies featured in the mentioned review is the pollutants with the strongest adverse effects on human health, and notably particulate matter.

After the passing of the Clean Air Act in 1970, the United States Environmental Protection Agency (USEPA) has both produced and contracted the most visible and recognizable body of evidence in the application of CBA to air quality control. USEPA is required by Executive Order 12866 to estimate the benefits of major new pollution control regulations. Therefore, it is not surprising that most standards for costing and benefit estimation in the field are developed by this agency. Such standards are explained in detail in numerous reports (USEPA, 1995, 1997).

Other national governments have issued their own guidelines for the application of CBA to their air quality policies (DEFRA, 2004; Bin Jalaludin et al., 2009). One major effort in CBA of air quality policies has been recently undertaken by the European Union through its Clean Air for Europe Program, in which cost and benefit valuation of major air quality control strategies was a fundamental part (AEAT, 2005 and 2007). Table 2 summarizes selected CBA exercises for regulatory assessments.

Some patterns arise from these few, but comprehensive, examples. The three main valuation techniques used in air quality CBA are 1) market valuation of physical effects where possible (including human capital approach and cost of illness among others), 2) contingent valuation (i.e. stated preferences) and 3) revealed preferences based on surrogate prices (e.g. travel costs, hedonic salaries, aversive behaviors).

In addition, CBA is used both to prioritize alternatives and to evaluate benefits from a single alternative. Benefits of air pollution reduction consistently and strongly outweigh costs (although some specific metrics, like the value of a statistical life, significantly contribute to those large differences). The largest benefits in most CBAs of this type come from health benefits, time savings and energy savings, although there are also small gains derived from improved visibility and other "environmental services" types of improvements. Lastly, since typically health and other benefits are accrued over several years, annualized metrics are highly relevant, and so are considerations such as discount rates and uncertainty.

Table 2. Examples of regulatory CBA for major interventions to reduce air pollution

Study	Period	Area	Pollutants	Type of analysis	Results	Prioritized actions
Voorhees et al. (2000)	1994	Tokyo (Japan)	NO_x	Retrospective costs and benefits of air pollution control	Average benefit-to-cost ratio 6:1 (0.3:1 – 44:1)	1. Stationary NO_x sources 2. All NO_x sources 3. Motor vehicles
Austin et al. (1997)	Late 1990s	Virginia (USA)	NO_x	Prospective costs and benefits of air pollution control	Cost-effectiveness of pollution control (USD 1000/ton NO_x controlled) (2.7 – 5.6)	N/A
Blackman et al. (2000)	1999	Ciudad Juarez (MX) and El Paso (USA)	PM_{10}	Retrospective/Prospective costs and benefits of air pollution control from informal brick kilns	Annual net benefits of interventions from approx. 1 Million to 60 Millions 1999 USD	1. Use of natural gas 2. Improved kilns 3. Relocation 4. No-burn days
Eliasson (2009)	2005 - 2006	Stockholm (Sweden)	Pollutants from mobile sources	Retrospective analysis of costs and benefits of congestion charging	The congestion charges produce a net social benefit around EUR 70 Million/year)	N/A
USEPA (2000)	2000 - 2030	USA	PM and Ozone	Prospective analysis of benefits of the HD Engine/Diesel Fuel Rule	Monetary benefits range from USD 16 Billion to USD 165 Billion (1999 USD)	N/A
USEPA (2010a)	2005 - 2020	USA	SO_2	Review of the National Ambient Air Quality Standards for Sulfur Dioxide	Monetary estimates in Billions of 2001 USD Benefits (2.9 – 38.6) Costs (0.3 – 2.0)	N/A
USEPA (2010b)	2013	USA	$PM_{2.5}$	Regulatory Impact Analysis (RIA) for Existing Stationary Compression Ignition Engines	Monetary estimates in Millions of 2008 USD Benefits (940 – 2300) Cost (373)	N/A
EU (2005)	2005 - 2020	EU 25	All major pollutants	Evaluation of costs and benefits of the "Clean Air for Europe" program	Average EU 25 benefit-to-cost-ratio MTFR*: 4.6:1 St. C: 10.7:1 St. B: 13.7:1 St. A: 20.3:1	MTFR entails highest benefits but also high costs.

*MTFR: Maximum Technically Feasible Reduction

2.2.3. Limitations of CBA in the evaluation of environmental policies

There is obvious value in the use of CBA techniques for the evaluation of environmental policy. The widespread use of these techniques attests to their merits in supporting transparent and accountable decision-making (Atkinson & Mourato, 2008; Pearce et al., 2006; Voorhees et al., 2001). However, their shortcomings and limitations are the frequent target of criticism on political, philosophic, economic and technical (practical) grounds. These include:

a) the measurement of societal welfare (Kaldor-Hicks principle);
b) individual preferences and social decision-making;
c) equity and distribution of impacts;
d) monetization of non-marketed goods;
e) discounting; and
f) uncertainty.

Such criticisms of CBA are succinctly explored below, with a focus on environmental policy:

a) *The Kaldor-Hicks compensation principle and deriving social welfare from individual preferences*. Many theoretical criticisms have been made to the Kaldor-Hicks criterion (a cornerstone of current practice in social CBA application) mainly on the grounds of resulting income distribution. Scitovsky (1941) showed that if a movement from utility point A to B was PPI, then moving back from B to A was also Pareto improving (a counterintuitive idea in the practice). Boadway (1974) demonstrated that policies or options with the largest net gain may not be the optimal, since income redistribution could result in changes in relative pricing, thus rendering PPI estimates of potential welfare gains incorrect. Bergson (1938) noted that a "social welfare function" (SWF) that could "adequately aggregate individual changes in welfare" would circumvent these problems. However, Arrow (1963) demonstrated that there is no method for deriving social preferences from individual preferences; that is, we cannot construct a SWF based on individual preferences without breaking some axioms that would make any voting system intrinsically fair.

b) *The individual preferences in social decision-making.* It is not clear whether individual preferences should guide social decision-making. Some authors argue that such preferences might not reflect societally desirable states, a fundamental consideration in environmental issues, where the individual awareness of the importance of ecosystems and environmental quality for welfare and survival may be poor. Others argue that the intrinsic rights of species others than humans are not commonly considered in human preferences, thus neglecting their value. Importantly, the model of preferences based on individual rational choice is much debated, even the existence of "exogenous preferences" independent of other individuals or society. Gowdy (2004) argues that most of the instruments that welfare economists use to elicit preferences are biased to make individuals' choices conform to the economists' assumptions.

c) *Equity and distribution of impacts*. A conventional cost–benefit analysis has traditionally not concerned itself with how costs and benefits are distributed, a matter of especial importance for cases of environmental impacts because degradation costs tend to disproportionately fall on disadvantaged groups. This has been often cited as one fundamental criticism to the technique, which in purity cannot say much about the social desirability of a given distribution of benefits and costs (Pearce et al., 2006; Randall, 1999). While conceptually sound, this criticism stands little scrutiny from the practical perspective of policy evaluation. For starters, CBA is just one input out of several possible in the process of policy-making, where equity or distributional considerations could be included through separate channels (Atkinson & Mourato, 2008). Furthermore, recognizing the potential value of including distributional considerations in the CBA framework, several authors have suggested ways to include equity considerations in which now constitutes the foundation of modern CBA theory (Little & Mirrlees, 1974; Squire & Van der Tak, 1974). The simplest way of including such considerations would be to document the distribution of impacts as far as possible during the CBA process and leave them to be integrated during the rest of the policy-making process. Another way would be to make the final decision based on a rule that takes distributional weights into account, be it implicitly or explicitly (Pearce, 2003). Conventional CBA assigns an equal weight to every unit of benefit regardless of who receives it. However, depending on society's distributional goals, an asymmetrical distribution of weights might be more appropriate. This is highly relevant for projects or actions aiming at improving environmental quality since environmental impacts tend to fall disproportionately on disadvantaged segments of the population.

d) *Monetization of non-marketed goods.* Some authors (and presumably several unpublished practitioners) believe that many items routinely included in CBAs cannot be truly monetized. One of the most important implications of these issues with monetization for CBA as applied to environmental policy evaluation is the valuation of health risks. Evidence consistently shows that the benefits of reducing health effects associated with environmental risks (including air pollution) exceeds by ample margins the cost of the policies that could provide those benefits (Cifuentes et al., 2005; Hutton, 2000; Kinney & Nori-Sarma, 2011). Whereas the benefits of avoiding morbidity (I.e. nonfatal health outcomes) can be substantial (and rather relevant, since they commonly entail solid healthcare usage savings), the highest and most studied willingness to pay value refers to mortality risk. WTP for reductions in mortality risk is used to estimate the value of a statistical life (VSL), which does not refer to any given life, but it is a statistical construct summarizing risk of premature death. Unsurprisingly, VSL depends on the considered risks, the age of respondents, income and many other context specific variables (Krupnick, 2007; Ortiz et al., 2009). Valuation of environmental health risks in children is particularly controversial; it is assumed that they cannot properly state their preferences (OECD, 2006) since their understanding of trade-offs is incomplete. In addition, evidence for VSL studies outside Europe and North America is still rather limited (Atkinson & Mourato, 2008). So whereas the application of this type of monetization is widely prevalent, the inherent uncertainties are large, including extrapolations between countries, settings, and ages. These error-compounding extrapolations build on top of the inherent uncertainties of contingent valuation and the valuation of non-marketed goods in general.

e) *Discounting.* The concept of discounting was coined and developed theoretically for financial purposes. Hence, the application of a pure discount rate barely fits the kind of goods and services dealt with in the evaluation of environmental issues. Because time scales of environmental issues are often much longer than those of financial investments, strictly applied discounting of, say, climate change in CBA would make the problem all but disappear from present concern (Atkinson & Mourato, 2008). This problem is known in the field as "the tyranny of discounting" and has been the focus of countless arguments in the field of welfare economics. Several types of specially designed discounting schemes for environmental policies have been applied and discussed, whereas others remain convinced that a pure rate of time preference best reflects people's actual behavior (Nordhaus, 2007). This is indeed a complex matter. While relevant for the focus of this study (any CBA needs to make a choice regarding discounting), the conceptual ramifications of discounting in environmental CBA have been explored elsewhere (Pearce et al., 2003) and will not be repeated here.

f) *Uncertainty and interactions.* Uncertainty is indeed an issue that CBA has to deal with, but is by no means exclusive to this technique. It has to do with the fact that certain benefits or costs do not have a known probability distribution, something which most risks do have (albeit frequently rather crude). This, in itself is not as much a criticism to CBA as a reminder that such considerations must be taken into account in the analysis (Pearce et al., 2006). Lastly, a typical shortcoming of CBA as applied to environmental issues is its inability to capture the effect of interactions between different impacts. This has been noted in the case of ecosystems services (Arrow et al., 2000) but it can easily be extrapolated to the case of air quality control policies. Certain combination of impacts that could cause strong opposition by stakeholders might not necessarily translate into an unfavorable ratio of economic benefits to costs. CBA as a tool is simply not designed to account for such synergistic effects.

2.3. Multi-criteria analysis and environmental policy

2.3.1. Fundamentals of multi-criteria analysis

While cost–benefit analysis is the most widely used technique to assess economic costs and benefits of environmental policies, multi-criteria analysis (MCA) might well be the "most commonly used technique to compare unvalued costs and benefits" (HMT, 2011). A MCA is a set of techniques designed to establish preferences between options with regard to an explicit set of objectives that the decision-making body or the stakeholders have identified. For each one of those objectives, measurable criteria are established so that the performance of different options can be compared (CLG, 2009).

As a tool for decision-making, MCA is a means of simplifying and structuring complex problems that can involve many stakeholders, diverse outcomes and several criteria by which to assess these outcomes, of which some may or may not be translated into economic terms (Garmendia et al., 2010). Due to its shorter history and the very nature of the tool, MCA is far less standardized than CBA, resulting in a wide range of variation in its design and implementation. However, despite its flexibility MCA is not an informal judgment. Objectives and criteria need to be explicit and transparent, and can be cross-referenced or changed if felt inadequate by the stakeholders. Moreover, the performance of different options is commonly assessed by parties external to the decision-makers, so as to remove subjectivity. In essence, the use of MCA in environmental policy stems from the frequency with which monetary evidence cannot provide an accurate picture of the costs and benefits entailed by a given policy (DEFRA, 2011). This section reviews the framework, process and examples of application of MCA to environmental issues.

Originally designed as a mathematical aide for decision-making (based on a well-defined problem worked through an algorithm), multi-criteria decision techniques were mostly a form of optimization up until the 1970s (Munda, 2006). Key contributions like the incorporation of the concept of creative approach, that is, assigning importance to the decision process itself rather than only to the outcome (Roy, 1985) and the incorporation of participatory mechanisms (Banville et al. 1998) contributed to its evolution into decision-support tools useful for policy-making.

Lessons from the application of the technique in several sectors (e.g. Transport, Urban planning, Water resource management, etc.) have contributed to the wealth of MCA varieties that are available today. A generic process of MCA might take a structure like the one described below (CLG, 2009).

1. Identifying objectives, which can be categorized in several ways, but are often divided loosely as immediate and ultimate. Desirable characteristics of the objectives to be evaluated are Specificity, Measurability, Attainability, Relevance and Time-boundedness, as stressed by the commonly used acronym SMART first used in corporate environment about three decades ago (Doran, 1981).

2. Identifying options for achieving the objectives: after the objectives are defined, options that can contribute to the achievement of these objectives need to be identified. These options can range in scope, from broad policy action to specific ground-level projects.
3. Identifying the criteria to be used to compare the options considered for each objective: options are compared regarding their performance in meeting the immediate and ultimate objectives. To allow such comparison, key criteria need to be selected. That selection can be based on expert judgment, a consultation with stakeholder or other procedures, but in any case criteria must be measurable in order to assess how each option is likely to perform regarding each particular criterion.
4. Analysis of the options, based on the comparison of the measures of performance of each option regarding the selected criteria, either taken separately or jointly. Multiple techniques, either based on monetary evaluation or otherwise, can be used. In any case, a performance matrix of options versus criteria will most usually be part of the analysis.
5. Making choices: the actual decision on which option/s will be chosen relies on, but is necessarily somewhat independent of, the previous analysis. The obvious reason is that a formal analysis, whether based or not on economic tools, cannot possibly incorporate all judgments, sensitivities or fine distributional points. Decisions will typically be taken by high officials or ministers based on evidence and, inevitably, political considerations.
6. Feedback: appropriate decision-making requires post-hoc reassessments of past decisions. This formal process of institutional learning is especially important to avoid repeating mistakes and spending public resources in strategies of unlikely success.

This framework and process might apply to the development of a policy, a program or a project. Although the specific shape and components of an MCA might vary widely, the key concepts of a typical MCA are often the same (see Table 3).

Table 3. Key concepts in a multi-criteria analysis process

Key concepts	Brief description
Objectives	The effect/s that the policy, plan or project intends to achieve
Alternatives	Possibilities to achieve the required objectives; these options will be ranked, rated or otherwise assessed
Criteria	Principles or standards that the alternatives are judged by. Typically defined by a set of dimensions to be measured.
Indicators	Variable or component used to infer the status of a particular criterion
Weights	Numerical values assigned to criteria when assessing alternatives, reflecting the strength of preference or relative importance of each.
Performance matrix	A table in which each row describes an option and each column describes the performance of the options against each criterion.

Source: CLG (2009) and CIFOR (1999)

The establishment of objectives and criteria can be done in multiple ways, depending on the goals and scope of the object of evaluation. In any case, clarity in those objectives and criteria, as well as in the relative importance of each is vital to the usefulness of MCA as a decision-support technique. Economic criteria such as those derived from a CBA or a CEA can be included in a MCA which in turn is open to combine non-monetary criteria for a comprehensive approach to evidence-based decision-making.

In theory, a MCA can be applied to processes or outcomes with an infinite amount of alternatives. However, when applied to decision-making by public sector agencies, typically the alternatives are few. For these finite alternatives, a common appraisal method includes performance matrices, that is, a finite number of alternatives are ranked or rated against a predefined set of criteria. A performance matrix is built by assessing the performance of any given option considered through a key indicator for each criterion. Data in the performance matrix may be summarized into numerical values through specific functions based on the weighting and scoring of the criteria by stakeholders. In that way, stakeholders' preferences are incorporated into the evaluation. The analysis that follows largely defines the type of MCA - a brief overview of the main methods is provided in the next subsection. After the analysis is complete, the analyst/s needs to present the results to the end users of the system, the decision-makers. Often, the performance matrix itself is presented, as it can illustrate trade-offs between criteria in each alternative. Most commonly, both the performance matrix and the ranked list of alternatives are the visible end products of the MCA for the decision-makers. However, MCA results can be presented in several ways, some of which are also explored in the next subsection.

2.3.2. MCA models

Several researchers (e.g. Brans, Vincke, & Mareschal, 1986; Goicoechea et al., 1982; Pomerol & Romero, 2000; Roy, 1985; Saaty, 2001) have proposed methodologies to represent a real world multi-criteria decision system into a manageable numerical or conceptual model. With many decades of developments and growth, the multiplicity of available multi-criteria decision methodologies has led to the need of formal classifications and full studies on the suitability of different methods for different problems or areas. The majority of MCA methods belong to one of the following categories (Velasquez & Hester, 2013): 1) Multi Attribute Utility Theory and Linear Additive Models; 2) Analytical Hierarchy Processes; and 3) Outranking methods. There are, of course, several MCA methodologies that do not belong fully into these three categories (a notable example are the MCA methods based on fuzzy sets) but they are a minority and their use marginal compared with the ones that belong. Since the focus on this study is not exclusively on MCA methods per se, only these three mainstream categories are considered.

Albeit formally and practically quite different, the three mentioned categories of MCA methods are based on early and standard economic theory assumptions about rational choice (Von Neumann & Morgenstern, 1947) under which an individual, presented with a set of options with different expected utilities and having to choose one, should normally choose the option with the maximum subjective expected utility (SEU) value. In this process, the decision maker is thus considering a limited set of options, along with a set of future "states of the world" of which the probability of occurring can just be guessed, and then obtaining an expected utility for a given option. This underlying logic underpins all MCA methods mentioned and, although it has its critics, it is accepted in this dissertation as a given. Crucially, however, there is no hard and clear way to evaluate the utilities mentioned, so each methodology tries to deal with this problem in different ways. In order to choose a methodology that is suited to the case study at hand (i.e. the prioritization of air quality management policies in an urban setting) this subsection explores the basics of the main categories of MCA methodologies:

- Multi attribute utility theory (MAUT) models
- Analytic hierarchy process models
- Outranking methods

Multi attribute utility theory (MAUT) models

Keeny and Raiffa (1976) published the first comprehensive study on how to formalize the quantification of the utilities associated with options in the context of future states of the world. Basically, the authors set out to assign numerical values to different options in a MCA choice problem with *m* alternatives and *n* attributes. Once these are established, the preferences of the decision maker are modelled through subsequent tests intended to determine the minimum necessary difference in any given attribute that makes one choice preferable to other. Those preferences essentially determine the difference in utility with regard to one attribute between two choices. Subsequently, the relative weights of all considered attributes are determined.

Given a set of mutually utility independent attributes, Keeny and Raiffa (1976) specify the conditions under which the utility of the consequences of a decision adopts either a multiplicative or an additive form. Let x_{ij} be the outcome for attribute *j* in alternative *i*, $u_j(x_{ij})$ be the utility of outcome x_{ij} for attribute *j*, with values between 0 and 1, and k_j be the relative weight of attribute *j*. The mathematical formalization of the multiplicative form is:

$$U(x_i) = \frac{\prod_{j=1}^{n}\left[K \times k_j \times u_j\left(x_{ij}\right) + 1\right] - 1}{K} \qquad (1)$$

$$1 + K = \prod_{j=1}^{n}\left(1 + K \times k_j\right) \qquad (2)$$

Where $\sum_{j=1}^{n} k_j \neq 1.0$ and *K* is a scaling constant, which must be found iteratively. The multiplicative form allows for preference interaction among attributes, which can be either preference substitutes or preference complements. If such interactions do not exist, then the problem can be simplified into a linear additive form. The interaction in preferences among attributes is described by K: if it is positive, all attributes are preference complements; if it is negative, the attributes are all preference substitutes.

The additive form is given by:

$$U(x_i) = \sum_{j=1}^{n} k_j \, u_j\left(x_{ij}\right) \qquad (3)$$

With $\sum_{j=1}^{n} k_j = 1.0$. The linear additive form allows for no preference interactions among attributes, so the change in utility caused by a variation in one attribute does not depend on whether there are any variations in other attributes. All attributes must be mutually utility independent. Thus essentially, when the right conditions are met, this model requires only weights and scores to come up with overall scores that are consistent with the axioms of rational choice.

Analytic hierarchy process models

Originally developed by Thomas L. Saaty in the 1970s as a non-prescriptive MCA technique, the Analytic Hierarchy Process (AHP) was intended as a way to provide decision-makers with insights as to the structure of the problem and trade-offs involved. To describe the problem as a hierarchy with an overall goal, alternatives to achieve that goal, and the criteria to measure how far each alternative goes towards achieving the goal.

While more complex hierarchies can be built, the simplest form is best for illustrative purposes. To establish priorities between elements of the hierarchy by means of pairwise comparisons, where elements in the same level of the hierarchy are compared in pairs with regard to one criterion or attribute. Since such comparisons are completely subjective, a suitable psychometric scale can be introduced (Saaty, 1980). Hence, by comparing each alternative with the rest, one at a time, we can obtain a numerical *n* by *n* comparison matrix, where *n* is the number of criteria evaluated.

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{k1} & \cdots & a_{nn} \end{pmatrix} \quad (4)$$

Where $a_{jj} = 1$ and $a_{jj^*} = 1/a_{j^*j}$. The next step is to derive the scale of priorities (or weights). It has been shown that this scale is obtained by solving for the principal eigenvector w of the matrix and then normalizing the result (Saaty, 1987). Then, options are compared with one another in a pairwise comparison for each criterion, resulting in n matrices B_j of dimensions m by m, where m is the number of alternatives. Each entry of B_j represents the performance of option i compared with option k with respect to criterion j.

The same procedure applied to matrix A is applied in order to obtain the weights of each option for each criterion. A global matrix of dimensions m by n of option scores S is constructed with these weights. Each column gives the priorities of all options for a given attribute. To rank the alternatives the weight vector w is multiplied by the score matrix S, thus obtaining a global scores vector v where the i^{th} entry of v represents the global score assigned by the AHP to the i^{th} option. Note that this step is in essence a linear additive model.

Outranking methods

Both multi attribute utility models (including the linear additive ones) and analytical hierarchy process models ultimately rely on the concept of a value function that multiplies weights by performance across criteria. Due to the axioms underlying the model, the preference relation built by this function is supposed to be complete and transitive. In turn, that requires detailed information on all attributes and a deep analysis of the trade-offs between them. For this reason, this model met with frequent issues in its practical application; the "outranking" family of methods was developed to overcome practical difficulties of the "value function" approach. Let us consider an example of a multi-criteria problem, with a set I of potential alternatives i, to be evaluated according to a set of n criteria g_j. Let us take two alternatives, i and i', to be compared; $g_j(i)$ and $g_j(i')$ are numbers representing the evaluation according to criterion j of the alternatives i and i', respectively; and let us further suppose a natural dominance relation of the type:

$$\begin{aligned} &\begin{cases} \forall j: g_j(i) \geq g_j(i') \\ \exists j^*: g_{j^*}(i) > g_{j^*}(i') \end{cases} && \Leftrightarrow && iPi' \\ &\forall j: g_j(i) = g_j(i') && \Leftrightarrow && iIi' \\ &\begin{cases} \exists j^*: g_{j^*}(i) > g_{j^*}(i') \\ \exists j^{**}: g_{j^{**}}(i) < g_{j^{**}}(i') \end{cases} && \Leftrightarrow && iRi' \end{aligned} \quad (5)$$

Where P stands for "preferred to", I for "indifferent from" and R for "incomparable to".

If one option is better than other in all criteria, then it is preferred. If it is equal in all criteria, then it is indifferent. But if it is better in some criteria and worse in others, then we cannot decide which is preferable with the available information. As they stand, the options are incomparable. This is one key issue that outranking methods try to address.

Like MAUT and AHP models, OMs build a preference relation (also called "outranking" relation) among alternatives evaluated on several attributes or criteria (Roy, 1985). Such relation is built through pairwise comparisons. Going back to our example, a new "outranking relationship" S can be then defined i and i' (iSi') if i is at least as good as i', while there are no strong reasons to refute that statement. In other words, i is at least as good as i' by outperforming it on enough criteria of enough importance (as per weighted criteria addition) and not underperforming significantly against it in any given criterion. This logic is summarized in the concordance-discordance principle, whereby i is at least as good as i' if: 1) a majority of the attributes (criteria) support this assertion, and b) the opposition of the other –minority- attributes is not too strong. These conditions are obviously less stringent than those of MAUT and AHP models, and result in binary relations that are not necessarily either complete or transitive. It can, however, work in problems with low availability of information. In addition, OMs tend to downgrade options that perform badly on any one criterion (which could be politically sensitive). This regard for specific aspects and the "voting" flavor of it can make it more appropriate for some specific policy evaluation areas. Beyond generalities, the model must define what the dominance of one alternative versus another means in the practice within one criterion. Since criteria can take multiple forms (continuous, binary, ordinal, qualitative, etc.) and their evaluation may be uncertain, the overall principle "i is at least as good as i'" might have an unclear application for one given criterion. Moreover, let us consider a case where two options rank only slightly differently in one criterion; mathematically, even the slightest difference would simply render the better performing preferable. However, in the mind of the decision-maker these two alternatives may be indifferent. This leads to the need to establish an indifference threshold (Roy, 1991) mathematically:

$$iI_j i' \quad \text{iff} \quad g_j(i) - g_j(i') \leq q_j \tag{6}$$

Where the real positive number q_j defines the decision-maker's indifference threshold. Therefore, to each criterion, g_j, it is possible to associate a restricted outranking relation S_j. By definition, S_j is a binary relation: iS_ji' holds if the values of the performances $g_j(i')$ and $g_j(i)$ give a sufficiently strong argument for considering the following statement as being true in a decision-maker's model of preferences: i with respect to the j^{th} criterion only, is at least as good as i'.

$$i\, S_j i' \text{iff} \;\; g_j(i) \geq g_j(i') - q_j \tag{7}$$

Given the complete set F of criteria in the problem, if for all $j \in F$, i S_j i', then "i is at least as good as i' with respect the n criteria" and it is denoted by iSi'. That is, if iSi' is met, then it is a criterion in concordance with the overall assertion iSi'. It belongs to the "concordant coalition" denoted by $C(iSi')$. Any criterion j that is in discordance with this assertion means that option i' is strictly preferred to i with respect to criterion j, or $i'P_ji$ ("discordant coalition" $C(i'Pi)$).

The strict preference restricted to the j^{th} criterion of i over i' is given by the condition:

$$iP_j i' \text{ iff } g_j(i) - g_j(i') > p_j \tag{8}$$

Where p_j is the preference threshold. Lastly, there can be a criterion which is neither concordant nor discordant with the assertion iSi', mathematically:

$$g_j(i') - p_j \leq g_j(i) < g_j(i') - q_j \tag{9}$$

The subset of criteria satisfying condition (9) is denoted by $C(i'Qi)$. The model is completed with the inter-criteria information: the importance coefficients k_j and the veto thresholds v_j, for each $j = 1, \ldots, n$. In this method, in Roy's words, "the importance of the j^{th} criterion is taken into account by means of, at most, two characteristics: 1) its importance coefficient k_j (≥ 0) which is intrinsic: the k_j's only intervene in the definition of the concordance degree (see below); and 2) its veto threshold v_j ($\geq p_j$) which only intervenes in the definition of the discordance degree of criterion g_j (see below). By definition, the concordance index $c(i,i')$ characterizes the strength of the positive arguments able to validate the assertion iSi'. The strongest among them come from the criteria of $C(iSi')$ since they are all in favor of the assertion considered. Some weaker positive arguments can also come from criteria of $C(i'Qi)$ since such criteria reflect a hesitation between the two following possibilities: $iI_j i'$ (which is in favor of iSi') and $i'P_j i$ (which is not in favor of iSi')" (Roy, 1991, p60)." Based on these concepts, a "concordance index" $c(i,i')$ can be defined based on the criteria in the concordant coalition $C(iSi')$ and those criteria neither concordant nor discordant $C(i'Qi)$:

$$c(i,i') = c_1(i,i') + c_2(i,i') \tag{10}$$

$$c_1(i,i') = \frac{1}{\sum_{j\in F} k_j} \sum_{j\in C(iSi')} k_j \tag{11}$$

$$c_2(i,i') = \frac{1}{\sum_{j\in F} k_j} \sum_{j\in C(i'Qi)} \varphi_j k_j \quad \text{with} \quad \varphi_j = \frac{p_j + g_j(i) - g_j(i')}{p_j - q_j} \tag{12}$$

Where the following conditions are met:

$$\begin{gathered} 0 \leq c(i,i') \leq 1 \\ c(i,i') = 0 \text{ if } C(i'Pi) = F \\ c(i,i') = 1 \text{ if } C(iSi') = F \end{gathered} \tag{13}$$

The ratio $k_j/\Sigma k_j$ reflects the relative strength in F assigned to each g_j when this criterion is in the concordant coalition, or in Roy's words, "k_j can be viewed as the number of representatives supporting the point of view synthesized by the j^{th} criterion in a voting procedure". So if j is in the concordant coalition, all the k_j representatives contribute to $c(i,i')$; if j is in the discordant coalition, this contribution is null; and if it is in the "neither/nor" category, it contributes only a fraction. It is worth noting that this methodology assigns exceptional power to any discordant criterion, a principle which leads Roy to define a "veto effect". To reflect the power given to a discordant criterion, a veto threshold $v_j > 0$ is defined, so that if:

$$g_j(i') - g_j(i) > v_j \tag{14}$$

Then the assertion iSi' does not hold, regardless of the other performances. For "fuzzy" outranking relationships (highly uncertain, in which a "credibility index" is attached to each comparison) a "discordant index" is defined to modulate from 0 to 1 the opposition to the assertion iSi'.

Strengths and weaknesses of different MCA models

The main strengths and weaknesses of the main types of MCA models have been analyzed by several authors. Below is a summary based on Moffett & Sarkar, (2006) and Linkov & Steevens (2008).

Multi attribute utility theory advantages:

- It allows for a clear comparison between alternatives through the difference in value between any two alternatives to be quantified.
- Theoretically sound, ranking based exclusively on utilitarian considerations.
- No assumptions are made with respect to the form of the single criterion value functions.
- Because it uses value functions, its results are independent of the number of alternatives. Therefore, unlike the AHP, the method is not susceptible to rank reversal.

Multi attribute utility theory disadvantages:

- It requires the assumption that the performance of the alternatives with respect to each of the criteria can be evaluated on the basis of a common scale. This is a strong assumption and in many decision scenarios it will prove to be unfounded. As such, the applicability of this method is somewhat limited.
- The decision maker must be able to assign a scaling constant to each criterion. This requirement can be difficult to fulfill and serves also to increase the subjectivity of the method.
- The requirement that a numerical value be assigned to each performance evaluation limits the applicability of this method to those decision scenarios in which the performances of the alternatives can be readily quantified.
- The construction of the aggregate value function can be quite challenging.

Analytical hierarchy process advantages:

- It allows for the difference in value between any two alternatives to be quantified.
- The methodology by which the AHP assigns weights to criteria is simpler than that used by most of the other MCD methods that rely upon an aggregate value function.
- It does not assume the complete transitivity of the decision maker's preferences. A certain degree of inconsistency is allowed, which in most decision scenarios is realistic.
- The pairwise comparison methodology of the AHP is very intuitive and easy for most decision-makers to understand.

Analytical hierarchy process disadvantages:

- Requires that comparisons between both alternatives and criteria can be quantified.
- Possibility of rank reversal, whereby the introduction of another non-optimal option can change the overall ranking, including the best ranked option.
- Compensation: due to its additive component, good scores in some criteria can compensate for bad scores in others, depriving the end user of potentially important information on trade-offs.
- With a large number of options and/or criteria, the pairwise comparison may become a very lengthy task.
- The use of the AHP requires the assumption that the performance of the alternatives with respect to each of the criteria can be evaluated on the basis of a common ratio scale.

Outranking methods advantages:

- Does not require the reduction of all criteria to a single unit.
- Does not require the assumption that the set of criteria under consideration are commensurable.
- Guaranteed to produce an ordering of the alternatives.
- Clear visualization of trade-offs between alternatives and relative performance in criteria.

Outranking methods disadvantages:

- Some OMs require the calculation of a value function for each criterion. As such, the method assumes substantially more of a decision maker than do most other methods.
- The algorithms used may be difficult to understand for the decision maker.
- Requires the attribution of quantitative weights to the criteria; however, it does not provide a clear method by which to assign these weights.

These advantages and disadvantages must be taken into account when selecting a specific method to conduct a MCA – see introduction of Chapter 5 (MCA).

2.3.3. Application of MCA in air quality management

MCA has been used extensively for local planning and decision-making, with special relevance in areas where spatial considerations are of importance, like transportation (Macharis et al., 2009) or post-disaster urban reconstruction (Opricovic & Tzeng, 2002). This increasing use of MCA in local settings probably has to do with the alignment and fitting of MCA processes to the characteristics of local urban decision-making, mentioned in section 1.1. However, the instances in which MCA has been used regarding environmental issues in a context of local urban decision-making are few and far apart, and almost exclusively from developed countries. Moreover, studies featuring the application of MCA to air quality management policies are still rather scarce. The most relevant published references are briefly summarized hereafter.

In a mainly "theoretical" type of prioritization, Pisoni, Carnevale and Volta (2009) applied a mathematical approach, proposing PM_{10} reduction strategies for the region of Lombardy (Italy) through a multi-objective optimization. The researchers established all parameters, including control objectives (i.e. the final emissions levels desired), decision variables and problem constraints. The main result is a set of maps depicting the efficient trade-off between the air quality benefit and the internal costs (emission reduction technology costs). Other studies from this research group (Carnevale et al., 2008; Guariso et al., 2004) utilize similar mathematical optimization techniques and produce similar results.

In 2003, the United Kingdom Department of the Environment, Food and Rural Affairs (DEFRA) explored in depth the issue of the utilization of multi-criteria analysis in the appraisal of air quality policies. It did so through two different processes involving separate theoretical frameworks, but basically relying on the inputs from experts designated by key stakeholders. The team of analysts at DEFRA put an emphasis on the simplification of both the process and the outcomes of the MCA, utilizing pairwise comparisons (the simplest approach) and performance matrices as the main output. This maximized the usefulness of the analysis for decision-makers and stakeholders (DEFRA, 2003).

In two studies by Tzeng and collaborators, MCA techniques are applied to environmental planning and policy prioritization in Taipei, Taiwan. In the first study (Tzeng et al., 2002) priorities regarding environmental issues are derived from a survey among residents. The responses to the survey indicated that the main concern for the residents was air pollution. Air pollution was then considered as the problem for which optimal solutions had to be found. The researchers obtained a list of policy options through a panel of experts. Subsequently, a group of stakeholders put together by the local authorities reached a consensus on a list of criteria and their relative importance. With this information, a second expert panel evaluated the policy alternatives against the chosen criteria. For the prioritization itself, the researchers used straightforward techniques: weighted sums with Analytical Hierarchy Process (AHP). The ranking was done through two systems (VIKOR and TOPSIS) providing slightly different results that were combined for presentation in performance matrices and a final ranked list of options. The second study (Tzeng et al., 2005) is different mainly on the topic of evaluation (alternative fuels for the bus fleet in Taipei) and in which priorities are not ascertained from the public. Moreover, the researchers did not rely on committees of stakeholders or experts, but rather established themselves the alternatives (based on state of the science at the time of the study), the criteria and the weights. Finally, they prioritized the options with a system similar to the one used in the previous study. More modern and exclusively focused on the topic at hand (MCA applied to air quality policy prioritization) are recent studies carried out in Thessaloniki, Greece, by a combined research group of local universities. The study by Achillas et al. (2011), like that by Tzeng et al. (2002), takes on the issue with a bottom-up approach, ascertaining the public's environmental concerns through face-to-face interviews designed after a preliminary scoping phase. Subsequently, the research team enunciates a set of policy alternatives and requests a group of experts and key stakeholders to propose criteria and assign relative weights. The final ranking of alternatives is obtained through weighted sum and binary outranking, and results are presented along with a sensitivity analysis.

In another study (Vlachokostas et al., 2011) the same research group takes specifically on the prioritization of policies to improve air quality in the greater Thessaloniki area. Alternatives and criteria are proposed by a panel comprising experts, local authorities and other relevant stakeholders. Like in the study of Achillas et al. (2011), the ranking of alternatives is obtained through weighted sum and binary outranking, and results (a ranked list of alternatives) are presented along with a sensitivity analysis. This small pool of studies, summarized in Table 4, is characterized by the diversity in the ascertainment of criteria, techniques for the prioritization and presentation of the results. Their overall approach, ranges from the purely theoretical to a full involvement of the stakeholders. In general, the research team plays a large role regarding the ascertainment of criteria, even though it is often recognized that the criteria should ideally be set by the end users of the system. Managing the prioritization without stakeholder participation is regarded as a convenience alternative, accounting for the practical difficulties of participatory processes. Weighted sum and pairwise comparisons using Analytical Hierarchy Processes are the most common tools, although alternative and combined approaches are also used. Lastly, simplicity in the presentation of results is regarded as an adequate strategy to gain relevance through legibility, which translates into a preference for performance matrices and ranked lists as displays of prioritization results. These characteristics are not fundamentally different from those derived from the application of MCA to other fields like transport (Macharis et al., 2009). The importance of the process in MCA, the choice of AHP for alternative analysis and an emphasis in the involvement of stakeholders in the generation of criteria and weight are also common features in the application of MCA to other urban decision-making problems.

Table 4. Main characteristics of studies on the application of MCA to air quality policies

Case	Location	Objective	Criteria setting	Weighting	Analysis details	Result
Pisoni et al. (2009)	Lombardy, Italy	Evaluate alternate strategies for PM_{10} emissions reduction	Researchers (expert judgment)	Researchers (computer modelling)	Multi-objective optimization, dominance analysis through Sequential Quadratic Programming	Pareto boundary maps reflecting cost, PM_{10} levels and relative reductions
Guariso et al. (2004)	Lombardy, Italy	Evaluate alternate strategies for ground level ozone concentration control	Researchers (expert judgment)	Researchers (computer modelling)	Multi-objective optimization, dominance analysis through conjugate gradient method	Emission abatement strategy cost curves
Carnevale et al. (2008)	Milano, Italy	Evaluate alternate mesoscale strategies for ground level ozone concentration control	Researchers (expert judgment)	Researchers (computer modelling)	Multi-objective optimization plus weighted sum strategy	Pareto boundary maps reflecting cost and relative reductions of Ozone and precursors
DEFRA (2003)	United Kingdom	Evaluate policy options to reduce NO_x and SO_x emissions	Workshop with key stakeholders	Workshop with stakeholders (scale 0 -100)	MCDA with qualitative weighting of criteria (Macbeth software) and pairwise comparisons	Performance matrices; marginal benefit/cost curves with weighted preference values
Tzeng et al. (2002)	Taipei, Taiwan	A. Obtain a prioritized list of environmental concerns B. Evaluate options to improve air quality	A. Survey of a sample of residents B.- Researchers (expert judgment)	Expert panel (scale 0 - 1)	A. MCDA with weighted sum B. Ranking through Analytical Hierarchy Process with VIKOR system	A.- Performance matrix B.- Ranked list of options
Tzeng, Lin & Opricovic (2005)	Taipei, Taiwan	Evaluate alternative fuels for public bus fleets	Researchers (expert judgment)	Expert panel	MCDA with weighted sum, ranking through AHP with VIKOR and TOPSIS systems	Performance matrix plus ranked list of options
Achillas et al. (2011)	Thessaloniki, Greece	Evaluate alternatives to confront urban deterioration	Survey of local residents, face-to-face interviews and expert judgment	Expert panel (scale 1 - 10)	MCDA with weighted sum, ranking through binary outranking with ELECTRE III	Performance matrix plus ranked list of options
Vlachokostas et al. (2011)	Thessaloniki, Greece	Selection of best strategies to improve air quality in the metropolitan area	Expert panel	Expert panel (scale 1 - 10)	MCDA with weighted sum, binary outranking with ELECTRE III and sensitivity analysis of options	Performance matrix plus ranked list of options

2.3.4. Limitations of MCA in the evaluation of environmental policies

MCA has been widely subjected to scrutiny and critique. Several reviews have listed pitfalls and shortcomings of MCA. Some of them are listed by Dobes & Bennett (2009) as applied to the prospective evaluation of government policies:

- Perspective/standing: the very selection of stakeholders can add a critical bias to the determination of criteria, scores and weights. The representativeness of this group of individuals with regard to their society, as well as the presence of vested interests can make MCA less fair about the standing of different groups and perspectives.
- Inclusion of impacts and criteria: there is no overriding principle for the inclusion of impacts and criteria, so an MCA evaluation is inherently partial and incomplete.
- Alternative policies: multi-criteria analysis is incapable of comparisons between unrelated programs, because their impacts or attributes are so different.
- Dimensionality: by breaching the barrier of the addition or product of different dimensions, MCA results become logically flawed.

One key limitation, not listed by these researchers, is that multi-criteria analysis is not designed to or capable of measuring welfare changes. Therefore, the Potential Pareto Improvement rules cannot be applied; a policy ranked best by MCA does not necessarily accrue more benefits than costs as a whole, so it could be inconsistent with the best-ranked alternative by CBA (CLG, 2009). This entails a crucial philosophical implication: a MCA-preferred option might be in theory openly inconsistent with society's best interest.

A further consequence of this dichotomy has to do with the possibility of compensation and more deeply, with the issue of comparability. This discussion is often synthesized in the concept of weak comparability (Martinez-Alier et al., 1998) which in very simple and loose terms implies that indifference/preference cannot be properly established between two alternatives because they would rank differently in different, non-comparable criteria. Rather than a shortcoming, some authors point at the weak comparability as a further argument for the application of MCA to environmental issues (Munda, 2004, 2006). Since the simultaneous optimization of different criteria is impossible, MCA offers the possibility of adequate compromise solutions.

Perhaps the most often cited limitation of MCA is its purported subjectivity. Rather than a degeneration or misapplication of the technique, subjectivity is built in its core. In the most common forms of MCA, decision criteria are ranked in importance through subjective weights (although sometimes objective weighting is possible). Some analysts and decision-makers feel that the introduction of such subjectivity taints the analysis with arbitrariness, thus diminishing its value as a decision-support tool (Belton & Stewart, 2001; Henig & Buchanan, 1996). To navigate around this apparent pitfall, sometimes analysts identify Pareto-optimal solutions within the MCA. Such solutions, however, tend to be multiple. Therefore, a subjective choice from between them is required in the end. A core component of subjectivity is thus inherent to the use of MCA, a matter frowned upon in operational research and used by some (e.g. Dobes & Bennett, 2009) to downplay the value of the technique in contrast with the perceived rationality of CBA.

There are, however, a rising number of analysts and researchers that advocate for the notion that modeling should accommodate human decision-making rather than the opposite, and that MCA is a good starting point (Kersten & Noronha, 1996; Olson, 2006; Wenstop, 2005). They argue that the values and beliefs of the decision maker play a key role in final decisions, regardless of the theoretical objectivity of decision aides. Unlike other tools, MCA can consider values explicitly, whether aligned or competing, and offers solid methods to handle conflicts among objectives (Wenstop, 2005). Moreover, some researchers use ample scientific evidence to argue that emotions play a vital role in reliable and effective decision-making (Follesdal, 1982, 2004; Wenstop & Seip, 2001). Outside the field of operational research and economics, much has been written about rationality and the fitting of current support tools to human mind and decision-making. However, for the purposes of this study we shall only retain the notion that subjectivity is not necessarily something to avoid in decision-support modeling. Rather, it should be acknowledged and managed in ways that are reasonable, replicable and fair.

2.4. Combination of cost–benefit and multi-criteria Analyses

Prospective systematic evaluation should always be part of the decision-making process in policies regarding urban sustainability, and very especially in far-reaching environmental issues like air pollution. CBA and MCA are the best suited comprehensive analysis models for such evaluation. Moreover, far from replacing each other, they are suited to be complementary, thus providing a more complete portrait of the evidence relevant for decision-making. Drawing from different sources cited in Chapters 1 and 2, their main comparative merits in a practical application for policy evaluation are summarized as:

- Comparative advantages of CBA
    - Comprehensive accounting of costs and benefits
    - Less prone to double counting
    - Theoretically it includes societal preferences
    - Unified metric
    - Reflects time preferences
    - Clear solution
    - Theoretically free of bias
- Comparative advantages of MCA
    - More intuitive for the end-user
    - The process involves stakeholders
    - Transparency in criteria
    - Flexible to include considerations that do not admit monetization
    - Can be combined with other techniques
    - It offers tools for compromise solutions
    - Includes non-monetizable items
    - Trade-offs between criteria are explicit

The question remains, how to draw out the best characteristics of each type of analysis for one decision problem? Although the concept is hardly new (Manheim et al., 1975) and has been analyzed periodically in connection with infrastructure and transport policies, the literature on actual operational approaches is scarce.

An obvious alternative is to conduct both types of analysis simultaneously and keep the results separate, as distinct inputs for the decision-making process. Such approach has been used in the evaluation of sustainable forestry (EFORWOOD, 2011) and for evaluation purposes by DEFRA (2003), through the simultaneous presentation of physical impacts from MCA assessment and monetized impacts from the CBA by criterion in an "Appraisal Summary Table". The main potential problem with this approach, aside from what the definition of "integration" may be, is dealing with two sets of potentially conflicting prioritizations, as well as the unacknowledged overlap and double counting of the analytical results. Another approach has been to use the tools in different stages of the policymaking process, most often with MCA used for screening of alternatives and CBA for pre-implementation analysis of the chosen alternatives. One inconvenience of such approach, though, is its liability to consume far more time that what's available in the typical policy cycle, in which simultaneous processes of prospective appraisal are the norm. Moreover, the set of alternatives may be constrained or limited from the beginning, so screening would not be necessary. Regardless of the technical complexities of combining both tools, the set of options for the combination is basically limited to:

a) Combining the outputs of CBA and MCA and reporting a unified metric;
b) Expanding or modifying a CBA with MCA attributes (i.e. non-monetizable items);
c) Integrating CBA components into a MCA framework; and
d) Creating a hybrid system.

Each option has advantages and disadvantages, and none of them is backed by a very large body of scientific literature. Below is a brief overview of each option, illustrated by published examples where available.

2.4.1. Combining the outputs of CBA and MCA

In this approach, conceptually separate CBA and MCA analyses would provide specific outputs that would then be combined into one integrated result. Thus, the informative value of both the CBA and MCA outputs are kept, and added value is provided through the combination of both types of outputs. Moreover, the relative weight of each type of output in the final evaluation can be decided by the stakeholders by assigning a weight coefficient. The decision-support research group at the Technical University of Denmark (DTU) proposed a composite decision-support system based on the idea that MCA criteria evaluations can be added to the CBA outputs through a value function computed using a weighting procedure for the criteria (Barfod et al., 2011; Salling et al., 2007). Neither additive value functions (explored extensively in their own right) or the idea of applying them to decision-making are novel; the latter has been featured frequently in management and decision-making research (Keelin, 1981; Keeney & Raiffa, 1993). As applied to MCA, the linear additive model is used to combine many criteria into one overall value, by multiplying the value score on each criterion by the weight of that criterion, and then adding the weighted scores together. However, the idea of adding non-monetary MCA criteria to monetized CBA results in a relatively direct way is the novelty proposed by this group. This additive decision-support system is called Composite Model for Assessment (COSIMA) and it consists of a CBA part and a MCA part.

As applied to the assessment of a project alternative *i* the outputs of both analyses are combined into a Total Value (TV) for that given alternative:

$$TV_i = CBA_i + MCA_i \quad (15)$$

This addition cannot be done directly; the units of one of the terms (either CBA or MCA) have to be translated into the unit of the other. Once it is done, an alternative that is not profitable by CBA standards might still be preferable overall on account of its MCA ranking, depending on the relative weight of each term. In other words, the total rate of return (TRR) for a given alternative *i* would be:

$$TRR_i = \frac{TV_i}{C_i} = \frac{1}{C_i}\left[\sum_{k=1}^{K} V_k(X_{ik}) + \alpha\left[\sum_{j=1}^{J} w_j VF_j(Y_{ji})\right]\right] \tag{16}$$

Where:

C_i are the total costs or expenses of alternative i

X_{ik} is the quantity of the CBA impact k for alternative i

$V_k(X_{ik})$ is the value in monetary units for the CBA impact k for alternative i

α is an indicator that expresses the balance between the CBA and MCA parts in the model

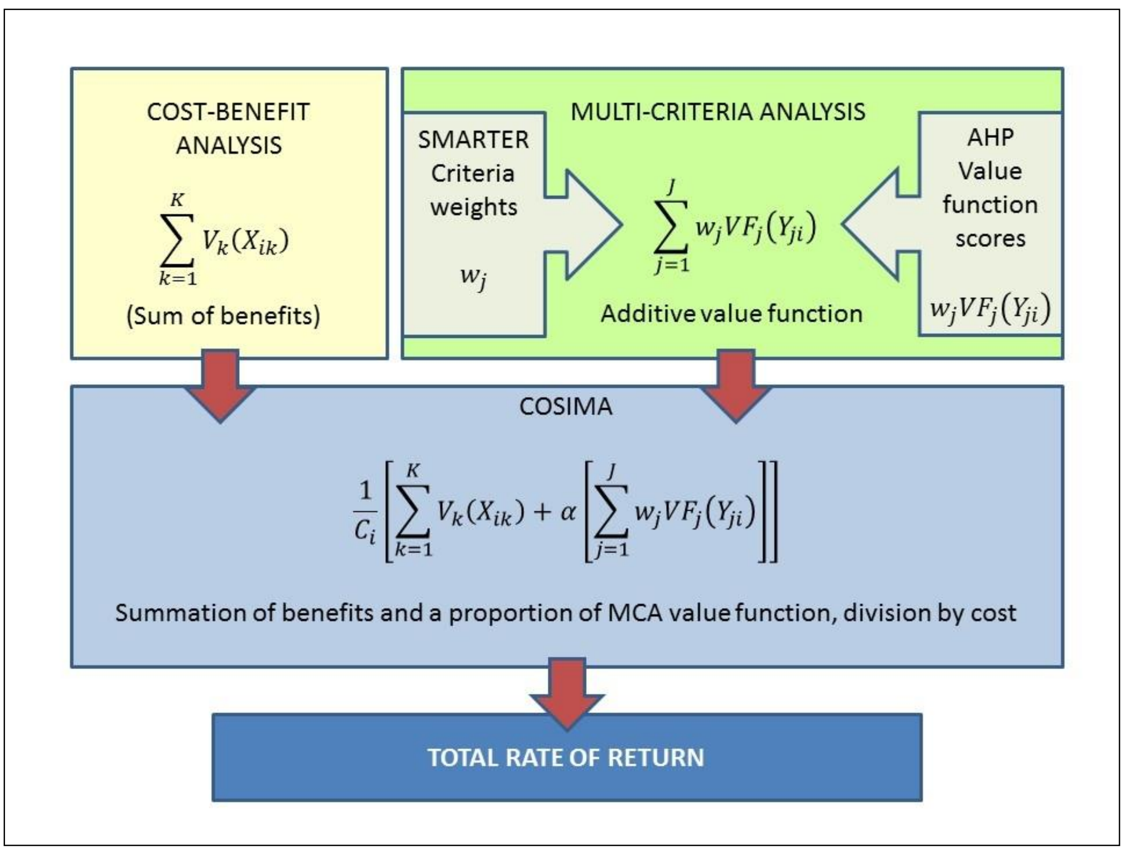
w_j is a weight which reflects the importance of MCA criterion j

Y_{ji} is a parameter value for MCA criterion j for alternative i

$VF_j(Y_{ji})$ is a value-function score for MCA criterion j for alternative i

Thus, an alternative with a high BCR could be weighed down by its MCA performance, depending on the value of the parameter α. It is noteworthy thus the capital importance of that "balance indicator" α and the "value function scores" $VF_j(Y_{ji})$ – see process in Figure 5.

Figure 5. COSIMA process for combining CBA and MCA outputs



Source: adapted from Barfod et al., (2011) Besides the combination of the two analyses, both the CBA and MCA in the COSIMA approach are standard. In their case study, the CBA was run as a typical public project investment one. The scoring for alternatives in the MCA evaluation component was based on the Analytical Hierarchy Process (as described in Saaty (2001)) and weighting of criteria is based on Ranked Order Distribution (ROD) weights (Roberts & Goodwin, 2002), a type of surrogate weighting conceived as a way to circumvent the difficulties of eliciting weights from stakeholders or decision-makers. The COSIMA approach has only been applied to the evaluation of alternatives in transportation planning in Denmark, but as a framework it has in theory the potential to be applied to the appraisal of alternatives in sustainability practices and policies, specifically in the prioritization of alternatives for air quality management in urban areas.

2.4.2. Expanding or modifying a CBA with MCA attributes

The extension of CBA results by adding non-monetized effects or criteria organized through MCA has been proposed as an optimized combination, benefiting from the objective perspective of CBA and the inclusiveness of MCA (OECD, 2008; Pearce et al., 2006). Some EU country governments issued guidance in the mid-1990s on how to use (but not combine) both jointly at the strategic level, either giving more preponderance to MCA (e.g. Belgium and the Netherlands) or a balanced contribution of both techniques (e.g. France, Germany, Italy, UK) (EC, 1996). The combination of monetary and non-monetary criteria was proposed formally by the UK government in Transportation planning (DETR, 1998a). In the practice (DETR, 1998b) it is unclear whether the core of the analysis would be that of an MCA or a CBA; the proposed system utilizes five criteria (Environmental, Safety, Economy, Accessibility, Integration) but translates results into three scales for analysis (Monetary value, Non-monetized quantitative units and qualitative assessment scale). The results can be combined in various fashions, depending on the analytical and presentation needs, but the defining characteristics of CBA and MCA are difficult to ascertain in the summary outputs.

In the context of the "External Costs of Energy" (ExternE) project (http://www.externe.info), Diakoulaki & Grafakos (2004) proposed a method to monetize certain MCA criteria that did not admit ready monetization, so that they could be included in a CBA. The policy scenarios evaluated consisted or strategies to reduce emissions of certain pollutants (specifically NO_x, NH_3, SO_2 and VOCs) in Europe. The strategies were evaluated on their prospective physical impacts for acidification, eutrophication, mortality, and changes in agricultural production. Of these criteria, acidification and eutrophication (two forms of ecosystem degradation) do not admit ready monetization, so the researchers propose that since the elicited weights in a MCA are meant to represent the preferences of stakeholders, these can in turn be used for an indirect monetization of environmental goods. The monetization is done by relating the weight calculated for each physical impact to the weight of the "cost" (i.e. the cost of abating emissions of the pollutants) criterion, such that:

$$\frac{w_i}{IR_i} = \frac{w_c}{IR_c} \tag{17}$$

Where:

w_i represents the normalized weight of the criterion representing the physical impact i

w_c represents the normalized weight of the criterion representing the cost

IR_i represents the impact range (i.e. the maximum potential gain) of the physical impact i, and

IR_c represents the impact range (i.e. the maximum potential gain) of the cost criterion

The impact ranges are obtained by calculating the potential gains achievable through the maximum feasible reduction of pollutant emissions resulting from the full application of currently available control technologies and subtracting the minimum projected performance of each criterion. Based on this relation, the per unit monetary value of the physical impact i would be calculated as:

$$m_i = \frac{IR_c \times w_i}{IR_i \times w_c} \tag{18}$$

This "indirect monetization" would create a new set of values that would be added to the benefits tally in the CBA, thus increasing in various degrees the BCR of the different policy alternatives.

2.4.3. Integrating CBA components into a MCA framework

Given the flexibility and inclusive nature of MCA, one obvious possibility is to integrate the CBA results into the MCA framework. Indeed, this approach is explicitly advocated for in authoritative guides on the matter (CLG, 2009). The European Union's EUNET methodology for transportation appraisal (EUNET/SASI, 2001) treats the CBA results as one of the criteria within the MCA, by applying scores to the investment criterion (e.g. the Benefit-Cost ratio). A similar approach has been used in the UK, again in the realm of transportation planning (Sayers et al., 2003). A relatively recent paper (Gühnemann et al., 2012) illustrates this alternative for the evaluation of a road infrastructure improvement investment scheme in Ireland. The researchers included the CBA results into the MCA framework through a scoring function. Thus, given an alternative *i*, its value for money could be incorporated into the MCA through a score of the form:

$$Score_{ji} = 4 + 3\left(\frac{PV_{ji} - PV_j^{Do\ minimum}/PVC_i}{\theta\alpha_j}\right) \tag{19}$$

Where:
$Score_{ji}$ is the score for each sub-criterion j included in the CBA;
PV_{ji} is the present value of sub-criterion j in the alternative i;
PV_j Do minimum is the present value of sub-criterion j in the baseline situation;
PVC_i is the present value of the cost of the intervention proposed;
θ is a benefit-cost ratio threshold, and
a_j is the average proportion that sub-criterion j contributes to an scheme's present value of benefits

Through this function, CBA sub-criteria are scored like the rest of criteria (from 1.0 to 7.0 in this specific case) and thus included fully into the MCA analysis. Scores below 1.0 are rounded to 1.0 and scores above 7.0 are rounded to 7.0. The reason for the unusual form of the function (4 + 3(f)) lies in which it had to fit the appraisal scale of the National Roads Authority of Ireland, a requirement specific to the case study and with no particular conceptual basis in economics. That scale ranks each sub-criterion considered in the evaluation of an investment from 1 to 7, in which a score of 4 is considered neutral, whereas the value 1 is highly negative and 7 highly positive (NRA, 2008). The scoring function is linear between 1.0 and 7.0, symmetrical around 4.0 (neutral score) and truncated at 1.0 and 7.0. The benefit by sub-criterion is, according to the authors, "normalized" through division by the present value of costs to prevent the scores from being biased in favor of larger projects of which both benefits and costs are higher – thus increasing the net benefit term:

$$\left(PV_{ji} \; - PV_j^{Do\ minimum}\right) \tag{20}$$

This “normalization” is in effect the means through which the CBA information is integrated into the MCA, through a Benefit-to-cost ratio. The term $\theta\alpha_j$ whereby a threshold is established that permits the translation of Benefit-to-cost ratios into scores. This threshold term defines the level of impact that is considered “highly positive” (i.e. it is allocated a score of 7.0). It is calculated as the product of a “highly positive” benefit cost ratio (BCR) and an expected proportion that j would contribute to total benefits. In this application θ was set to 2.5, a value that the authors say is considered a strong BCR by the Irish National Roads Authority, or NRA, again with no specific economic basis to it. The authors do not explain why the NRA adopted 2.5 as its BCR threshold instead of, say, 1.1. Moreover, the NRA’s own published project appraisal guidelines (NRA, 2011) do not mention any particular BCR threshold, so the choice of value can only be speculated about. In turn α_j was determined by analyzing the contribution of each of the sub-criteria to total benefits. Since θ is fixed to one specific value, the term $\theta\alpha_j$ is proportional to the median contribution of each sub-criterion to the benefits of the option appraised. Effectively, what is being done is to translate monetized criteria into units suitable for an MCA, and both the threshold term and the shape of the function are intended to make the score fit into the pre-set scale.

In the example featured in the study, the authors combine monetized impacts (air and climate, noise, transport efficiency, wider economic impacts and accident reduction) along with non-monetized impacts (environment, accessibility, safety and integration) to come up with an overall MCA appraisal of each project between 1.0 and 7.0. The “economic” component in which a BCR of 2.5 is deemed worthy is then translated into an overall MCA “investment worthiness” threshold through the relationship between costs and θ. The resulting cutoff threshold beyond which a project is assumed to provide net value (including value for money) is 5.2. Thus, only projects beyond this threshold are considered for further evaluation, with greater value as their score approaches the top of the scale (7.0). The results of the prioritization of projects itself are of limited interest for this dissertation.

2.4.4. Creating a hybrid system

Lastly, the only other combined approach that could be found in this review was to integrate both tools into a hybrid type. The concept, mentioned in a few policy and general documents (Migliavacca, 2011; Moylan, 2009; Papaemmanouil & Andersson, 2010) is usually vaguely described and not as yet standardized. A dissertation presented at the University of Groningen (Sijtsma, 2006) attempted to standardize steps for a “multi-criteria cost–benefit analysis”, or MCCBA. This approach has been advocated periodically by the same researcher in subsequent articles and studies.

The proposed steps of the MCCBA are: 1) Identify function, project alternatives and scale of the evaluation; 2) Involve a broad group of stakeholders; 3) Organize judgment criteria on Triple E impacts; 4) Quantify impacts physically; 5) Aggregate monetary scores; 6) Aggregate non-monetary scores; 7) Interpret trade-offs; and 8) Perform sensitivity analysis and reconsider project alternatives.

These steps are reminiscent of both a standard CBA and a standard MCA and indeed their implementation is not different from the same steps in either tool.

The main differences concern:

- The "Triple E impacts": the MCCBA approach is in theory designed for evaluations concerning environmental impacts, and the author proposes to disaggregate impacts so that three metrics can be used. Health impacts are measured in the standard Disability-Adjusted Life Years (DALYs), monetary impacts are measured through a net present value (NPV) and "Nature's wellbeing" is measured in "Threat weighted Ecological Quality Area" units.
- Aggregate monetary scores consensus based: this stage is based on a "limited CBA" in which the stakeholders determine which monetizable impacts are included. Other than this, the CBA used is standard, and its output a discounted NPV.
- Aggregate non-monetary scores consensus based: this stage is a MCA in which the analyst guides the discussion of the stakeholders into a minimum set of criteria fitting the "Triple E impacts" metrics.

There is no further integration of MCA and CBA in the MCCBA approach. According to the author, "The best end result of the MCCBA is therefore an aggregation of the performance matrix, based on a broad consensus among stakeholders. This means that a MCCBA usually has more than one final criterion. To briefly illustrate this possibly vague formulation, a MCCBA can consist of, for example, the following criteria: (i) a net present value, which aggregates all the impacts included in the "limited" CBA; (ii) health impacts measured in DALYs; and (iii) biodiversity impacts measured in a composite nature indicator.

Essentially then, in the MCCBA, the CBA outputs and the MCA outputs are presented separately, with the additional step of conversion into the "Triple E impact" units. It is thus difficult to argue a conceptual difference between this approach and the one-analysis core (CBA or MCA) complemented with the other. For that reason, this approach is excluded from application to the case study.

2.4.5. Implications of CBA-MCA combination literature for case study

This dissertation focuses on the combination of CBA and MCA for decision support, and the trade-offs, weaknesses and strengths of each way in which this integration may happen. All of the combination alternatives found in the literature have strengths and weaknesses, although some are conceptually stronger than others. Issues to be assessed in each combination option include: whether the combination represents a comprehensive accounting of the variables in the policy problem; the possibility of double-counting of impacts; the replicability of the procedure; and the balance between societal and stakeholder preferences, among other considerations.

Part of the evaluation of the strengths and weaknesses of each combination mode can be done a priori. However, only the application of these methodologies to the same case study would allow a fair comparative assessment of strengths and weaknesses. The logical next step, then, would be to apply all possible combination CBA-MCA alternatives (except the "hybrid" approach, see previous subsection) to the case study of policies to improve air quality in Lima-Callao. Prior to that combination, it is necessary to: conduct an impact assessment (Chapter 3); CBA (Chapter 4); and a MCA of the actions proposed (Chapter 5).

Chapter 3. Case study policy impact assessment

3.1. Introduction

The majority of regulatory cost–benefit analysis for the evaluation of air pollution abatement utilizes for the economic valuation of benefits of regulatory action the same conceptual model (the impact pathway model) or variations thereof. Such is the case of the major governmental and international agencies that regularly engage in these exercises (Holland et al., 2005; Bin Jalaludin et al., 2009; USEPA, 2000, 2004, 2010a). The model describes the pathway of the pollutant in discrete steps; based on Holland (2014) these steps can be summarized as:

- Evaluation of emissions (PM, NO_x, SO_x, etc.);
- Evaluation of dispersion and the pollutants' atmospheric chemistry;
- Assessment of exposure (people, buildings, crops, etc.);
- Impact assessment (mortality, morbidity, building damage, reduced crop yields, etc.); and
- Economic evaluation of impacts.

Since the objectives of this dissertation do not include innovation in the field of air quality impact assessment, there is no reason to use any other conceptual model. The specific elements of the impact pathway model as applied to the present case study are largely determined by the nature of the problem evaluated. Similarly, each one of the steps is conducted in accordance with commonly used and accepted methodologies.

Prior to the impact modelling, it is important to note the limitations under which such modelling was undertaken. The most important challenge stemmed from the low level of definition of each one of the policies. Aside from the scarce information in the plan documents themselves (CGIAL, 2011b), little more could be obtained from the institutions and stakeholders involved, even from implementing agencies. This vagueness in policy is not uncommon in developing country settings, and it adds to the challenges facing the analyst aiming at supporting decision-making with evidence. Another important caveat is the fact that the prospective fleet and emissions modelling was entirely conducted by an external partner, the Clean Air institute for Latin American Cities (CAI-LAC). The technicalities of this type of modelling are not in my field of expertise, and so its results were taken at face value for the purposes of this dissertation. For context, however, it should be mentioned that this think tank has successfully conducted analysis and advised policy to governments, international organizations and development banks for air quality improvement in Latin America and the Caribbean for over two decades. Nevertheless, all these factors meant that we had to make a number of assumptions in the modelling for these impact assessments. All assumptions regarding fleet trends, retirement rates and emissions factors were made by CAI-LAC. Thereafter, in the health impact assessment and subsequent economic evaluation, the necessary assumptions were made by me, based on the available information and on previous studies.

3.2. Methods

3.2.1. Emissions scenarios

The four actions considered in the policy program to reduce particulate matter were briefly presented in the case study introduction in subsection 1.2.2. They can be summarized as follows:

- Action A1: Private vehicle fleet renewal (used vehicles import ban and additional measures);
- Action A2: Bonus for scrapping older public transportation vehicles;
- Action A3: Promotion of Vehicular Natural Gas (VNG) in public transportation vehicles; and
- Action A4: Enforcement of better vehicle maintenance.

All these actions involve changes or improvements in the local vehicle fleet. Therefore, the first step in the assessment of the effects of such policies required an inventory of the circulating vehicle fleet in the MALC. The main indicators of interest were:

- The types of vehicle circulating, with a break-up by class.
- The fuel used by those vehicles, including gasoline, diesel, Compressed Natural Gas (CNG) and VNG.
- Age of the vehicles, which aside from the technology used represents an average status of maintenance. Each combination of fuel and type of vehicle, along with age, results in the application of one specific set of emissions factors.
- Kilometers traveled on average by type of vehicle; each type of vehicle is assigned an average occupancy and travel length by year, in order to complete a scenario of emissions produced by each individual characterization of vehicle.

The transportation authority of Lima and Callao (PROTRANSPORTE) provided such indicators to the Clean Air Institute, which then calculated estimates of vehicle-km travelled by type of vehicle, and retirement rates per type of vehicle. These estimates, combined with emissions factors by vehicle technology, allowed to calculate emissions of particulate matter attributable to mobile sources. The model, the specifics of which are beyond the scope of the present dissertation, was run ad hoc for the study, but it is available upon request.

Once a baseline was obtained, the Clean Air Institute modelled the effect of the proposed actions on the vehicle fleet of the MALC, and consequently on the emissions by mobile sources in the area and resulting concentrations. Based on those results, I elaborated the illustrative graphs for subsections 3.2.1.1 through 3.2.1.4, and then conducted a health impact assessment (later in this chapter) and the subsequent economic evaluation (Chapter 4). The proposed actions and the methods for the estimation of their consequences in terms of particulate matter impacts are succinctly explained in the following subsection.

3.2.1.1. Action A1 (private vehicle fleet renewal)

The declared goal of this measure is for the private vehicle fleet of the MALC to become newer. Specifically, the goal is to reduce the average age of circulating vehicles in the MALC from 17 years to 10 years by the end of the program.

In order to do so, the measures considered are:

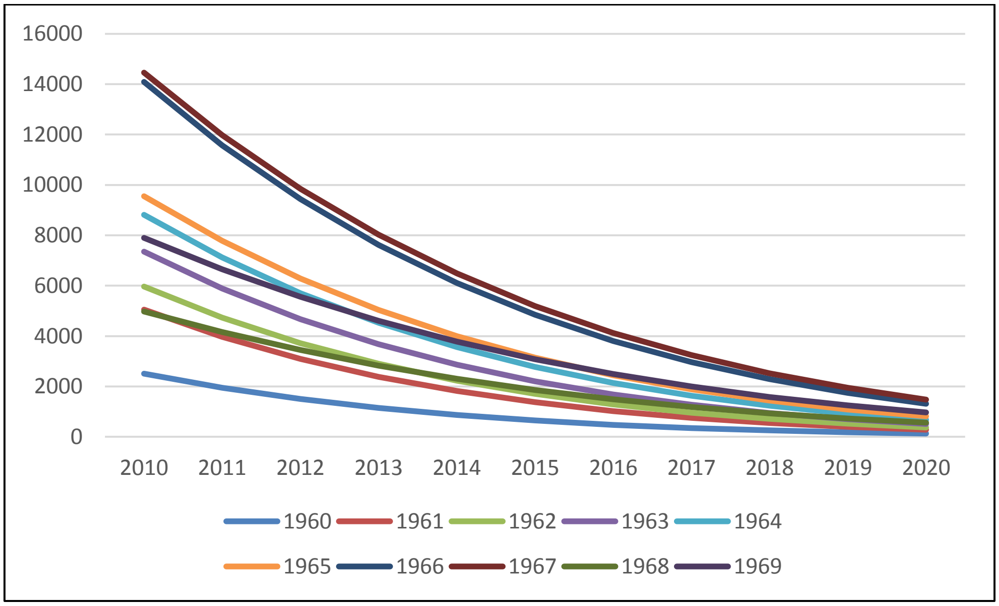
- A1.1 – a complete ban on the import of second hand cars into Peru;
- A1.2 – a decrease in the import tax of new vehicles;
- A1.3 – diminishing the tax on new car ownership and
- A1.4 – a tax on vehicle age.

It is important to note that A1 does not imply any difference in the final size of the private vehicle fleet compared to the baseline scenario, but rather a difference in its composition. Specifically, older cars would retire from the fleet and be substituted by new cars under A1 at a higher rate than under BAU. This would theoretically make sense in the current situation of Peru, where motorization rates are relatively low (Concepcion, 2014) but the current road transport infrastructure of the MALC is simply not capable of absorbing the current vehicle load, let alone an increase. As it is, congestion rates in the area are untenable, even for the Latin American context (Thomson & Bull, 2002). For the sake of analysis, the policy is modeled as it is formulated in the policy description. Below is an explanation of the specific measures.

A1.1: the background of this measure (the ban on import of second hand cars) is the complete liberalization of used car imports in Peru in the early 1990s. From 1992 onwards, there were massive imports of old, heavily used Japanese cars with hundreds of thousands of kilometers travelled, destined for the junkyard in their home country, but instead restored for sale in Peru. Moreover, these cars were heavily retrofitted (e.g. the steering wheel in on the right hand side in Japan) without any effective technical clearance or standard. The import of used vehicles heavily outweighed the import of new ones in Peru for most of the 1990s and early 2000s (Vial et al., 2010). In the emissions model, measure A1.1 is reflected in the comparative composition of the fleet between the Business As Usual (BAU) scenario and the A1 scenario in vehicles from the year 2010 onwards. In the A1 scenario, every car added to the pool from 2011 to 2020 is a new one, with corresponding lower emissions of particles.

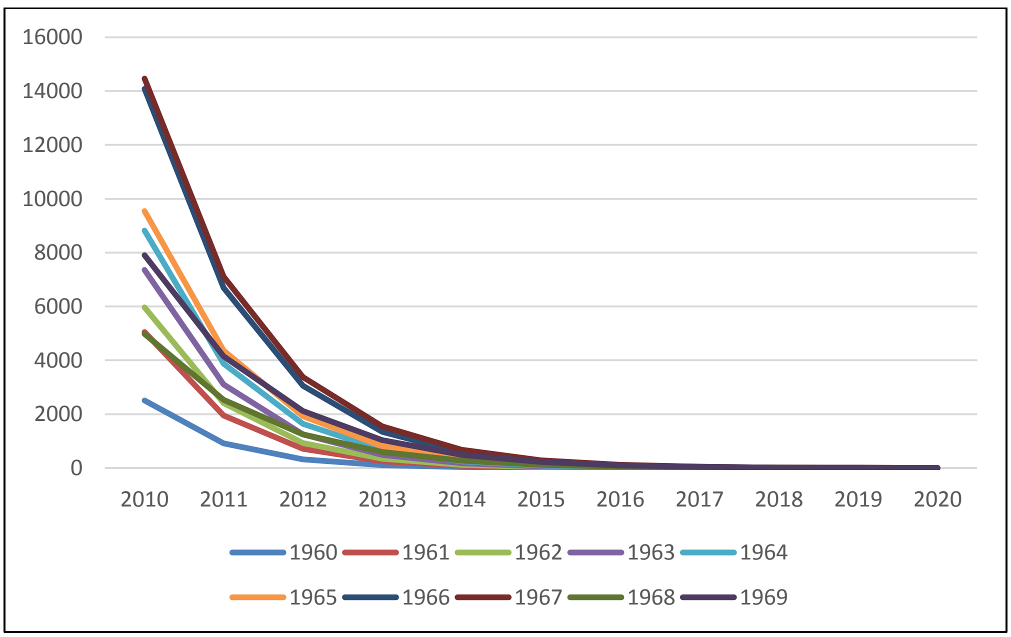
A1.2, A1.3 and A1.4 are best analyzed jointly, as they are meant as a joint system to incentivize the purchase of new vehicles in comparison with used vehicles, by reducing tax charges for importers (a price decrease that would presumably benefit the end customer) and for vehicle owners themselves. With regard to the import tax, currently set at 12% of the declared value of the imported vehicle (SUNAT, 2012), the policy document does not specify how much this tax would be lowered, an omission perhaps explained by the political context – Peru has recently signed numerous free trade agreements under which import tax reductions or elimination are mandatory, but may be reluctant to refer to tariff scrapping outside the context of those negotiations. Regarding A1.3, aside from the PISA-2 document, there is no mention whatsoever in the Peruvian tax policy plans (CIES, 2011) to diminishing the tax on vehicle ownership. By contrast, action A1.4 is explained as an expansion of the tax on car ownership towards older cars, expanding the pool of cars subject to tax by one year of vehicle age each passing year. Overall, the dynamics in the modelling of the A1 scenario consist of an increase in the slope of the retirement rates of older vehicles and their substitution with new vehicles from 2011 to 2020. Figure 6 and Figure 7 illustrate the differences from the BAU scenario to the A1 scenario in terms of retirement rate for vehicles manufactured between 1960 and 1969.

Figure 6. Private car retirement rates by construction year under BAU scenario



Source: Graph elaborated by the author based on Clean Air Institute data. Y axis: # of vehicles

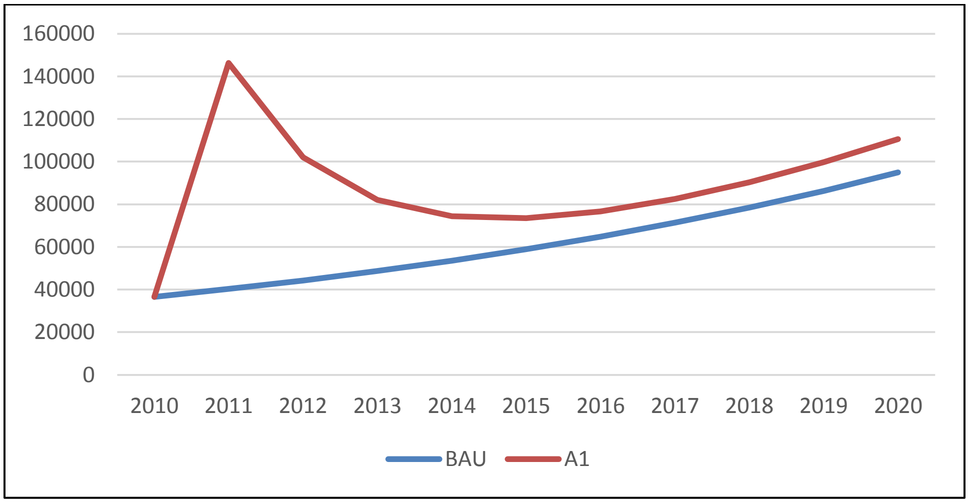
Figure 7. Private car retirement rates by construction year under A1 scenario



Source: Graph elaborated by the author based on Clean Air Institute data. Y axis: # of vehicles

In addition, the evolution of the incorporation of new vehicles marks the main difference between BAU and A1. As the measure is rolled in, a spike in the incorporation of vehicles compensates for the severe downsizing of the oldest part of the fleet. As time goes by, the effect between the A1 incorporation of new cars and BAU decreases. The effect is illustrated in Figure 8.

Figure 8. Evolution of new vehicles incorporated to the private fleet

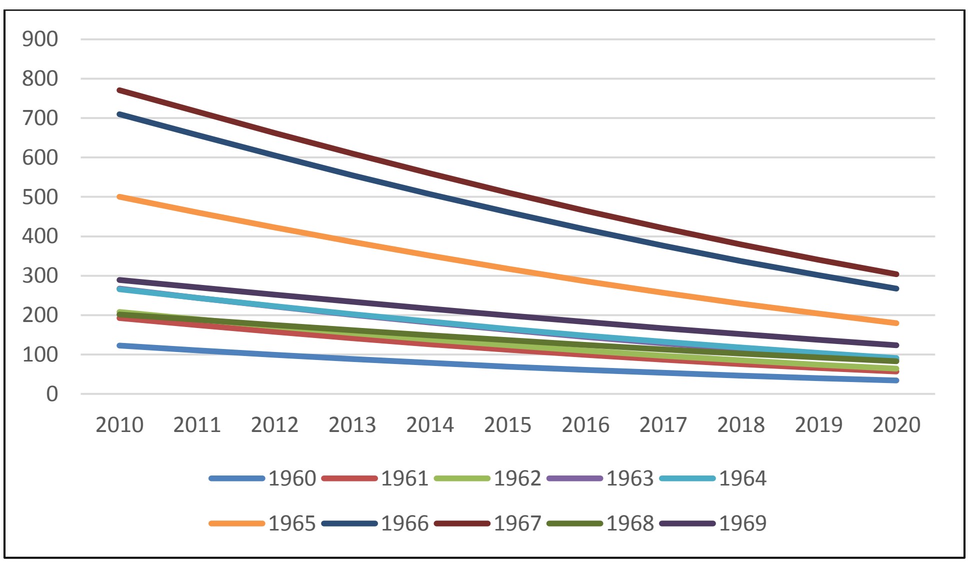


Source: Graph elaborated by the author based on Clean Air Institute data. Y axis: # of vehicles

3.2.1.2. Action A2 (renewal of public bus transportation fleet)

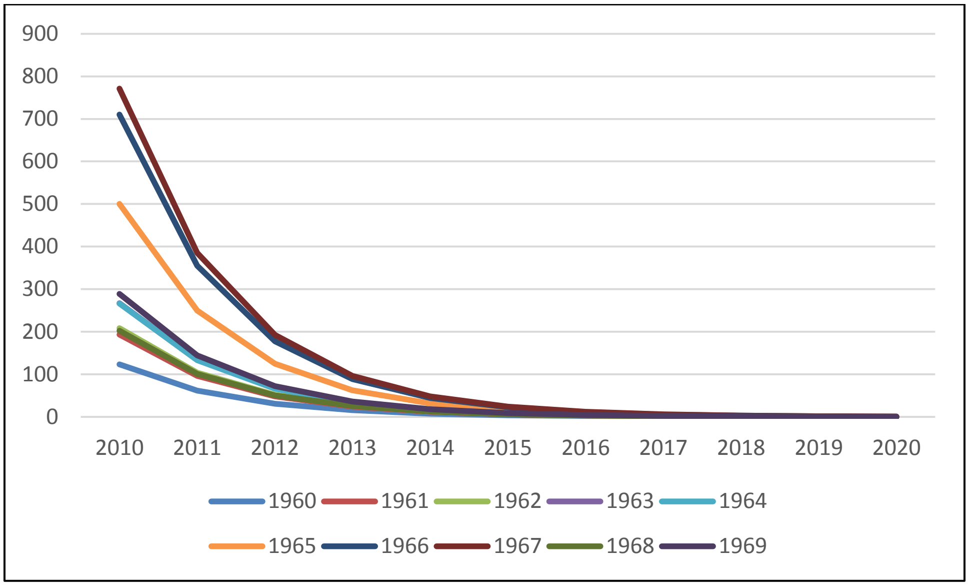
This action intends to renew the fleet of public transportation vehicles by speeding the retirement of older vehicles through a scrapping bonus. It also intends to promote the substitution of older bus units with smaller capacity with new buses with higher occupancy levels. This is accomplished through stringent licensing requirements resulting in decommissioning of smaller units. Again, there is no detail regarding goals and indicators in the policy description. The effects of this policy were modelled through a higher rate of retirement of older units under the A2 policy scenario than under the BAU scenario. Without specific stated goals, CAI established a decreasing rate according to the age of the vehicle: buses from 1950 to 1959 retire at 60% per year, buses from 1960 to 1969 retire at 50% per year… and so on until 2010. Again, the 1960-1969 cohort of circulating vehicles is presented in Figure 9 (BAU scenario) and 10 (A2 scenario) below for illustrative purposes.

Figure 9. Bus retirement rates by construction year under BAU scenario



Source: Graph elaborated by the author based on Clean Air Institute data. Y axis: # of buses

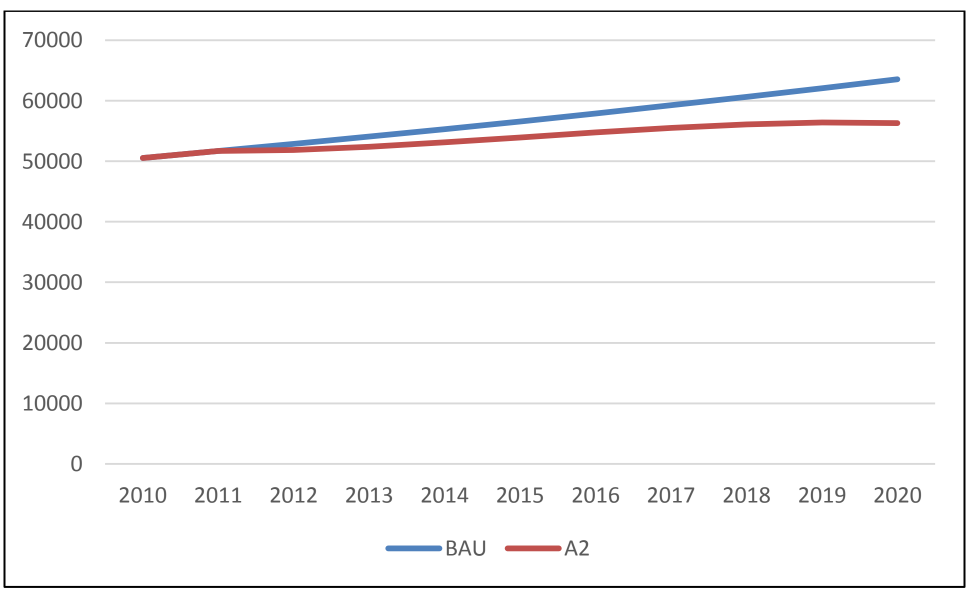
Figure 10. Bus retirement rates by construction year under A2 scenario



Source: Graph elaborated by the author based on Clean Air Institute data. Y axis: # of buses

From 2010 onwards, under A2 more new buses are purchased every year, but less in terms of units than those retired. Specifically, in 2011 all of the buses retired through A2 intervention are replaced by bigger, newer ones; but the next year the proportion is 90%; and it continues decreasing at 10% per year thereafter. This is a reasonable assumption; the bus fleet in the MALC is oversized in terms of units, and the current and projected carrying capacity can be handled by a more efficient fleet of larger buses (CAI executive director, personal communication). The evolution of the bus fleet under BAU and under A2 is presented in Figure 11.

Figure 11. Bus fleet in the MALC under alternative scenarios



Source: Graph elaborated by the author based on Clean Air Institute data. Y axis: # of buses

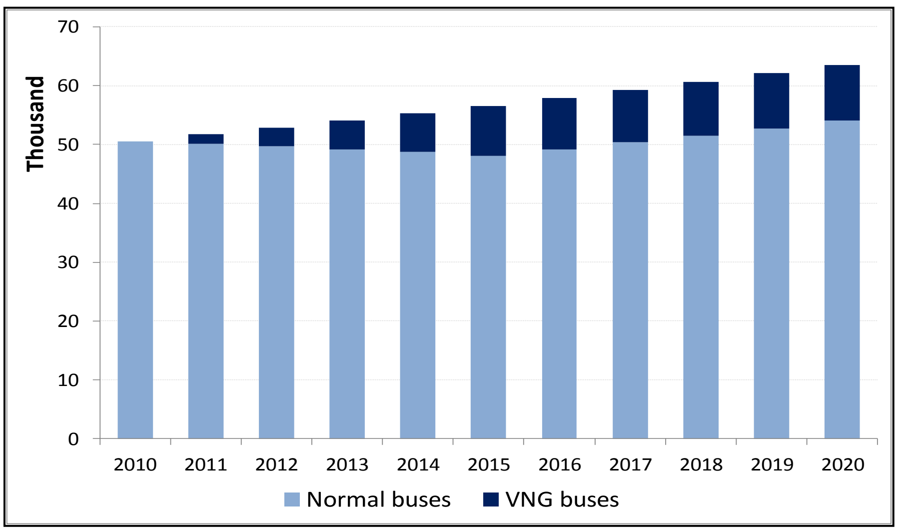
3.2.1.3. Action A3 (promotion of vehicular natural gas in the bus fleet)

This action intends to promote the use of VNG (a cleaner fuel with virtually no PM emissions) in the fleet of public transportation bus vehicles. The tools for doing so are:

- A3.1. A fund for the conversion of conventional vehicles to VNG-using vehicles; and
- A3.2. A line of credit available for public transportation operators to finance the purchase of native VNG vehicles.

The goal and intended mechanics of this action were modelled by CAI to reach a 15% of the buses working on VNG after five years and maintain it at least at that level thereafter. This penetration is achieved at a rate of 3% per year, staying at 15% for the whole considered timeframe thereafter (see Figure 12). In action A3, therefore, there is no change in the relative number of units according to age (i.e. retirement rates) or in the resulting final fleet compared with BAU. The only change is one of technology, with its associated emissions factors and knock-on effects on emissions.

Figure 12. Bus fleet under A3 scenario



Source: Graph elaborated by the author based on Clean Air Institute data. Y axis: # of buses

3.2.1.4. Action A4 (promotion and enforcement of vehicle maintenance)

The goal of this action is reducing emissions through the following interventions:

- A4.1. Enforcement of regulation compliance by vehicle inspection centers; and
- A4.2. Random emissions checks in the streets.

The objective after five years is that 80% of the circulating fleet is inspected periodically and complies with the established emissions limits. At baseline, only 22.4% of the vehicles mandated to undergo yearly inspection actually did so (INVERMET, 2009). Increasing that proportion to 80% would mean a net 58% increase over five years, which CAI assumed as an uniform yearly increase means an extra 11.52% every year. Like the case of A3, this measure does not entail a change in circulating fleet with respect to BAU. This has a caveat, since some vehicles repeatedly failing inspections would ultimately be deemed not fit for circulation. However, given the current compliance, this is not likely to be a significant difference. The strength of this measure is in the fact that it applies to the whole vehicle fleet, inducing improvements in emissions factors within the existing fleet.

3.2.2. Exposure to particulate matter at baseline (2010)

The pollutant of interest is determined by the scope of the actions evaluated, and thus restricted exclusively to inhalable particulate matter (both PM_{10} and $PM_{2.5}$). Data on environmental concentrations are available for Lima and El Callao – see section 1.2. The measured average annual concentrations for PM_{10} and $PM_{2.5}$ and associated population in each one of the stations and the four-year average available for both indicators are presented in Table 5 and Table 6. The air pollution areas are those stated by the relevant authorities (DIGESA, 2012) as covered by the five monitoring stations.

The total population nominally covered by monitoring of inhalable particles is approximately 9.29 million. Since data for 2010 (baseline year of reference) lack one of the stations and is incomplete for the rest, the value for 2010 is calculated as the average from the three previous years.

Table 5. Annual average concentrations of PM_{10} and $PM_{2.5}$ in the MALC (μg/m3)

MS	Callao		South		North		East		Center	
Year	PM_{10}	$PM_{2.5}$	PM_{10}	$PM_{2.5}$	PM_{10}	$PM_{2.5}$	PM_{10}	$PM_{2.5}$	PM_{10}	$PM_{2.5}$
2007	42	37	78	63	94	55	91	60	128	76
2008	36	18	105	59	103	51	108	56	126	76
2009	34	25	79	39	121	46	78	60	114	70
Average	37	27	87	54	106	51	92	59	123	74

Source: DIGESA, (2012). MS: monitoring sites

Table 6. Population and PM concentrations in the MALC

Monitoring sites	Callao	South	North	East	Center
Population (millions) 2010	1.74	2.1	3.11	0.96	1.38
Annual average PM_{10} (μg/m3) 2007-2009	37	87	106	92	123
Annual average $PM_{2.5}$ (μg/m3) 2007-2009	27	54	51	59	74

Source: INEI, (2011) and DIGESA, (2012)

From these data we can derive a population-weighted average concentration:

$$C_{pwa} = \frac{1}{\sum_i P_i} \sum_{i=1}^{5} C_i \times P_i \qquad (21)$$

Where:
C_{pwa}= population-weighted average concentration in micrograms per cubic meter;
C_i= Concentration measured in station *i* in micrograms per cubic meter;
P_i= Population under measurement area of station *i* in millions;

So measured, the population-weighted average concentrations of $PM_{2.5}$ and PM_{10} for the MALC are, respectively: C_{pwa} (PM_{10}) = 89.98 μg/m^3; and C_{pwa} ($PM_{2.5}$) = 51.14 μg/m^3.

These indicators are to be interpreted cautiously for various reasons: firstly, on account of the uncertainty regarding temporal and spatial representativeness; secondly because of the assumption that average concentrations as measured in ambient air represent real exposures in the population; thirdly because of the very few measurement data for a very large city like Lima; and lastly because of the possible differences in measuring methodologies and status of operation and maintenance between stations. With these caveats, we can accept the population-weighted average concentrations as a rough measure of exposure to inhalable particulate matter in the MALC. Without projections of population changes by catchment area, the projections of concentrations of PM in the MALC are applied uniformly to all the population. Therefore, annual average concentrations in micrograms per cubic meter continue to be the exposure assessment indicator throughout the considered timeframe (2011-2020).

3.2.3. Health impact assessment

The physical impacts of particulate matter, the pollutant of concern, consist mainly of an increase in premature mortality and morbidity, and damage to the built environment and heritage through soiling (Watkiss, 2006). No information was found on pollution soiling effects in Peru. Nevertheless, the actual magnitude of these economic costs compared to those of the health effects is in the practice often negligible (Holland, 2014). Therefore, the benefits of the policies proposed for the reduction of particulate matter in the MALC accounted for in this CBA will be those of the potential health effects avoided through the reduction of pollution concentrations. The calculation of the mortality and morbidity due to particulate matter pollution relative to each proposed action was done through an attributable fraction model (Ostro, 2004). The complete process is detailed below, and based completely on generally accepted and published methods, largely those used for the World Bank's "Country Environmental Assessments"; examples of application can be found at the published studies by Golub et al. (2014); Larsen & Strukova (2005); and Larsen (2004), among others. A comprehensive discussion on air pollution health impact assessment methodologies is beyond the scope of this dissertation. Therefore, only the basics of the health impact assessment are summarized here. Very simply put, the strength of the association between the exposure of the pollutant of interest and the selected health outcomes (in this case inhalable particles of an aerodynamic diameter of less than 10 micrometers and a subset of these with aerodynamic diameter of less than 2.5 micrometers) is obtained through empirical observation and comparison of populations exposed to different levels of the pollutant. This allows studying the association between concentrations of PM and counts of deaths or morbidity measures. That association is then summarized into concentration-response curves describing the increase in a given health outcome, or set of them, as the concentration of pollutants increases. Most of these studies assume linear associations between air pollution and health outcomes, although logarithmic transformations are often used for high concentrations of pollutants (Katsouyanni et al., 1997).

From the β coefficients describing the rate of change in the concentration-response curve, a relative risk can be obtained, describing the ratio of the risk of developing a certain health outcome between those exposed to the pollutant/s and those unexposed. The relationship between the β coefficient and relative risk varies, typically being an exposure-dependent function of either a linear or a log-linear form. While the choice of form does not usually have large effects in the final estimates, the available evidence seems to suggest a better fit of either for different outcomes (Ostro, 2004).

With the concentration-response curve and associated relative risks, we can calculate the fraction of mortality and other outcomes in the population that is attributable to inhalable particulate matter. The relationship between this attributable fraction (AF) and relative risk (RR) can be expressed as:

$$AF = \frac{\sum P_i \, RR_i - 1}{\sum P_i \, RR_i} \qquad (22)$$

Where:
P_i is the proportion of the population at exposure category i including the unexposed category (i.e. P_i RR_i becomes P_1 RR_1 + P_2 RR_2 +…+ $P_{unexposed}$ x1).
RR_i is the relative risk at exposure category i compared to the reference level.

This equation takes into account different population groups who are exposed to different levels of pollutants, for example populations in various cities of a country. In the case of one setting (e.g. one city) and with only one exposure level, everyone is considered exposed (P_i becomes 1) and only one relative risk value applies, so the attributable fraction is calculated as:

$$AF = \frac{RR - 1}{RR} \qquad (23)$$

While for several pollutants, concentrations that are below a certain level do not have observable adverse effects, in the case of PM there is no epidemiological evidence of a safe threshold in concentration (WHO, 2006). However, a counterfactual lower threshold is usually established in health impact assessment because of the physical impossibility of attaining zero concentrations of particles in an urban environment. A recent review (Krewski et al., 2009) suggests the threshold to be 5 μg/m³ of annual average ambient air concentration of $PM_{2.5}$, or 10 μg/m³ of annual average ambient air concentration of PM_{10}. This threshold is subtracted from the concentration used as exposure indicator, in our case the annual average concentration of $PM_{2.5}$ or PM_{10}. The expected number of cases of the relevant outcome can then be calculated as:

$$AC_i = (X - X_o) \times AF_i \times B_i \times P \qquad (24)$$

Where:
AC_i is the expected number of cases of the outcome of interest.
X is the annual average concentration of the pollutant ($PM_{2.5}$ or PM_{10}) in μg/m³.
X_o is the pollutant-specific counterfactual threshold.
AF_i is the attributable fraction
B_i is the population incidence of the given health effect (e.g. deaths per 1000 people).
P is the population exposed.

The direct application of this model to the case of the MALC is challenging. Epidemiological research on air pollution in Peru is still in its early stages. Thus, for the association between exposure to inhalable particulate matter and health effects, results from widely accepted international studies were used. In a recent review of the evidence (WHO, 2013), the World Health Organization listed the recommended studies from which to derive relative risks for health impact assessments. Those estimates are the ones used for the health outcomes included in the present evaluation.

Mortality

For the estimation of mortality, the outcome selected is all-cause (non-accidental) mortality attributable to long-term exposure to $PM_{2.5}$ in adults older than 30 years of age as reflected in the latest recommendations of WHO (2013). The relative risk for this outcome is derived from a meta-analysis of the best available evidence for the association between long term exposure to inhalable particulate matter and mortality (Hoek et al., 2013). The relative risk estimate is 1.062 (95%CI: 1.040-1.083). In order to ascertain the share of mortality that is attributable to air pollution, baseline data on population and causes of mortality by age are required. In Peru these data are collected by the National Institute of Statistics (INEI) and compiled regularly in reports (INEI, 2011). The total non-accidental annual average mortality rate in adults over 30 years of age for 2010 in Lima-Callao is 0.006. The population over 30 years of age in 2010 in Lima-Callao was 4,319,850. Through direct application of equation (24), attributable mortality at baseline (2010) can be calculated.

Morbidity

Regarding nonfatal outcomes with known air-pollution associations, perhaps the most burdensome for patients and health systems is chronic bronchitis (CB). The outcome of interest in this case can be defined as incident cases of chronic bronchitis attributable to PM10 pollution. The WHO recommendations for population exposed is adults 18 years of age and older, and the recommended relative risk and 95% confidence interval is 1.117 (1.040–1.189) (WHO, 2013). Lacking studies on prevalence or incidence of the disease in Peru, the rates applied are those from previous estimates (Shibuya, Mathers, & Lopez, 2001; WHO, 2001) for the AMR-D WHO sub-region[4] which includes Peru. The annual incidence rate of chronic bronchitis is 2.05 cases per 1000 population. Since such incidence estimate is available, attributable CB can thus be calculated via equation (24). Regarding other morbidity outcomes like pollution-related hospital admissions, no reliable estimates of baseline annual incidence rates are available for Peru yet. Without that information, the burden of morbidity associated with particulate matter cannot be calculated with the standard burden of disease model expressed by equations (22), (23) and (24). The alternatives available are extrapolating base rates from elsewhere, or using empirical functions from pooled studies. Both options have their own set of disadvantages. In this case, we are using the second alternative. The methodology applied (Ostro, 1994) was developed for estimations of environmental burden of disease in settings with low data availability. The generic formula applied is:

$$AC_i = P_i \times \beta_i \times X \tag{25}$$

Where:

AC_i are the cases of health outcome i (see list in Table 7) attributable to PM_{10}

P_i is the total population susceptible (e.g. adults over 30 years of age, children under 5 …)

β_i is the empirically derived rate of increase of outcome i for a population group per 1 μg/m3 increase in annual average ambient concentration of PM_{10}

X is the average annual concentration of PM_{10}

No threshold is applied in these formulas, on account of the pooling of various studies with diverse or no thresholds in them. The key therefore lies in the empirically derived coefficients linking

[4] Member states of WHO are divided into six geographical regions. These regions are further subdivided into sub-regions according to child and adult mortality from A (lowest) to E (highest). The Americas are one region (AMR) and Peru is one of the countries in the sub-region D.

concentrations and health outcomes. Those coefficients, along with the relevant categories in the International Statistical Classification of Diseases and Related Health Problems 10th Revision (ICD-10) are listed in Table 7.

Table 7. Coefficients for morbidity outcomes attributable to PM_{10}

Health outcome	ICD-10 codes	Dose-response coefficient*	Pollutant
Hospital admissions	I00 – I99, J00 – J99	1.2 per 100,000 population per year	PM_{10}
Emergency room visits	I00 – I99, J00 – J99	24 per 100,000 population per year	PM_{10}
Restricted activity days	N/A	5,750 per 100,000 adults over 30 per year	PM_{10}
Lower respiratory illness in children	J09 – J22, J40 – J47	169 per 100,000 children under 5 per year	PM_{10}
Respiratory symptoms	N/A	18,300 per 100,000 adults over 30 per year	PM_{10}

Source: (Larsen, 2004; Larsen & Strukova, 2005; Ostro, 1994). * Increase per 1 μg/m^3 annual average ambient concentration

These effect categories are frequently used in economic evaluations of air pollution, although there are others that could be used with enough reliable baseline information. In order to facilitate comparisons with other environmental risk factors, health effects can be converted to disability adjusted life years (DALYs, a measure of burden of disease). To do so, disability weights and average duration of each outcome are assigned to each health effect. The formula for the calculation of a DALY (Prüss-Üstün et al., 2003) is:

$$DALY = YLL + YLD \tag{26}$$

Where:
YLL are the Years of Life Lost (YLL) due to premature mortality in the population, and
YLD are the Years of Life Lost due to Disability (YLD) for people living with the health condition

The calculation of each term is described elsewhere (WHO, 2015) and lies beyond the scope of this dissertation. All health end-points considered in this analysis are listed in Table 8, along with a predetermined calculation of DALYs lost per 10,000 cases of the considered outcome.

Table 8. DALYs lost per selected category of health effects

Health outcome	DALYs lost per 10,000 cases
Mortality	75,000
Chronic bronchitis (adults)	22,000
Hospital respiratory admissions	160
Emergency room visits	45
Lower respiratory illness: children	65
Restricted activity days (adults)	3
Respiratory symptoms (adults)	0.75

Source: Adapted from Larsen (2004)

The disability weights and average duration of illness that have been used in this report to calculate the DALYs were determined by previous studies (Larsen, 2004; Larsen & Strukova, 2005) and are presented in Table 9. The average duration of CB is estimated based on age distribution in Peru and age-specific CB incidence by Shibuya et al., 2001. Years lost to premature mortality from air pollution were estimated from age-specific mortality data for cardiopulmonary and lung cancer deaths, discounted at three percent per year.

Table 9. Calculation of DALYs per case by category of health outcomes

Health outcome	Disability weight	Average duration of illness
Mortality	1.00	(7.5 years lost)
Chronic bronchitis	0.20	20 years
Hospital admissions	0.40	14 days*
Emergency room visits	0.30	5 days
Lower respiratory illness: children	0.28	10 days
Respiratory symptoms: adults	0.05	0.5 days
Restricted activity days: adults	0.10	1 day

Source: (Larsen, 2004; Larsen & Strukova, 2005) * Includes recovery period after hospitalization.

3.3. Results

3.3.1. Pollution at baseline and under prospective scenarios

According to the model by CAI, all scenarios lead to higher emissions in the future than at baseline. Compared with the BAU scenario, all proposed policies will, by themselves, produce a comparative decrease in the PM_{10} emissions in the first four to five years, but eventually will regain an upwards tendency. Policy A4 (Inspection and maintenance of private fleets) has the largest emissions reduction effect, followed by A2 (scrapping of older buses), A3 (Natural Gas in public transportation vehicles) and A1 (private fleet renewal). Assuming homogeneous mixing and without the benefit of a baseline dispersion model, the ambient concentrations of inhalable particles would follow suit. The current ratio $PM_{2.5}$ to PM_{10} in the MALC is then applied to derive $PM_{2.5}$ concentrations. The results are listed in Table 10.

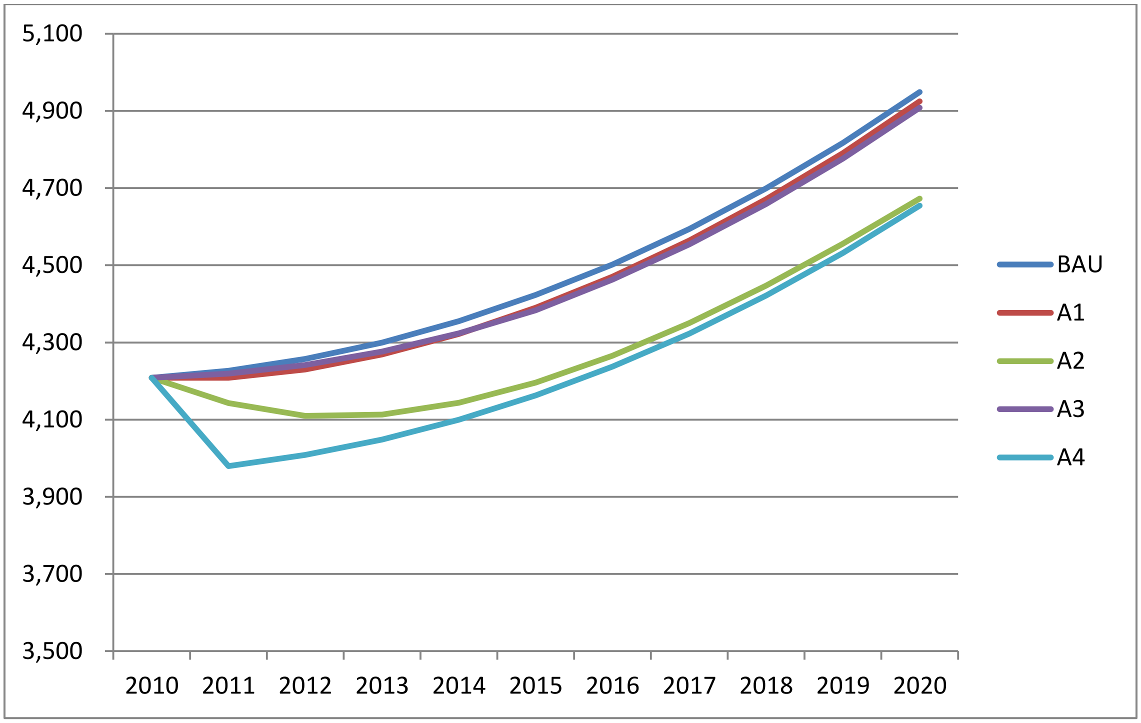
Table 10. Concentrations of PM_{10} and $PM_{2.5}$ in alternative scenarios, 2010-2020

Scenarios	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
BAU PM_{10}	89.98	90.37	91.03	91.95	93.12	94.57	96.27	98.24	100.48	103.01	105.82
A1 PM_{10}	89.98	89.98	90.45	91.27	92.42	93.87	95.60	97.59	99.87	102.44	105.30
A2 PM_{10}	89.98	88.59	87.88	87.95	88.60	89.73	91.21	93.02	95.10	97.41	99.91
A3 PM_{10}	89.98	90.21	90.70	91.45	92.45	93.73	95.42	97.38	99.62	102.14	104.94
A4 PM_{10}	89.98	85.09	85.71	86.57	87.67	89.01	90.60	92.44	94.54	96.90	99.53
BAU $PM_{2.5}$	51.14	51.36	51.74	52.26	52.92	53.75	54.71	55.83	57.11	58.54	60.14
A1 $PM_{2.5}$	51.14	51.14	51.40	51.88	52.53	53.35	54.33	55.47	56.76	58.22	59.85
A2 $PM_{2.5}$	51.14	50.35	49.94	49.99	50.36	51.00	51.84	52.87	54.05	55.36	56.79
A3 $PM_{2.5}$	51.14	51.27	51.55	51.97	52.54	53.27	54.23	55.35	56.62	58.05	59.64
A4 $PM_{2.5}$	51.14	48.36	48.71	49.20	49.83	50.59	51.49	52.54	53.73	55.07	56.57

Source: Clean Air Institute (unpublished) based on data provided by the Lima Transportation Authority

The comparative emissions reductions are illustrated in Figure 13.

Figure 13. Evolution of PM_{10} emissions (Tm) from 2010 to 2020 under different scenarios



Source: Author's own elaboration based on Clean Air Institute data

The difference between BAU and A1 and A3 in terms of pollution is small. In the case of A1, this is consistent with the notion that the private fleet in the MALC produces a small proportion of the particulate matter, and explaining the better performance of A2, which targets polluting heavy-load, diesel-equipped public transportation vehicles. The particulate matter emissions of inefficient and polluting public transportation vehicles are thus a factor to take into account in terms of policy levers. Regarding A3, the emission factors of the prevailing bus engine technology (Diesel) are much higher than those associated with VNG. However, the 15% VNG penetration does not affect all vehicles, so the final effect is diluted. The difference between BAU and A4 in terms of pollution is the largest of all considered policies, illustrating the potential for improvement even within the existing vehicle fleet. This result is consistent with expert opinions of stakeholders (COFIDE, ARAPER, personal communication) and reflects a strong deficit in compliance with large potential gains in the short term. Despite the reductions, the prospective concentrations in all scenarios remain very high compared not only with the current situation in European or North American countries, but also in the context of Latin America and the Caribbean (WHO, 2016). The estimates reflect a very clear result: these policies cannot by themselves reduce concentrations of particulate matter anywhere near safe levels such as the World Health Organization air quality guidelines (WHO, 2006). The proposed actions must be complementary to much larger interventions, such as a massive land use management, reorganization of public transportation with heavy stress on non-polluting modes like subway or trams, behavioural or modal shift interventions, or other long-term urban transformation options. While interesting on their own, the policy analysis consequences of these results are beyond the scope of this dissertation. These impacts are relevant here only insofar as they provide a real-life basis for an environmental CBA and MCA, and possibilities for a combination thereof.

3.3.2. Health impacts at baseline and under prospective scenarios

The estimated health impact of urban air pollution in the Metropolitan Area of Lima-Callao at baseline (year 2010), that is, before any intervention, is presented in Table 11. These estimates are consistent with previous ones for the city of Lima, albeit slightly increased (Larsen & Strukova, 2005). Decreases in air pollution concentrations are likely compensated with a marked growth in exposed urban population to explain this modest increase.

Table 11. Estimated health impacts of PM pollution in the MALC at baseline (2010)

Health categories	Cases	DALYs
Premature mortality	6,982	52,363
Chronic bronchitis	9,899	21,778
Hospital admissions	10,031	160
Emergency room/outpatient hospital visits	196,774	885
Lower respiratory illness in children	127,143	826
Restricted activity days	22,350,256	6,705
Respiratory symptoms	71,132,119	5,335
Total	N/A	88,052

Source: Authors' estimates, based on Larsen & Strukova (2005) and Larsen (2004)

These results need to be interpreted cautiously. Some important health outcomes attributable to air pollution (e.g. asthma attacks related to air pollution) are not included in this analysis. Hence, these estimates should be treated as an underestimate. By applying iteratively this model, and using respectively the concentrations of PM for each proposed action and a "Business as usual" (BAU) scenario we obtain the burden of mortality and morbidity attributable to air pollution that each policy can avert over a 10 year period (Table 12). For ease of comparison, all health outcomes under each scenario are converted into DALYs (Disability-Adjusted Life Years, a widely used measure of burden of disease) in the last row. These numbers represent the health gains associated to each policy as compared to the BAU scenario.

Table 12. Total averted mortality and morbidity accrued over 10 years by each action

Health outcomes averted	A1 renewal	A2 scrapping	A3 VNG	A4 maintenance
Mortality	583	4,426	667	5,474
Chronic bronchitis	839	6,370	959	7,878
Hospital admissions	755	5,738	864	7,096
Emergency room visits	14,818	112,554	16,950	139,190
Lower respiratory illness in children	9,574	72,725	10,952	89,936
Restricted Activity Days	1,683,038	12,784,220	1,925,253	15,809,697
Respiratory Symptoms	5,356,451	40,687,170	6,127,329	50,316,079
Burden of disease (DALYs)	7,263	55,169	8,308	68,226

Source: Authors' estimates. DALYs: Disability-Adjusted Life Years

The attributable mortality, morbidity outcomes and overall burden of disease are roughly proportional to the concentrations expected under each scenario, a result consistent with the form of the health impact assessment model. Consequently, the largest health gains are expected under action A4, followed by A2, and much less under A3 and A1. The cases of premature mortality and incident cases of each morbidity outcome are the basis for the calculation of the health cost of air pollution under each proposed measure. In turn, the monetized averted morbidity and mortality of each measure as compared to the BAU scenario essentially constitutes the benefits associated to each measure. Such benefits are measured against the social cost of the measures in the following chapter (Cost Benefit Analysis).

Chapter 4. Case study cost–benefit analysis

4.1. Introduction

Cost–benefit analysis is routinely used for the prospective evaluation of public policies to reduce air pollution for several reasons, mainly for its ability to summarize complex bodies of evidence in support of decision-making. The largest benefits in most CBAs of this type come from averted impacts and to a much lesser extent from time savings and energy savings, improved visibility and other "environmental services". At the same time, air quality improvement policies frequently involve long timeframes and large investments, and are usually shrouded by considerable uncertainty.

These are all important considerations to take into account when engaging in a CBA of air quality improvement policies, particularly if, as it is the case with this CBA, data availability is low. In turn, this determines the reach of the analysis in this case; whereas costs and benefits are considered from a whole-of-society perspective, not all possible social costs and benefits of the actions considered are included, simply because there is no information to model them soundly. In this regard, the present CBA cannot be a complete social CBA. Instead, the benefits and costs for which a relatively solid basis of evidence can be found are included. On the other hand, costs and benefits are estimated as homogeneously as possible across actions, in order to maximize fair comparability. The methods and details of this estimation are detailed below.

4.2. Methods

4.2.1. Scope, timeframe and discounting

The scoping of the CBA entails geographical and temporal limits in the accrual of costs and benefits. The geographical limit of this evaluation is restricted to the Metropolitan Area of Lima and El Callao (MALC), as determined by the scope of the actions evaluated. On the timeframe for analysis, regardless of the nominal duration of the specific policy or program, air pollution reductions are meant to be sustained beyond the policy implementation timeframe. Moreover, the consulted stakeholders assessed that the implementation timeframes of all the proposed activities would all go beyond their nominally established five year period (see Chapter 5, MCA). Thus, the CBA timeframe is set for ten years, from 2011 to 2020, hoping to capture a better portrait of costs and benefits of the proposed policies. As discussed in chapter 2, imposing a high discount rate for health damages projected far into the future may be unfairly detrimental to their value. For that reason, the prevailing scientific *status quo* tends to support the notion of small (3% or less) and/or declining discount rates (Goulder & Williams, 2012). For the purposes of this CBA, two discount rates will be applied to both benefits and costs: 0% and 3%.

4.2.2. Benefits estimation

Though not the only impacts of air pollution in cities, health effects are by far the most important and best studied. There are no obvious time savings associated to the actions evaluated in this CBA, and while there may be some cost savings from improved fuel efficiency, these would be conditional to many other variables (e.g. actions on congestion) thus making them challenging to estimate. Therefore, the benefits of the policies proposed for the reduction of particulate matter in the MALC accounted for in this CBA will be those of the potential health effects avoided through the reduction of pollution concentrations. The steps required for the calculation of benefits (accrued over a period of ten years) are thus:

a) A forecast of the potential PM emissions scenarios under a BAU scenario and under each one of the actions considered. This step was kindly performed by the Clean Air Institute for Latin American Cities, based on a fleet inventory and appropriate technology-specific emissions factors.
b) A forecast of the resulting concentrations and exposure of the population to PM under a BAU scenario and with each one of the actions considered.
c) The calculation of the attributable fraction of mortality and morbidity due to the levels of particulate matter under BAU scenario and with each one of the actions considered.
d) The calculation of the economic cost due to excess premature mortality and morbidity associated with PM under a BAU scenario and with each one of the actions considered.
e) The comparison of avoided health and economic impacts due to PM between each action and the BAU scenario. The avoided economic costs between an action and the BAU scenario, appropriately discounted, constitute in this CBA the economic benefits of that action.

All the steps up to c) are reflected in Chapter 3 (Case study policy impact assessment) and constitute the basis for the estimation of the benefits of the actions considered. There is an extensive body of literature dealing with the valuation of the health costs of air pollution, the details of which have been discussed at length elsewhere (OECD, 2010 and 2014). However, in the overwhelming majority of applications there are two basic types of costs considered in the valuation: 1) the cost of mortality attributable to air pollution; and 2) the cost of illness attributable to air pollution, as well as of its knock-on consequences.

The economic value of mortality risk reduction can be calculated in several ways. Each option has advantages and disadvantages that have been extensively discussed elsewhere (OECD, 2010). In our case, "value of a statistical life" (VSL) estimates obtained through "willingness to pay" techniques are used for pragmatic reasons: 1) consistency with previous studies (Larsen, 2004; Larsen & Strukova, 2005); 2) the human capital approach is considered to vastly underestimate mortality cost at present in Peru; and 3) most large regulatory and international organizations use it. In practice, we use a VSL to calculate mortality costs (OECD, 2012). There are currently no local studies that provide this value for Peru in a context of air pollution risks, although there is a relevant benefit transfer study conducted by the National Energy Office (Vasquez, 2006) for use in the regulation of the Energy sector. This study takes the datasets (in USD of 2000) used by Viscusi and Aldy (2003) in a meta-analysis of US and global studies and estimates the VSL for chronic outcomes (reference year 2005) for Peru adjusting by local income (GDP PPP), Consumer Price Index, education and levels of occupational risk. We will use this value, of which the central estimate is 1,841,135 Soles of 2005 with a 90% confidence interval [730,798 - 4,638,458].

Besides the mortality cost, the welfare cost of morbidity is often measured by the willingness to pay (WTP) to avoid or reduce the risk of illness. This WTP is often higher than the cost of medical treatment and the value of time losses (Cropper & Oates, 1992) reflecting the additional value that individuals may assign to avoiding pain and discomfort (Sanchez Triana, Ahmed, & Awe, 2007). However, there are no relevant WTP studies from Peru, so the cost-of-illness (COI) approach (mainly medical cost and value of time losses calculated from local data of the MALC) is used, even though this is likely a severe underestimate of the true costs of morbidity. Both the cost of morbidity itself and of the lost productivity (through lost salaries) resulting from morbidity are included.

The calculation of the health cost of air pollution in the MALC takes into account two main components: the premature mortality[5] cost of particulate matter in the MALC (or PMC) and the total cost of illness attributable to particulate matter (COI_{tot}). Then, full health cost (FHC) is:

$$FHC = PMC + COI_{tot} \tag{27}$$

In turn, premature mortality cost results from:

$$PMC = AD_{pm} \times VSL \tag{28}$$

Where:
AD_{pm} are the premature deaths attributable to particulate matter in the MALC.
VSL is the value of a statistical life.

And total cost of illness results from:

$$COI_{tot} = \sum_{i=1}^{n} COI_i \tag{29}$$

Where COI_i is the cost of illness of the proportion of outcome *i* attributable to inhalable particulate matter in the MALC.

As listed in Chapter 3 the health outcomes (n=6) considered in the evaluation are: 1) incident chronic bronchitis; 2) hospital admissions for cardiovascular and respiratory causes in adults over 30 years of age; 3) emergency room and outpatient visits for cardiovascular and respiratory causes in adults over 30 years of age; 4) acute lower respiratory infections in children under 5 years of age; 5) Restricted activity days in adults over 30 years of age; and 6) respiratory symptoms not requiring medical attention in population over 30 years of age attributable to air pollution.

The calculation of the cost of illness for each morbidity outcome considered (six of them) all follow the same pattern. The cost of illness approach, applied equally to all outcomes, is decomposed into four components:

$$COI_i = CIT_i + COT_i + CLT_i + CD_i \tag{30}$$

[5] Defined as all-cause (non-accidental) mortality attributable to long-term exposure to $PM_{2.5}$ in adults older than 30 years of age as reflected in the latest recommendations of WHO (2013); see subsection 3.1.2.5

Where:

CIT_i is the cost of inpatient treatment due to outcome *i* given by:

$$CIT_i = Hospital\ days \times Average\ cost\ of\ hospital\ da \tag{31}$$

COT_i denotes the cost of outpatient treatment due to outcome *i:*

$$COT_i = Outpatient\ consultations \times Average\ cost\ of\ outpatient\ consultation \tag{32}$$

CLT_i is the cost of time lost to outcome *i:*

$$CLT_i = Number\ of\ days\ lost\ to\ sickness \times Average\ daily\ wage \tag{33}$$

And CD_i is the private unsubsidized cost of pharmacy drugs purchased due to outcome *i.*

$$CD_i = (UD_a \times UC_a) + \cdots + (UD_n \times UC_n) \tag{34}$$

Where UD_a… UD_n are the units of all drugs required for one-off treatment and UC_a… UC_n are the respective drug unit costs. Depending on the specific outcome, some of the terms in (30) may have a zero cost; for instance, in outpatient[6] outcomes, CIT_i would not apply and its value would be 0; in symptoms not requiring medical attention, both CIT_i and COT_i would be zero; and so on. The number of days lost due to illness, needed to calculate CLT_i, include the case of the sick person's own time and also the case of the time lost to caretaking. The valuation of opportunity cost of time lost to illness at 75 percent of the average urban wage in Peru reported by the national institute of statistics (INEI, 2012b) was used and applied this cost both to working and nonworking individuals, based on the assumption of an equivalent opportunity cost for both categories. These and the rest of assumptions regarding duration of illness, rate and length of hospitalization, average time lost per health end-point, frequency of doctor visits and discount rate were taken from previous studies and are listed in Annex 1. The calculated cost of treatment was based on consultations with private practitioners, health authorities and the upper bound of the publicly listed prices that public insurers pay healthcare providers.

4.2.3. Costs estimation

Since the CBA in this dissertation is undertaken from a social perspective, costs of the different actions evaluated need to be considered in a broad sense, including not only the cost of their implementation for the concerned agencies but also costs for other stakeholders and opportunity costs for society at large. However, this conceptual requirement often meets the reality of a scarce evidence base to calculate such broad-based costs. Two main cost components are taken into account in each action: 1) The cost of implementation for the government, and 2) The private cost of the action. The governmental costs of implementation for each action were provided by the implementing agencies through the CGIAL, while the actions' private costs were estimated following some assumptions. Below are the rationales and assumptions in each case.

[6] Outpatient healthcare refers to patients that attend a healthcare facility without staying overnight. Inpatient healthcare is the care given to a patient who is admitted into a healthcare facility and stays for one or more nights.

A1: Private vehicle fleet renewal (used vehicles import ban and additional measures)

The intended goal of this measure is for the private vehicle fleet of the MALC to become newer. In order to do so, the measures considered are: 1) A complete ban on the import of second hand cars into Peru; 2) A decrease in the import tax of new vehicles; 3) Diminish the tax on new car ownership and 4) A tax on vehicle age.

Governmental implementation cost: the allocated implementation cost (excerpted from CGIAL from governmental budgets) for this activity is fifteen Million Soles for compliance enforcement and monitoring, to be spent uniformly across five years. The CGIAL insisted that the measures A1.2, A1.3 and A1.4 are meant to be revenue-neutral as a whole for the Peruvian government, but as long as there are variations in taxes and governmental revenues, it is clear that some of these costs are externalized either to individuals (taxes on ownership and age) or both to firms and individuals (variations in import taxes).

Private cost: the case of A1.1 is relevant from the social CBA private cost standpoint, since there is a clear cost to individuals in the decrease of availability of used cars. Assuming the same rate of motorization as in the BAU scenario, a forced substitution of incoming used cars by brand new cars would suppose a heavy increase in private cost for the brand new ones (see Figure 6). The costing of sub-actions A1.2, A1.3 and A1.4 is far less clear from the perspective of a social CBA. The relative decrease or increase of certain taxes would, all else being equal, not automatically imply net costs or benefits for society as a whole. It is clear, however, that different groups in society bear different costs and benefits from wealth redistribution, such as that promoted by tax variations. While CBA is notoriously ill-suited to analyze such distributional effects, various techniques exist to account for them. For instance, Yitzhaki (2003) proposes the use of inequality indicators to analyze the consequences of redistribution. This approach, however, requires far more and better information than what is available for the present analysis, and is anyway usually not included in CBAs. On the whole, aside from sub-action A1.1, it is unclear whether the other components represent a loss or a win for society.

The restrictions in the availability of second hand cars may certainly have a regressive component, with a net cost on the poorest. New taxes on age may represent for the government a net increase in revenue that will hopefully revert secondarily into the citizens, but for the citizens this represents a new cost in the short run, with a probably aggravated regressive effect. In this regard, Yitzhaki (2003) relates this judgment to the effect of a tax in economic growth and in redistribution. At face value, it is difficult to see redistribution resulting from this measure; rather, there could be a regressive effect since owners of older cars probably tend to have lower incomes. However, if the growth component of this measure can be proven in a magnitude large enough to offset possible negative re-distributional effects, it can still be argued that the measure is worthwhile for society as a whole. In conclusion, from all of action A1, only the governmental implementation cost and the private cost incurred by individuals due to sub-action 1.1 are clear, the former predefined in budgets and the latter consisting of a net increase in the expenditure on private fleet (see Chapter 3 – impact assessment). Lacking model-specific projections of purchases, we may assume an average model for pricing. According to the association of Peruvian car importers and retailers (ARAPER, 2011), the approximate weighted average cost of the most popular light private vehicles (Toyota Yaris and Chevy Astra) with taxes included in the MALC is slightly under 30,000 Soles. In this case, we use the full price of a best-selling car at the time (Chevrolet Spark), listed at 27,381 Soles, dealership price (Chevrolet, 2010).

A2: Bonus for scrapping older public transportation vehicles

The intended goal of this action is to renew the fleet of public transportation vehicles by speeding the retirement of older vehicles through a scrapping bonus.

Governmental implementation cost: the budget that the different agencies intend to dedicate to this activity is of 54 Million Soles, distributed over five years (see Table 13).

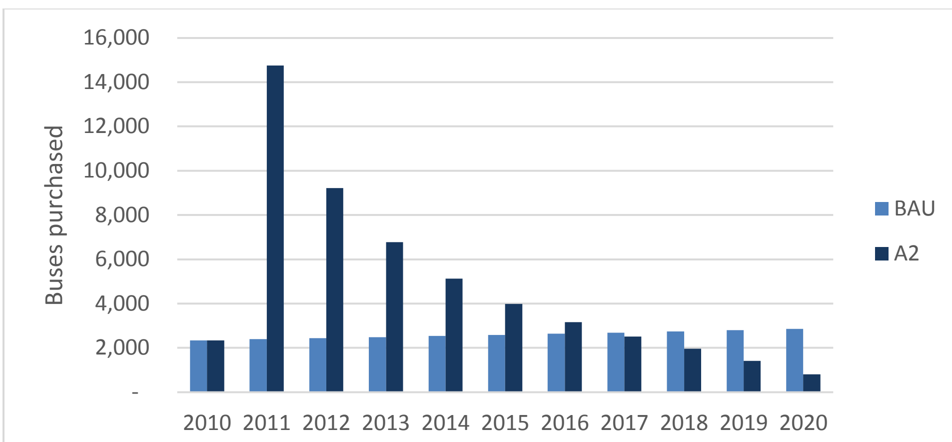
Table 13. Governmental budget allocated to action A2, by year

Year	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Budget allocation (M. Soles)	29.7	13.5	5.4	2.7	2.7	0	0	0	0	0

Source: CGIAL

Private cost: there is a clear element of private costs from the increased purchase of new buses, as the rate of retirement of older units increases under the A2 policy scenario but the fleet continues to grow, meeting the public demand (see Figure 11). As better performing (i.e. more passenger-km per year) and relatively larger buses replace smaller, older ones, the fleet is reduced somewhat; and after a spike in new bus purchases, the trend stabilizes over the ten years considered (see Figure 14).

Figure 14. New public transportation buses purchased



Source: author's estimates, see Chapter 3, section 3.2.1.2

The average price of an omnibus, as estimated by the government for tax purposes, was in 2010 of 215,290 Soles (SUNAT, 2012). While the cost of A2 is greatly increased in comparison with BAU during the first years, it becomes negative (i.e. smaller than in BAU) towards the end of the period considered.

A3: Promotion of VNG use in public transportation vehicles

This action intends to promote the use of VNG (a cleaner fuel with virtually no PM emissions) in the fleet of public transportation vehicles.

Governmental implementation cost: the total budget allocated to this activity (excerpted from CGIAL from governmental budgets) is of around 80 Million Soles over ten years.

Private cost: from the private perspective, there is also a clear cost attached to the VNG conversion, albeit one that can be recovered somewhat depending on prospective pricing of VNG versus diesel. As mentioned in Chapter 3 (impact assessment), the goal of this action is to reach a 15% of the buses working on VNG after five years and maintain at least that level afterwards. This penetration is achieved at a rate of 3% per year, staying at 15% for the whole considered timeframe thereafter (see Figure 12). It is assumed that under BAU, no VNG penetration occurs; therefore, under BAU, all buses are "normal". Since there is no change in the relative number of units according to age (i.e. retirement rates), the resulting final fleet, or the purchase of new units, the only costs are those associated to the retrofitting of the buses. According to governmental data (CPGNV, 2014), the average price of such retrofitting is of 4,400 Soles per unit.

A4: Enforcement of better vehicle maintenance

The intended goal of this action is reducing emissions through the following interventions: 1) Enforcement of regulation compliance by vehicle inspection centers; and 2) Random emissions checks in the streets. The objective after five years is that 80% of the circulating fleet is inspected periodically or complies with the established emissions limits.

Governmental implementation cost: the allocated budget by the authorities is as follows:

- Compliance enforcement budget (commissioned company) = 1,698,000 Soles
- Audits to inspection centers = 141,500 Soles
- Compliance enforcement (Ministry of Transport) = 2,099,294 Soles
- Random checks = (73580 Soles/year) times 5 years = 367,900

Total budget is 4.3 Million Soles over five years.

Private cost: in this case too, a sizeable part of the cost is being externalized. That is, from the current situation, where 22.4% of the vehicles undergo inspection (INVERMET, 2009), to the goal of 80% of vehicles undergoing inspection, at least the private cost of the inspections needs to be added to the implementation costs of the measure. Increasing that proportion to 80% would mean a net 58% increase over five years, with yearly increase of 11.52% every year. The assumption is that inspection levels would remain at least at 80% thereafter. Regarding the cost of inspection, the current tariffs by Lidercon (the largest inspection agency in Lima) are in Table 14.

Table 14. Official vehicle inspection fees in the MALC, 2010

Type of vehicle	Motorcycle	Light vehicle	Heavy vehicle
Cost in Soles	17.70	82.60	132.70

Source: http://www.liderconperu.com/tarifas.htm

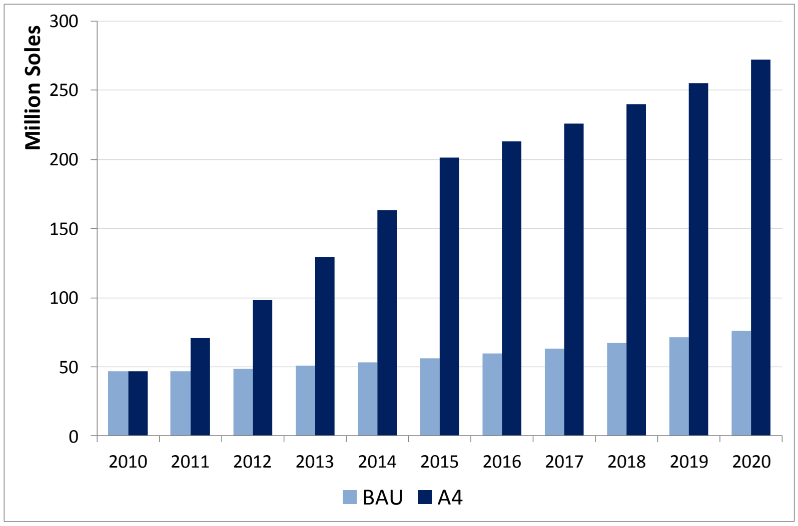
Assuming that inspection costs stay the same, the increase in costs in concept of paid inspections would increase significantly, both for light vehicles and heavy vehicles. The difference between the new scenario (A4 – 80% of vehicles undergo inspection) and a BAU scenario (22.4% of vehicles undergo inspection) in terms of vehicles inspected is reflected in Table 15 and the increase in annual private expenditure in inspections in Figure 15.

Table 15. Projection of additional vehicles inspected by year under A4 compared with BAU scenario

Year	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Light vehicles	953,018	910,767	891,647	880,680	877,680	882,548	895,271	915,918	944,643	981,684	1,027,360
Heavy vehicles	488,917	498,945	537,248	578,689	623,251	670,935	721,762	775,772	833,020	893,573	957,512

Source: author's estimates, see Chapter 3, section 3.2.1.4

Figure 15. Annual expenditure in inspections



Source: author's estimates, see subsection 3.2.1.4

4.2.4. CBA outputs and discount rates

The two most typically used output indicators for cost–benefit analysis are the Net Present Value (NPV) and the Benefit to Cost Ratio (BCR). The NPV is simply the difference between the currency -valued benefits and the currency-valued costs, with discounting applied to both benefits and costs as appropriate; mathematically:

$$NPV = \sum_{t=0}^{n} \frac{(Benefits - Costs)_t}{(1+r)^t} \tag{35}$$

Where:
n is the analytic time horizon (in years)
t is time, expressed in years
r is the discount rate

The BCR is the ratio of total benefits divided by total costs, often presented as a ratio of present value of benefits ($PV_{benefits}$) and present value of cost (PV_{costs}), with discounting applied to both as appropriate; that is:

$$BCR = \frac{PV_{benefits}}{PV_{costs}} \tag{36}$$

As discussed in the methods section (4.2.1), for this CBA we will consider two alternatives: no discounting (i.e. a rate of 0%) and a low, social discount rate of 3%, typically favored by agencies engaging in economic evaluations of health prevention programs (WHO, 2008).

4.3. Results

4.3.1. Benefits of proposed actions

The estimates of the cost per case (year 2010, before the start of the plan) attributable to urban air pollution in the MALC are presented in Table 16.

Table 16. Estimated annual cost of health impacts (Soles, reference year 2010)

Health categories	Cost per case
Mortality (Soles 2005)	1,841,135
Chronic bronchitis	20,803
Hospital admissions	2,540
Emergency room visits	277
Lower respiratory illness in children	156
Restricted activity days (adults)	21
Respiratory symptoms (adults)	5

Source: author's estimates, see sections 3.2 and 3.3 and Table 14

Acknowledging the inherent uncertainties in this estimation, we can now calculate the health benefits and associated savings accrued through the analyzed actions. For that purpose, the avoided mortality and morbidity as estimated for each action (Table 12) is multiplied by the cost per case in Table 16. The results by year are listed in Table 17 below.

Table 17. Monetized annual health benefits (Soles) actions versus a "business as usual" scenario

Benefits of	A1 renewal vs. BAU	A2 scrapping vs. BAU	A3 VNG vs. BAU	A4 maintenance vs. BAU
Year 0	0	0	0	0
Year 1	68,903,280	310,825,175	28,568,302	922,414,129
Year 2	104,233,233	561,609,890	58,668,765	947,472,067
Year 3	121,858,416	725,599,720	90,377,355	976,441,884
Year 4	129,587,630	836,323,216	123,782,118	1,009,507,374
Year 5	131,545,026	914,611,737	158,984,191	1,050,148,781
Year 6	130,031,821	974,385,893	163,415,742	1,092,141,428
Year 7	126,390,348	1,026,082,532	168,040,248	1,138,967,514
Year 8	121,434,751	1,079,235,482	172,876,781	1,190,938,560
Year 9	115,677,015	1,143,704,493	177,945,437	1,248,401,294
Year 10	109,450,209	1,232,141,875	183,267,297	1,311,739,283
Total	1,159,111,731	8,804,520,014	1,325,926,236	10,888,172,315

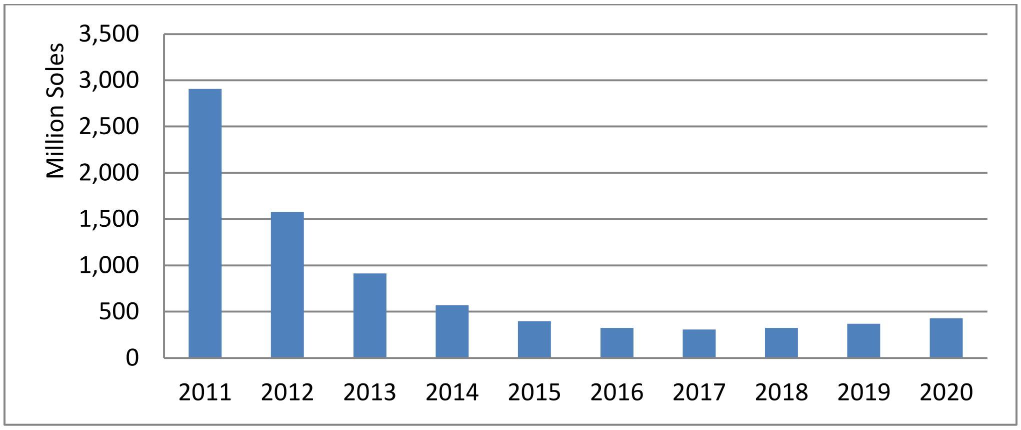
Source: author's estimates, section 3.3 and Table 12

The addition of these averted costs represents in this CBA the benefits of each policy in comparison with the BAU scenario. Typically for these evaluations, most of the benefits (98% on average across actions) accrued account for avoided mortality, reflecting the high value of VSL compared with the monetary costs of health care for morbidity.

4.3.2. Costs of proposed actions

Action A1: Based on the projection of new vehicles purchase (see Chapter 3, section 3.2.1.1) and the average price per vehicle, the net private cost of new vehicles is illustrated in Figure 16.

Figure 16. Net private cost of new vehicles by year (constant prices, reference year 2010)

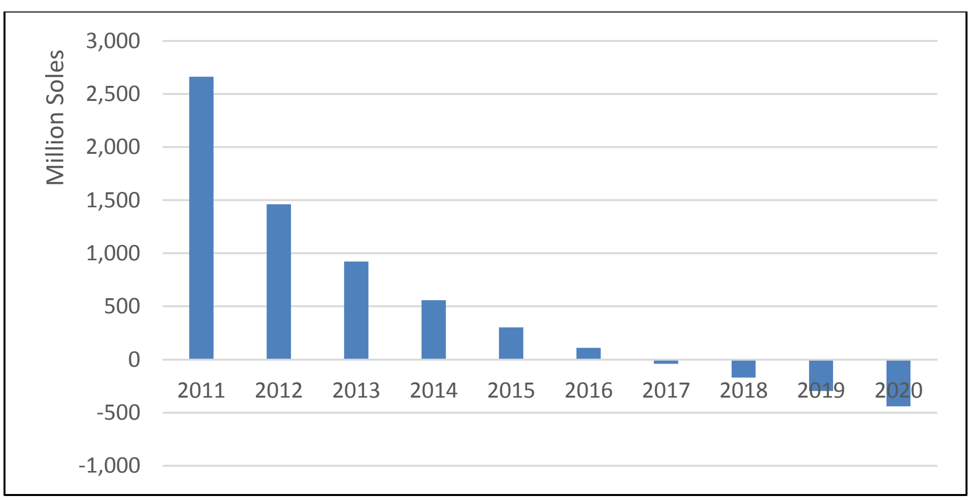


Source: author's estimates, see subsections 3.2.1.1 and 4.2.3

The undiscounted 10-year cost is 8,116 Million Soles. With a discount rate of 3% the present value of the 10-year cost is 7,156 Million Soles. Adding the allocated budget for this activity is 15 Million Soles for compliance enforcement and monitoring, to be spent at a rate of 3 Million Soles a year from 2011 to 2015, the final resulting costs are 8,131 Million Soles without discounting, and 7,169 Million Soles with a 3% discount rate.

Action A2: Under this action, private expenditures in bus purchases are highly positive in the first years of the timeframe, decreasing thereafter and becoming negative in the last four years – that is, more new buses are purchased in those years under BAU (see Figure 17 below). A standard price for a bus unit is assumed based on the value tables of the tax authorities, for a resulting overall 10-year cost of 5,069 Million Soles without discount, and 4,921 Million Soles with a 3% discount rate.

Figure 17. Private expenditure in A2 relative to BAU (constant prices, reference year 2010)

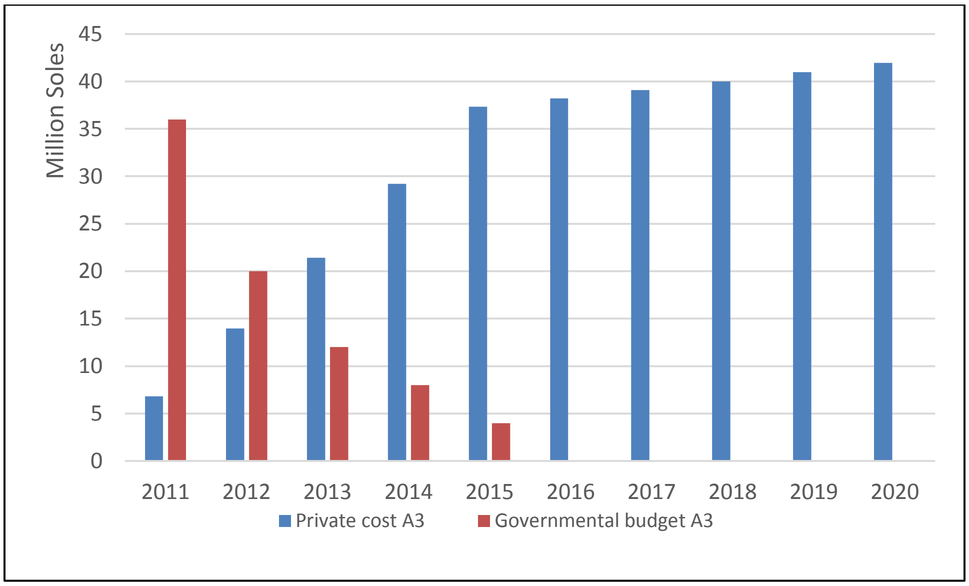


Source: author's estimates, see subsections 3.2.1.2 and 4.2.3

Adding the governmental budget allocations (see previous subsection), the total resulting cost for A2 is 5,122 Million Soles (no discount), or 4,984 Million Soles with a 3% discount.

Action A3: In this action, the proportion of the costs borne privately and by the government are relatively comparable (see Figure 18), with a governmental allocation of 80 Million Soles over five years and a private total cost around 309 million Soles over ten years, assuming an unchanging unit cost of VNG retrofitting (4400 Soles, see section 4.2.3). The final total cost of action A3 is nearly 339 Million Soles (no discount) or 331 Million Soles (3% discount).

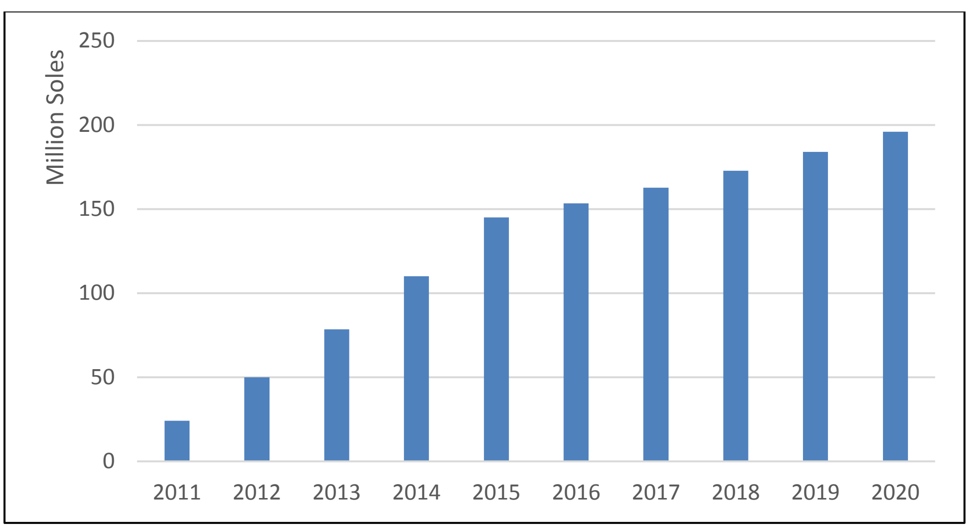
Figure 18. Private and public costs A3 compared to BAU, (constant prices, reference year 2010)



Source: author's estimates, see subsections 3.2.1.3 and 4.2.3

Action A4: Consistently with the increase in inspections, the private net cost of action A4 rises along with the proportion of the fleet subject to inspection (see Figure 19)

Figure 19. Private net cost of action A4, (constant prices, reference year 2010)



Source: author's estimates, see subsections 3.2.1.4 and 4.2.3

In addition, the governmental allocated budget for compliance and enforcement is listed in Table 18. The ten-year undiscounted total cost for action A4 is 1,280 Million Soles, and the cost under a 3% discount rate is 1,053 Million Soles.

Table 18. Governmental budget in Soles allocated to action A4, by year

Year	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Allocation	1,292,008	861,339	1,076,674	646,004	430,669	0	0	0	0	0

Source: (CGIAL, 2011a)

4.3.3. CBA outputs

We can compare now the costs and benefits of each action. In principle, these are meant to reflect the most important social costs and benefits, accrued over a period of ten years. The main indicators of the CBA under the assumption of no discounting are listed in Table 19.

Table 19. Cost–benefit analysis output indicators with no discounting

Action	A1 renewal	A2 scrapping	A3 VNG	A4 maintenance
PV Benefits	1,159,111,731	8,804,520,014	1,325,926,236	10,888,172,315
PV Costs	8,130,810,554	5,122,336,396	388,960,694	1,280,409,769
NPV	-6,971,698,823	3,682,183,618	936,965,542	9,607,762,546
BCR	0.14	1.72	3.41	8.50

The CBA results offer interesting insights as to the optimality of each action. In terms of value for each monetary unit invested, Action A1 is not cost-beneficial, with a BCR of well under 1.0. The rest of actions are all cost-beneficial, with A2 being clearly but not greatly cost-beneficial (BCR 1.72), a solid BCR of 3.41 for A3 and a large BCR of 8.50 for A4. These BCRs are not uncommon in air quality policy CBAs, mainly on account of the large social benefits brought about by mortality risk reductions, in turn highly dependent on the use of metric, VSL in this case.

In terms of NPV the prioritization changes, however, due to the differences in the scale of costs and benefits. A4 has the largest net benefits, followed by A2, then A3. A1 has negative benefits. The prioritization of the actions either according to BCR or to NPV does not change by applying a 3% discount rate (see Table 20) albeit the net benefits decrease substantially under discount.

Table 20. Cost–benefit analysis output indicators with a discount rate of 3%

Action	A1 renewal	A2 scrapping	A3 VNG	A4 maintenance
PV Benefits	954,270,891	7,109,570,898	1,063,664,449	8,930,817,722
PV Costs	7,169,438,445	4,983,578,276	330,855,598	1,052,880,692
NPV	-6,215,167,554	2,125,992,623	732,808,851	7,877,937,031
BCR	0.13	1.43	3.21	8.48

Whether the BCR or the NPV better reflects the priorities of the evaluating agency is largely a matter of choice. Ideally, the scope and scale of compared alternatives would be comparable enough for both metrics to provide comparable, if not necessarily identical, prioritizations.

In our case, given the small differences between the larges and smallest costs and benefits, BCR will be the primary output. In practical terms, the prioritization of actions can be interpreted as follows:

- Vehicle maintenance yields the largest benefits per money invested, largely due to the large scale of the transformation represented by adequate inspections of a vehicle fleet as that of the MALC; the large potential for emissions reductions heavily outweighs the additional cost incurred by the inspections. This balance may change slightly with the addition of the cost of time allocated to the inspection, but it is unlikely to radically affect the relative position of this action, given its high BCR value;
- Retrofitting existing buses with VNG capabilities also yields large returns per money invested; this has to do with the critical role of the public transportation fleet in the MALC regarding PM emissions. In addition, the cost of the action is relatively minor compared to the rest;
- Scrapping older public transportation vehicles is cost-beneficial, for the same reason mentioned for the VNG retrofitting; however, the cost of new vehicles purchases cancels out a substantial proportion of the accrued benefits; and
- The renewal of the private fleet is not cost-beneficial. This reflects various aspects: firstly, the relatively minor importance of the private fleet in the overall emissions in the MALC compared with the large and loosely regulated diesel-fueled public transportation fleet; secondly, the fast motorization rate under which the action takes places, which dilutes the effect of the measure; and thirdly, the large additional cost imposed by the increasing difficulty of finding "fresh" (i.e. imported) second hand cars in the market.

These insights may (and did) have some interest in their own right for the evaluating authorities; however, as far as this dissertation is concerned, they are of relative importance only, since the focus is on exploring the integration of the CBA and MCA analyses, elements or outputs.

4.3.4. Sensitivity analysis

Several possibilities can be analyzed in terms of the CBA results, and how they would change with variation in selected parameters.

An obvious one would be a change in the main determinant of accrued benefits, the point estimate of VSL, by replacing the average VSL with the minimum and maximum values of the VSL 90% confidence interval (730,798 - 4,638,458). This change would result in variations in the BCR of each action. The results are listed in Table 21 and 22.

Table 21. Sensitivity of benefit-to-cost ratios to VSL variation, under no discounting

Action ID	A1 renewal	A2 scrapping	A3 VNG	A4 maintenance
PV Benefits - low	512,097,276	3,889,849,958	585,796,171	4,810,410,624
PV Benefits - center	1,159,111,731	8,804,520,014	1,325,926,236	10,888,172,315
PV Benefits - high	2,789,164,963	21,186,274,008	3,190,569,900	26,200,156,482
PV Costs	8,130,810,554	5,122,336,396	388,960,694	1,280,409,769
BCR - low	0.06	0.76	1.51	3.76
BCR - center	0.14	1.72	3.41	8.50
BCR - high	0.34	4.14	8.20	20.46

Table 22. Sensitivity of benefit-to-cost ratios to VSL variation, under a 3% discount rate

Action ID	A1 renewal	A2 scrapping	A3 VNG	A4 maintenance
PV Benefits - low	421,598,290	3,141,018,933	469,928,526	3,945,648,472
PV Benefits - center	954,270,891	7,109,570,898	1,063,664,449	8,930,817,722
PV Benefits - high	2,296,257,439	17,107,726,133	2,559,490,629	21,490,183,575
PV Costs	7,169,438,445	4,983,578,276	330,855,598	1,052,880,692
BCR - low	0.06	0.63	1.42	3.75
BCR - center	0.13	1.43	3.21	8.48
BCR - high	0.32	3.43	7.74	20.41

The uncertainty brought about by the estimation of the VSL would not change their prioritization, although using the low end of the VSL estimate would render action A2 not cost-beneficial. In other words, the preferred order of the actions would remain, in so far as the same VSL variation is used for all actions, although their BCR absolute values vary heavily depending on the VSL used. The same, however, may not necessarily be true for a case where only morbidity, not mortality, would be counted in benefits. If that were the case, most or all of the actions may no longer be cost-beneficial. This, however, is not standard practice, mainly since it would no longer consider social benefits as a whole. In addition, avoided mortality was explicitly considered by the stakeholders a priority goal for the programs being evaluated. Therefore, CBA results have to consider avoided mortality in the monetization of benefits.

Chapter 5. Case study multi-criteria analysis

5.1. Introduction

The multi-criteria analysis problem at hand is the prioritization of a set of actions designed to reduce the concentrations of particulate matter in the Metropolitan Area of Lima and El Callao (MALC). The actions are detailed in Chapter 3 (Impact Assessment), but for clarity, they are listed again:

- A1: Private vehicle fleet renewal (used vehicles import ban and additional measures)
- A2: Bonus for scrapping older public transportation vehicles
- A3: Promotion of VNG use in public transportation vehicles
- A4: Enforcement of better vehicle maintenance

These actions are not mutually exclusive, and the purpose of the evaluation was not to discard any; that is why they are not designated “alternatives”. Rather, the decision-makers wished to have an idea of the trade-offs and potential benefits of each, so that their implementation could be prioritized or postponed in a context of challenging governance, competing public demands and constrained resources. The process of this MCA evaluation was long and complex, and for the sake of clarity only the most relevant information is presented. The structure of this chapter is as follows:

- Firstly, an overview of the MCA method used and its conceptual and mathematical basis is provided (section 5.2).
- Secondly, I explain the process by which I obtained the information necessary to apply the MCA method to this specific MCA evaluation (section 5.3).
- Lastly, I present the main results and briefly reflect upon their policy implications (section 5.4)

The results presented will feed into the combination of MCA and CBA results explored in Chapter 6.

5.2. Methods

The relative strengths and weaknesses of different MCA methodologies are explored in the literature review. In this particular case study, some constraints determined the choice of model:

1) The face to face interaction time with the stakeholders was quite limited;
2) The experience of stakeholders in evidence-based decision processes was almost negligible;
3) There were severe problems of data availability, specifically regarding means of public policy implementation at different institutional levels (national and local) and also regarding socio-economic status variables;
4) A final prioritization of the actions was a must for policy decision-making; and
5) The visualization of trade-offs was important for the stakeholders.

The scarce data availability and lack of experience of the stakeholders advised against the use of multi attribute utility theory methods, reliant in a rather quantitative and objective dataset, and to a lesser extent against AHP. Because of the time required by the pairwise comparison methodology, AHP could be ruled out. On the other hand, outranking models fit the problem at hand: they do not require that the set of criteria under consideration are commensurable, or that they are measured in the same unit. They are guaranteed to produce an ordering of the actions and, crucially, they provide a clear visualization of trade-offs between alternatives and relative performance in criteria.

The specifics of the outranking models and their underlying assumptions have been explored extensively, both theoretically (Boyssou, 1996) as well as in its applications (De Leeneer & Pastijn, 2002; Geldermann, Spengler & Rentz, 2001). Thus, only the information needed to understand the case study is presented here. After considering various options I opted for a widely applied and published one: Preference Ranking Organization METHods for Enrichment Evaluations (PROMETHEE) and Geometrical Analysis for Interactive Aid (GAIA) (Brans, 1982; Brans et al., 1986).

What follows is a description of the model and its application (subsections 5.2.1. through 5.2.5.), based on a collection of edited excerpts of explanatory papers by its creators (Brans, 1982, 1996; Brans et al., 1986) as well as in the operational manual of the related software PROMETHEE (Brans & Mareschal, 2012). Extensive information on every aspect of it is available in those publications but not reproduced here on account of conciseness.

5.2.1. The PROMETHEE-GAIA model

The PROMETHEE-GAIA method provides a structured framework for decision support in problems in which multiple criteria must be considered, for instance comparing various complex policy actions in terms of their performance across selected variables simultaneously. It belongs to the category of outranking methods (Velasquez & Hester, 2013), which generally could be described by the assertion "x is at least as good an action as y" if 1) a majority of the criteria supports this assertion (a condition known as "concordance"), and 2) the opposition of the rest of the criteria is "not too strong" (a condition known as "non-discordance"). The exact meaning of "not too strong" is determined within the specific method, usually meaning thresholds of preference and/or indifference. The structure of the preferences thus determines the overall form of the model and its application to decision problems. The conceptual basis for such structure, presented generically within the outranking models in subsection 2.3.1.1, is explored below.

5.2.2. Preferences in PROMETHEE-GAIA

An underlying model of individual preferences is necessary for the prioritization of actions under any multi-criteria analysis model, including outranking ones. Let us consider a finite set of m actions A, n real-valued function-criteria $g_1,\dots, g_n$ of which the importance for the decision-maker is represented by their respective weights. The weight of each criterion j is a positive real number noted w_j so that the higher the weight, the more important the criterion and the sum of all criteria equals the unit:

$$\sum_{j=1}^{n} w_j = 1 \tag{37}$$

That the weights add up to 1 is not strictly necessary in the PROMETHEE-GAIA method, albeit it is advisable, and since it is done in our case study, it is also assumed for the sake of the formalization of the methods here. Under these assumptions, every action i can be compared to any other i' based on the information provided by the established criteria, their weights, the performance of each action within each criterion and crucially, additional information regarding the preferences of the decision maker in the context of the given decision problem.

In the case of PROMETHEE, the additional information reflecting preferences is translated into information between the criteria (that is, weights) and information within each criterion. The latter refers to the underlying assumption that a decision-maker would assign a small preference to one action over other if, everything else being equal, the deviation of them over one criterion is small, and a larger preference if such deviation is large. This can be expressed mathematically:

$$P_j(i,i') = F_j\left[d_j(i,i')\right] \quad \forall i,i' \in A, \tag{38}$$

$$d_j(i,i') = g_j(i) - g_j(i') \tag{39}$$

$$0 \leq P_j(i,i') \leq 1 \tag{40}$$

Where:
$P_j(i, i')$ is the preference function between actions on criterion j,
$d_j(i, i')$ is the difference in performance between action i and action i' on criterion j,
$g_j(i)$ and $g_j(i')$ are the performances of action i and action i' on criterion j, respectively.

That is, the preference between actions i and i' with respect to criterion j is measured by $P_j(i,i')$, a function of the difference between the performance of i and i' in this criterion which takes values between 0 and 1. In the current version of PROMETHEE (version 1.3.1.0) there are six different types of preference functions available (see Table 23 below).

In the preference functions, thresholds can be defined for indifference and for preference between actions. The threshold of indifference q_j, a positive real number, is the largest value considered negligible (that is, it does not make a difference in the decision maker's preference) when comparing two actions in any given criterion.

If the difference in performance in one criterion between two actions is smaller than q_j the corresponding uni-criterion preference degree is equal to zero. The threshold of preference p_j, also a real positive number, is the smallest value of a difference considered decisive to prefer one action over other in any given criterion. These parameters are computed based on the views of the decision maker for each criterion. If the difference in performance between two actions in one criterion is greater than p_j the corresponding uni-criterion preference degree is equal to one (the maximum possible value). Parameter s is used only for the case of a Gaussian preference criterion, where the preference function remains increasing for all deviations and has no discontinuities either in its shape or its derivatives; s_j defines the inflection point of the preference function, and its value (to be decided by the analyst or the decision-maker) is by definition between p_j and q_j and it can be computed using linear interpolation.

Table 23. Types of preference functions in PROMETHEE

Name	Structure	Parameters	Tips for application
Usual	$P(d) = \begin{cases} 0 & \text{if } d \leq 0 \\ 1 & \text{if } d > 0 \end{cases}$	-	Simple optimization (the larger, the better). Recommended for quantitative and qualitative criteria where any improvement is considered relevant.
U-shape	$P(d) = \begin{cases} 0 & \text{if } d \leq q \\ 1 & \text{if } d > q \end{cases}$	q	Simple optimization principle holds (the larger, the better) but an indifference threshold q can be established. Recommended for quantitative criteria in which a relatively small difference may be considered negligible by the decision maker.
V-shape	$P(d) = \begin{cases} 0 & \text{if } d \leq 0 \\ \frac{d}{p} & \text{if } 0 < d \leq p \\ 1 & \text{if } d > q \end{cases}$	p	The V-shape preference function is a special case of the linear preference function where the q indifference threshold is equal to 0. It is thus well suited to quantitative criteria when even small deviations should be accounted for.
Level	$P(d) = \begin{cases} 0 & \text{if} \quad d \leq q \\ \frac{1}{2} & \text{if} \quad q < d \leq p \\ 1 & \text{if} \quad \text{if } d > p \end{cases}$	p, q	The level preference function is better suited to qualitative criteria when the decision maker wants to modulate the preference degree according to the deviation between evaluation levels.
V-shape with indifference	$P(d) = \begin{cases} 0 & \text{if} \quad d \leq q \\ \frac{d-q}{p-q} & \text{if} \quad q < d \leq p \\ 1 & \text{if} \quad \text{if } d > q \end{cases}$	p, q	The V-shaped with indifference preference is the best choice for quantitative criteria when a q indifference threshold is wished.
Gaussian	$P(d) = \begin{cases} 0 & \text{if } d \leq 0 \\ 1 - e^{-\frac{d^2}{2s^2}} & \text{if } d > 0 \end{cases}$	s	An alternative to the linear function where there is high uncertainty about the preference and indifference thresholds p and q, so it relies instead on a single threshold s. It is seldom used.

Source: adapted from (Brans & Mareschal, 2005, 2012)

In the present case study (prioritization of actions to reduce particle pollution in the MALC), all information regarding preferences, criteria and weights was provided by the stakeholders through successive interactions in the form of face-to-face interviews and a workshop. The information on performance of qualitative criteria was also obtained from the stakeholders (See section 5.3).

5.2.3. Rankings in PROMETHEE-GAIA

The main result of PROMETHEE is a ranking of the actions evaluated. In this case, there are two types of ranking: PROMETHEE I and PROMETHEE II.

The PROMETHEE I provides a partial ranking of the actions considered (P^I, I^I R^I)[7] based on pairwise comparisons of every action. The PROMETHEE I is based on two functions, or preference flows, denominated Phi+ (φ^+) and Phi- (φ-), or "positive preference" and "negative preference" flows, respectively. The positive preference flow $\varphi^+(i)$ measures how much an action *i* is preferred to the other *m*-1 ones. It is a global measurement of the strengths of action *i*. The larger $\varphi^+(i)$ the better the action. The negative preference flow φ-(*i*) measures how much the other *m*-1 actions are preferred to action *i*. It is a global measurement of the weaknesses of action *i*. The smaller φ-(*i*) the better the action.

The preference flows are defined as:

$$\varphi^+(i) = \frac{1}{m-1}\sum_{i'\in A}\pi(i,i') \tag{41}$$

$$\varphi^-(i) = \frac{1}{m-1}\sum_{i'\in A}\pi(i',i) \tag{42}$$

Where $\pi(i, i')$ is the PROMETHEE aggregated or multi-criteria preference index, defined as:

$$\pi(i,i') = \sum_{j=1}^{n} w_j P_j(i,i') \tag{43}$$

The resulting structure of preferences for option *i* is as follows:

- *i* is preferred to *i'*, that is, $iP^I i'$ iff $\begin{cases}\varphi^+(i) > \varphi^+(i') & \text{and} \quad \varphi^-(i) < \varphi^-(i'), \text{or} \\ \varphi^+(i) = \varphi^+(i') & \text{and} \quad \varphi^-(i) < \varphi^-(i'), \text{or} \\ \varphi^+(i) > \varphi^+(i') & \text{and} \quad \varphi^-(i) = \varphi^-(i');\end{cases}$

- *i* is indifferent to *i'*, that is, iIi' iff $\varphi^+(i) = \varphi^+(i')$ and $\varphi^-(i) = \varphi^-(i')$;

- *i* is incomparable to *i'*, that is, iRi' iff $\begin{cases}\varphi^+(i) > \varphi^+(i') & \text{and} \quad \varphi^-(i) > \varphi^-(i'), \text{or} \\ \varphi^+(i) < \varphi^+(i') & \text{and} \quad \varphi^-(i) < \varphi^-(i')\end{cases}$

The partial nature of PROMETHEE I matters for our purposes, since we want to be able to visualize actions with very different criteria performance profiles. For instance, action *i* can rank better than action *i'* in φ^+ and worse in φ-. Such a situation (denominated "incomparability" in outranking parlance) entails useful information, for it reflects trade-offs and flags difficult choices that the decision-makers may wish to consider further.

[7] P stands for preference, I for indifference and R for incomparability.

The PROMETHEE II complete preorder (P^{II}, I^{II}) is obtained considering a net outranking flow φ that combines φ^+ and φ- thus providing a single summary score:

$$\varphi(i) = \varphi^+(i) - \varphi^-(i) \tag{44}$$

The higher the net flow, the better the action. When considering PROMETHEE II, all actions are comparable. In this ranking, action i is preferred to action i' if it is preferred to it according to the net preference flow, with the following structure of preferences.

$$\begin{aligned} iP^{II}i' \text{ iff } \quad & \varphi(i) > \varphi(i'), \\ iI^{II}i' \text{ iff } \quad & \varphi(i) = \varphi(i'). \end{aligned} \tag{45}$$

The following properties hold:

$$-1 \leq \varphi(i) \leq 1, \tag{46}$$

$$\sum_{i \in A} \varphi(i) = 0. \tag{47}$$

The net outranking flow φ in (43) can be expressed as a weighted sum of simple indices in order to evaluate each criterion separately. Thus a "single criterion net flow" can be defined for option i, $\varphi_j(i)$ that defines the ranking of option i in criterion j as:

$$\varphi_j(i) = \frac{1}{m-1} \sum_{i' \in A} [P_j(i, i') - P_j(i', i)] \tag{48}$$

Using the definitions of positive and negative outranking flows given by (41) and (42), and the aggregated preference index (4), it can be shown that the net outranking flow (43) can be decomposed as:

$$\varphi(i) = \sum_{j=1}^{n} \varphi_j(i) w_j \tag{49}$$

While it is tempting to think of PROMETHEE II as superior, it can actually lead to more disputable results, especially in the presence of strongly conflicting criteria, since some information get lost in the process of obtaining the net flow. Therefore, both rankings should be seen as complementary.

5.2.4. Weight stability intervals in PROMETHEE-GAIA

Once the final preference structure (i.e. ranking) has been obtained, it is worth investigating the robustness of the results to variations in weights, since these are widely considered the most subjective element of the multi-criteria analysis. Instead of the usual sensitivity analysis, Mareschal (1988) proposed a procedure to determine the range of values for the weights leading to the same results.

In order to do this, stability intervals for the weights are defined. The researcher distinguishes among three types of "weight stability": full stability (where all actions are considered simultaneously); partial stability (where the stability of a subset of actions is studied, while the rest are included but disregarded); and subset stability (where the stability of a group of the best actions is studied in isolation from the rest). In our case of study only full stability is of interest, since only four actions are considered; partial and subset stability are relevant when a large number of actions are being prioritized.

Let us consider this definition: a MCA method is additive of order *r* if and only if *r* real valued functions exist of the form:

$$V_l(i) = \sum_{j=1}^{n} w_j V_{lj}(f_j(i), f_j(A)), \quad i \in A; l = 1, \dots, r, \qquad (50)$$

Such that for all $i, i' \in A$

$$\begin{cases} iPi' \text{ iff } \quad V_l(i) \geq V_l(i'), l = 1, \dots, r, \text{ (with at least one >)}, \\ iIi' \text{ iff } \quad V_l(i) = V_l(i'), l = 1, \dots, r, \\ iRi' \text{ Otherwise.} \end{cases} \qquad (51)$$

In this dissertation, the weight stability interval (WSI) process is applied to the PROMETHEE II method, which belongs to the group of MCA additive methods of order *r* = 1 and which provides a complete pre-order *(P, I)* in which all options are comparable. Therefore, the rest of the WSI methods description is restricted to this case, thus eliminating the subscript *l* for simplicity.

Let us then consider an initial set of weights $w = (w_1 \dots w_{j'} \dots w_n)$ normalized to 1. We can analyze the change in the initial *(P, I)* structure that would occur when applying new normalized weights $w^* = (w_1^* \dots w_{j'}^* \dots w_n^*)$ in which:

- Only the weight of criterion $w_{j'}^* = w_{j'}(1+\beta), \; -1 \leq \beta \leq 1/w_{j'} - 1$ is modified, and
- The structure of the rest remains stable, that is, $w_j^* = \alpha_{j'} w_{j'}, \; 0 \leq \alpha_{j'} \leq \frac{1}{1-w_{j'}}, \; \forall j \neq j'$.

It can then be demonstrated that $\alpha_{j'} = \frac{1-(1+\beta)w_{j'}}{1-w_{j'}}$. A weight stability interval under the full stability condition can be then defined as the interval of values $[w_{j'}^-, w_{j'}^+]$ that does not alter the initial preference structure *(P,I)*.

If the actions are initially not indifferent, that is, iPi' or $i'Pi$, then, the minimum value $w_{j'}^-$ and maximum value $w_{j'}^+$ can be calculated as:

$$w_{j'}^- = 1 - \alpha_{j'}^+(1 - w_{j'}) \quad \text{and} \quad w_{j'}^+ = 1 - \alpha_{j'}^-(1 - w_{j'}) \qquad (52)$$

where $\alpha_{j'}^+$ y $\alpha_{j'}^-$ are, respectively, the maximum and minimum values of $\alpha_{j'}$ for the stability:

$$\alpha_{j'}^{+} = \min_{(i,i')\in \Omega^+} \frac{\Delta(i,i')\,\Delta_{j'}(i,i')}{\Delta(i,i')\,\Delta_{j'}(i,i') - \Delta^2(i,i')} \quad \text{and} \quad \alpha_{j'}^{-} = \max_{(i,i')\in \Omega^-} \frac{\Delta(i,i')\,\Delta_{j'}(i,i')}{\Delta(i,i')\,\Delta_{j'}(i,i') - \Delta^2(i,i')} \qquad (53)$$

These values are a function of:

- The difference in the initial total performance of the actions $\Delta(i,i') = V(i) - V(i')$. In PROMETHEE II, for instance, the difference is calculated as $\Delta(i,i') = \varphi(i) - \varphi(i')$ (with the initial weights *w*). If the actions are initially indifferent (iIi'), then $\Delta(i,i') = 0$ and if one action is preferred to other (iPi' or $i'Pi$), then $\Delta(i,i') \neq 0$.
- The initial difference in performance of each pair of actions within the considered criterion, that is, $\Delta_{j'}(i,i') = V_{j'}(i) - V_{j'}(i')$. In PROMETHEE II, for instance, that difference is the difference in the single criterion net flow, $\Delta_{j'}(i,i') = \varphi_{j'}(i) - \varphi_{j'}(i')$, and it does not depend on the weights *w*.
- The set of pairs of actions Ω^+ for which criterion g_c is highly concordant with the general ranking for those actions, defined by the condition
  $\Omega^+ = \{(i,i') \epsilon\, A\times A,\ s.t.\ \Delta(i,i')\,\Delta_{j'}(i,i') > \Delta^2(i,i')\}$
  As the weight of g_c decreases and the rest of the weights increase ($\uparrow \alpha_{j'}$), the preference order can change. Thus, with Ω^+ an upper limit $\alpha_{j'}^{+}$ can be calculated for the preference stability.
- The set of pairs of actions Ω^- for which criterion g_c is in discordance with the general ranking for those actions, defined by the condition: $\Omega^- = \{(i,i') \epsilon\, A\times A,\ s.t.\ \Delta(i,i')\,\Delta_{j'}(i,i') < 0\}$. As the weight of g_c increases and the rest of the weights decrease ($\downarrow \alpha_{j'}$), the preference order can change. Thus, with Ω^- a lower limit $\alpha_{j'}^{-}$ can be calculated for the preference stability.

If the actions are initially indifferent, that is, (iIi'), the preference order is stable for any w^* if the actions are also indifferent in criterion *j'*, that is, $iI_{j'}i$. If the indifference does not hold, then the interval of stable weights is restricted to w, and any change in weights alters the preference order. The set of pairs of actions meeting this requirement is designated Ω^0, and defined as: $\Omega^0 = \{(i,i') \epsilon\, A\times A,$ $s.t.\ \Delta(i,i') = 0 \text{ y } \Delta_c(i,i') \neq 0\}$

5.2.5. The GAIA plane

The GAIA plane is a graphical representation of the multi-criteria decision problem as addressed by the PROMETHEE methods. It is important because it reflects visually conflicts between criteria and thus potential trade-offs in final decisions. It is based on the notion that the information relative to a decision problem including *n* criteria can be represented in an n-dimensional space. Let us consider the (*m* × *n*) matrix *M* of net flows by criterion and action, where each column represents the set of values of a single criterion net flow $\varphi_j(i)$ for all the actions. Matrix *M* represents a pool of information in which each action *i* is a point in a space of *n* dimensions (as many as criteria analyzed). The GAIA plane is the plane for which as much information as possible is preserved after projection of the *n*–dimensional space on the plane. The principal components analysis is used. So, the GAIA plane is defined by the two eigenvectors corresponding to the two largest eigenvalues of the covariance matrix of the single criterion net flows.

Since some information gets lost in the projection, if the information preserved is high enough (a consideration to be judged by the analyst) the following properties hold (Brans & Mareschal, 2012):

1. The longer a criterion axis in the GAIA plane, the more discriminating this criterion.
2. Criteria expressing similar preferences are represented by axes oriented in approximatively the same direction.
3. Criteria expressing conflicting preferences are oriented in opposite directions.
4. Criteria that are not related to each other in terms of preferences are represented by orthogonal axes.
5. Similar actions are represented by points located close to each other.
6. Actions being good on a particular criterion are represented by points located in the direction of the corresponding criterion axis.

With regard to the role of the weights in this projection, since the PROMETHEE net flow of an action *i* is the scalar product between the vector of its single criterion net flows $\left(\varphi_1(i), \dots, \varphi_n(i)\right)$ and the vector of weights $w = (w_1, \dots, w_k)$., then this means that the PROMETHEE net flow of *i*, $\varphi(i)$, is the projection of the vector of its single criterion net flows on w. Thus, the relative positions of the projections of all the actions on w provide the PROMETHEE II ranking. The projection of w on the GAIA plane is denominated by the authors as the "PROMETHEE decision axis". The PROMETHEE decision axis appears as a weighted resultant of all the criterion axes. When the weights are modified, the positions of the actions and of the criteria remain unchanged in the GAIA plane. The weight vector appears as a decision stick that the decision-maker can move according to his preferences in favor of particular criteria.

The importance of the GAIA plane, the decision axis and the decision stick lie in the possibilities for visualization of a sensitivity analysis: when weights are modified, the decision stick and the decision axis provide clues regarding the trade-off consequences of decision-making.

5.3. Ascertainment of criteria, weights and preferences

The set of criteria, or attributes, needed for a MCA can be obtained in several ways. Frequently, discussion rounds of expert panels are used (Achillas et al., 2011; Vlachokostas et al., 2011) or the research team proposes a set (Tzeng et al., 2002; Wang, Jing, Zhang, & Zhao, 2009). However, researchers frequently recognize that the criteria should ideally be proposed by the end users of the system, that is, the stakeholders. The reasons why this is seldom the chosen strategy has to do mainly with the inherent difficulties and time requirements of participatory processes.

Despite such practical difficulties we (the CGIAL, the World Bank and I) chose to consult stakeholders on the criteria because the opportunity to interact with them through interviews, workshops and surveys was readily available. In addition, we felt that this strategy added validity to the process for the prospective end-users of the decision-support systems. We consulted a limited sample of stakeholders through face-to-face interviews, a workshop and an online questionnaire. This process was used to derive the list of criteria and weights for the MCA problem of our case study.

5.3.1. Interviews with key stakeholders

The technical coordination unit of CGIAL (Comité Gestor de la Iniciativa de Aire Limpio, the policy promoting agency) identified a sample of 30 key stakeholders, including elected politicians, appointed officials and civil servants at various levels, industry association representatives, experts, media and consumer association representatives. They were all requested to assist the CGIAL through their participation, with the following objectives:

1. Ensuring that all of them were aware and understood the relevant policy process, the prioritization exercise, their expected inputs and the expected outputs and outcomes of the project;
2. Individually first, and then collectively, ascertaining criteria they considered important for the prospective evaluation and/or prioritization of the proposed policies; and
3. Identifying sources of relevant data and information, as well as other resources, which could be useful in this and other forthcoming policy processes.

All 30 stakeholders agreed to assist in the process, and related activities started shortly thereafter. They were asked to participate in an individual, face to face interview in order to collect baseline qualitative information, including criteria for an MCA evaluation. The type of interview was a semi-structured one, commonly used in social sciences to collect qualitative data by using open-ended questions on a particular subject. It is used when the aim is to understand a point of view, rather than making generalizations (Gregory & Brierley, 2010; Naylor & Appleby, 2013), and it is suitable for small samples. An interview script (available in Spanish in Annex 2) was developed for an average duration of about 45 minutes. All interviews were recorded and the selection of topics was e-mailed to the interviewees in advance, along with a reminder of day, time, place and the name of the interviewer. From a total of 30 planned interviews, 26 were conducted in the months of June and July of 2012, of which 24 were validated (i.e. the interview had been administered in person by a trained interviewer; all questions had been answered and the interview had been recorded) and processed. Two interviews had been transcribed but not recorded, so they were not included in the sample. The topics explored were:

- Specific roles of competencies regarding air quality management and with regard to the current policy process in particular.
- Opinions on the current situation, trajectory and future perspectives of air quality and its management in the area of Lima-Callao.
- Perceived barriers and success factors for an effective air quality management in the area of Lima-Callao.
- Criteria for the evaluation and/or prioritization of air quality management policies.
- Relevant data sources and other resources for the current and forthcoming policy processes.

Albeit all results from the interviews were interesting on their own, the only results highlighted in this dissertation are for those related to the criteria for the MCA for the sake of conciseness. Through the criteria-specific questions, a preliminary list of 96 criteria was obtained. Several of them were identical, so we could narrow the list down to 26 different criteria (see Table 24 in next subsection) without losing substantial information. We categorized those 26 criteria into six categories: 1) Environmental, 2) Health, 3) Economics, 4) Social, 5) Political and 6) Implementation. Each criterion was assigned a "label" (that is, a descriptive name) and a tentative short description based on the interviewees responses.

5.3.2. Workshop with key stakeholders

After all the interviews were completed, all contacted stakeholders were gathered for a workshop on the matter. All 30 stakeholders were present, and the whole process was explained again (some had received the information but had not been interviewed). The results of the interviews were presented to them, including the original and refined list of criteria. They agreed to the downsizing of the original list, as well as the categories chosen. We then proceeded to the ranking of the criteria according to their importance for the stakeholders. The list of criteria was presented and their tentative definition explained to ensure a common understanding of the meaning of each criterion. The criteria were then evaluated via a "traffic light" system: a set of three cards (green, amber and red) was given to each stakeholder, with the following color key:

- Green: "this is a key criterion, in my opinion";
- Amber: "this is a relevant criterion, but does not categorize as *key*";
- Red: "this is not a relevant criterion."

Each color was assigned a numerical value: 3 for green, 1 for amber and 0 for red, with the leap between the numerical values of green and red reflecting the comparative importance of a "key" criterion. The use of color-coded cards for weight elicitation has many variations Figueira & Roy, 2002; Pictet & Bollinger, 2008) on which this system was inspired. This exercise served at once to prioritize the criteria (with a view to downsizing the list to a manageable number) and to assign a weight to those criteria. The weight of a given criterion is a positive number that represents the relative importance of that criterion. Weights can be obtained in several ways; the most common include: the analysts' (researchers') expert judgment; the judgment of a panel of external experts, stakeholders, or both; a survey of stakeholders affected by the actions evaluated; and computer models, among others. There is surprisingly little discussion in the articles featuring the application of MCA to air quality policy regarding the advantages and disadvantages of different weighting strategies (Tzeng et al., 2005, 2002; Vlachokostas et al., 2011). Indeed, even the mechanisms through which discussions turned into weight estimates are absent in several available studies. More generally, there is a considerable amount of literature regarding weight elicitation techniques, including extensive reviews (Riabacke, Danielson, & Ekenberg, 2012). In a preparatory meeting with the policy-promoting agency (CGIAL), we explored the possibility of utilizing a complete standardized weight elicitation model like SWING (von Winterfeld & Edwards, 1986) or TRADE-OFF (Keeney & Raiffa, 1993), but the commission concluded these were too complex and cumbersome, and advised instead to use a simple points-based system resulting in ratio-based weights. In that context, the "traffic light" evaluation exercise is at least a transparent methodology that the stakeholders agreed to. The mechanics of the exercise were as follows:

- Presentation of category, by the facilitator;
- Presentation of each criterion, by the facilitator;
- Evaluation of each criterion, by the stakeholders.

The stakeholders were then (virtually) divided into three groups of ten people each and a team member counted green, amber and red cards within each group for each criterion.

The point count of each potential criterion is calculated as:

$$Point\ count = (N_{green} \times 3) + (N_{amber} \times 1) + (N_{red} \times 0) \quad (54)$$

Where:
N_{green} is the number of green cards,
N_{amber} is the number of amber cards,
N_{red} is the number of red cards.

The results of the point count are listed in Table 24 (below).

Table 24. Categories, criteria and point evaluation of their importance according to stakeholders

Categories and potential criteria	Code	Weights as Points
Environmental		
Reduction in PM_{10} emissions	AM1	85
Reduction in people's exposure to particle pollution	AM3	72
Reduction in the emissions of other pollutants	AM2	53
Economics		
Economic benefit concentrates in the poor	EC7	59
Cost of implementation (budgetary)	EC2	50
Economic incentives available	EC4	48
Available financing	EC5	48
Economic benefit for all society	EC1	46
Net gain of person-hours due to reduced travel times	EC3	46
Internal rate of return of the measure	EC6	35
Implementation		
Fast implementation	IM4	83
Ease of enforcement	IM7	69
Maximize number of affected sources	IM5	58
Ease of progress monitoring	IM2	55
Improvement of competitiveness	IM3	55
Ease of implementation	IM1	47
"Bottleneck" measures (allow others to occur)	IM6	38
Immediate real possibilities of enforcement	IM8	36
Political		
Acceptance of the measure by general public	PO1	50
Acceptance of the measure by affected sectors	PO2	40
Health		
Reduction of illness attributable to air pollution	SA3	85
Reduction of mortality attributable to air pollution	SA2	79
Reduction of traffic accidents	SA1	66
Social		
Promotion of social inclusion	SO5	60
Benefit to vulnerable groups	SO4	55
Creation or recovery of urban public spaces	SO1	45
Creation of employment	SO2	36
Promotion of economic activities	SO3	28

After the ranking rounds were over, data were collected into a spreadsheet and presented in summary graphs and tables to the present stakeholders. A brief discussion on the context is in order to understand the wording of some of these criteria; below is a brief explanation on some of the most ambiguous:

- The criteria in the "Environmental" category are unequivocal and self-explanatory
- In the "Economics" category, some criteria deserve additional context explanation:
  - Economic benefit concentrates in the poor: in this criteria, the concern of the stakeholders was whether the action maximized net benefit to the population groups with lowest socio-economic status (i.e. minimized the cost externalized to them and provided them with the highest potential for economic benefit) either through savings (e.g. in fuel expenditures) or through increased income (e.g. through the ability to engage in revenue-generating activities related to transport). This was of concern to the stakeholders because of the large proportion of the low socio-economic status population of the MALC that live off the fully deregulated (and inefficient through oversupply) public transportation industry.
  - Cost of implementation (budgetary): in this criterion, the cost of implementation for the government was the concern for the stakeholders. Asked whether they would not rather consider the cost for all society, they confirmed that the allocation of governmental resources in this context was much more important for them in this context. This is noteworthy, since the stakeholders included non-governmental actors, and matters significantly for the interpretation of both the MCA and CBA results.
  - Economic incentives available: this criterion relates to the sectoral politics of transportation in the MALC, where the availability of subsidies in various forms, grants, and low- or no-interest loans matters greatly for both the uptake of the measures by the sector and, crucially, the non-opposition of unions and industry interest groups.
- In the "Implementation" category, some criteria deserve additional context explanation:
  - Fast implementation: the stakeholders were concerned, based on previous experiences, about whether the measures could be implemented within reasonable timeframes and would rather favor actions that could get done reasonably fast.
  - Ease of enforcement/ Ease of implementation/ Ease of progress monitoring: these criteria reflected the stakeholders' concern about the real possibilities of governmental actors to enforce compliance and implementation of the actions, and monitor progress, given the still relatively weak governance structures regarding environmental and energy matters in the country and in the city. Other equally ambiguous criteria in this category reflect this concern.
- The criteria in the "Political" category are self-explanatory, if rather subjective.
- The criteria in the "Health" category are unequivocal and self-explanatory.
- In the "Social" category, some criteria deserve additional context explanation:
  - Promotion of social inclusion/ Benefit to vulnerable groups/ Promotion of economic activities: these criteria reflected the governmental and non-governmental actors' willingness to favor actions that facilitated "inclusion", then understood as a catch-all label reflecting the priorities of the government at the time. When pressed for more precision, the stakeholders agreed that it meant basically the facilitation of the access of slum dwellers to employment opportunities, and of people employed informally into the formal sector.

Once the list of criteria was collected, it became clear that it was simply too large and overlapping for an operational MCA. In order to rationalize the list of criteria, the stakeholders were presented with desirable characteristics that a list of MCA criteria should ideally have: 1) Manageability: too many criteria make MCA complex, thus obscuring the process. Most MCA include from five to fifteen criteria; 2) Independence: criteria should be independent from each other, in order to include the maximum amount of explained variability without expanding the model unnecessarily; and 3) Measurability: some criteria might not have quantitative data readily available. While that is not a problem in itself (qualitative information is also valuable) it may strongly influence the interpretation of the results of the analysis.

In the ensuing comprehensive discussion, the stakeholders extensively discussed the groups of criteria and decided the following by consensus:

- The category "Environmental" could be eliminated altogether since all the relevant information is summarized in the "health" category and environmental concerns in this context are a proxy for health ones. Reductions of PM_{10} (AM1) and of other pollutants (AM2), and population exposure (AM3) were not deemed by the stakeholders important *per se*, but because of their potential health benefits. Consistently, criterion IM5 (maximize number of sources affected) was also considered to be summarized into the criteria of the "health" category.
- In the category "Economics" the criteria EC4 (Economic incentives) and EC5 (Available financing) referred to the same overall concept and should be merged into one, and EC1, EC3 and EC6 should be dropped because they were considered by the stakeholders to be of limited relevance in this case (even after noting the decision-making significance of consciously dropping social costs out of the evaluation). EC7 (Economic benefit concentrates in the poor) is included into SO4 (Benefit to vulnerable groups).
- In the category "Implementation" criteria IM2 (ease of progress monitoring) and IM8 (immediate real possibilities of enforcement) were, in the view of the stakeholders, to a large extent described by IM7 (ease of enforcement). They also thought that criteria IM4 (Ease of implementation) and IM1 (Fast implementation) relate to the same topic and can be merged into one, and IM6 ("Bottlenecks") should be dropped off.
- "Political" should not be a standalone category, since it is mostly included in "Social".
- In the category "Health" criterion SA1 (Reduction of traffic accidents) should be dropped because it is not a matter dealt with by health or environmental authorities. If there is a way to combine all health benefits into one unified metric, it should be done for clarity (The standard "Burden of Disease" measured in DALYs -Disability-Adjusted Life Years, a measure of the burden of disease attributable to a risk factor- was found acceptable).
- The category "Social" should include only criteria PO1 (Acceptance of the measure by general public), and SO5 (Promotion of social inclusion). SO5 is deemed to be highly related to SO4 (benefit to vulnerable groups) and EC7, so only SO5 is to be kept. The rest of criteria are to be dropped off.
- The weight assigned to any merger of criteria should be the maximum weight among the weights of all the individual criteria merged.

It is worth noting that these decisions strongly determine the comprehensiveness of the final multi-criteria analysis. Some of the criteria dropped or merged have important societal relevance. For instance, the elimination of EC1 ("Economic benefit for all society") essentially means obviating the social economic benefit as it would be captured by a CBA.

That the stakeholders would rather consider the governmental costs than social cost in the analysis is very significant, since the idea of not considering private costs in the analysis undermines the notion of benefiting vulnerable groups or favoring actions of which the economic benefit concentrates in the poor. This perspective seems more fitting to an analysis strictly from the organizational point of view of the government where impacts on the administration are at the forefront. Given the large representation of non-state actors in the sample (about two thirds), this may seem counterintuitive[8]. Similarly, considering EC7 ("Economic benefit concentrates in the poor") outside the "Economic" category and into the "Social" strips it from its welfare economics component.

The societal perspective is thus partially absent from the MCA of this case study. However, albeit clearly advisable from the collective and public policy decision-making perspective, such societal perspective is not, strictly speaking, a formal requisite for a MCA. Moreover, this particular MCA process was requested by the implementing agency to be as stakeholder-driven as possible, in a position ostensibly geared towards ensuring representation of societal interest in decision-making; thus the analyst should clearly avoid influencing something as fundamental as the choice of criteria.

Finally, an agreement on a definitive set of criteria and categories was reached (see Table 25, below).

Table 25. Final set of criteria and weights

Categories	Code/s	Criteria
Health	SA2/SA3	Reduction of PM_{10}-related burden of disease (DALYs)
Social	SO5/ SO4/EC7	Promotion of social inclusion
	PO1	Acceptance of measure by general public
Implementation	IM2/IM7/IM8	Monitoring and enforcement
	IM4	Speed of implementation
Economics	EC2	Cost of implementation
	EC4/EC5	Economic incentives (including financing)

Some of the criteria were quantitative in nature and their performance could be measured in standard units through available data; such was the case for the reduction of PM_{10}-related burden of disease, speed of implementation, and cost. The rest were qualitative in nature, and it was agreed that the performance of the analyzed actions against these criteria would be measured through a questionnaire to be sent to the stakeholders.

[8] A possible explanatory hypothesis for the alignment of stakeholders in this case with pro-governmental positions could be the "revolving door" (i.e. the movement of personnel) between regulators and the regulated industries and associations, a reportedly widespread phenomenon in Peru; however, this is a mere conjecture.

5.3.3. MCA model parameters

Several parameters defining criteria, weights and preferences must be specified in the MCA model, based on the input from the stakeholders. Below are the specifications and rationales for these.

Criteria specification

The criteria indicators, measurement units and groups were agreed upon with the stakeholders in the workshop. They are presented in Table 26, below.

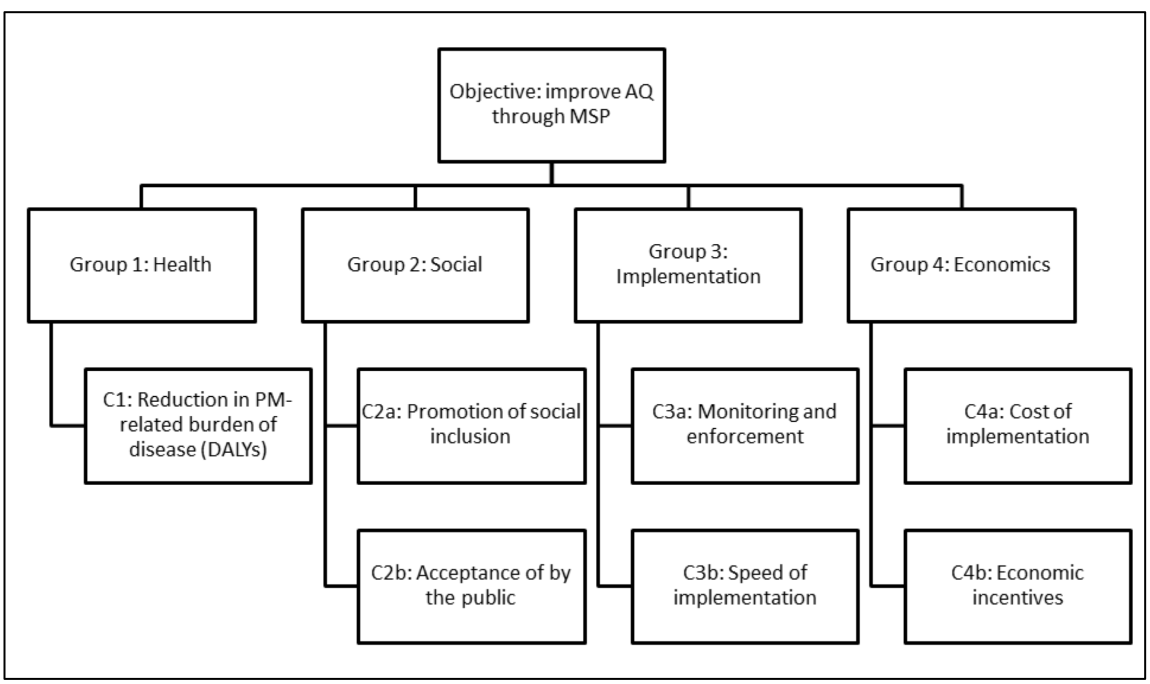
Table 26. Criteria units

Group	Criteria indicators	Units
Health	Reduction of PM10-related burden of disease (BoD)	DALYs avoided
Social	Promotion of social inclusion	N/A*
	Acceptance of measure by general public	N/A*
Implementation	Monitoring and enforcement	N/A*
	Speed of implementation	Years
Economics	Cost of implementation	Million Soles
	Economic incentives (including financing)	N/A*

*unit-less Likert psychometric scale (1: min – 5: max); thus "not applicable" (N/A)

The criteria and categories established by the stakeholders can be readily translated into a working hierarchy for a MCA (see Figure 20), with an overarching objective (Improving the air quality situation in Lima-Callao through mobile source policies), and the criteria organized into four groups (Health, Social, Implementation and Economics) equivalent to the previously established criteria categories. This is not required by the model, but can provide valuable insights in the analysis.

Figure 20. Hierarchy and criteria groups for the MCA



Source: author, based on stakeholders' inputs. PM: particulate matter. DALYs: Disability-Adjusted Life Years.

Weights specification

Ideally, weights in PROMETHEE should add up to the unit, and be in the format of percentages, or proportions of a unit. However, the weights obtained from the stakeholders (CS_j) are in the format of point scores from the traffic light exercise. To transform them into proportions, the point score is simply divided by the sum of all criteria scores:

$$w_j = \left(\frac{CS_j}{\sum_{j=1}^{n} CS_j} \right) \tag{55}$$

It could be argued that this transformation adds a layer of uncertainty to the original subjectivity of the weighting exercise; the weight assessment by the stakeholder could have been different, had they been aware of this particularity of the model. However, the actions for weight assessment from stakeholders available in the literature required a time available for interaction that was simply not available in this case. The final criteria weights are listed in Table 27, below.

Table 27. Criteria and their weights, as points and as a proportion of 1

Code	Criteria	Weight as points	Weight (adding up to 1)
C1	Reduction of PM-related burden of disease	85	0.1910
C2a	Promotion of social inclusion	60	0.1348
C2b	Acceptance of measure by general public	50	0.1124
C3a	Monitoring and enforcement	69	0.1551
C3b	Speed of implementation	83	0.1865
C4a	Cost of implementation	50	0.1124
C4b	Economic incentives (including financing)	48	0.1079

Preferences specification

A preference function must be specified for each one of the criteria in the MCA problem at hand; whether the objective is maximization or minimization must also be specified. The concept of preferences and thresholds has been discussed previously in this methodology section. Regarding the question of the direction of optimization (minimize/maximize), the stakeholders agreed unanimously that all criteria were to be maximized, except for "speed of implementation" and "cost". In order to ascertain the preferences, we went jointly during the workshop over each criterion asking the key questions suggested by the PROMETHEE model itself: 1) "do you think that the minimum difference in performance between two actions in this criterion is negligible?" and 2) "is the absolute difference in performance between two actions more important than the relative difference?"

The stakeholders responded respectively "no" and "yes" to the questions for all criteria. That is, they considered any difference significant, and a higher score better in any case, which translates into a choice of "Usual" for the preference type in PROMETHEE. However, given the stakeholders' scarce experience in evidence-based processes, I proceeded afterwards to evaluate independently the soundness of those choices for each criterion, using the PROMETHEE software assistant. The resulting choices of preference functions for each criterion were the same as those specified by the stakeholders. They are presented in Table 28, below.

Table 28. Criteria preference functions choice and rationale

ID	Criteria	Minimum difference negligible?	Absolute more important?	Optimization	Preference function and rationale
C1	Reduction of PM10-related burden of disease	NO (negligible amounts of lost DALYs entails clear ethical issues)	YES	Maximization	Usual: more is better, and a small difference could be considered significant
C2a	Promotion of social inclusion	NO (a small difference in a sample average could be important)	YES	Maximization	Usual: more is better, and a small difference could be considered significant
C2b	Acceptance of measure by general public	NO (a small difference in a sample average could be important)	YES	Maximization	Usual: more is better, and a small difference could be considered significant
C3a	Monitoring and enforcement	NO (a small difference in a sample average could be important)	YES	Maximization	Usual: more is better, and a small difference could be considered significant
C3b	Speed of implementation	NO (any difference in implementation time seemed to matter to the stakeholders)	YES	Minimization	Usual: less is better, and a small difference could be considered significant
C4a	Cost of implementation	NO (any difference in cost seemed to matter to the stakeholders)	YES	Minimization	Usual: less is better, and a small difference could be considered significant
C4b	Economic incentives (including financing)	NO (a small difference in a sample average could be important)	YES	Maximization	Usual: more is better, and a small difference could be considered significant

There is no reason to suppose that most criteria in this MCA should be governed by any preference function other than "Usual", thus discarding the need to specify indifference and preference thresholds.

5.3.4. Performance evaluation of actions

Quantitative criteria (C1 and C4a)

Criteria C1 (Reduction of PM10-related burden of disease) is a quantitative input that required ad hoc modeling based on the health impact assessment featured in Chapter 3. Details of the calculation can be found there (specifically, the translation of attributable mortality and morbidity into Disability-Adjusted Life Years, or DALYs). First, a baseline situation was drawn regarding the status of inhalable particle pollution in the MALC. Then, an attributable reduction of PM emissions was calculated for each policy package as a basis for the estimation of reductions in environmental concentrations of the pollutant. These were calculated by the Clean Air Institute in the vehicle fleet and emission factors model referred to in the calculation of benefits of the CBA. After emissions reductions are extrapolated into changes in exposure, attributable mortality and morbidity were calculated through established assessment methodologies (Cohen et al., 2004; Ostro, 1994). The result is an estimation of the lost Disability-Adjusted Lost Years (DALYs) avoided through each specific policy action, when compared over ten years to a "business as usual" scenario. The information on Costs (criterion C4a, referring exclusively to governmental budgetary cost) was obtained from the authorities, and the assumptions and caveats considered for the costing are explained in the CBA chapter. The units of the criterion C4a are Millions of Peruvian Soles of 2010.

Qualitative criteria (C2a, C2b, C3a, C3b and C4b)

Qualitative assessments of criteria C2a, C2b, C3a, C3b and C4b were obtained through an online survey completed by the stakeholders. The full survey is available (in Spanish) in Annex 3. The results are also available in the supplemental materials to this dissertation. The survey was developed and administered through the online platform Survey monkey ® (http://www.surveymonkey.com/) and sent to the stakeholders via e-mail. It was sent to the full list of 30 stakeholders that had participated in the stakeholder workshop, and 24 responded – a 80% response rate.

The introduction of the survey explained its purpose, then each one of the scenarios was briefly explained and the respondent was asked to evaluate the performance of each action regarding each one of the criteria in a verbal Likert-style psychometric scale of five points (Completely disagree/Disagree/Neither agree nor disagree/Agree/Completely agree). The utilization and value of psychometric scales in policy evaluation and countless other areas has been explored extensively elsewhere (Allen & Seaman, 2007; Laerhoven, Zaag-Loonen & Derkx, 2004; Lee & Jones, 2002). The verbal responses were then translated into a numerical score from 1 to 5, which constitutes the scale of criteria C2a, C2b, C3a and C4b. Criterion C3b (speed of implementation) is measured in years as judged by the respondent.

5.4. Results

5.4.1. Problem statement

The MCA problem is represented in Table 29. From here onwards this original MCA specification will be denominated Scenario 0. In the columns are the MCA criteria, as decided by the stakeholders. In rows, the preference functions of the criteria, the preference and indifference criteria p_j and q_j (not applicable in this case, since all criteria are "usual" preference form), units of the criteria, the definition of criteria optimization, criteria weights as a proportion of one, and the evaluations of the actions to be prioritized. In the evaluations of the actions are the values obtained by measuring each action according to each criterion. Green figures signal "best performing" and red "worst performing" within each criterion.

Table 29. Original MCA problem, with criteria, weights and performance of each option

Criterion	BoD C1	Inclusion C2a	Pub.op C2b	Enforce. C3a	Implem. C3b	Cost C4a	Incentives C4b
Preference fn.	Usual	Usual	Usual	Usual	Usual	Usual	Usual
p_j, q_j	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Units	DALYs	N/A	N/A	N/A	Years	Million S./	N/A
Optimization	MAX	MAX	MAX	MAX	MIN	MIN	MAX
Weights	0.1910	0.1348	0.1124	0.1551	0.1865	0.1124	0.1079
A1 renewal	7263	1.80	1.60	3.80	5.77	15.00	2.40
A2 scrapping	55169	2.60	3.80	2.40	6.63	53.96	2.00
A3 VNG	8308	4.20	4.40	3.80	6.36	80.00	4.80
A4 maintenance	68226	2.40	3.00	3.00	6.51	4.31	2.40

Green: best performing; Red: worst performing

Action “A3 VNG” attains the highest scores in several criteria (4 out of 7) albeit it is the worst performer in the cost of implementation (i.e. for the government). On the other extreme, action “A2 scrapping” does not perform best in any criterion and performs worst in three. However, the true value of each choice is not immediately apparent from their criteria-specific performance only, but rather their overall performance when compared across all relevant criteria.

5.4.2. Rankings and prioritization

We can now proceed to calculate the PROMETHEE rankings. To establish preference relations between actions, each action must be compared with the rest.

PROMETHEE I ranking

The preference value of one action over another is in this model between 0 and 1. Given that all criteria follow the “usual” preference structure, if one action performs worse than, or equal to other, it will not be preferred and the assigned preference value will be zero. If it performs better than the other, the preference value will be one. That means that the PROMETHEE preference index $\pi(i, i')$ will essentially be the sum of weights of the criteria in which one action is preferred to another.

Table 30 illustrates the mechanics of the calculation of the preference flows. Since one action cannot be compared with itself, the diagonal of the table is all zeros. Both the negative φ^- and the positive φ^+ preference flows are calculated as the weighted average of the indexes (not including the comparison of actions with themselves). The resulting PROMETHEE I partial ranking is given by the φ^+ column and φ^- row.

Table 30. Calculation of PROMETHEE I partial rankings

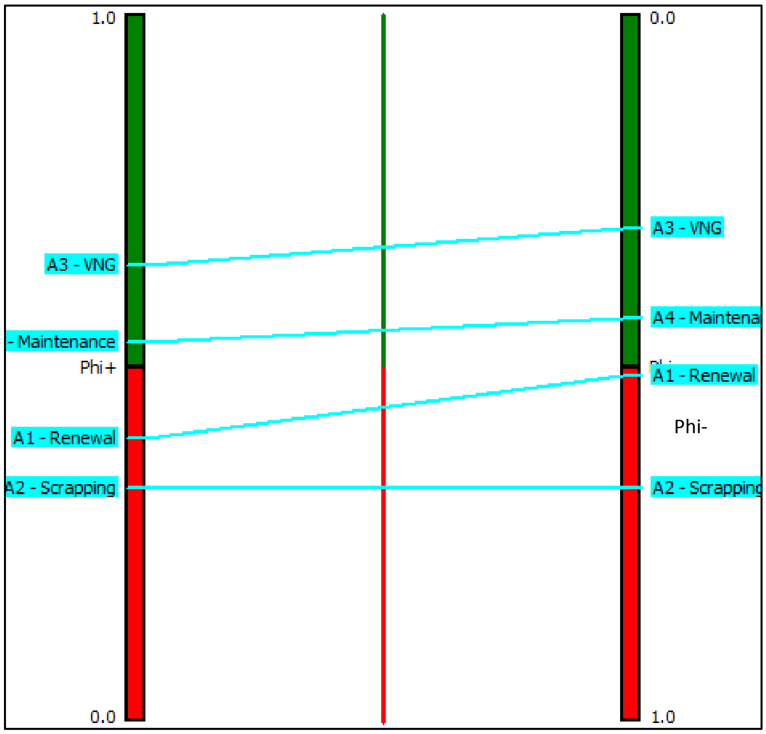
	A1	A2	A3	A4	φ^+ (Phi+)
A1 renewal	0.0000	0.5618	0.2989	0.3416	0.4007
A2 scrapping	0.4382	0.0000	0.3034	0.2472	0.3296
A3 VNG	0.5461	0.6966	0.0000	0.6966	0.6464
A4 maintenance	0.5506	0.7528	0.3034	0.0000	0.5356
φ- (Phi-)	0.5116	0.6704	0.3019	0.4285	

No two actions score the same value in either φ^+ or in φ^- so no two actions are indifferent. According to the φ^+ values, action A3 (the promotion of vehicular natural gas) is preferred to action A4 (promotion of better vehicle maintenance), which in turn is preferred to action A1 (renewal of private fleet), which is preferred to A2 (scrapping old public transportation vehicles). φ^- shows that all other actions are not preferred to A3, and that A3 is not least preferred than A4, A4 not least preferred than A1, and A1 not least preferred than A2.

This preference structure can be represented graphically in a partial ranking diagram (Figure 21). The left bar represents the φ^+ parameter scores, the right bar the φ^- parameter scores, and the middle bar, the net scores for each action. Actions are represented by lines crossing the bars. The higher a line crosses a bar, the more preferable an action is compared with lines crossing at lower heights.

Reading from left to right, if a line stays consistently above another, then it is better in terms of on both φ^+ and φ^- and preferable overall. If, however, a line starting above crosses another and ends up in a lower position, those two actions are "incomparable" in the partial ranking. "Incomparability" in PROMETHEE means that one action compares better than other in some criteria and worse in some other criteria, thus entailing difficult trade-offs in the decision-making.

Figure 21. Partial ranking PROMETHEE I



Source: author

PROMETHEE II ranking

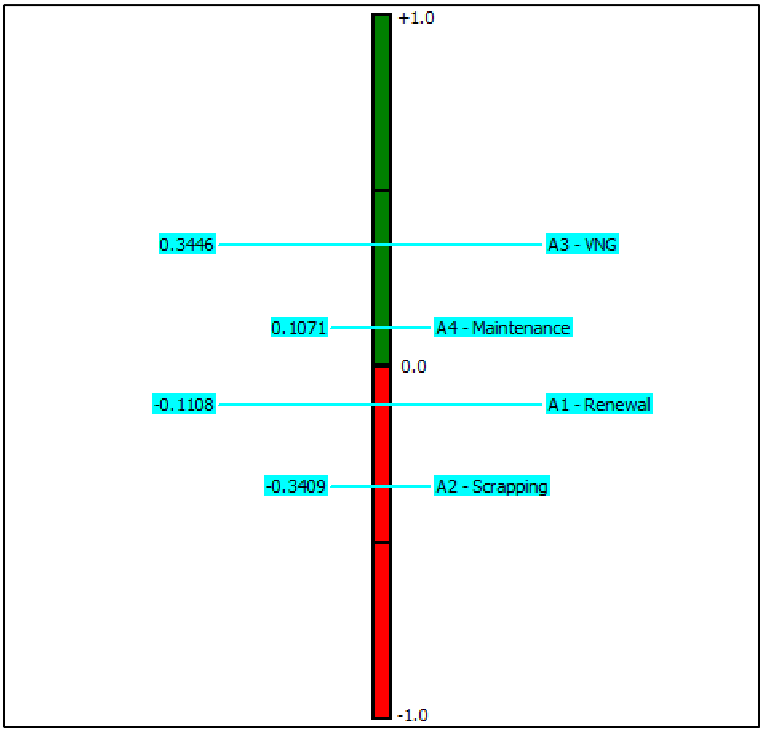
The PROMETHEE II complete ranking is calculated by subtracting φ^- from φ^+. Complete PROMETHEE rankings display an absolute rather than relative ranking, by merging performance across criteria into an overall index (see Table 31).

Table 31. PROMETHEE flows and resulting rankings

Ranking	Action	$\varphi(i)$ (Phi)	$\varphi^+(i)$ (Phi+)	$\varphi^-(i)$ (Phi-)
1	A3 - VNG	0.3446	0.6464	0.3019
2	A4 - Maintenance	0.1071	0.5356	0.4285
3	A1 - Renewal	-0.1108	0.4008	0.5116
4	A2 - Scrapping	-0.3409	0.3296	0.6704

This preference structure can be represented graphically in a partial ranking diagram (Figure 22). The green and red middle bar represents the total scale of φ (from -1 to +1). The higher an action bar crosses the φ bar, the better ranked it is in PROMETHEE II.

Figure 22. Complete ranking PROMETHEE II



Source: author

In this case, PROMETHEE II offers a maximum value of φ for A3, followed by A4, A1 and A2. The partial and complete rankings are thus completely consistent: VNG promotion ranks first, second is vehicle maintenance, third vehicle fleet renewal and last is public transportation vehicle scrapping.

5.4.3. Robustness of prioritization to weight variation

We can now study how variations in the weights affect the stability of the prioritization. The mathematical definition of the stability region for a given criterion (that is, the weight interval of that criterion for which the preference structure does not change) is featured in Section 5.2.4. Since the structure of preferences is complete and no action is indifferent to other, we can obtain the weight stability interval (w_j^-, w_j^+) for each criterion, listed in Table 32, below:

Table 32. Weight stability intervals for all criteria

Criterion	Code	Weight w_j	w_j^-	w_j^+	Range
BoD	C1	0.1910	0.0920	0.3100	0.2180
Inclusion	C2a	0.1348	0.0000	0.2621	0.2600
Pub.op.	C2b	0.1124	0.0000	0.2430	0.2430
Enforce.	C3a	0.1551	0.0198	0.3063	0.2865
Implem.	C3b	0.1865	0.0807	0.3008	0.2201
Cost	C4a	0.1124	0.0000	0.2066	0.2066
Incentives	C4b	0.1079	0.0000	1.0000	1.0000

The interpretation of a weight stability interval (WSI) in PROMETHEE is not extensively explored, either in the methodological publications or in the MCA applications; however, there are some clear guidelines (Mareschal, 1988) of importance to our example:

1) If the WSI of one criterion includes zero, then that criterion can be removed and, everything else being equal, the prioritization would not change;
2) If a criterion displays maximum weight stability (in our case, with a WSI from 0 to 1, both included), then it meets the previous condition, and, furthermore, no change (increase or decrease) in its weight would change the prioritization;
3) Actions that remain highly ranked with relatively important changes in weight across most criteria are stable good choices that are resilient to weight variation (argued to be the most subjective component of MCA). What constitutes "important" changes is not consistently interpreted in the literature.

From the ranges of the weight stability intervals in Table 32, we can see that the criterion "Incentives" displays the maximum possible weight stability (0-1). This means that variation in the criterion weight does not affect the prioritization. In addition, the criteria "Inclusion", "Public opinion", and "Cost" include zero in their WSIs, so removing any one of them alone would not change the prioritization, everything else remaining equal. On the other hand, "Burden of Disease", "Enforcement" and "Implementation" do not include zero in their WSIs, so the elimination of any of them would change the prioritization, as would criteria-specific changes in the weights. A visual display of the stability intervals in each criterion (available in Annex 4) can further help interpret the robustness of the ranking to weight variations.

Importantly, we can observe that the first MCA-ranked action A3 (VNG) remains first regardless of weight variations in four out of seven criteria. The other three criteria, in which a change in weight could theoretically modify the ranking, are "Burden of Disease", "Cost" and "Implementation". But how much should weighting change in these criteria to change the prioritization? In the absence of guidelines on what constitutes high weight stability, I checked whether the WSIs for these three criteria are beyond an arbitrary interval $w_j \pm 0.5w_j$. The WSIs of the three criteria are well beyond this interval (see Annex 4, table A4.1) which means that it would take a fairly large weight variation (of over 50%) in any one of them to change the prioritization. We can then accept that the ranking provided is relatively solid.

Once this is established, we can take again into account the monetizable *vs.* non-monetizable classification. From the WSI graphs (Annex 4) can be observed that the ranking of action A3 falls sharply as the weight of the monetizable criteria (i.e. "Burden of Disease" and "Cost of implementation") increases, since it scores poorly in them. However, the first position of A3 tends to increase its lead when the weight of the non-monetizable criteria (i.e. "inclusion", "public opinion", "enforcement", and "incentives") increases.

What this implies is that the choice of A3 as first ranked action entails a trade-off, whereby bad performance in monetizable criteria is compensated by high performance in non monetizable criteria. This is, of course, a particular result of our case study. However, the fact that there can be a systematic difference in the ranking of one action depending on whether it is considered from the monetary point of view or otherwise is an important conceptual implication, further elaborated on below.

5.4.4. Case study GAIA plane

The concordance of preferred actions with the different criteria can be visualized further through the GAIA plane, a representation of a principal component analysis in which U is the first principal component, containing the maximum possible quantity of information, and V is the second principal component, providing maximum additional information orthogonal to U. The formalization of the plane and featured vectors is briefly summarized in section 5.2.5 and explained extensively elsewhere (Brans, 1996; Brans & Mareschal, 2005; Brans et al., 1986).

Actions are represented in the plane by points (free standing small squares) whose positions are related to their evaluation according to the criteria; thus actions with similar profiles are close to each other. Criteria are represented by axis drawn from the center of the plane, whose orientation shows how closely they relate to each other; thus it is possible to identify groups of criteria expressing similar preferences. In this case, the U-V plane gathers 89% of the total information. Regarding the concordance of the criteria:

- Related criteria (similar preferences): "Inclusion" and "Public opinion" are perfectly concordant, so that their axis overlay completely in the GAIA plane. "Incentives", "Enforcement" and "Implementation" are all within the same quadrant, and relatively aligned with the decision axis, in the direction of action A3.
- Opposite criteria (conflicting preferences): "BoD" is clearly in opposition with "Incentives", "Enforcement" and "Implementation". In addition, "Cost" is in opposition with "Inclusion" and "Public opinion".
- Uncorrelated or low-correlated criteria: "Implementation" and "Cost", "Enforcement" and "Public opinion"/ "Inclusion", "BoD" and "Public opinion"/ "Inclusion" and to a lesser extent, "BoD" and "Cost" are lowly correlated.

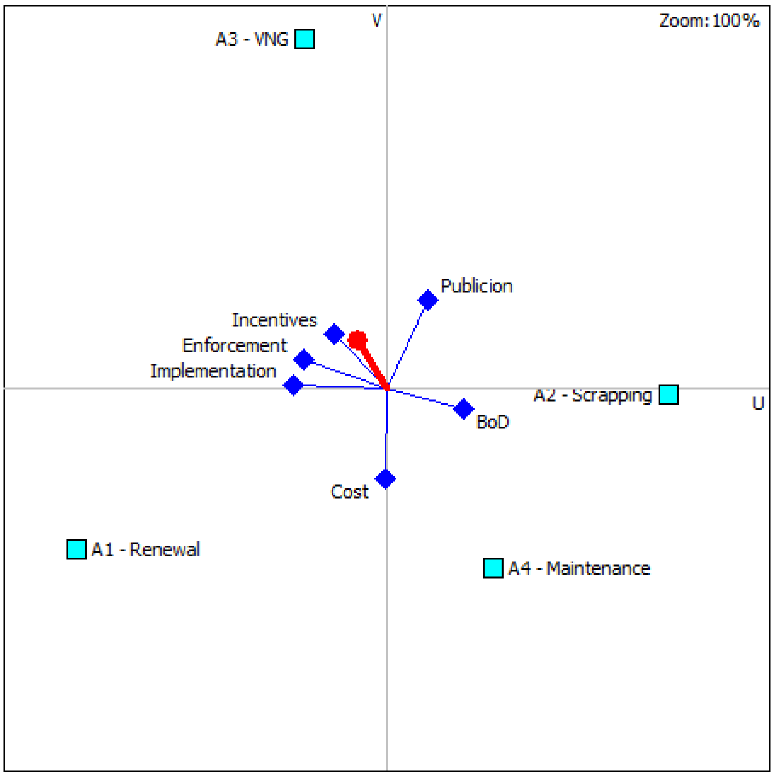
Regarding the relative performance of actions in different criteria:

- A1 has less cost, less implementation time, better enforcement chances and more incentives than the average; and lower public opinion acceptance and inclusion potential than the average;
- A2 has higher reduction of BoD and a higher public opinion acceptance and inclusion potential than the average, but it has a longer implementation time, higher cost, less incentives and worse enforcement chances than the average.
- A3 has higher public opinion acceptance and inclusion potential, as well as more incentives, better chances of enforcement and less implementation time than the average, but it reduces less BoD and costs the government more than the average.
- A4 has the highest BoD reduction and costs less that the average, but it has a longer implementation time, less incentives and worse enforcement chances, as well as lower public opinion acceptance and inclusion potential than the average.

The grouping of criteria in the GAIA plane with respect to the actions considered is qualitatively interesting and it suggests a possibility of interest for this dissertation. The criteria seem to be grouped in two opposite directions, one of which includes the two criteria that are part of both the MCA and the CBA (See Figure 23; please note the axes of "Inclusion" and "Public opinion" overlay; thus the software-generated portmanteau "Publicion").

Let us look at the list of criteria from the standpoint of monetizability, that is, whether they can be readily converted to economic values. The criteria “inclusion”, “public opinion”, “enforcement”, “implementation” and “incentives” are here measured through either a psychometric scale or a subjective evaluation, based on the perceptions of the stakeholders, and do not admit ready monetization. On the other hand, “Cost of implementation” is itself a monetary quantity. While “Burden of Disease” is not an economic variable, it can be readily translated into economic costs, as seen in the “Benefits” section of the CBA chapter.

Figure 23. GAIA plane of the MCA problem Scenario 0



The orientation of the “decision axis” indicates which criteria are in agreement with the PROMETHEE rankings and which are not. In this case, non-monetizable criteria (i.e. “inclusion”, “public opinion”, “enforcement”, “implementation” and “incentives”) are better aligned with the most preferred option (A3) than the monetizable criteria. Indeed, action A3 scores poorly in terms of BoD reduction and worst in terms of cost to the government, so its overall ranking as first preferred option masks a discrepancy in performance between monetizable and non-monetizable criteria.

This discrepancy (or rather, the possibility of it happening) is potentially important when considering, as is the case with this dissertation, the integration of CBA and MCA. This matter is further explored in the conclusions of this chapter, below.

5.4.5. Conclusions of the case study MCA

The best action in this MCA problem is A3 – Promotion of VNG use in public transportation vehicles, followed by A4 – Enforcement of better vehicle maintenance, then A1 – Private vehicle fleet renewal, and lastly A2 – Bonus for scrapping older public transportation vehicles. The ranking of A3 as most preferred option is independent of weight variation in a majority (four out of seven) of criteria and relatively resilient (WSI larger than $w_j \pm 0.5w_j$ interval) in the rest. However, the highest ranking of A3 masks a discrepancy in performance between monetizable and non-monetizable criteria. The action performs well in criteria that do not admit monetization, but very poorly in criteria that admit monetization. As monetizable criteria gain weight, the prioritization changes in favor of other actions.

This discrepancy illustrates the possibility that stakeholders may inadvertently accept large trade-offs between monetizable and non-monetizable variables in MCAs, which could result in decisions leading to inequities or other adverse outcomes. If we accept that individual and societal preferences can be described in terms of "willingness to pay" or "willingness to accept", then the monetizable criteria tend to better describe society's best interest. In other words, the best option according to the subjective criteria of the decision maker may be openly misaligned with the option/s in the best interest of society. That this is a possibility was predictable; the risk of MCA not capturing societal costs and benefits is well known and one of the main criticism to its use in public policy evaluation. However, the open visualization of the trade-offs between the preferences of the stakeholders or decision-makers and those of society (measured through costs and benefits), and further consequences for decision-making, can be a relevant and so far barely explored dimension of the use of decision-support tools in evidence-based policy-making.

There is additional value in the fact that this conclusion resulted from a real life case study, instead of a theoretical exploration of alternative models. For instance, this MCA was largely driven by the stakeholders in Lima and El Callao, as specifically requested by the agency driving the process. This is relatively uncommon largely on account of time and resource considerations, particularly in low income settings like the MALC. The exceptional nature of this case (a governmental agency looking to innovate in its decision-making and an international organization willing to finance the process) allowed a strong degree of stakeholder involvement, with its advantages and disadvantages. The main positive aspect was that there was a direct input of the most crucial parameters in the MCA, namely criteria and weight distribution, by the people driving the policy process and ultimately, decision-making. The involvement of the stakeholders increased understanding, engagement and ownership of the process, and reduced the possibilities of additional bias being introduced by the analyst, a common criticism in the field of decision support (Brown, 2005). On the other hand, the stakeholder-driven MCA shows obvious shortcomings that an analyst-driven one would probably have avoided. The most obvious one is the minimal consideration of society-wide effects, notably costs (only governmental costs were considered).

These imbalances point at the need for stakeholder education in participatory decision-support processes, particularly in low income settings, as well as the importance of the choice of stakeholders in MCA processes (Harrison & Qureshi, 2000). Crucially, they further support the notion that the integration of CBA considerations into a MCA framework can contribute to a more complete and balanced picture of the trade-offs (particularly in society-wide issues) involved in the various possible decisions. This point is explored in Chapter 6, based on the results from the CBA and the present chapter.

Chapter 6. Integrating MCA and CBA – application to the case study

6.1. Introduction

The literature on possible combinations of MCA and CBA has been explored and discussed in section 2.4. In summary, the various options can be grouped into four categories:

1) Combining the outputs of CBA and MCA;
2) Expanding or modifying a CBA with MCA attributes;
3) Integrating CBA outputs or components into a MCA framework; and
4) Creating a hybrid system.

The main objective of this dissertation is basically to provide a background to understand the implications of the combined use of MCA and CBA and to develop a conceptual framework for the integration of the two methods. Upon scrutiny of the specificities, strengths and weaknesses of each approach, not all four categories of MCA-CBA combination are relevant to this objective. Specifically, category 4) "hybrid systems" does not seem to represent a true integration of CBA and MCA; rather, it consists in the practice of a different presentation of the results of both techniques (see subsection 2.4.4). Thus, it is not applied here to the case study of the MALC. Similarly, other non-integrative ways to jointly use MCA and CBA are not considered in this chapter, whether they consist of joint presentation in an "Appraisal Summary Table" (DEFRA, 2007) or in the successive application of one technique or the other for screening purposes. On the other hand, categories 1) to 3) provide some measure of real integration of MCA and CBA. After having conducted a CBA and a MCA of the case study decision problem, it is now possible to attempt these integrative approaches, comparing the process as well as the results on an equal footing. The specific approaches applied in this section are:

1. Combining the outputs of a MCA and CBA (section 6.2), by applying the COSIMA approach proposed by Barfod, Salling, & Leleur (2011).
2. Expanding or modifying a CBA with MCA attributes (section 6.3), by applying the approach proposed by Diakoulaki & Grafakos, (2004).
3. Integrating a CBA in a MCA framework (section 6.4), by applying two proposed experimental approaches:
    a. Introducing social economic costs and benefits separately into a MCA framework; and
    b. Introducing a "BCR" criterion into a MCA framework.

It is important to note the practical orientation of this dissertation with regard to the proposed approaches. That is why there is no discussion of possible effects in conceptual matters underlying MCA like commensurability or strength of comparability (Munda, 2004) or other underlying CBA like welfare change measurements. Rather, what is sought here is to identify ways to exploit jointly the capacity of CBA and MCA to inform policy based on the information available to planners and decision-makers. Below are the proposed methods and the results.

6.2. Combination of MCA and CBA outputs

According to Barfod et al. (2011) in the context of prioritization of transportation projects, we could derive for a given alternative A_i a total rate of return (TRR) considering both CBA and MCA components (see equation 16). To do so, the authors turned the two addition terms (CBA and MCA value function score) within the brackets into consistent units through the use of an fictitious currency ("assessment m DKK"). To obtain a total value for an examined alternative, m DKK and "assessment m DKK" are added, and the resulting mix expressed by the unit "attractiveness m DKK". The authors do not specify exactly how criterion performance is translated into "assessment DKK"; this (crucial) process is implicitly suggested to have taken place through the interaction with the stakeholders. Despite the obvious replicability flaw, it is worth applying the approach to our case study for the sake of discussion.

For such application, the original MCA has to be adapted and trimmed. Double counting by Barfod et al. (2011) is eliminated from the start of the analysis, since the MCA criteria are chosen by the analysts and analyzed in a MCA specifically because they are not monetizable. In our case study, double counting is addressed by extracting all monetizable criteria into the CBA and leaving the MCA purely composed of non-monetizable ones. The resulting MCA problem has five criteria: "Inclusion", "Public opinion", "Enforcement", "Implementation", and "Incentives". All criteria except "Implementation" share the same scale (1 to 5) and preference structure (better is more, and there is no indifference threshold), and are unit-less. By contrast, the "implementation" criterion is measured in years and its preference structure is reversed (less is better) although it does not have indifference threshold either.

For the purposes of baseline homogeneity as a prerequisite for the translation into "Assessment million Soles", our equivalent of mDKK, we can transform the preference structure of "Implementation" into a maximization one simply by subtracting the implementation time from an arbitrary maximum of 10 years. As it happens, all the resulting values are within the limits of 1 to 5. While normalization to a 1-5 scale may perhaps be more formally correct, it may also artificially inflate differences in performance that would disproportionately add "assessment currency" value to the highest "implementation" performing action. Moreover, normalization in this criterion results in no differences in the final prioritization and minimal differences in the TRR values (see Annex 6). For simplicity, we will leave the new "Time left" criterion untransformed, with the values listed in Table 33.

Table 33. Scaling of "implementation" criterion

Action	Implementation (years)	Time left from a 10 year horizon
A1 renewal	5.77	4.23
A2 scrapping	6.63	3.37
A3 VNG	6.36	3.64
A4 maintenance	6.52	3.48

We can now calculate the new weights proportionally to the previous so as not to change the order, by dividing the individual weight of each remaining criterion by the sum of weights of the remaining ones. The resulting MCA problem is summarized in Table 34.

Table 34. MCA problem with all criteria in the same scale and preference function

Criterion	Inclusion	Pub.op.	Enforce.	Implem.	Incentives
Preference fn.	Usual	Usual	Usual	Usual	Usual
P, Q	N/A	N/A	N/A	N/A	N/A
Units	N/A	N/A	N/A	N/A	N/A
Optimization	MAX	MAX	MAX	MAX	MAX
Weights	0,1935	0,1613	0,26	0,2677	0,1548
A1 renewal	1.80	1.60	3.80	4.23	2.40
A2 scrapping	2.60	3.80	2.40	3.37	2.00
A3 VNG	4.20	4.40	3.80	3.64	4.80
A4 maintenance	2.40	3.00	3.00	3.48	2.40

Green: best performing; Red: worst performing

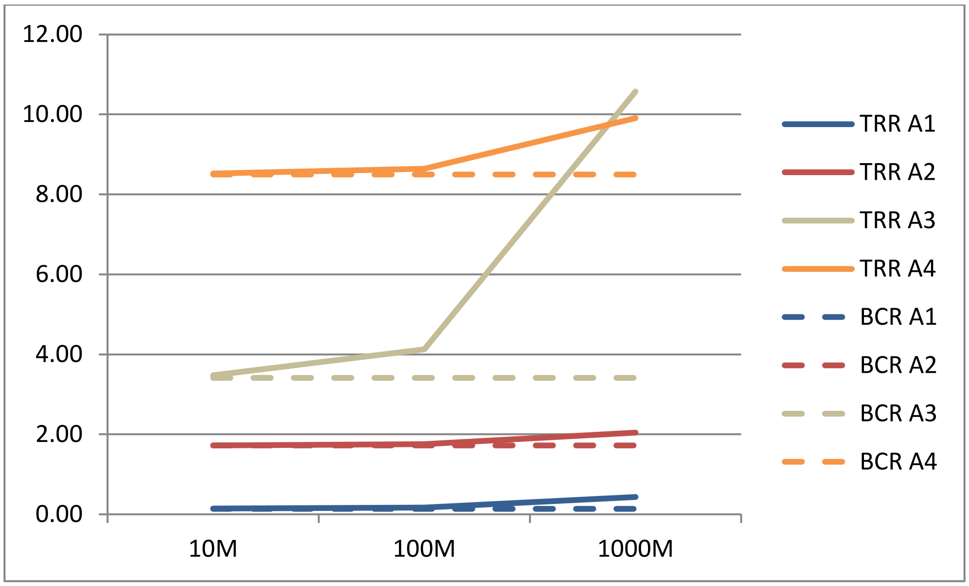
The MCA prioritization itself is not of particular interest for the argument, so the resulting PROMETHEE rankings I and II and consequent prioritization are not listed. However, the COSIMA-specific MCA problem is necessary to address the question of how to translate MCA scores into "assessment currency", in our case "Millions assessment Peruvian Soles". Without guidance, a reasonable option is a sensitivity analysis. We can choose an arbitrary rule by which a point of difference in the five point scale of all the criteria translates into 10 Million Soles, 100 Million Soles, and 1000 Million Soles, respectively. Taking into account the Barfod et al. (2011) formula (see Figure 5), we can calculate the $VF_j(Y_{jk})$ term for each action considered, and analyze the effect of the change in α in the resulting TRR. Table 35 illustrates the effect of different values of α on the TRR of each action, when the values per scale point are 10, 100 and 1000 Million Soles, respectively. For simplicity, the first term of the sum within the brackets of equation (16) is denoted "CBA benefit" and the second term "VF".

Table 35. Change in VF and TRR with increasing values of a and score point

Value of α	Items	A1 renewal	A2 scrapping	A3 VNG	A4 maintenance
N/A	CBA cost (ND) M. S/	8,131	5,122	388	1,280
N/A	CBA benefit (ND) M. S/	1,159	8,804	1,325	10,888
0.25	VF with SP= 10 M. aS/	7.90	5.56	9.28	5.98
	100 M. aS/	79.06	55.69	92.89	59.81
	1000 M. aS/	790.65	556.94	928.90	598.12
	TRR with SP= 10 M. aS/	0.14	1.72	3.43	8.51
	100 M. aS/	0.15	1.73	3.65	8.55
	1000 M. aS/	0.24	1.83	5.80	8.97
0.75	VF with SP= 10 M. aS/	23.72	16.71	27.87	17.94
	100 M. aS/	237.19	167.08	278.67	179.44
	1000 M. aS/	2371.94	1670.81	2786.69	1794.35
	TRR with SP= 10 M. aS/	0.15	1.72	3.48	8.52
	100 M. aS/	0.17	1.75	4.13	8.64
	1000 M. aS/	0.43	2.05	10.57	9.91

In this case, the effect of the MCA addition in the TRR does not change the prioritization in the first two orders of magnitude tested for the value function (10 and 100 and Million Soles per point valuation) but at 1000 Million Soles per point valuation and a value of α=0.75, action A3 (VNG) surpasses action A4 (maintenance). This is illustrated in Figure 24 below; in general, by adding to the benefit component, the value of TRR increases as the score point valuation rises, ultimately altering the prioritization.

Figure 24. BCR and COSIMA TRR of actions at α=0.75 and increasing values of score point



Source: author

This analysis reveals a downside of the approach: the valuation of the differences in performance determines greatly the effect in the final prioritization. Only in the case of very similar BCR values would different values of α change the ranking. That is clearly not the case in our example. In addition, the MCA scores increase only the "Benefit" component of the TRR, whereas such scores do not contribute "Assessment currency" to the costs in the TRR. Thus, alternatives that are not cost-beneficial in the original analysis can obtain larger TRRs when going through the COSIMA process, but not the other way around. So an alternative that is cost-beneficial in the original CBA but performs very poorly in the MCA will only be penalized in relative terms, by minimizing the MCA component of the TRR. But it will not turn into non-cost-beneficial.

6.3. Expanding or modifying a CBA with MCA attributes

The CBA process or its outputs can be modified with MCA attributes. This modification would essentially imply either 1) modifying the valuation of the items monetized in the CBA according to their weighting in the MCA, if they happened to be considered in it; or 2) assigning indirect monetary values (such as through the COSIMA value functions) to the criteria, and then adding them to the costs and/or benefits to be compared.

Modifying CBA elements or outputs according to MCA variables

The most obvious method for the modification of a CBA according to MCA variables would be through the MCA criteria weighting, either by: a) adding to or subtracting from the importance of each monetized component of the CBA according to their weighting in the MCA, or b) adjusting the output indicators of the CBA (whether IRR, BCR or NPV) based on the stakeholders' preferences, reflected in their weight distribution. For instance, if economic benefits and costs are included in the original MCA (a condition not guaranteed in general) and weighted differently, their weights could act as modifiers through a different multiplier of the benefits and costs as accounted in the CBA all actions. Following this lead, in our case study, the weight of "BoD" would then be a multiplier for the benefits. However, the weight of the "Cost" criterion could hardly be a valid multiplier for the CBA costs, since it only represents governmental costs and the multiplication would essentially unfairly diminish the estimation of the social costs in the final BCR. In addition, again, there is no guarantee that any given MCA will include monetizable criteria - it simply cannot be expected in general.

Another plausible way forward in this integration could consist of a method for "penalization" of the CBA outputs of actions that perform badly in an MCA, and a "reward" of those that perform well. The "penalization" of a given BCR would be proportional to the lacking in performance in the specific action analyzed when compared with the average performance of all actions. Similarly, the "reward" to the BCR of an action that performs well in the MCA would be proportional to the above-average performance. For this approach, a MAUT-type MCA would be more adequate, since one single number could summarize the performance of each alternative taking into account all criteria and weights. In that regard, the application to the case study is simply unfeasible, since it would require a whole new framework of interaction with stakeholders. More importantly, it is difficult to see decision-support value in an approach that ultimately boils down to tampering with CBA outputs in a way that makes them less understandable than the originals. An additional problem with this type of approaches is that we cannot assume that there will be any monetizable criterion in any given MCA, and so this could not be a generalizable methodology, which may explain its absence in the published literature. On the other hand, the other alternative for modifying a CBA with MCA attributes, namely the forced monetization, can be generalized so, for the sake of completeness, the method is applied to the case study.

Indirect monetization of MCA criteria

The forced indirect monetization of items that do not readily admit monetary valuation[9] can be done in at least two ways:

- A post hoc monetization by the analyst similar to that of the COSIMA method; and
- A monetization based on the notion that the elicited weights in a MCA can truly represent the preferences of stakeholders, and that in turn these can be used for the indirect monetization of otherwise non-monetizable goods or services. This method, of which the methodology is laid out in subsection 2.4.2, was first explored by Diakoulaki & Grafakos (2004) and used as part of the "External Costs of Energy" (ExternE) project (http://www.externe.info) in the context of the evaluation of possible scenarios for reducing emissions of NOx, NH3, SO2 and VOC in Europe, relative to their levels in 1990. The reductions were described in terms of the corresponding physical impacts for acidification, eutrophication, mortality, and changes in agricultural production; and the indirect monetization was used to evaluate acidification and eutrophication, two forms of ecosystem degradation.

[9] If the item would admit monetization it would have been included in the CBA

The second option requires our baseline MCA to be modified for application to our case study. Since interacting with the stakeholders to elicit weights in this new framework is not possible, the application will simply be a post hoc modification of the analyses completed. Firstly, monetizable criteria (i.e. BoD) susceptible of providing benefits are dropped, since the idea of the method is to account for non-monetizable benefits, and their weight redistributed so as not to alter the balance of the remaining criteria. Because the CBA to be modified is a social CBA, the original Scenario 0 MCA (governmental) cost criterion will be substituted by the social cost, with the new adjusted weight. Finally, the preference structure of the "implementation" criterion is reversed from minimization to maximization by simply subtracting its performance (years) from an arbitrary upper bound of ten years. The resulting MCA is expressed in Table 36 below; the method is named "ExternE" for conciseness. Applying the equations in subsection 2.4.2 we obtain the unit valuation of these criteria (see Table 37 below).

Table 36. MCA modified for application of ExternE CBA expansion

Criterion	Inclusion	Pub.op.	Enforce.	Implem.	Social Cost	Incentives
Preference fn.	Usual	Usual	Usual	Usual	Usual	Usual
P, Q	N/A	N/A	N/A	N/A	N/A	N/A
Units	N/A	N/A	N/A	Years left	Million Soles	N/A
Optimization	MAX	MAX	MAX	MAX	MIN	MAX
Weights	0.17	0.14	0.19	0.23	0.13	0.14
A1 renewal	1.80	1.60	3.80	4.23	8,131	2.40
A2 scrapping	2.60	3.80	2.40	3.37	5,122	2.00
A3 VNG	4.20	4.40	3.80	3.64	389	4.80
A4 maintenance	2.40	3.00	3.00	3.48	1,280	2.40

Green: best performing; Red: worst performing

Adding these benefits to the social benefits of the original CBA (no discount), we can observe the effect in Table 38.

Table 37. Unit valuation of indirectly monetized criteria

Criteria	Unit valuation	Units
Inclusion	4,032	Million Soles per 1 to 5 psychometric scale point
Public opinion	2,880	Million Soles per 1 to 5 psychometric scale point
Enforcement	7,949	Million Soles per 1 to 5 psychometric scale point
Implementation	15,566	Million Soles per year gained
Incentives	2,880	Million Soles per 1 to 5 psychometric scale point

Table 38. Variation in BCR due to indirect monetization of MCA criteria

CBA items	A1 renewal	A2 scrapping	A3 VNG	A4 maintenance
PV (original benefits)	1,159	8,805	1,326	10,888
PV (extended benefits)	56,924	65,940	85,375	68,468
PV (social costs)	8,131	5,122	389	1,280
Original BCR	0.14	1.72	3.41	8.50
Extended BCR	**7.00**	**12.87**	**219.45**	**53.47**

In this case, the indirect monetization of these qualitatively assessed criteria changed the original prioritization of the CBA, and the change is largest for A3 because of its high performance in the newly monetized criteria. However, the application of this method to the case study, and indeed to any general case, entails some difficulties. Firstly, the method was developed for use with MCA methods following Multi Attribute Value Theory, assuming the possibility of full compensation among criteria, something that could not apply to MCA problems with any qualitative criteria. Secondly, it assumes the existence of a "cost" criterion, mainly because in this case the MCA process is driven by the analyst, who established the criteria and determined performances. And thirdly, the method does not account for the bias that the knowledge that all criteria will be monetized may induce in the process of weight elicitation; this in turn leads to the larger question of whether the stakeholders would be comfortable at all with the monetization of every criterion considered of importance for a policy decision. This idea is counterintuitive in view of the existing literature. On the other hand, when compared with other alternatives for "indirect monetization", namely the COSIMA approach, this method provides replicability and reduces arbitrariness by relating the magnitude of the value to the impact range, or performance of the criteria.

6.4. Integration of CBA into a MCA framework

The integration of a CBA into a MCA problem requires careful consideration to avoid practical pitfalls, especially double counting. A social CBA of a set of actions would ideally include every monetizable criterion considered for a MCA for those actions. Therefore, the addition of one or more "CBA criteria" would in effect double-count to various extents monetizable criteria, disproportionately adding to their importance. Obviously, the potential for double counting of project effects exists also within a MCA, across criteria; but such double counting would ideally have been eliminated during the MCA problem statement phase, and so it is not considered as an issue to focus on in this case. An additional consideration is the redistribution of weights when changing the considered criteria, especially if the weights are relative. Also, the analyst must be mindful (and make the stakeholders and decision-makers also aware) of the fact that the introduction of the CBA would usually bring into the MCA new criteria that they may not to be considered in it, and for which no formal, separate weighting has been conducted. Moreover, a "CBA criterion" could in effect be represented in either a benefit-to-cost ratio or a net present value, with important consequences embedded in this choice.

There are various examples in the literature of integration of CBA into a MCA framework:

1. In the approach taken by Gühnemann, Laird, & Pearman (2012), the researchers included the CBA results into the MCA framework through a scoring function (see subsection 2.4.3). The generalization of this methodology, however, is difficult. Firstly, the specific form of the function (4+3X) is determined by an evaluation scale set by the case study and it would an arbitrary element to apply in other settings. Similarly, whereas in this approach a threshold of 2.5 is proposed, with significant impact on the scoring function, there is no reason in general to choose a benefit-cost ratio threshold larger than 1.0 (the breakeven point). Thus, the formula would in a general case adopt the form:

$$Score_j = \frac{\left(PV_j^{Do\ something} - PV_j^{Do\ minimum}\right)/PVC}{\alpha_j} \tag{56}$$

In our case study, the CBA benefits are exclusively derived from the reduction of burden of disease achievable through air pollution reduction, so its contribution to the monetized benefits is 100%, thus allowing only for a value of 1.0 for α_j. The case would be different if the benefits of the actions stemmed from different sources aside from air pollution reduction, for instance noise reduction, or reduction of fatal accidents. However, the ultimate goal of the MCA prioritization in our case study was air pollution reduction, and those other considerations were either not mentioned (noise) or mentioned and discarded (accidents). Thus, there is simply no value in applying this methodology to the case study considered.

2. The EUNET project (Nellthorp, Mackie, & Bristow, 1998) addresses the explicit separation of CBA criteria and other non-monetizable MCA criteria for the purposes of avoiding double counting, with two particularities: 1) the benefit-to-cost ratios are converted into a 0 to 100 scale common with the rest of criteria; and 2) the weight allocation of CBA versus the rest of the criteria is decided by the researchers. However, in the EUNET approach the design of MCA criteria and weight elicitation is analyst driven. So, albeit double-counting can be controlled from the inception, this happens at the expense of stakeholder involvement, effectively stripping the MCA approach from arguably its greatest strength – process engagement (Estévez, Walshe, & Burgman, 2013).

3. Schutte & Brits (2012) proposed a fixed MCA framework for transport investments including an "efficiency" criterion based on a "narrow CBA" (i.e. strictly from the transport cost and efficiency perspective) which could rely on various indicators, including Equivalent Uniform Annual Cost (EUAC), IRR, NPV, and BCR. Similarly, the idea of a fixed MCA framework diminishes the stakeholder engagement value of the MCA process by imposing predefined categories that may or may not be in line with the decision-makers' views. Granted, if such fixed structure is the result of a wide consensus and experience in application in a specific setting, there would be value in a consistent approach. That situation is, however, rather far from the current situation of policy evaluation in most countries, particularly those of middle- or low-income. Specific regulatory authorities may already be systematically including a similar "efficiency" criterion in their fixed MCA frameworks, albeit no indication of that was found in the literature review and the resources and time required to review all existing guidelines on the matter are beyond the scope of this dissertation. In addition, while the choice of a narrow CBA –when applicable- may preclude some double-counting, it unfairly reduces the scope of the CBA back to a purely utilitarian "return on investment".

Given these practical drawbacks, there is value in additional proposals for approaches to integrate CBA elements or outputs into a MCA framework. In the following two subsections two such approaches are proposed and applied to our case study. What is being proposed, in summary, is:

1) To conduct a cost–benefit analysis including all MCA criteria that could be monetized, if any;
2) To substitute in the MCA problem all monetizable criteria, if they are present, by one or more "CBA criteria" that would represent them (and potentially additional ones); and
3) To implement a second iteration of the MCA with the "CBA criteria" in addition to all non-monetizable criteria.

In our case study (air quality improvement policies in the MALC), there are criteria in the original MCA problem that admit monetization: "Burden of disease" and "Cost" (reflecting cost of implementation for the government, not social costs). On this basis, the following scenarios were defined:

- Scenario 0: the original MCA problem, all criteria and weights were the ones originally determined by the stakeholders.
- Scenario 1: "Burden of Disease" and "Cost" (i.e. governmental, as defined by stakeholders) both disappear as criteria, and they are substituted by two new criteria – "Social economic benefits" and "Social economic costs". These new criteria have the weights of the criteria they substituted.
- Scenario2: "Burden of Disease" and "Cost" both disappear as criteria, and they are substituted by one new criterion – tentatively BCR. This new criterion has a weight equivalent to the sum of weights of "Burden of Disease" and "Cost".

The scenarios, with the resulting set of valid criteria and their weights are summarized in Table 39.

Table 39. Criteria and their weights in each considered scenario

Criteria	Scenario 0 (original)	Scenario 1 (SEC/SEB)	Scenario 2 (BCR)
BoD	0.1910	---	---
Inclusion	0.1348	0.1348	0.1348
Public Opinion	0.1124	0.1124	0.1124
Enforcement	0.1551	0.1551	0.1551
Implementation	0.1865	0.1865	0.1865
Cost	0.1124	---	---
Incentives	0.1079	0.1079	0.1079
SEB	---	0.1910	---
SEC	---	0.1124	---
BCR	---	---	0.3034

SEC: Social Economic Cost; SEB: Social Economic Benefit

In principle, both Scenario 1 and Scenario 2 would capture the information of the MCA and the CBA of our case study without double counting. The differences are: 1) whether benefits and costs are different criteria, or whether they are merged into one; 2) the weight variations derived from such difference, and 3) the final number of criteria.

6.4.1. Introducing social economic costs and benefits into a MCA framework

In this scenario, the social costs and benefits of the evaluated actions are introduced into the MCA problem as "Social economic benefits" and "Social economic costs", substituting the monetized and/or monetizable criteria, in this case "Burden of Disease" and "Cost" (i.e. governmental, as defined by stakeholders). The new criteria have the weights of the criteria they substituted.

Scenario 1 problem statement

The baseline (Scenario 0) is summarized in Table 29. Using that as a reference, this scenario is defined by the substitution of the criterion "Burden of Disease" by its monetized counterpart "SEB" and of the criterion "Cost" (i.e. governmental) by the social cost "SEC". The new criteria have the weights of the criteria they substituted. The problem statement is laid out in Table 40.

Table 40. MCA Scenario 1 problem

Criterion	Inclusion	Pub.op.	Enforce.	Implem.	Incent.	SEB	SEC
Preference fn.	Usual	Usual	Usual	Usual	Usual	Usual	Usual
P, Q	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Units	N/A	N/A	N/A	Years	N/A	M S/.	M S/.
Optimization	MAX	MAX	MAX	MIN	MAX	MAX	MIN
Weights	0.1348	0.1124	0.1551	0.1865	0.1079	0.1910	0.1124
A1 renewal	1.80	1.60	3.80	5.77	2.40	1159	8131
A2 scrapping	2.60	3.80	2.40	6.63	2.00	8805	5122
A3 VNG	4.20	4.40	3.80	6.36	4.80	1326	389
A4 maintenance	2.40	3.00	3.00	6.52	2.40	10888	1280

Green: best performing; Red: worst performing

Scenario 1 MCA prioritization

The modification of the prioritization of the actions under the new criteria may change the overall MCA prioritization of the actions from Scenario 0 to Scenario 1. In order to test that possibility we need to examine the PROMETHEE flows (Table 41):

Table 41. PROMETHEE ranking and flows in Scenario 0 and Scenario 1

		Scenario 0			Scenario 1		
Ranking	**Action**	**Phi**	**Phi+**	**Phi-**	**Phi**	**Phi+**	**Phi-**
1	A3 - VNG	0,3446	0,6464	0,3019	0,5693	0,7588	0,1895
2	A4 - Maintenance	0,1071	0,5356	0,4285	0,0322	0,4981	0,4659
3	A1 - Renewal	-0,1108	0,4008	0,5116	-0,2607	0,3258	0,5865
4	A2 - Scrapping	-0,3409	0,3296	0,6704	-0,3409	0,3296	0,6704

Scenario 1 discussion

Some of the characteristics of Scenario 1 do not differ essentially from Scenario 0. Firstly, the number of criteria and the relative weight of each one of them is unchanged. And secondly, the monetary value of the "Burden of Disease" criterion is roughly proportional to it and yields the same uni-criterion prioritization (see Table 42), so the substitution of the criterion "Burden of Disease" by a new criterion "SEB" (Social Economic Benefit) with equivalent weight is, by itself, unlikely to make a difference in outputs under the PROMETHEE method.

Table 42. Burden of disease (DALYs) avoided by each action, and its economic valuation

Action	BoD (DALYs)	SEB (Million S/.)
A1 - Renewal	7623	1159
A2 - Scrapping	55169	8805
A3 - VNG	8308	1326
A4 - Maintenance	68226	10888

Green: best performing; Red: worst performing

On the other hand, the criterion "Cost" as specified in Scenario 0 includes only governmental costs of the interventions but not their social costs. Therefore, the new criterion "SEC" represents a true modification of the original MCA problem (see Table 43, below)

Table 43. Governmental and social cost of each action, in Millions S/.

Action	Cost (Scenario 0)	SEC (Scenario 1)	Difference SEC-Cost	Ratio Cost/SEC (%)
A1 - Renewal	15.00	8131	8116	0.18
A2 - Scrapping	53.96	5122	5068.04	1.05
A3 - VNG	80.00	389	309	20.56
A4 - Maintenance	4.31	1280	1275.69	0.33

Green: best performing; Red: worst performing

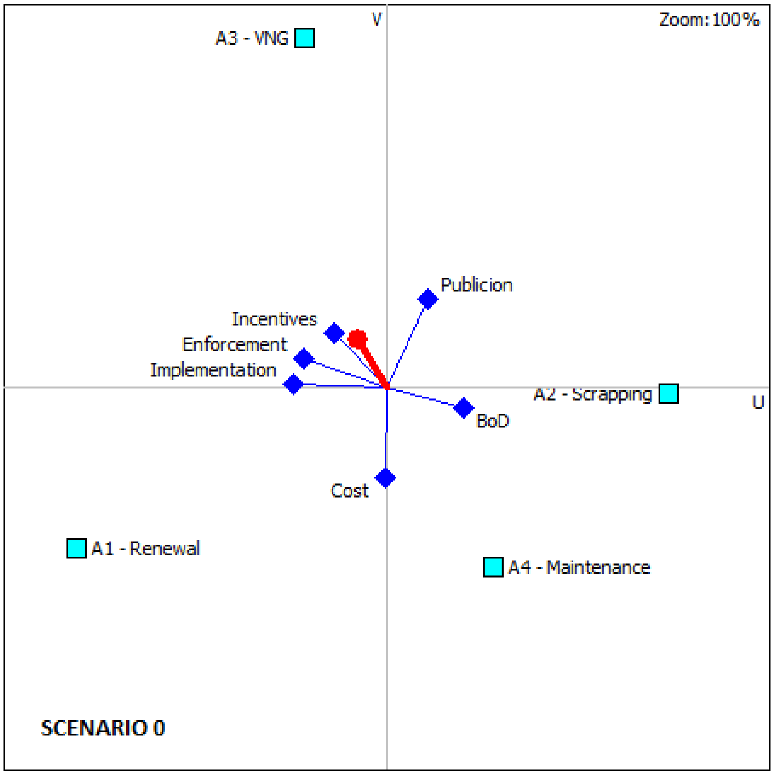
What does this entail for the prioritization of actions in the application of Scenario 1?

As mentioned, the number of criteria and the relative weights of each are unchanged, so the differences in preference flow are strictly due to differences in performance. Performance-wise, the substitution of health benefits (BoD) by their monetized equivalent (SEB) cannot change the prioritization of the actions in Scenario 1 compared with Scenario 0. On the other hand, the shift from governmental cost (Scenario 0) to Social cost (Scenario 1) modifies the prioritization of the actions under the new "SEC" criterion. However, the effect of this modification is not sufficient to change the overall MCA prioritization of actions. That is, the substitution of governmental costs by social cost and of health benefits by their monetized equivalent does not change the prioritization of the actions.

The realization of the large proportion of the social cost that is externalized to individuals (Table 43) is, in itself, a valuable input for decision-making in this case. Given that the governmental costs are from the societal point of view mainly a redistribution of resources, social costs were used in the CBA. However, in the MCA, the stakeholders placed the utmost importance to the cost borne by the government. Using the MCA results alone would have clear consequences in terms of social costs unaccounted for: A1 would be very cheap for the government but the absolute cost for society would be comparatively the highest; the government would bear just 0.18% of the social cost of the action. On the other hand, Action A3 (VNG) would be even more desirable after considering its social costs. Albeit comparatively expensive for the government, its social cost would be the lowest. Moreover, the government could claim to be bearing over 20% of the social cost of A3.

While the proportion of the social cost of the actions that should be borne by the government is debatable, only the evaluation of social vs. governmental costs would allow a fair comparison of the alternatives. Also of interest is the fact that the subjective evaluation of the actions by the stakeholders (reflected in this case by the non-monetizable criteria) is more consistent with their optimality in terms of (low) social cost that it is with the governmental cost. This is illustrated in Figure 25, which shows the GAIA planes of the MCA scenarios 0 and 1.

Figure 25. GAIA planes of Scenario 0 (top) and Scenario 1 (bottom)



In the original MCA (Scenario 0, top) the decision axis confirmed A3 as the best preferred option even though it performed badly in "Cost" and "BoD", namely the monetizable criteria. In Scenario 1 the social costs are better aligned with some of the subjective criteria as evaluated by the stakeholders than the governmental cost was. Despite these insights, it is arguable that including social benefits and costs as separate criteria in a MCA framework actually captures the CBA information wholly, since the value of the technique lies in the synthetic power of reporting benefits per unit invested, or net benefits.

6.4.2. Introducing a "BCR" criterion into a MCA framework

Scenario 2 problem statement

For the purposes of illustrating the combination of MCA and CBA, BCR will be used as the CBA representative indicator. The problem statement is summarized in Table 44, below.

Table 44. MCA Scenario 2 problem

Criterion	Inclusion	Pub.op.	Enforce.	Implem.	Incentives	BCR
Preference fn.	Usual	Usual	Usual	Usual	Usual	Usual
P, Q	N/A	N/A	N/A	N/A	N/A	N/A
Units	N/A	N/A	N/A	Years	N/A	N/A
Optimization	MAX	MAX	MAX	MIN	MAX	MAX
Weights	0.1348	0.1124	0.1551	0.1865	0.1079	0.3034
A1 renewal	1.80	1.60	3.80	5.77	2.40	0.14
A2 scrapping	2.60	3.80	2.40	6.63	2.00	1.72
A3 VNG	4.20	4.40	3.80	6.36	4.80	3.41
A4 maintenance	2.40	3.00	3.00	6.52	2.40	8.50

Red: worst performing; Green: best performing

Scenario 2 MCA prioritization

The prioritization of actions under Scenario 2 can be visualized in Table 45, below.

Table 45. Scenario 2 PROMETHEE flows and resulting rankings

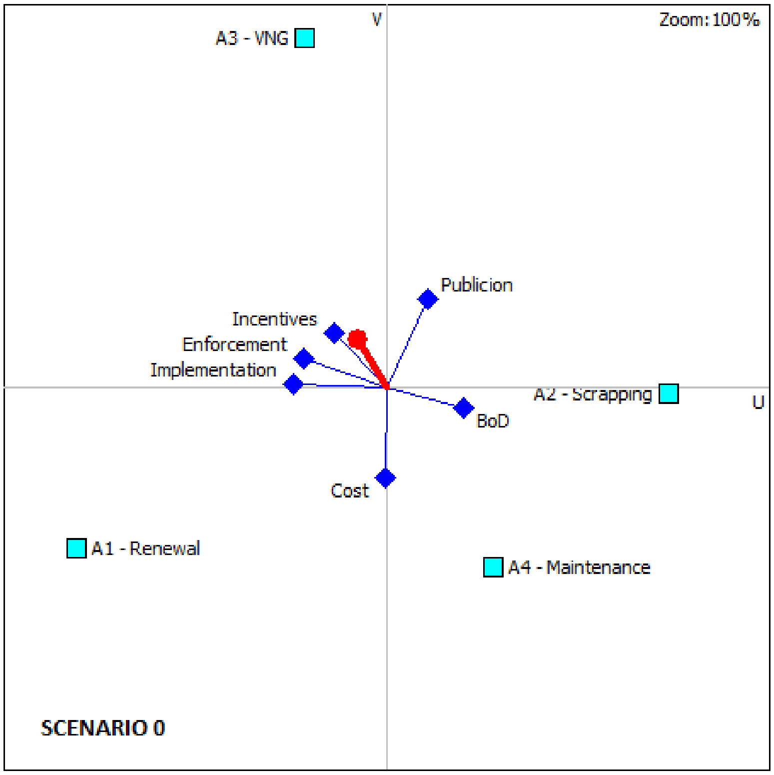
Ranking	Action	Phi	Phi+	Phi-
1	A3 - VNG	0,6217	0,7850	0,1633
2	A4 - Maintenance	0,1071	0,5356	0,4285
3	A1 - Renewal	-0,2607	0,3258	0,5865
4	A2 - Scrapping	-0,4682	0,2659	0,7341

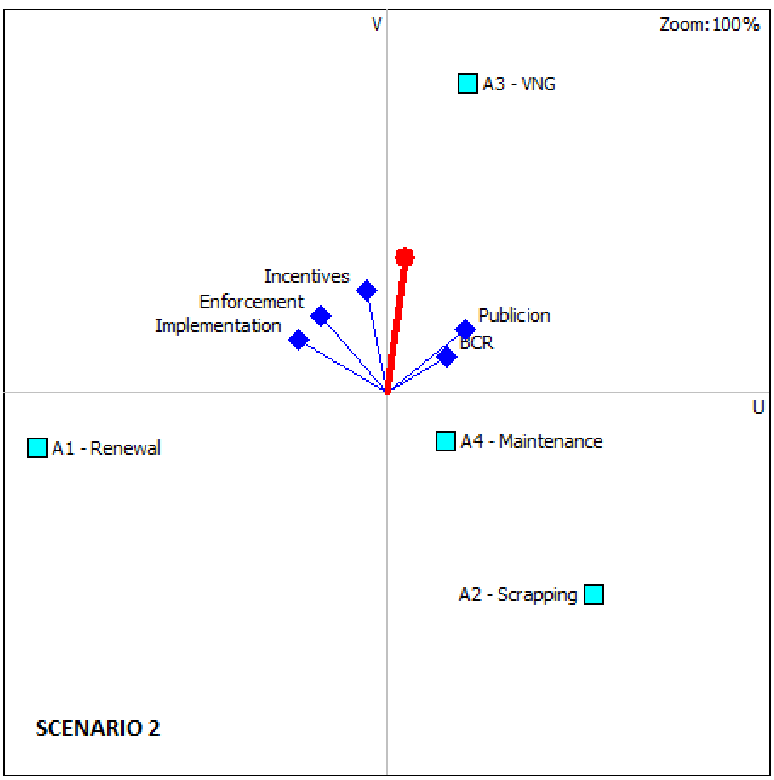
The resulting prioritization of actions does not change from Scenario 2 (BCR criterion) to Scenario 0 (original MCA).

Scenario 2 discussion

The unchanged prioritization of actions between Scenario 0 and 2 is easily explained: the prioritization of actions according to their BCR matches that of one of the criteria it substitutes (BoD), and the only discrepant original criterion (Cost) does not weigh enough to make a difference. This relative alignment can be visualized through the GAIA plane of Scenario 2 (see Figure 26, below).

Figure 26. GAIA planes of Scenario 0 (top) and Scenario 2 (down)





Note the similar grouping of the common criteria (Incentive, Enforcement …) taking action A3 as a reference and the small difference in orientation between BoD in scenario 0 and BCR in Scenario 2. It must be noted, however, that the comparison from Scenario 0 to Scenario 2 should be cautious, since the number of criteria and their relative weights have changed. Moreover, the effect of the BCR performance alone would be enough to change the prioritization of the actions (see annex 6).

In fact, the alignment observed in this case study cannot be generalized to the overall case of including a BCR criterion in an MCA problem. There is no reason why the non-monetizable criteria of the MCA representing the views of the stakeholders or objective but non-monetizable variables should consistently be aligned with the BCR criterion. In any case, the discrepancy itself (or its absence) between non-monetizable criteria and the CBA is potentially informative for decision-makers and deserves further scrutiny. It is interesting because of its ability to reveal potential mismatches or alignments between subjective and/or non-monetizable criteria, and society's best interest as summarized by, say, the BCR of a set of actions. For instance, in our case, when monetizable criteria are included into a BCR, the additional information (namely the addition of private costs to the governmental costs) results in a closer alignment of non-monetizable and monetizable criteria, with a clearer further advantage of action A3.

Let us step back and summarize the method implicit in Scenario 2:

1) Categorizing all MCA criteria into monetizable and non-monetizable;
2) Eliminating monetizable criteria from the MCA;
3) Incorporating a new criterion featuring a synthetic CBA indicator (e.g. BCR);
4) Assigning this new criterion an "adequate" weight, ideally by the stakeholders; in this case we used instead the cumulative weight of monetizable criteria;
5) Analyzing trade-offs between the CBA-obtained criterion and the rest of criteria.

In this modified MCA we have, on the one hand the non-monetizable (objective or subjective) criteria; on the other are the monetizable criteria plus additional monetizable information summarized into a CBA and included in the MCA through a "BCR" criterion.

Let us assume that, like in the MALC case study, part of the non-monetizable criteria are subjective and represent, along with the criteria weighting, the preferences of stakeholders. Then, it could be argued that an alternative that is well ranked both in the BCR and in the non-monetizable criteria (NMC) is desirable in all accounts and that the preferences of society and decision-makers are aligned. If, on the other hand, an action ranks at or near the top and scores well in the BCR but badly in the NMC, then the decision-makers' considerations are losing in the final decision. In the opposite case, societal preferences would be partially sacrificed in favor of the preferences of decision-makers.

These trade-offs could be analyzed by altering the relative weight of the BCR criterion compared with the combined weight of the NMC. In our case study, the weight of the BCR criterion (0.30) is the added weight of the monetizable criteria that were absorbed into the new criterion, namely "BoD" and "Cost". The question is then, what relative weight of the CBA would alter the prioritization of the MCA alone? In our case study, the weight stability interval (WSI) for the BCR criterion is from 14.64% to 46.87%. That is, increasing the weight of the "CBA component" beyond 46.87% at the expense of equally diminishing all non-monetizable criteria would change the prioritization of the actions considered.

In addition, the WSI does not include zero, so the BCR criterion could not be eliminated from the MCA without affecting the prioritization.

It should be noted, however, that there is more information in the BCR criterion compared with the monetizable criteria; in our case study, this additional information is the societal cost that was not considered in the original "Cost" criterion, which referred to governmental expenditures. In other cases, the newly included BCR criterion may also include monetizable criteria that are not part of the original MCA. Whether that additional information should translate into additional relative weight, and if so how much should ideally be the stakeholders' decision.

In fact, the possibilities of stakeholder engagement are probably the core added value of this post hoc integration of a BCR criterion into a MCA that has been reduced to non-monetizable criteria only. Assuming the possibility that the prioritization given by the BCR and that by NMC may be different, it could be agreed upon a priori by the stakeholders the extent to which this misalignment is or not acceptable. The idea of a priori veto thresholds was explored by Stewart & Losa, (2003) both for outranking methods and MAUT, but it was focused on the definition of the alternatives, rather than a post hoc analysis.

In the case of PROMETHEE, this veto could be a proportion of the maximum possible variation of flows due to BCR performance, and it could be useful to screen out alternatives from a large pool. Embedded in this whole line of reasoning is a fundamental question, namely whether good performance in BCR should be allowed to compensate for bad performance in the rest of criteria, and vice versa. Pictet & Belton (2000) proposed the concept of analyzing extreme compensation in general, and formalized it mathematically in the context of Multi Attribute Value Theory MCA. Notwithstanding this type of formalization, there would be value, in my view, in helping the stakeholders address this question in a guided discussion with the analyst.

The possibility of testing out these questions would have been of great interest for this dissertation. However, by the time these became apparent, there was no further possibility of interaction with the stakeholders, so these issues will remain open for other case studies or settings.

The choice of BCR to represent the CBA inclusion in the MCA is also noteworthy. In purity, only a synthetic CBA indicator like the benefit-to-cost ratio (BCR), the internal rate of return (IRR) or the net present value (NPV) would yield the CBA decision-support power wholly. Even then, it is arguable whether considerations of scale should be accounted for (in which case NPV would be the indicator of choice) or only value for money is the desired output (reflected better by IRR and BCR). Ultimately, the choice of indicator desired from the CBA is up for the analyst, and the discussion beyond the scope of this dissertation.

6.5. Conclusions of the combined application to the case study

The application of various approaches to integrate MCA and CBA to a real-life policy evaluation case study is useful to reveal their relative strengths and weaknesses, and to assess the added value of the integration itself. For instance, the combination of the outputs of a CBA and MCA through the application of the COSIMA approach proposed by Barfod, Salling, & Leleur (2011) provides a clear and easy to understand framework to assess the final relative contribution of MCA and CBA to the outputs. On the other hand, it reveals flaws in the valuation of the differences in performance, which in turn determine greatly the effect in the final prioritization. An additional problem of the approach, also shared by the second alternative (expanding a CBA with MCA attributes, ExternE approach) concerns the incorporation of non-monetizable items to benefits, but not to costs. While this may change the value for money of the alternatives (its intended effect) it may also turn poor performing alternatives into cost-beneficial ones, a transformation of unclear decision-support value.

The integration of CBA into MCA frameworks can naturally avoid some of these weaknesses while having others of its own. Given its flexibility, it is not surprising that it is the approach that is more represented in the literature. In this regard, the application of the two proposed experimental approaches: introducing social economic costs and benefits separately; and introducing a "BCR" criterion into a MCA framework, may add value in a promising area or application. Common strengths to both approaches include: 1) the realization of the difference in social vs governmental costs and benefits, which is of great interest for the decision process; 2) they can be applied both to decision-support processes that are stakeholder-driven and to those that are driven by the analyst or planner; and 3) it can be applied regardless of whether the original MCA includes economic and monetizable criteria or not.

Specifically for the separate introduction of social economic costs and a clear strength is that it clearly spells out those criteria for stakeholders so that, for instance, different weighs can be applied to either. On the other hand, the introduction of costs and benefits separately does not include a scaling effect, so differences in net benefit or value for money would not translate in a direct proportion into differences in alternatives' performance. Moreover, with an "outranking" MCA methodology, the actual difference between costs or benefits across alternatives would not matter for the prioritization, since only the relative order would matter for the preference flows. Conversely, the proposed introduction of a "BCR" criterion into a MCA framework allows for an integrated visualization of the CBA performance of the actions compared with their nonmonetary performance, and the potential related trade-offs embedded in policy decisions. As a price to pay for this synthetic value, the MCA+BCR is no longer in accordance to the original priority setting by the stakeholders. In the MALC case study this drawback cannot be avoided, short of presently unfeasible successive rounds of iterative involvement of the stakeholders in the MCA analysis. However, in subsequent applications, the entire process of the MCA, CBA and combination could be agreed upon with the stakeholders from the beginning, who would be involved crucially at a second step in the reallocation of weights in the integrated MCA+BCR. The entire process is illustrated in Annex 7. The comparison of the prioritization of alternatives under MCA, CBA and the new MCA+BCR can further provide insights for decision-making, and the sequential application of CBA and MCA is not incompatible with the integration. For instance, either the MCA or the CBA can be "rescued" back after having been used in the preliminary screening, and become part of the integration.

Chapter 7. Conclusions

After a theoretical exploration and application to a real-life case study, it can be concluded that a number of valid methodologies exist to integrate the processes, elements, or outputs of a CBA and a MCA conducted on a given set of policy actions. Each has strengths and weaknesses, and none can be claimed to preserve all the desirable characteristics of both types of analysis. In fact, if the welfare economics foundations of CBA are fully and strictly accepted, there is no possible integration with MCA that would preserve the policy decision value of CBA outputs, as descriptors of social optimality.

However, in the real-life practice of decision support in public policy, there are reasons why that argument should not prevent an integration of CBA with other types of analysis, including MCA. The existing empirical evidence suggests that the impact of CBA into the final decision is limited at best, particularly if the decision is in the hands of politicians rather than planners. Eliasson & Lundberg (2010) found that the predictive value of BCRs was roughly as useful as a random pick of public transportation investments in Sweden, a finding confirmed in other countries with a long history of CBA application such as Norway and the UK. While planners find CBA useful, ultimate decision-makers place limited value on their results, considering it one of many decision inputs.

How this input could be used along with others is completely debatable, with the obvious potential pitfall of conflicting prioritizations. This phenomenon was evaluated by Tudela, Akiki & Cisternas (2006), who found in a reasonable large sample of projects that MCA prioritizations did not match CBA prioritizations – although they did roughly match the final decisions made by the competent authority, a conclusion confirmed later on by Khaki & Shafiyi (2011). This does not necessarily mean that the MCA results are a better descriptor of the decision-makers preferences in the context of the prioritization, since it could well be that those preferences are shaped by the MCA process. Either way, all this presents a nuanced picture: the usefulness of CBA in policy appraisal is clear, but it cannot credibly be the only decision-support input. In other words, CBA results should be considered as one important criterion among several others, and MCA allows a flexible way to integrate CBA elements or outputs with other criteria, with advantages in terms of stakeholder engagement.

The main integration approaches that are conceptually solid enough for a plausible real-life application are:

- The post hoc inclusion of a new "CBA criterion" (e.g. BCR, or NPV) in a MCA framework stripped of monetizable criteria (Scenario 2 in this chapter). This approach is referred to as "MCA+BCR" for conciseness;
- The integration of monetized CBA elements as criteria in a MCA framework. This approach is referred to as "Gühnemann" for conciseness;
- The combination of MCA and CBA outputs in a final unified metric (e.g. the Total Rate of Return in the COSIMA approach).This approach is referred to as "COSIMA" for conciseness;
- The forced indirect monetization of MCA criteria that add to the benefits of the CBA. This approach is referred to as "ExternE" for conciseness.

How well do these approaches keep the advantages of each technique (CBA and MCA)? In the literature review chapter, a few of the relative advantages of each technique were listed based on various sources (section 2.4). Based on the case study application, it can be subjectively evaluated whether those comparative advantages are kept in each modality of MCA-CBA integration. Because of the practical approach to this evaluation, one more characteristic has been added: to avoid "indirect monetization"; insofar as this technique is not widely accepted, using it can be considered a conceptual disadvantage. Below is an overview of how well each mode of MCA-CBA integration keeps each desirable trait:

1. Comprehensive accounting: one of the main conceptual strengths of CBA is its ability of account for all (monetizable) social costs and benefits, something that is not a strict requisite for an MCA to be formally valid. Insofar as the full CBA is preserved in the MCA-CBA integration, it can be argued that the comprehensive accounting is also preserved. All modes of integration thus have this advantage, except the Gühnemann approach which does not include all CBA elements fully.
2. Minimization of double counting: the possibility of double counting is in theory minimized in a CBA if it is well conducted, whereas most MCA problems do not deal with this issue. In that regard, any MCA-CBA integration would carry the potential double counting internal to the MCA analysis. But what about double counting *between* CBA and MCA? MCA+BCR derives all possible overlap (i.e. monetizable criteria) to the new BCR criterion; the Gühnemann approach substitutes –does not duplicate- MCA elements with monetized counterparts; COSIMA considers only non-monetizable criteria in its MCA; and ExternE brings into the CBA elements that were not in it originally since they were not really monetizable. Therefore, this issue is accounted for and controlled in all modes of integration considered.
3. Low data requirements: though it is not always automatically the case, MCA is typically less data-intensive than CBA. For instance, a social CBA would frequently rely on quantitative data from external sources (e.g. prices, time use surveys, etc.) whereas a MCA may range from a strict focus on quantitative data-driven criteria performance measurements to a strong reliance in qualitatively assessed stakeholder preferences. However, as long as the CBA-related process of accounting and monetization has to be done, every MCA-CBA combination will be burdened by high data requirement.
4. Reflects societal preferences: as mentioned, CBA would in theory better represent societal preferences than MCA in general. Insofar as the MCA-CBA combination would not yield the exact same prioritization of alternatives than the CBA alone, the combination could not be said to fully represent social preferences. And vice versa, if the prioritization of the MCA-CBA combination differs from that of the MCA alone, the combination cannot be said to fully represent the preferences of stakeholders. This suggests a partial representation of social preferences, although it is unclear by how much in each combination. For instance, in the Gühnemann et al. (2012), monetized variables (theoretically reflecting societal preferences) enter the MCA as normal criteria with weights that are not assigned by society, but rather by the researchers or a group of stakeholders. Moreover, these monetized criteria are combined in the MCA with non-monetized ones, further diluting the purported representation of societal preferences. In short, by mixing together societal preferences and those of the stakeholders, the final evaluation may reflect neither, or partly both. This is the case for all combinations explored.
5. Reflects stakeholders' preferences: the same reasoning can be applied to this characteristic. As long as there is a clear influence of criteria and weights in the final decision, the preferences of

stakeholders can be said to be partly represented. That would be the case even in the absence of qualitative stakeholder-assessed criteria through the role of criteria weighting, so it applies to all combinations.

6. Clear solution through a complete one-metric prioritization of alternatives: this characteristic is essential to both MCA and CBA, and it is kept in all combinations tested.
7. Reflects time preference: CBA can reflect rates of time preference in their partial and final outputs, something MCA does not do. As long as these time preferences are carried on with the CBA element into the MCA-CBA combination (e.g. through alternative scenarios with different discount rates) a certain degree of time preference can be assumed to be kept in all combinations, although the extent is unclear. In the ExternE method, however, time preferences are fully kept since the final form of the analysis is a CBA that can accommodate discounting.
8. Visualization of trade-offs / compromise solutions: the ability of MCA to offer a compromise solution is closely related to its ability to display trade-offs, something that CBA is able to do only at the cost of its power of synthesis. In this regard, combination modes that keep the MCA framework (i.e. MCA+BCR and Gühnemann) keep the ability to visualize the trade-offs albeit with a different focus – MCA+BCR between criteria in general and between BCRs and non-monetizable criteria, Gühnemann between criteria, monetized or not. COSIMA can be argued to partially reflect trade-offs through the significance of the alpha parameter in the balance between CBA and MCA. By monetizing MCA variables into a joint CBA process, ExternE loses that ability to display trade-offs.
9. Minimizes risk of bias: as long as weights and criteria influence the prioritization of alternatives, no MCA-CBA combination can claim to keep the theoretical CBA's minimal risk of bias.
10. Replicability: whereas both CBA and MCA can be conducted in a fully replicable manner, some of that replicability can be lost in the MCA-CBA combination. Such is the case in the process of obtaining the "desirability currency" in COSIMA, and to a lesser extent on the reasons for the choice of the parameter θ in the Gühnemann approach. While conceptually debatable, the MCA-CBA combinations in MCA+BCR and ExternE are fully replicable.
11. Stakeholder involvement: as long as there is an MCA process that allows for it (i.e. is not fully driven and conducted by the analyst), there is in principle potential for stakeholder involvement in any MCA-CBA combination.
12. Accepts monetizable and non-monetizable variables: the main reason to attempt the combination of MCA and CBA is to account for all types of variables regardless of whether they can be monetized; in consequence, all combinations accept in principle all types of variables.
13. Avoids forced monetization: a pure social CBA would typically include only variables for which widely agreed upon monetization techniques exist, whereas a MCA does not formally require the monetization of any criterion. Because of its lack of conventional economics basis, forced monetization is arguably best avoided. The separation between (traditionally) monetizable and non-monetizable is visible in both MCA+BCR and Gühnemann. However, the distinction is blurred in COSIMA through the "desirability currency" conversion. While ExternE forces monetization in a replicable manner through the weight-driven calculation, it similarly circumvents conventional monetization techniques.

This comparative evaluation is summarized in Table 46 below:

Table 46. Integration of MCA and CBA – comparative assessment of strengths and weaknesses

Desirable trait	CBA	MCA	MCA+BCR	Gühnemann	COSIMA	ExternE
Comprehensive accounting	✓	✗	✓	✗	✓	✓
Minimizes double counting	✓	✗	✓	✓	✓	✓
Low data requirements	✗	✓	✗	✗	✗	✗
Reflects societal preferences	✓	✗	?	?	?	?
Reflects stakeholders' preferences	✗	✓	?	?	?	?
Complete one-metric prioritization	✓	✓	✓	✓	✓	✓
Reflects time preference	✓	✗	?	?	?	✓
Visualization of trade-offs	✗	✓	✓	✓	✓	✗
Minimizes risk of bias	✓	✗	✗	✗	✗	✗
Replicability	✓	✓	✓	?	✗	✓
Stakeholder involvement	✗	✓	✓	✓	✓	✓
Accepts monetizable and NM variables	✗	✓	✓	✓	✓	✓
Avoids forced monetization	✓	✓	✓	✓	✗	✗

The comparative evaluation of MCA-CBA integration approaches confirms the notion that some of the comparative advantages of each technique are lost in the main types of integration considered thus far in the literature and published practice. This suggests that a mode of integration may best be chosen depending on the necessary desirable traits and the acceptable drawbacks in the specific decision-making problem at hand.

Is there, then, potential value in the integrative combination of MCA and CBA in the context of public environmental policy-making?

In light of the literature and the comparative application of approaches to a case study, the answer is relative and best considered in reference to the separate values of CBA and MCA for decision-making.

The existing evidence suggests that the impact of CBA into the final decision is limited at best, particularly if the decision is in the hands of politicians rather than planners. The main reason is that CBA cannot capture important non-monetizable dimensions of policy alternatives. Therefore, despite its usefulness in policy appraisal, CBA cannot be the only evidence-based decision-support input. In other words, CBA results should be considered as one important criterion among several others.

Based on that premise, it can be argued that an MCA framework provides a replicable and potentially transparent means to enumerate and simultaneously consider CBA-related criteria and others, while facilitating stakeholder engagement in the process. In that regard, there would be added value to the MCA-CBA integration in making explicit all nominally important considerations, regardless of whether they can be monetized or not.

Furthermore, the simultaneous consideration of societal preferences, as reflected in a CBA, along with non-monetizable variables and the subjective preferences of stakeholders may show hidden trade-offs that would not be acceptable or change their views once revealed. If the decision that suits best the most important criteria for the stakeholders contradicts the prioritization based on benefit to cost ratios, this may affect the final decisions of the decision-maker.

None of the available approaches for the combination of CBA and MCA may credibly diminish the often cited drawback of subjectivity in MCA. In fairness, however, it is arguable whether subjectivity in decision-making could or should be reduced; the best aspiration of a decision-support tool may be to manage subjectivity (Olson, 2006) in a way that is as transparent as possible. By highlighting the alignment or misalignment of the stakeholders' decisions to those that a CBA deems best for society, MCA-CBA combinations may contribute to shed light on some of that subjectivity.

In summary, insofar as both MCA and CBA provide inherently partial views of the implications of policy actions, there is room for their simultaneous application. While their separate application (either simultaneously or sequentially) may be useful for screening purposes or a broader overview of impacts, integrative approaches can provide added value by revealing important trade-offs of potential decisions.

REFERENCES

Achillas, C., Vlachokostas, C., Moussiopoulos, N., & Banias, G. (2011). Prioritize strategies to confront environmental deterioration in urban areas: Multicriteria assessment of public opinion and experts' views. Cities, 28(5), 414–423. http://doi.org/10.1016/j.cities.2011.04.003

AEAT. (2007). Evaluation of the costs and benefits of the implementation of the IPPC Directive on Large Combustion Plants. Retrieved from http://www.cafe-cba.org/assets/ippc_ec_lcplant.pdf

Allen, I., & Seaman, C. (2007). Likert scales and data analyses. Quality Progress. Retrieved from http://mail.asq.org/quality-progress/2007/07/statistics/likert-scales-and-data-analyses.html

Anderson, J. E. (1994). Public policymaking (2nd Ed.). Princeton, NJ: Houghton Mifflin.

ARAPER. (2011). Peruvian Association of Automotive Sellers. Retrieved December 10, 2012, from http://www.aap.org.pe/

Arrow, K. (1963). Social Choice and Individual Values (2nd Ed.). New York, NY: Wiley.

Arrow, K., Daily, G. C., Dasgupta, P., Levin, S., & Maler, K.-G. (2000). Managing ecosystem resources. Environ. Sci. Technol., (34), 1401–6.

Atkinson, G., & Mourato, S. (2008). Environmental Cost-Benefit Analysis. Annual Review of Environment and Resources, 33(1), 317–344. http://doi.org/10.1146/annurev.environ.33.020107.112927

Austin, D. H., Krupnick, A. J., & McConnell, V. D. (1997). Efficiency and Political Economy of Pollution Control with Ancillary Benefits: An Application to NOX Control in the Chesapeake Bay Airshed. Discussion Paper 97-34. Washington, D.C., USA: Resources for the Future. Retrieved from http://www.rff.org/files/sharepoint/WorkImages/Download/RFF-DP-97-34.pdf

Banks, G. (2009). Evidence-based policy making: What is it? How do we get it? Canberra, Australia: (ANU Public Lecture Series, presented by ANZSOG, 4 February), Productivity Commission. Retrieved from http://www.pc.gov.au/__data/assets/pdf_file/0003/85836/cs20090204.pdf

Banville, C., Landry, M., Martel, J. M., & Boulaire, C. (1998). A stakeholder approach to MCDA. Systems Research and Behavioral Science, (15), 15–32.

Barfod, M. B., Salling, K. B., & Leleur, S. (2011). Composite decision support by combining cost-benefit and multi-criteria decision analysis. Decision Support Systems, 51(1), 167–175. http://doi.org/10.1016/j.dss.2010.12.005

Bateman, I., Carson, R. T., Day, B., Hanemann, M., Hanley, N., Hett, T., … Swanson, J. (2002). Economic Valuation with Stated Preference Techniques: A Manual. (Edward Elg). Cheltenham.

Bell, M. L., Cifuentes, L. A., Davis, D. L., Cushing, E., Telles, A. G., & Gouveia, N. (2011). Environmental health indicators and a case study of air pollution in Latin American cities. Environmental Research, 111(1), 57–66. http://doi.org/10.1016/j.envres.2010.10.005

Bell, M. L., Davis, D. L., Gouveia, N., Borja-Aburto, V. H., & Cifuentes, L. A. (2006). The avoidable health effects of air pollution in three Latin American cities: Santiago, São Paulo, and Mexico City. Environmental Research, 100(3), 431–440. http://doi.org/10.1016/j.envres.2005.08.002

Belton, V., & Stewart, T. (2001). Multiple Criteria Decision Analysis: An Integrated Approach. (Kluwer Aca).

Bergson, A. (1938). A Formulation of Certain Aspects of Welfare Economics. Quarterly Journal of Economics, 52(2), 310–334.

Blackman, A., Newbold, S., Shih, J., & Cook, J. (2000). The Benefits and Costs of Informal Sector Pollution Control : Mexican Brick Kilns. Washington, D.C., USA. Retrieved from http://www.rff.org/files/sharepoint/WorkImages/Download/RFF-DP-00-46.pdf

Boadway, R. (1974). The Welfare Foundations of Cost-Benefit Analysis. Economic Journal, 84, 926–939.

Boyssou, D. (1996). Outranking relations: do they have special properties? Journal of Multi-Criteria Decision Analysis, 5, 99–111.

Brans, J. (1982). L'ingénièrie de la décision; Elaboration d'instruments d'aide à la décision. La méthode PROMETHEE. In R. Nadeau & M. Landry (Eds.), L'aide à la décision: Nature, Instruments et Perspectives d'Avenir (pp. 183–213). Québec, Canada: Presses de l'Université Laval.

Brans, J. (1996). The space of freedom of the decision maker modelling the human brain. European Journal of Operational Research, 92(3), 593–602.

Brans, J., & Mareschal, B. (2005). PROMETHEE Methods. In J. Figueira, S. Greco, & M. Ehrgott (Eds.), Multiple Criteria Decision Analysis: State of the Art Surveys (pp. 163–196). Boston, Dordrecht, London: Springer-Verlag.

Brans, J., & Mareschal, B. (2012). Visual PROMETHEE 1.0 Manual, version 1.0.1 - November 27, 2012. Retrieved from http://www.promethee-gaia.net/files/VPManual.pdf

Brans, J., Vincke, P., & Mareschal, B. (1986). How to select and how to rank projects: the PROMETHEE method. European Journal of Operational Research, 24, 228–23.

Brekke, K., & Howarth, R. (2000). The social contingency of wants. Land Economics, 76(November), 493–503.

Brook, R. (2010). Particulate Matter Air Pollution and Cardiovascular Disease : An Update to the Scientific Statement From the American Heart Association. Circulation, 121, 2331–2378.

Broome, J. (1992). Counting the Cost of Global Warming. Cambridge: White Horse Press.

Brown, R. (2005). The Operation Was a Success but the Patient Died: Aider Priorities Influence Decision Aid Usefulness. Interfaces, 35(6), 511.

Bullock, H., Mountford, J., & Stanley, R. (2001). Better policy-making. London. Retrieved from http://www.civilservant.org.uk/library/policy/2001_cmps-better_policy_making.pdf

Carnevale, C., Pisoni, E., & Volta, M. (2008). A multi-objective nonlinear optimization approach to designing effective air quality control policies. Automatica, 44(6), 1632–1641. http://doi.org/10.1016/j.automatica.2008.04.001

Carson, R. T. (2000). Contingent valuation: an user's guide. Environ. Sci. Technol., 34, 1413–18.

CGIAL. (2011a). II PLAN INTEGRAL DE ATMOSF{É}RICO PARA LIMA - CALLAO PISA 2011-2015. Lima.

CGIAL. (2011b). Segundo plan integral de atmosférico para Lima - Callao. Lima, Callao. Retrieved from http://eudora.vivienda.gob.pe/OBSERVATORIO/PISA_MUNICIPALIDADES/LimaCallao/II_Plan_Integral_de_Saneamiento_Atmosferico_Lima_Callao_PISA_2011_2015.pdf

Chevrolet. (2010). Peru retail website. Retrieved September 12, 2014, from http://www.chevrolet.com.pe/

Chiabai, A., Galarraga, I., Markandya, A., & Pascual, U. (2012). The Equivalency Principle for Discounting the Value of Natural Assets: An Application to an Investment Project in the Basque Coast. Environmental and Resource Economics. http://doi.org/10.1007/s10640-012-9589-8

CIES. Politica tributaria para 2011-2016 (2011). Lima, Perú. Retrieved from http://elecciones2011.cies.org.pe/documentos-de-politica/politicas-tributarias.html

CIFOR. (1999). Guidelines for Applying Multi-Criteria Analysis to the Assessment of Criteria and Indicators. The Center for International Forestry Research, Jakarta, Indonesia. Retrieved from http://www.cifor.org/livesinforests/publications/pdf_file

Cifuentes, L. A., Krupnick, A. J., Ryan, R. O., & Toman, M. A. (2005). Urban Air Quality and Human Health in Latin America and the Caribbean. Washington, D.C., USA. Retrieved from https://publications.iadb.org/handle/11319/2988

CLG. (2009). Multi-criteria analysis : a manual. United Kingdom Government, Department for Communities and local governments. London, UK. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/7612/1132618.pdf

Cohen, A. J., Anderson, H. R., Ostro, B., Pandey, K. D., Krzyzanowski, M., Künzli, N., … Smith, K. R. (2004). Urban Air Pollution. In Comparative Quantification of Health Risks (Vol. 77, pp. 1353–1434). Retrieved from http://cdrwww.who.int/publications/cra/chapters/volume2/1353-1434.pdf

Concepcion, M. (2014). Estrategia nacional para combustibles y vehiculos mas limpios en el peru. Lima, Perú. Retrieved from http://www.minam.gob.pe/calidadambiental/wp-content/uploads/sites/22/2014/06/Presentaci%25C3%25B3n_estrategia_transporte_limpio_Peru_final-MINAM-2.pdf

CPGNV. (2014). Analisis del parque vehicular: ventajas del GNV. Lima, Perú.

Crabbé, A., & Leroy, P. (2008). The Handbook of Environmental Policy Evaluation. London: Earthscan.

Cropper, M., & Oates, W. (1992). Environmental Economics : A Survey. Journal of Economic Literature, 30(2), 675–740.

De Leeneer, I., & Pastijn, H. (2002). Selecting land mine detection strategies by means of outranking MCDM techniques. European Journal of Operational Research, 139(2), 327– 338.

DEFRA. (2003). Use of Multi-Criteria Analysis in Air Quality Policy: A Report Prepared for the Department for Environment, Food and Rural Affairs by Catalize, Inc. London, UK. Retrieved from http://uk-air.defra.gov.uk/reports/cat09/0711231556_MCDA_Final.pdf

DEFRA. (2004). Valuation of health benefits associated with reductions in air pollution. Retrieved from http://www.defra.gov.uk

DEFRA. (2007). Integrating Cost-Benefit Analysis and Multi-Criteria Analysis of Flood and Coastal Erosion Risk Management Projects. Joint Defra/EA Flood and Coastal Erosion Risk Management R&D Programme. R&D Project Record FD2018/PR2. London, UK. Retrieved from http://evidence.environment-agency.gov.uk/FCERM/Libraries/FCERM_Project_Documents/FD2018_5184_PR_pdf.sflb.ashx

DEFRA. (2011). Social Impacts and Wellbeing : multi-criteria analysis techniques for integrating non-monetary evidence in valuation and appraisal. Defra Evidence and Analysis Series Paper 5. London, UK.

DETR. (1998a). A New Deal for Trunk Roads in England. Department of the Environment, Transport and the Regions. London, UK. Retrieved from http://www.semmms.info/wp-content/uploads/2016/06/A-new-deal-for-Trunk-Roads-in-England-1998-PDF-479Kb.pdf

DETR. (1998b). A New Deal for Trunk Roads in England: Guidance to the New Approach to Appraisal. Department of the Environment, Transport and the Regions. London, UK. Retrieved from http://www.semmms.info/wp-content/uploads/2016/06/A-new-deal-for-Trunk-Roads-in-England-1998-PDF-479Kb.pdf

Diakoulaki, D., & Grafakos, S. (2004). ExternE-Pol Multicriteria Analysis. Final report on Work Package 4. Retrieved from http://www.externe.info/externe_d7/sites/default/files/expolwp4.pdf

DIGESA. (2012). Estudio de saturación de la calidad del aire en Lima y Callao 2011: vigilancia confiable para la protección de la salud de las personas y de su entorno. Dirección General de Sanidad Ambiental, Ministerio de salud, Lima, Peru.

Dobes, L., & Bennett, J. (2009). Multi-Criteria Analysis: "Good enough" for government work? Agenda, 16(3), 7–29.

Dockery, D. W., & Pope, C. A. (2006). Health effects of fine particulate air pollution: lines that connect (2006 critical review). Journal of the Air & Waste Management Association, 56, 709–742.

Dockery, D. W., Pope, C. A. I., Xu, X., Spengler, J. D., Ware, J. H., Fay, M. E., … Speizer, F. E. (1993). An association between air pollution and mortality in six US cities. New England Journal of Medicine, 329, 1753–1759.

Doran, G. (1981). There's a S.M.A.R.T. way to write management's goals and objectives. Management Review, 70(11), 35–36.

EC. (1996). Cost-benefit and multi-criteria analysis for new road construction, Transport Research EURET Concerted Action 1.1. Brussels, Belgium. Retrieved from http://orbit.dtu.dk/en/publications/transport-research-euret-concerted-action-11(75f5f650-592d-4f63-800e-2c712b23631a).html

Eckstein, O. (1958). Water Resource Development: the Economics of Project Evaluation. Cambridge, Mass.: Harvard University Press.

EFORWOOD. (2011). A technical report documenting the results of the MCA and CBA procedures for a regional-defined single chain in Baden-Württemberg. Joensuu, Finland. Retrieved from http://www.efi.int/files/attachments/publications/eforwood/efi_tr_49.pdf

Eftim, S., Samet, J., Janes, H., McDermott, A., & Dominici, F. (2008). Fine particulate matter and mortality: a comparison of the six cities and American Cancer Society cohorts with a medicare cohort. Epidemiology, 19(2), 209–216.

Eliasson, J. (2009). A cost–benefit analysis of the Stockholm congestion charging system. Transportation Research Part A: Policy and Practice, 43(4), 468–480. http://doi.org/10.1016/j.tra.2008.11.014

Eliasson, J., & Lundberg, M. (2010). Do Cost – Benefit Analyses Influence Transport Investment Decisions ? Experiences from the Swedish Transport Investment Plan 2010 –2021. Transport Reviews : A Transnational Transdisciplinary Journal, 32(August 2012), 37–41.

EU. (2005). The Communication on Thematic Strategy on Air Pollution and The Directive on "Ambient Air Quality and Cleaner Air for Europe" Impact Assessment. European Commission, Brussels, Belgium. Retrieved from The Communication on Thematic Strategy on Air Pollution and The Directive on "Ambient Air Quality and Cleaner Air for Europe" Impact Assessment.

EU. (2008). Guide to cost benefit analysis of investment projects. Brussels, Belgium.

EUNET/SASI. (2001). Final Report – Executive Summary, 4th RTD Framework Programme of the European Commission. European Commission, Brussels, Belgium. Retrieved from http://www.transport-research.info/sites/default/files/project/documents/eunet.pdf

Fehr, E., & Gachter, S. (2000). Cooperation and punishment in public good experiments. American Economic Review, 90(September), 980–994.

Figueira, J., & Roy, B. (2002). Determining the weights of criteria in the ELECTRE type methods with a revised Simos' procedure. European Journal of Operational Research, 139(2), 317–326. http://doi.org/10.1016/S0377-2217(01)00370-8

Follesdal, D. (1982). The status of rationality assumptions in interpretation and in the explanation of action. Dialectica, 36(4), 301–316.

Follesdal, D. (2004). Ethical aspects of risk. In C. Streffer (Ed.), Low Dose Exposure in the Environment. Risk Assessments and Regulatory Processes. Berlin: Springer.

Garmendia, E., Gamboa, G., Franco, J., Garmendia, J. M., Liria, P., & Olazabal, M. (2010). Social multi-criteria evaluation as a decision support tool for integrated coastal zone management. Ocean & Coastal Management, 53(7), 385–403. http://doi.org/10.1016/j.ocecoaman.2010.05.001

Geldermann, J., Spengler, T., & Rentz, O. (2001). Fuzzy outranking for environmental assessment. Case study: iron and steel making industry. Fuzzy Sets and Systems, 115(1), 45–65.

Goicoechea, A., Hansen, D., & Duckstein, L. (1982). Introduction to multiobjective analysis with engineering and business applications. New York, NY, USA.: John Wiley publishers.

Golub, E., Klytchnikova, I., Sanchez Martinez, G., & Belausteguigoitia, J. C. (2014). Environmental Health Costs in Colombia: The Changes from 2002 to 2010. World Bank Occasional Paper Series, (92956). Retrieved from https://openknowledge.worldbank.org/bitstream/handle/10986/21096/929560WP0P14940s0occ.0paper0series0.pdf;sequence=1

Goulder, L., & Williams, R. (2012). Discussion paper: The Choice of Discount Rate for Climate Change Policy Evaluation. Washington, D.C., USA. Retrieved from http://rff.org/RFF/Documents/RFF-DP-12-43.pdf

Gowdy, J. M. M. (2004). The Revolution in Welfare Economics and its Implications for Environmental Valuation and Policy. Land Economics, 80(2), 239–257.

Gregory, C. E., & Brierley, G. J. (2010). Development and application of vision statements in river rehabilitation: the experience of Project Twin Streams, New Zealand. Area, 42(4), 468–478. http://doi.org/10.2307/40890904

Guariso, G., Pirovano, G., & Volta, M. (2004). Multi-objective analysis of ground-level ozone concentration control. Journal of Environmental Management, 71(1), 25–33. http://doi.org/10.1016/j.jenvman.2003.12.015

Gühnemann, A., Laird, J. J., & Pearman, A. D. (2012). Combining cost-benefit and multi-criteria analysis to prioritise a national road infrastructure programme. Transport Policy, 23, 15–24. http://doi.org/10.1016/j.tranpol.2012.05.005

Harrison, S. R., & Qureshi, M. E. (2000). Choice of stakeholder groups and members in multicriteria decision models. Natural Resources Forum, 24(1), 11–19. http://doi.org/10.1111/j.1477-8947.2000.tb00925.x

Henig, M. I., & Buchanan, J. T. (1996). Solving MCDM problems: Process concepts. Journal of Multi-Criteria Decision Analysis, 5(1), 3–11.

HMT. (2011). The green book: appraisal and evaluation in central government. (2nd ed.). Her Majesty´s Treasury, London, UK. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/220541/green_book_complete.pdf

Hoek, G., Krishnan, R. M., Beelen, R., Peters, A., Ostro, B., Brunekreef, B., & Kaufman, J. D. (2013). Long-term air pollution exposure and cardio- respiratory mortality: a review. Environmental Health : A Global Access Science Source, 12(1), 43. http://doi.org/10.1186/1476-069X-12-43

Holland, M. (2014). Cost benefit analysis of final policy scenarios for the EU clean air package. Laxemburg, Austria.

Holland, M., Watkiss, P., Pye, S., de Oliveira, A., & Van Regemorter, D. (2005). Cost-Benefit Analysis of Policy Option Scenarios for the Clean Air for Europe programme. Oxford. Retrieved from http://www.cafe-cba.org/assets/thematic_strategy_analysis_v3.pdf

Hutton, G. (2000). Considerations in evaluating the cost-effectiveness of environmental health interventions. Geneva, Switzerland. Retrieved from http://apps.who.int/iris/bitstream/10665/66744/1/WHO_SDE_WSH_00.10.pdf

INEI. (2011). Publicaciones de estadísticas vitales. Retrieved November 5, 2013, from http://www.inei.gob.pe/biblioteca-virtual/publicaciones-digitales/#url

INEI. (2012a). Comportamiento de la Econom{í}a Peruana. Web Del Instituto Nacional de Estadistica E Inform{á}tica Del Per{ú}.

INEI. (2012b). Comportamiento de la Economía Peruana. Retrieved November 5, 2013, from https://www.inei.gob.pe/biblioteca-virtual/boletines/pbi-trimestral/1/

INVERMET. (2009). Estudio inspecciones vehiculares 2007-2008. Lima, Callao. Retrieved from http://www.invermet.gob.pe/archivo/CUADRO-01.htm

Jalaludin, B., Salkeld, G., Morgan, G., Beer, T., & Bin Nisar, Y. (2009). A Methodology for Cost-Benefit Analysis of Ambient Air Pollution Health Impacts. Canberra, Australia. Retrieved from http://olr.npi.gov.au/atmosphere/airquality/publications/pubs/cost-benefit-analysis.pdf

Kahneman, D. (2003). A psychological perspective on Economics. American Economic Review, 93(May), 162–68.

Katsouyanni, K., Touloumi, G., Samoli, E., Gryparis, A., Le Tertre, A., Monopolis, Y., … Schwartz, J. (2001). Confounding and effect modification in the short term effects of ambient particles on total mortality: results from 29 European cities within the APHEA2 project. Epidemiology, 12(5), 521–531.

Katsouyanni, K., Touloumi, G., Spix, C., Schwartz, J., Balducci, F., Medina, S., … Anderson, H. R. (1997). Short-term effects of ambient sulphur dioxide and particulate matter on mortality in 12 European cities: results from time series data from the APHEA project. Air Pollution and Health: a European Approach. BMJ (Clinical Research Ed.), 314(7095), 1658–63.

Keelin, T. W. (1981). A Parametric Representation of Additive Value Functions. Management Science, 27(10), 1200–1208.

Keeney, R.L. Raiffa, H. (1993). Decisions with Multiple Objectives – Preferences and Value Tradeoffs. Cambridge University Press, UK.

Keeny, R. L., & Raiffa, H. (1976). Decisions with Multiple Objectives: Preferences and Value Trade-Offs. New York, NY, NY: Wiley.

Kersten, G. E., & Noronha, S. J. (1996). Comments to "Solving MCDM problems: Process concepts" by Henig and Buchanan. Journal of Multi-Criteria Decision Analysis, 5(1), 12–15.

Khaki, A., & Shafiyi, S. (2011). Comparison Between the Output of Cost Benefit Analysis and Multi-Criteria Analysis in Urban Transportation Investments. Australian Journal of Basic and Applied Sciences, 5(11), 667–677.

Kinney, P. L., & Nori-Sarma, A. (2011). Health and Economic Benefits of Clean Air Regulations. Joint Center for political and economic studies, Washington DC. USA. Retrieved from http://jointcenter.org/sites/default/files/White Paper_121611.pdf

Krewski, D., Jerrett, M., Burnett, R., Ma, R., Hughes, E., Shi, Y., … Thunt, M. (2009). Extended follow-up and spatial analysis of the American Cancer Society study linking particulate air pollution and mortality. In HEI Research Report 140. Boston, MA: Health Effects Institute.

Krupnick A. (2007). Mortality-risk valuation and age: stated preference evidence. Rev. Environ. Econ. Policy, 1, 261–82.

Laerhoven, H., Zaag-Loonen, H., & Derkx, B. (2004). A comparison of Likert scale and visual analogue scales as response options in children's questionnaires. Acta Paediatrica. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/j.1651-2227.2004.tb03026.x/abstract

Larsen, B. (2004). Cost of environmental damage in Colombia: A Socio-Economic and Environmental Health Risk Assessment. Washington, D.C., USA. Retrieved from http://www.bvsde.paho.org/texcom/cd050996/larsen.pdf

Larsen, B., & Strukova, E. (2005). Cost of Environmental Damage in Peru: An Analysis of Environmental Health and Natural Resources. Washington, D.C., USA. Retrieved from http://www.bjorn-larsen.com/reports_and_publications

Lee, J., & Jones, P. (2002). Cultural differences in responses to a Likert scale. Research in Nursing & …. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/nur.10041/abstract

Linkov, I., & Steevens, J. (2008). Chapter 35 Appendix A. Multi Criteria Decision Analysis. In Cyanobacterial Harmful Algal Blooms: State of the Science and Research Needs: State of the Science and Research Needs (pp. 815–830).

Little, I., & Mirrlees, J. (1974). Project Appraisal and Planning for Developing Countries. Oxford: Oxford Univ. Press.

Maass, A. (1962). Design of Water Resource Systems. New York, NY: Macmillan.

Macharis, C., De Witte, A., & Ampe, J. (2009). The Multi-Actor, Multi-Criteria Analysis methodology (MAMCA) for the evaluation of transport projects: Theory and Practice. Journal of Advanced Transportation, 43(2), 183–202.

Manheim, M. L., Suhrbier, J. H., Bennett, E. D., Neumann, L. A., Colcord Jr., F. C., & Reno Jr., A. T. (1975). Transportation Decision-Making: A Guide to Social and Environmental Considerations.National Cooperative Highway Research Program Report 156. Transportation Research Board, Washington, D.C.

Mareschal, B. (1988). Weight stability intervals in multicriteria decision aid. European Journal of Operational Research, 33, 54–64.

Martinez-Alier, J., Munda, G., & ONeill, J. (1998). Weak comparability of values as a foundation for ecological economics. Ecological Economics, 26, 277–286.

McKean, R. (1958). Efficiency in Government through Systems Analysis. New York, NY: Wiley.

Migliavacca, G. et al. (2011). The REALISEGRID cost-benefit methodology to rank pan-European infrastructure investments. In EPRI Workshop – Manchester 8 December 2011. Retrieved from http://realisegrid.rse-web.it/content/files/File/News/REALISEGRID at the ISGT Europe 2011/REALISEGRID EPRI.pdf

Moffett, A., & Sarkar, S. (2006). Incorporating multiple criteria into the design of conservation area networks: a minireview with recommendations. Diversity and Distributions, 12, 125–137. Retrieved from http://uts.cc.utexas.edu/~consbio/Cons/Minireview-Appendix.pdf

Molina, M. J., & Molina, L. T. (2005). Improving Air Quality In Megacities: Mexico City Case Study. Retrieved September 15, 2012, from http://peace-foundation.net.7host.com/file/Mexico City Project -Molina.pdf

Moylan, K. (2009). The Future of EU Cohesion Policy and its implications for Irish Regional. Paper presented at the Regional Science Association International Conference, Limerick, Ireland 3rd September 2009. Retrieved from http://www.bmwassembly.ie/publications/other_reports/The_Future_of_EU_Cohesion_Policy_and_its_Implications_for_Irish_Regional_Policy_Sept 2009.pdf

Munda, G. (2004). Social multi-criteria evaluation: Methodological foundations and operational consequences. European Journal of Operational Research, 158(3), 662–677. http://doi.org/10.1016/S0377-2217(03)00369-2

Munda, G. (2006). Social multi-criteria evaluation for urban sustainability policies. Land Use Policy, 23(1), 86–94. http://doi.org/10.1016/j.landusepol.2004.08.012

Murphy, J., Allen, P., Stevens, T., & Weatherhead, D. (2005). A meta-analysis of hypothetical bias in stated preference valuation. Environ. Resour. Econ., 30, 313–25.

Naylor, C., & Appleby, J. (2013). Environmentally sustainable health and social care: Scoping review and implications for the English NHS. Journal of Health Services Research & Policy, 18(2), 114–121. http://doi.org/10.1177/1355819613485672

Nellthorp, J., Mackie, P., & Bristow, A. (1998). Measurement and Valuation of the Impacts of Transport Initiatives. Retrieved from http://ec.europa.eu/environment/archives/tremove/pdf/measurement_maintext.pdf

Niederman, M., McCombs, J., Unger, A., Kumar, A., & Popovian, R. (1999). Treatment cost of acute exacerbations of chronic bronchitis. Clinical Therapy, 21(3), 576–591.

Nordhaus, W. (2007). A review of the Stern Review on the economics of climate change. J. Econ. Lit., 45, 686–702.

NRA. (2008). National Roads Authority Project Appraisal Guidelines. National Roads Authority, Dublin. Retrieved from http://www.tollcompare.ie/policy-publications/project-appraisal-guideli/

NRA. (2011). Unit 6.0 Cost Benefit Analysis. Retrieved March 2, 2012, from http://www.nra.ie/policy-publications/project-appraisal-guideli/

O'Neill, M. S., Bell, M. L., Ranjit, N., Cifuentes, L. a, Loomis, D., Gouveia, N., & Borja-Aburto, V. H. (2008). Air pollution and mortality in Latin America: the role of education. Epidemiology (Cambridge, Mass.), 19(6), 810–819. http://doi.org/10.1097/EDE.0b013e3181816528

OECD. (2006). Economic Valuation of Environmental Health Risks to Children. Environmental Health. Paris, France. Retrieved from http://www.oecd-ilibrary.org/environment/economic-valuation-of-environmental-health-risks-to-children_9789264013988-en

OECD. (2008). Introductory Handbook for Undertaking Regulatory Impact Analysis (RIA). Paris, France. Retrieved from https://www.oecd.org/gov/regulatory-policy/44789472.pdf

OECD. (2010). Valuing lives saved from environmental, transport and health policies: a meta-analysis of stated preference studies (Vol. 33). Retrieved from http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?doclanguage=en&cote=env/epoc/wpnep(2008)10/final

OECD. (2012). Mortality risk valuation in environment, health and transport policies. Paris, France. Retrieved from http://www.oecd.org/env/tools-evaluation/mortalityriskvaluationinenvironmenthealthandtransportpolicies.htm#Table_of_Contents

OECD. (2014). The Cost of Air Pollution: Health Impacts of Road Transport. Paris, France, France. Retrieved from http://www.oecd-ilibrary.org/environment/the-cost-of-air-pollution_9789264210448-en

Olson, D. (2006). Subjectivity in Multiple Criteria Decision Analysis. Working paper. Retrieved June 12, 2014, from cbafiles.unl.edu/public/cbainternal/facStaffUploads/subjitor.doc

Olson, M., & Bailey, M. (1981). Positive Time Preference. Journal of Political Economy, 89(1), 1–25.

Opricovic, S., & Tzeng, G.-H. (2002). Multicriteria Planning of Post-Earthquake Sustainable Reconstruction. Computer-Aided Civil and Infrastructure Engineering, 17(3), 211–220. http://doi.org/10.1111/1467-8667.00269

Ortiz, R. A., Markandya, A., & Hunt, A. (2009). Willingness to Pay for Mortality Risk Reduction Associated with Air Pollution in São Paulo. RBE, 63(1), 3–22.

Ostro, B. (1994). Estimating the Health Effects of Air Pollution: A Method with an Application to Jakarta. Washington, D.C., USA. Retrieved from http://documents.worldbank.org/curated/en/355391468752348015/Estimating-the-health-effects-of-air-pollutants-a-method-with-an-application-to-Jakarta

Ostro, B. (2004). Outdoor Air Pollution - Assessing the environmental burden of disease at national and local levels. In Environmental Burden of Disease Series (Vol. 5). Geneva, Switzerland: World Health Organization.

Papaemmanouil, A Andersson, G. S. (2010). Towards Future Electricity Networks: Project Phase 2. Federal Office of Energy, Grid Research Program. Zurich, Switzerland. Retrieved from http://www.bfe.admin.ch/forschungnetze/01246/03569/index.html?dossier_id=03324&lang=en

Pearce, D. (1998). Cost-Benefit Analysis and environmental policy. Oxford Review of Economic Policy, 14(4), 84–100.

Pearce, D. (2003). The Social Cost of Carbon and its Policy Implications. Oxford Review of Economic Policy, 19(3), 362–384. http://doi.org/10.1093/oxrep/19.3.362

Pearce, D., Atkinson, G., & Mourato, S. (2006). Cost-benefit analysis and the environment: recent developments. Analysis. Paris, France: Organisation for Economic Co-operation and Development (OECD).

Pearce, D., Groom, B., Hepburn, C., & Koundouri, P. (2003). Valuing the future: recent advances in social discounting. World Economics, 4(2), 121–141.

Pictet, J., & Belton, V. (2000). ACIDE. Analyse de la compensation et de l'incomparabilité dans la décision. In A. Colorni, M. Paruccini, & B. Roy (Eds.), A-MCD-A Aide Multicritère à la Décision - Multiple Criteria Decision Aiding (pp. 245–256). Luxembourg: Joint Research Center, European Commission.

Pictet, J., & Bollinger, D. (2008). Extended use of the cards procedure as a simple elicitation technique for MAVT. Application to public procurement in Switzerland. European Journal of Operational Research, 185(3), 1300–1307. http://doi.org/10.1016/j.ejor.2006.05.051

Pisoni, E., Carnevale, C., & Volta, M. (2009). Multi-criteria analysis for PM10 planning. Atmospheric Environment, 43(31), 4833–4842. http://doi.org/10.1016/j.atmosenv.2008.07.049

Pomerol, J., & Romero, S. (2000). Multicriterion decision in management: principles and practice. Dordrecht, Netherlands.: Kluwer Academic Publishers.

Pope, C. A. I., Burnett, R. T., & Thun, M. J. (2002). Lung cancer, Cardiopulmonary mortality, and Long-term exposure to fine particulate air pollution. Journal of the American Medical Association, 287, 1132–1141.

Pope, C. A. I., Thun, M. J., Nambudiri, M. M., Dockery, D. W., Evans, J. S., Speizer, F. E., & Heath, C. W. J. (1995). Particulate air pollution as a predictor of mortality in a prospective study of US adults. American Journal of Respiratory and Critical Care Medicine, 151, 669–674.

Portney, P., & Weyant, J. (1999). Discounting and Intergenerational Equity. Resources for the Future,Washington, D.C. Retrieved from http://www.rff.org/files/sharepoint/WorkImages/Download/RFF-Bk-discounting-eq-ch1.pdf

Prüss-Üstün, A., Campbell-lendrum, D., Corvalán, C., Woodward, A., Prüss-Ustün, A., Mathers, C., … Woodward, A. (2003). Introduction and methods. Assessing the environmental burden of disease at national and local levels. Geneva: World Health Organization. Retrieved from http://www.who.int/quantifying_ehimpacts/publications/en/9241546204.pdf?ua=1

Randall, A. (1999). Taking benefits and costs seriously. In H. Folmer & T. Tietenberg (Eds.), The International Yearbook of Environmental and Resource Economics (1999/2000). Cheltenham: Edward Elgar.

Riabacke, M., Danielson, M., & Ekenberg, L. (2012). State-of-the-Art Prescriptive Criteria Weight Elicitation. Advances in Decision Sciences, 1–24. http://doi.org/10.1155/2012/276584

Roberts, R., & Goodwin, P. (2002). Weight Approximations in Multi-attribute Decision Models. Journal of Multi-Criteria Decision Analysis, 11, 291–303.

Roy, B. (1985). Méthodologie multicritère d'aide à la décision. Paris, France: Economica.

Roy, B. (1991). The outranking approach and the foundations of ELECTRE methods. Theory and Decision, 31, 49–73. Retrieved from http://www.lamsade.dauphine.fr/~mousseau/pmwiki-2.1.5/uploads/Research/Roy91.pdf

Saaty, R. (1987). The Analytical Hierarchy Process - What it is and how it is used. Mathl Modelling, 9(3–5), 161–176.

Saaty, T. L. (1980). The Analytic Hierarchy Process. New York, NY, NY: McGraw-Hill.

Saaty, T. L. (2001). Decision Making for Leaders: The Analytical Hierarchy Process for Decisions in a Complex World (Second vol). Pittsburgh, PA, USA: RWS Pubs.

Sagoff, M. (1994). Should preferences count? Land Economics, 70(May), 127–44.

Salling, K. B., Leleur, S., & Jensen, A. V. (2007). Modelling Decision Support and Uncertainty for Large Transport Infrastructure Projects: The CLG-DSS model of the Øresund Fixed Link. Decision Support Systems, 43, 1539–1547.

Samet, J., Dominici, F., Curriero, F., Coursac, I., & Zeger, S. (2000). Fine particulate air pollution and mortality in 20 U.S. cities, 1987-1994. N Engl J Med, 323(24), 1742–1749.

Samet, J., & Krewski, D. (2007). Health effects associated with exposure to ambient air pollution. J Toxicol Environ Health A, 70(3–4), 227–242.

Sanchez Triana, E., Ahmed, K., & Awe, Y. (2007). Environmental Priorities and Poverty Reduction: A Country Environmental Analysis for Colombia. Washington DC, USA: World Bank.

Sayers, T. M., Jessop, A. T., & Hills, P. J. (2003). Multi-criteria evaluation of transport options – flexible, transparent and user-friendly. Transport Policy, 10, 95–105.

Schutte, I., & Brits, A. (2012). Prioritising transport infrastructure projects: towards a multi-criterion analysis. Southern African Business Review, (3), 97–117.

Scitovsky, T. (1941). A Note on Welfare Propositions in Economics. Review of Economic Studies, 9, 77–88.

Shibuya, K., Mathers, C., & Lopez, A. (2001). Chronic Obstructive Pulmonary Disease (COPD): Consistent Estimates of Incidence Prevalence and Mortality by WHO Region. In Global Programme on Evidence for Health Policy. Geneva, Switzerland: World Health Organization.

Sijtsma, F. J. (2006). Project evaluation, sustainability and accountability : combining cost-benefit analysis (CBA) and multi-criteria analysis (MCA). Retrieved August 25, 2013, from http://dissertations.ub.rug.nl/faculties/eco/2006/f.j.sytsma/

Squire, L., & Van der Tak, H. (1974). Economic Analysis of Projects. Baltimore: Johns Hopkins Univ. Press.

SRB. (2001). Confronting COPD in North America and Europe: A Survey of Patients and Doctors. New York, NY, USA.

Stavins, R. N., Wagner, A. F., & Wagner, G. (2003). Interpreting Sustainability in Economic Terms: Dynamic Efficiency Plus Intergenerational Equity. Economics Letters, 79, 339–343.

Stern, N. (2007). The Economics of Climate Change—The Stern Review. Cambridge: Cambridge Univ. Press.

Stewart, T. J., & Losa, F. B. (2003). Towards reconciling outranking and value measurement practice. European Journal of Operational Research, 145, 645–659. http://doi.org/10.1016/S0377-2217(02)00221-7

SUNAT. (2012). Website de orientacion aduanera para la importacion de vehiculos. Retrieved April 15, 2014, from http://www.sunat.gob.pe/orientacionaduanera/importacionvehiculos/tributos_aplicables.html

Sutcliffe, S., & Court, J. (2005). Evidence-Based Policymaking: What is it? How does it work? What relevance for developing countries? Overseas Development Institute (ODI), London, UK.

Thomson, I., & Bull, A. (2002). La congestión del tránsito urbano: causas y consecuencias económicas y sociales. Revista de La Cepal, 76(Abril), 109–121. Retrieved from http://www.cepal.org/publicaciones/xml/6/19336/lcg2175e_bull.pdf

Tudela, A., Akiki, N., & Cisternas, R. (2006). Comparing the output of cost benefit and multi-criteria analysis: An application to urban transport investments. Transportation Research Part A, 40(5), 414–423. Retrieved from http://www.worldtransitresearch.info/research/3418

Tzeng, G.-H., Lin, C.-W., & Opricovic, S. (2005). Multi-criteria analysis of alternative-fuel buses for public transportation. Energy Policy, 33(11), 1373–1383. http://doi.org/10.1016/j.enpol.2003.12.014

Tzeng, G.-H., Tsaur, S.-H., Laiw, Y.-D., & Opricovic, S. (2002). Multicriteria analysis of environmental quality in Taipei: public preferences and improvement strategies. Journal of Environmental Management, 65(2), 109–120. http://doi.org/10.1006/jema.2001.0527

UKCO. (1999). Modernising government. United Kingdom Cabinet Office, London, UK. Retrieved from http://webarchive.nationalarchives.gov.uk/20140131031506/http://www.archive.official-documents.co.uk/document/cm43/4310/4310-00.htm

USEPA. (1995). Sample City Cost Analyses of Alternative Particulate Matter National Ambient Air Quality Standards: Philadelphia. Research Triangle Park: U.S. Environmental Protection Agency.

USEPA. (1997). Final report to Congress on Benefits and Costs of the Clean Air Act, 1970 to 1990. Document Number EPA 450-R-97-002. Washington, D.C.: U.S. Environmental Protection Agency. Retrieved from https://nepis.epa.gov/Exe/ZyNET.exe/00000ENR.TXT?ZyActionD=ZyDocument&Client=EPA&Index=1995+Thru+1999&Docs=&Query=&Time=&EndTime=&SearchMethod=1&TocRestrict=n&Toc=&TocEntry=&QField=&QFieldYear=&QFieldMonth=&QFieldDay=&IntQFieldOp=0&ExtQFieldOp=0&XmlQuery=

USEPA. (2000). Final Heavy Duty Engine / Diesel Fuel Rule : Air Quality Estimation , Selected Health and Welfare Benefits Methods , and Benefit Analysis Results. Research Triangle Park, NC. Retrieved from https://www3.epa.gov/ttn/ecas/regdata/Benefits/tsdhddv8.pdf

USEPA. (2004). Final Regulatory Analysis: Control of Emissions from Nonroad Diesel Engines. EPA 420-R-04-007. Office of Transportation and Air Quality. Washington DC: US EPA. Retrieved from https://nepis.epa.gov/Exe/ZyNET.exe/P10003DE.TXT?ZyActionD=ZyDocument&Client=EPA&Index=2000+Thru+2005&Docs=&Query=&Time=&EndTime=&SearchMethod=1&TocRestrict=n&Toc=&TocEntry=&QField=&QFieldYear=&QFieldMonth=&QFieldDay=&IntQFieldOp=0&ExtQFieldOp=0&XmlQuery=

USEPA. (2010a). Final Regulatory Impact Analysis (RIA) for the SO2 National Ambient Air Quality Standards (NAAQS). Research Triangle Park: U.S. Environmental Protection Agency. Retrieved from https://www3.epa.gov/ttnecas1/regdata/RIAs/fso2ria100602full.pdf

USEPA. (2010b). Regulatory Impact Analysis (RIA) for Existing Stationary Compression Ignition Engines NESHAP. Research Triangle Park: U.S. Environmental Protection Agency. Retrieved from https://www3.epa.gov/ttnecas1/docs/ria/ic-engines_ria_final-existing-ci-engines_2010-02.pdf

USEPA. (2011). United States Environmental Protection Agency website on air pollution trends. Retrieved February 2, 2013, from http://www.epa.gov/airtrends/

USEPA. (2012). United States Environmental Protection Agency website on ground-level ozone. Retrieved February 2, 2013, from http://www.epa.gov/groundlevelozone/health.html

Varian, H. R. (2006). Revealed preference. Retrieved from http://people.ischool.berkeley.edu/~hal/Papers/2005/revpref.pdf

Vasquez, A. (2006). El Valor de la Vida Estadística y sus aplicaciones a la Fiscalización de la Industria de Hidrocarburos. Lima, Perú. Retrieved from http://www.osinergmin.gob.pe/newweb/uploads/Estudios_Economicos/DT18_OSINERG.pdf

Velasquez, M., & Hester, P. (2013). An Analysis of Multi-Criteria Decision Making Methods. International Journal of Operational Research, 10(2), 55–56.

Vial, J., Perea, H., Grippa, F., Deza, M., & Sanchez, R. (2010). Peru situacion automotriz. Lima, Perú: BBVA research. Retrieved from https://www.bbvaresearch.com/KETD/fbin/mult/peru_automotriz_2010_tcm346-274709.pdf

Viscusi, W. K., & Aldy, J. E. (2003). The Value of a Statistical Life: A Critical Review of Market Estimates Throughout the World. Journal of Risk and Uncertainty, 27(1), 5–76. http://doi.org/10.1023/A:1025598106257

Vlachokostas, C., Achillas, C., Moussiopoulos, N., & Banias, G. (2011). Multicriteria methodological approach to manage urban air pollution. Atmospheric Environment, 45(25), 4160–4169. http://doi.org/10.1016/j.atmosenv.2011.05.020

Von Neumann, J., & Morgenstern, O. (1947). Theory of Games and Economic Behaviour, second edition. Princeton, NJ, NJ: Princeton University Press.

von Winterfeld, D., & Edwards, D. (1986). Decision Analysis and Behavioural Research. Cambridge University Press.

Voorhees, A. S., Araki, S., Sakai, R., & Sato, H. (2000). An ex post cost-benefit analysis of the nitrogen dioxide air pollution control program in Tokyo. Journal of the Air & Waste Management Association, 50, 391–410.

Voorhees, S. S., Sakai, R., Araki, S., Sato, H., & Otsu, a. (2001). Cost-benefit analysis methods for assessing air pollution control programs in urban environments-A review. Environmental Health and Preventive Medicine, 6(2), 63–73. http://doi.org/10.1007/BF02897948

Wang, J.-J., Jing, Y.-Y., Zhang, C.-F., & Zhao, J.-H. (2009). Review on multi-criteria decision analysis aid in sustainable energy decision-making. Renewable and Sustainable Energy Reviews, 13(9), 2263–2278. http://doi.org/10.1016/j.rser.2009.06.021

Watkiss, P. (2006). Damage costs for air pollution. Final report to DEFRA. Oxford. Retrieved from http://archive.defra.gov.uk/environment/quality/air/airquality/panels/igcb/documents/dcs-report2006.pdf

Wenstop, F. (2005). Mindsets, Rationality and Emotion in Multi-criteria Decision Analysis. J. Multi-Crit. Decis. Anal., 13, 161–172.

Wenstop, F., & Seip, K. (2001). Legitimacy and quality of multi-criteria environmental policy analysis: a meta-analysis of five MCE studies in Norway. Journal of Multi-Criteria Decision Analysis, (10), 53–64.

WHO. (2001). Global Burden of Disease 2001. Geneva, Switzerland. Retrieved from http://www.who.int/healthinfo/global_burden_disease/estimates_regional_2001/en/

WHO. (2006). Air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide. Global update 2005. Summary of risk assessment. Geneva, Switzerland: World Health Organization. Retrieved from http://www.who.int/phe/health_topics/outdoorair/outdoorair_aqg/en/

WHO. (2008). WHO guide for standardization of economic evaluations of immunization programmes. Geneva, Switzerland: World Health Organization. Retrieved from http://apps.who.int/iris/bitstream/10665/69981/1/WHO_IVB_08.14_eng.pdf

WHO. (2013). Health risks of air pollution in Europe – HRAPIE project Recommendations for concentration–response functions for cost–benefit analysis of particulate matter, ozone and nitrogen dioxide. Copenhagen, Denmark: World Health Organization Regional Office for Europe. Retrieved from http://www.euro.who.int/__data/assets/pdf_file/0006/238956/Health-risks-of-air-pollution-in-Europe-HRAPIE-project,-Recommendations-for-concentrationresponse-functions-for-costbenefit-analysis-of-particulate-matter,-ozone-and-nitrogen-dioxide.pdf

WHO. (2015). Calculation of DALYs. Retrieved November 8, 2015, from http://www.who.int/healthinfo/global_burden_disease/metrics_daly/en/

WHO. (2016). Global Urban Ambient Air Pollution Database (update 2016). Retrieved from http://www.who.int/phe/health_topics/outdoorair/databases/cities/en/

Yitzhaki, S. (2003). Cost-benefit analysis and the distributional consequences of government projects. National Tax Journal, 56(2), 319–336.

Young, E., & Quinn, L. (2002). Writing Effective Public Policy Papers A Guide for Policy Advisers in Central and Eastern Europe. Local government and public service reform initiative. Open Society Institute, Budapest, Hungary.

ANNEXES

ANNEX 1. Baseline data to calculate health cost per case

Table A1. Resource use and cost data for the cost of illness assessment

Resource use and cost data	Estimate	Source
Cost Data for All Health End-Points:		
Cost of hospitalization (Soles per day)	355	Per consultations with medical service providers and health authorities (unpublished); rounded to the nearest thousand Soles
Cost of emergency visit (Soles): urban	195	
Cost of doctor visit (Soles) (mainly private doctors): urban	110	
Value of time lost to illness (Soles per day)	41	Based on urban wages in Peru (INEI, 2012a)
Chronic Bronchitis (CB):		
Average duration of Illness (years)	20	Based on Shibuya, Mathers, & Lopez (2001)
Percent of CB patients hospitalized per year (%)	1.5%	From SRB (2001), Niederman, et al. (1999), and personal communication: Pierpaolo Mudu (WHO) and Elena Strukova (WB)
Average length of hospitalization (days)	10	
Average number of doctor visits per CB patient per year (visits)	5	
Percent of CB patients with an emergency doctor/hospital outpatient visit per year (%)	15.0%	
Estimated lost workdays (including household workdays) per year per CB patient (days)	6	
Annual real increases in economic cost of health services and value of time (real wages) (%)	2.0%	Larsen (2004), Larsen & Strukova (2005)
Annual discount rate (%)	3.0%	WHO, (2008)
Hospital Admissions:		
Average length of hospitalization per case (days)	6	Larsen (2004), Larsen & Strukova (2005)
Average number of days lost to illness per case (after hospitalization)	4	
Emergency Room Visits:		
Average number of days lost to illness per case	2	Larsen (2004), Larsen & Strukova (2005)
Lower Respiratory Illness in Children:		
Number of doctor visits per case	1	Larsen (2004), Larsen & Strukova (2005)
Total time of caregiving by adult per case (days)	1	
Respiratory symptoms:		
Unsubsidized drug cost per case (Soles)	5	Larsen (2004), Larsen & Strukova (2005)

Note: the methodological framework, the template for data collection and assumptions, and the structure of this table come from the standard World Bank methodology to calculate the health cost of environmental degradation. Examples can be found at the published studies by Golub et al. (2014); Larsen & Strukova (2005); and Larsen (2004), among others.

ANNEX 2. Semi-structured interview (Spanish)

PARTE A.- GUION PARA LA ENTREVISTA
1.- Explicar contexto y motivo de la entrevista
(Saludo, presentación)

El Comité de Gestión de la Iniciativa de Aire Limpio aprobó recientemente el Segundo Plan Integral de Saneamiento Atmosférico para Lima-Callao (2011-2015) con el objetivo de reducir los niveles de contaminación hasta límites permisibles y minimizar los efectos de los contaminantes en la salud de los habitantes de la zona metropolitana. Tras la aprobación de este plan, y con vistas a futuros planes y políticas en el mismo sentido, el CGIAL requirió la asistencia del Instituto de Aire Limpio en Washington, DC (USA) para realizar una evaluación de las medidas y políticas del PISA-II. Como primer paso para esta evaluación, el CGIAL está conduciendo una serie de entrevistas con actores clave para el planeamiento e implementación de medidas para mejorar la calidad del aire. Usted ha sido indicado como uno de estos actores clave.

El objetivo de las entrevistas es obtener una panorámica de los puntos de vista de tomadores de decisiones, expertos y actores relevantes acerca de la calidad del aire en Lima-Callao y también (especialmente) de la gestión que se está haciendo de esa contaminación. También querríamos que, tomando como referencia algunas de las acciones contempladas en el PISA-II nos indique cuáles de estas medidas deberían ser implementadas con mayor urgencia y por qué. Por otra parte, estamos interesados en saber qué medidas pueden ser retrasadas y qué razones o criterios tiene usted en cuenta en este juicio.

Las entrevistas son parte de un proceso más amplio por el cual el CGIAL quiere 1) Hacer una evaluación objetiva de las políticas consideradas y 2) Establecer un sistema de toma de decisiones transparente y basado en la evidencia. Con las entrevistas se elaborara un diagnostico en base al cual el CGIAL le enviara información relevante para continuar este proceso. Para mas información puede ponerse en contacto con (datos de contacto de Gladis, Gerardo, Willy, Cristóbal).
Para efectos de comunicaciones con el CGIAL, ¿Tendría una tarjeta de contacto (business card)? (si no la tiene, anotar datos básicos de contacto)

La entrevista durará unos 45 minutos y será grabada para nuestra referencia[10] (mostrar la grabadora). A continuación encenderé la grabadora y procederé a preguntarle su Nombre, Institución y Cargo. Después comenzaremos con los temas específicos de la entrevista. ¿Está de acuerdo? ¿Alguna pregunta antes de comenzar?

(ENTREVISTADO ESTÁ DE ACUERDO; COMENZAR LA GRABACIÓN)

0. POR FAVOR, INDIQUE SU NOMBRE, CARGO Y ORGANIZACIÓN (¿Nivel Local, Provincial, Nacional, Internacional?)
1. Relación del entrevistado con la gestión de la calidad del aire en la zona de Lima Callao
    a) Interés, responsabilidad y/o aportación del entrevistado (o su organización) en la mejora de la calidad del aire en la zona de Lima-Callao
    b) Relación específica con PISA-II (¿Diseño, análisis, responsabilidad de implementación, capacidad de veto, necesidad de ser informado,…?

2. Percepción de la calidad del aire y de su gestión en Lima-Callao (trayectoria, presente, perspectivas)
    a) Sondear la opinión del entrevistado (Desde la perspectiva de su cargo y organización, o en su defecto personal) sobre el estado de la calidad del aire en la zona de Lima-Callao y su trayectoria histórica en los últimos años (7 a 10).

[10] Si el entrevistado pregunta sobre la grabación, aclarar que el único propósito de estas grabaciones es de investigación para el CGIAL y nunca serán cedidas, ni total ni parcialmente a terceros. La coordinación del CGIAL será el custodio de las mismas y serán eliminadas tras su análisis cualitativo.

b) Sondear la opinión del entrevistado sobre el efecto de la gestión de la calidad del aire sobre la contaminación atmosférica en la zona de Lima Callao

3. Barreras para una efectiva gestión de la calidad del aire (o si el entrevistado conoce el PISA-II, barreras para la efectiva implementación de las medidas del plan) ESCOGER UNA DE LAS OPCIONES LISTADAS DEPENDIENDO DEL PERFIL DEL ENTREVISTADO
   - Barreras para una efectiva gestión de la calidad del aire en la zona de Lima-Callao
   - Principales barreras a la implementación del Plan en su conjunto
   - Principales obstáculos para conjuntos específicos de medidas (FF, FM, C, I, IG). Escoger el/los sectores que por su perfil sean mas cercanos al entrevistado.

4. Factores determinantes para una exitosa gestión de la calidad del aire (o si el entrevistado conoce el PISA-II, para la efectiva implementación de las medidas del plan)
   - Factores o eventos fundamentales para que las medidas que mejoran la calidad del aire se puedan implementar de manera efectiva

5. Dimensiones, elementos y/o criterios clave para la priorización de medidas en PISA-II
   - ¿Qué sector seria para usted el prioritario y por qué? (insistir hasta obtener una razón o criterio de fondo)
   - Dentro de cada sector (ir sector por sector) ¿Qué medida es para usted la de mayor prioridad y por qué? (insistir hasta obtener una razón o criterio de fondo)
   - Respecto al empaquetado de las medidas (Fuentes Fijas, Fuentes Móviles, Combustibles, Impuestos, Gestión), le parece a usted a adecuado para su priorización, o agruparía usted las medidas de otra manera (P.ej. paquetes de implementación, etc.)
   - ¿Qué sector seria para usted el menos prioritario (el que se podría dejar para mas tarde) y por qué? (insistir hasta obtener una razón o criterio de fondo)
   - Dentro de cada sector (ir sector por sector) ¿Qué medida es para usted la menos prioritaria (el que se podría dejar para mas tarde) y por qué? (insistir hasta obtener una razón o criterio de fondo)

6. Preguntas complementarias
   - Actores relevantes adicionales

PARTE B.- PLANTILLA DE RECOLECCION DE INFORMACION CLAVE
NOMBRE ENTREVISTADOR:
NOMBRE Y CARGO DEL ENTREVISTADO:
FECHA Y LUGAR DE LA ENTREVISTA:
TEMA 1: Relación del entrevistado con la gestión de la calidad del aire en la zona de Lima Callao. Preguntas tipo: Por favor, describa su relación con la gestión de la calidad del aire en la zona de Lima-Callao. ¿Ha estado directa o indirectamente involucrado en la elaboración o implementación de los planes de saneamiento atmosférico, y en concreto del PISA-II?

Tipo de relación	Descripción
¿Miembro del CGIAL?	
¿Asesor del CGIAL?	
Actor sectorial (Empresas, instituciones públicas, etc.)	
Sociedad civil (ONGs, etc.)	
Otros (explicar)	

TEMA 2: Calidad y gestión del aire en Lima Callao

A.- Percepción de la calidad del aire (trayectoria, presente, perspectivas)

Preguntas tipo: ¿Cuál es su perspectiva institucional sobre el estado actual de la calidad del aire en la zona de Lima-Callao? ¿Ha cambiado (y si es así, cómo) esta situación en los últimos diez años?

Concepto evaluado	Evaluación u opinión	Observaciones
Estado presente de la calidad del aire	Mala	
	Ni buena ni mala	
	Moderadamente buena	
	Buena	
	Otros (especificar)	
Trayectoria reciente de la calidad del aire	Ha empeorado	
	Ni ha empeorado ni ha mejorado	
	Ha mejorado	
	Otros (especificar)	

B.- Opinión sobre la gestión de la calidad del aire en Lima-Callao

Preguntas tipo: Según su opinión, ¿Ha tenido la gestión que las autoridades han hecho de la calidad del aire efecto sobre la misma? Si es así, ¿en qué sentido? (mejorado, empeorado, otros)

Que efecto ha tenido la gestión de las autoridades sobre la calidad del aire en la zona de Lima-Callao	La gestion ha contribuido a empeorar la calidad del aire	
	La gestion no ha influenciado ni positiva ni negativamente a la calidad del aire	
	La gestion ha contribuido a mejorar la calidad del aire	
	Otros (especificar)	

TEMA 3.- Barreras para una efectiva gestión de la calidad del aire

Pregunta tipo: Usted ha sido designado como actor clave por el CGIAL por su relación o influencia en la gestión de la calidad del aire en la zona de Lima-Callao. Desde su perspectiva, ¿han existido barreras significativas para dicha gestión? Si es así, ¿de qué tipo? Por favor, describa dichas barreras en detalle.

Nota: importante hacer al entrevistado etiquetar el tipo de barrera. Algunos tipos de barreras frecuentes a la implementación de políticas ambientales son: 1) económicas 2) institucionales (P.ej. falta de voluntad política, falta de recursos, falta de cooperación, etc.) 3) regulatorias (P.ej. falta de un marco regulatorio adecuado) 4) Grupos de interés (empresas, sindicatos, lobbies, etc.)

TIPO DE BARRERA	DESCRIPCION

TEMA 4.- Factores determinantes para una exitosa gestión de la calidad del aire

Pregunta tipo: Desde su perspectiva, ¿Qué factores han sido o serían determinantes para una exitosa gestión de la calidad del aire en la zona de Lima-Callao? Por favor, describa dichos factores en detalle y su influencia sobre las iniciativas relevantes.

Nota: importante hacer al entrevistado poner un nombre a esos factores necesarios para implementar las medidas. Algunos factores que suelen incrementar el éxito o efectividad de políticas ambientales son: 1) Suficientes recursos humanos y económicos 2) Colaboración entre las instituciones 3) Consenso político sobre las medidas 4) Un marco regulatorio adecuado 5) Una ciudadanía consciente de la problemática ambiental, etc.

FACTOR DETERMINANTE	DESCRIPCION

Tema 5.- Criterio clave para la priorización de medidas que mejoran la calidad del aire[11]

Preguntas tipo: ver el guión del entrevistador.

(Dar al entrevistado la tabla de medidas y leer el siguiente texto). Supongamos ahora un caso hipotético. Usted ha sido encargado de implementar un plan de mejora de la calidad del aire para Lima-Callao. Sin embargo, por escasez de tiempo y recursos no es posible implementar todas las medidas en el periodo especificado y debe usted priorizar algunas de estas medidas, dejando el resto para después. Así, puede usted iniciar algunas medidas o sectores para implementar en el año 1, otras en el año 2, etc… dejando para el final otras. En ese contexto, le voy a plantear algunas preguntas

CRITERIO CLAVE	DESCRIPCION

TEMA 6.- Actores clave a involucrar

Pregunta tipo: Usted ha sido designado como actor clave por el CGIAL, el cual incorpora varios actores sectoriales relevantes para su función consultiva en materia de calidad del aire. Desde su perspectiva, ¿deberían incluirse a otros actores en los procesos de planeamiento e implementación de políticas de calidad del aire? Si es así, ¿podría decirnos qué personas, organizaciones o entidades representan a esos actores?

[11] Nota: las preguntas en esta sección quedan a criterio del entrevistador. El objetivo es obtener motivos últimos por los cuales ciertas medidas son consideradas prioritarias por parte del entrevistado.

PARTE C.- Ejemplo ilustrativo de medidas a priorizar (FUENTE: PROGRAMA TRES DEL PISA II)
Fuentes Móviles

- Prohibir la importación de vehículos usados y su ingreso al parque automotor
- Programa de Renovación del Parque automotor para vehículos livianos.
- Bonos de chatarreo para vehículos de Transporte Público
- Promoción del uso del GNV y GLP en transporte Público
- Fiscalización de los Centros de Inspección Técnica Vehicular.
- Control de emisiones vehiculares en vías públicas

Fuentes Fijas

- Regulación de emisiones a actividades de alta emisión de PM10, PM2.5, $SO2$ y NOx.
- Limitar el uso de tecnologías de producción obsoletas y agresivas al ambiente
- Instrumentos económicos y financieros que permitan la renovación tecnológica
- Implementación de sistemas de control de la contaminación atmosférica en industrias o actividades económicas localizadas en zonas de alta contaminación.
- Fiscalización de Fuentes Macroemisoras de PM10, PM2.5, SO2 y NOx

Combustibles

- Disminución del contenido de azufre en combustibles
- Fiscalización de la venta de combustibles de baja concentración de azufre
- Promover el cambio de uso de combustibles a nivel industrial
- Promover la eliminación del uso de aditivos metálicos en los combustibles

Tributación / Impuestos

- Disminuir el arancel de importación de vehículos nuevos
- Disminuir el Impuesto al Patrimonio Vehicular
- Generar y aplicar un Impuesto a la Antigüedad Vehicular
- Desarrollar instrumentos económicos y financieros que permitan el uso de GNV y GLP principalmente en vehículos de transporte público
- Establecer impuestos, contribuciones u otras compensaciones a las emisiones de contaminantes al aire.

Instrumentos de Gestión

- Establecer Normas que limiten las emisiones de contaminantes al aire
- Sistema de vigilancia y pronóstico de la Calidad del Aire en Lima-Callao
- Registro de emisiones de contaminantes del aire en Lima- Callao
- Desarrollar e implementar el sistema de vigilancia epidemiológica de enfermedades causadas por la Contaminación del Aire
- Desarrollar e implementar el sistema de fiscalización ambiental de nivel Metropolitano
- Fortalecimiento de capacidades locales responsables de la gestión de la calidad del aire
- Desarrollar e implementar programas de sensibilización de la población respecto a la problemática de la calidad del aire en Lima–Callao

ANNEX 3. Questionnaire for qualitative criteria in MCA (in Spanish)

En el marco de su asistencia técnica al Comité de Gestión de la Iniciativa de Aire Limpio (CGIAL), el Instituto de Aire Limpio (Washington, DC) le pide su ayuda con un breve cuestionario.

Queremos comprender mejor la perspectiva que actores clave como usted tienen respecto de los efectos de políticas frecuentemente propuestas para la mejora de la calidad del aire. Sus respuestas se utilizarán para una prueba de Análisis Multi Criterio de las políticas enunciadas. Sus consideraciones cualitativas se combinarán con insumos cuantitativos para una priorización teórica de las medidas. Por supuesto, usted recibirá los resultados principales de dicho ejercicio. A continuación enunciaremos cuatro paquetes de políticas de transporte que pueden resultar en la mejora de la calidad del aire. Le pedimos lea con atención la descripción del paquete de medidas y después conteste a la pregunta de acuerdo a su propia opinión o juicio.

Paquete 1: Modernización del parque automotor de Lima- Callao. El objetivo de esta política es reducir la antigüedad media de los vehículos circulantes desde 17 años (actual) hasta 10 años (meta) en un plazo aproximado de cinco años. Para ello se tomarán las siguientes medidas: 1) Prohibir el ingreso de vehículos de segundo uso al parque automotor 2) Disminuir el arancel de importación de vehículos nuevos 3) Disminuir el Impuesto al Patrimonio Vehicular 4) Generar y aplicar un Impuesto a la Antigüedad Vehicular

Señale su grado de conformidad con las siguientes afirmaciones

P1a. Esta medida será aceptada de buen grado por el público en general

Totalmente en desacuerdo	En desacuerdo	Ni de acuerdo ni en desacuerdo	De acuerdo	Totalmente de acuerdo

Observaciones:

P1b. Las autoridades podrán fiscalizar esta medida adecuadamente en su implementación

Totalmente en desacuerdo	En desacuerdo	Ni de acuerdo ni en desacuerdo	De acuerdo	Totalmente de acuerdo

Observaciones:

P1c. Esta medida puede beneficiar especialmente a grupos vulnerables de población (los más pobres, ancianos, niños, discapacitados, etc)

Totalmente en desacuerdo	En desacuerdo	Ni de acuerdo ni en desacuerdo	De acuerdo	Totalmente de acuerdo

Observaciones:

P1d. El tiempo que tomará implementar esta medida será de aproximadamente:

De 1 a 2 años	De 3 a 5 años	De 6 a 8 años	De 8 a 10 años	Más de 10 años

Observaciones:

P1e: Esta medida acarrea significativos incentivos económicos que ayudarán a su implementación

Totalmente en desacuerdo	En desacuerdo	Ni de acuerdo ni en desacuerdo	De acuerdo	Totalmente de acuerdo

Observaciones:

Paquete 2: Modernización de la flota de vehículos de transporte público de Lima- Callao: el objetivo de esta política es la renovación de los vehículos más antiguos de la flota de transporte público mediante el otorgamiento de Bonos de chatarreo para vehículos de Transporte Publico.

Señale su grado de conformidad con las siguientes afirmaciones

P2a. Esta medida será aceptada de buen grado por el público en general

Totalmente en desacuerdo	En desacuerdo	Ni de acuerdo ni en desacuerdo	De acuerdo	Totalmente de acuerdo

Observaciones:

P2b. Las autoridades podrán fiscalizar esta medida adecuadamente en su implementación

Totalmente en desacuerdo	En desacuerdo	Ni de acuerdo ni en desacuerdo	De acuerdo	Totalmente de acuerdo

Observaciones:

P2c. Esta medida puede beneficiar especialmente a grupos vulnerables de población (los más pobres, ancianos, niños, discapacitados, etc)

Totalmente en desacuerdo	En desacuerdo	Ni de acuerdo ni en desacuerdo	De acuerdo	Totalmente de acuerdo

Observaciones:

P2d. El tiempo que tomará implementar esta medida será de aproximadamente:

De 1 a 2 años	De 3 a 5 años	De 6 a 10 años	Más de 10 años	Otros (explicar)

Observaciones:

P2e: Esta medida acarrea significativos incentivos económicos que ayudarán a su implementación

Totalmente en desacuerdo	En desacuerdo	Ni de acuerdo ni en desacuerdo	De acuerdo	Totalmente de acuerdo

Observaciones:

Paquete 3: Promoción del Gas Natural Vehicular (GNV): desarrollar instrumentos económicos y financieros que incrementen el uso del GNV en la flota vehicular de Lima-Callao, particularmente en vehículos de transporte público. El objetivo final es una penetración del GNV del 15% de la flota de transporte público.

Señale su grado de conformidad con las siguientes afirmaciones

P3a. Esta medida será aceptada de buen grado por el público en general

Totalmente en desacuerdo	En desacuerdo	Ni de acuerdo ni en desacuerdo	De acuerdo	Totalmente de acuerdo

Observaciones:

P3b. Las autoridades podrán fiscalizar esta medida adecuadamente en su implementación

Totalmente en desacuerdo	En desacuerdo	Ni de acuerdo ni en desacuerdo	De acuerdo	Totalmente de acuerdo

Observaciones:

P3c. Esta medida puede beneficiar especialmente a grupos vulnerables de población (los más pobres, ancianos, niños, discapacitados, etc)

Totalmente en desacuerdo	En desacuerdo	Ni de acuerdo ni en desacuerdo	De acuerdo	Totalmente de acuerdo

Observaciones:

P3d. El tiempo que tomará implementar esta medida será de aproximadamente:

De 1 a 2 años	De 3 a 5 años	De 6 a 10 años	Más de 10 años	Otros (explicar)

Observaciones:

P3e: Esta medida acarrea significativos incentivos económicos que ayudarán a su implementación

Totalmente en desacuerdo	En desacuerdo	Ni de acuerdo ni en desacuerdo	De acuerdo	Totalmente de acuerdo

Observaciones:

Paquete 4: Promoción de un mejor mantenimiento y funcionamiento vehicular: se quiere reducir las emisiones del parque automotor mediante las siguientes medidas: 1) Fiscalización de los Centros de Inspección Técnica Vehicular 2) Control de emisiones vehiculares en vías públicas. El objetivo es conseguir que en un plazo de cinco años el 80% de la flota vehicular sea sometida periódicamente a inspecciones técnicas adecuadas.

Señale su grado de conformidad con las siguientes afirmaciones

P4a. Esta medida será aceptada de buen grado por el público en general

Totalmente en desacuerdo	En desacuerdo	Ni de acuerdo ni en desacuerdo	De acuerdo	Totalmente de acuerdo

Observaciones:

P4b. Las autoridades podrán fiscalizar esta medida adecuadamente en su implementación

Totalmente en desacuerdo	En desacuerdo	Ni de acuerdo ni en desacuerdo	De acuerdo	Totalmente de acuerdo

Observaciones:

P4c. Esta medida puede beneficiar especialmente a grupos vulnerables de población (los más pobres, ancianos, niños, discapacitados, etc)

Totalmente en desacuerdo	En desacuerdo	Ni de acuerdo ni en desacuerdo	De acuerdo	Totalmente de acuerdo

Observaciones:

P4d. El tiempo que tomará implementar esta medida será de aproximadamente:

De 1 a 2 años	De 3 a 5 años	De 6 a 10 años	Más de 10 años	Otros (explicar)

Observaciones:

P4e: Esta medida acarrea significativos incentivos económicos que ayudarán a su implementación

Totalmente en desacuerdo	En desacuerdo	Ni de acuerdo ni en desacuerdo	De acuerdo	Totalmente de acuerdo

Observaciones:

ANNEX 4. Weight stability intervals for original MCA

Figure A4.1. Criterion: Burden of Disease; WSI = [9.20%; 31.00%]

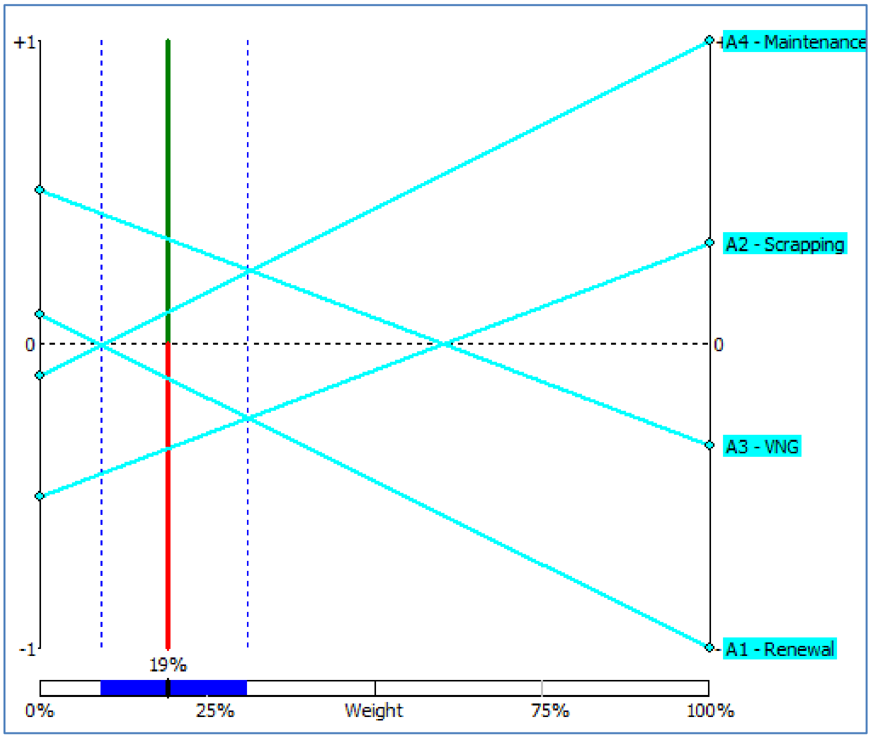


Figure A4.2. Criterion: Inclusion; WSI = [0.00%; 26.21%]

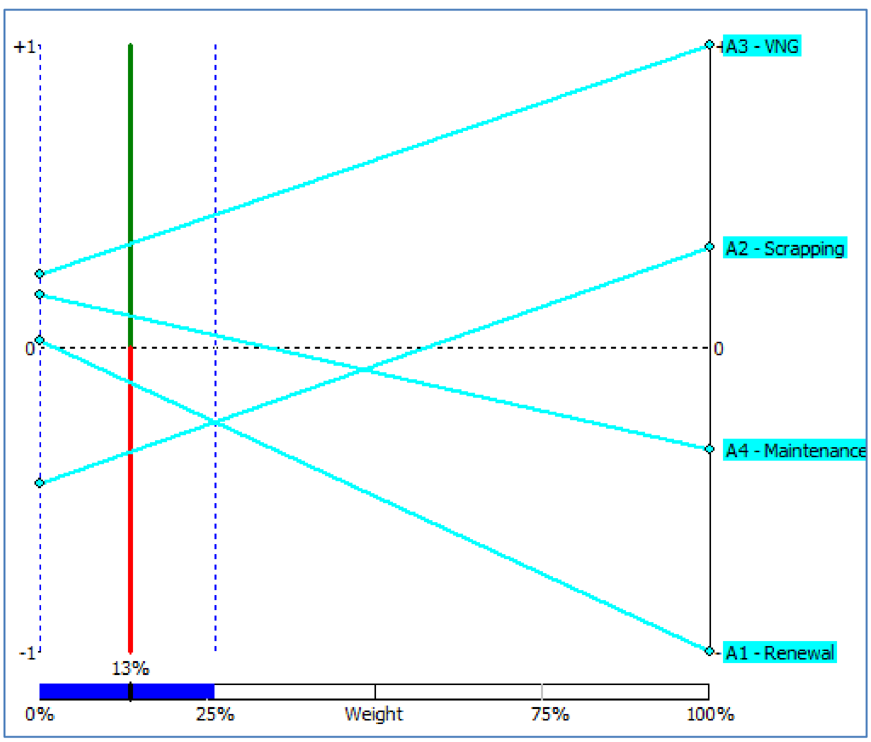


Figure A4.3. Criterion: Public Opinion; WSI = [0.00%; 24.30%]

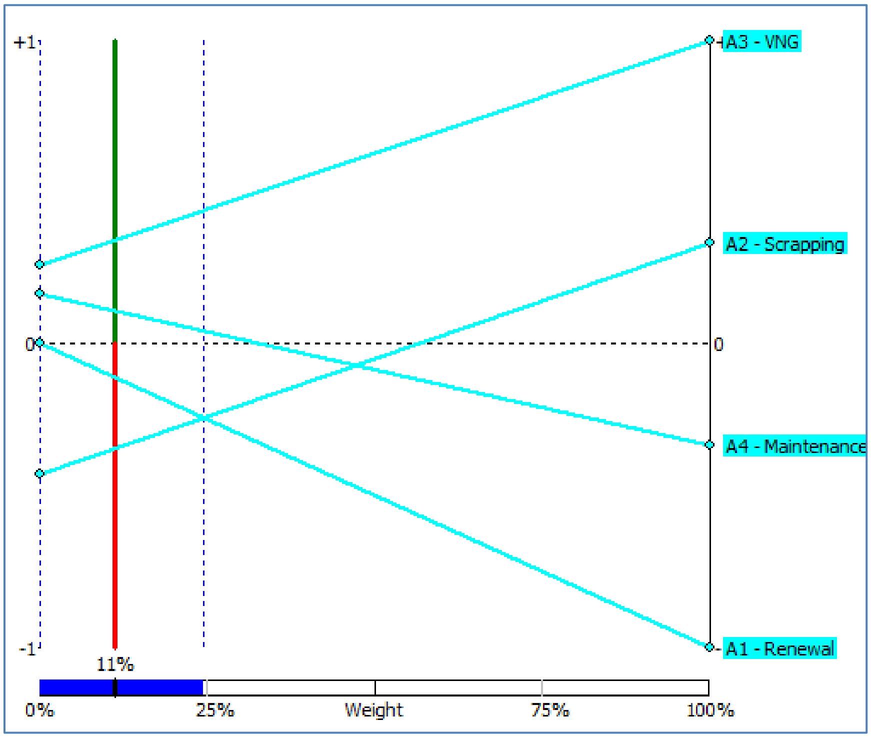
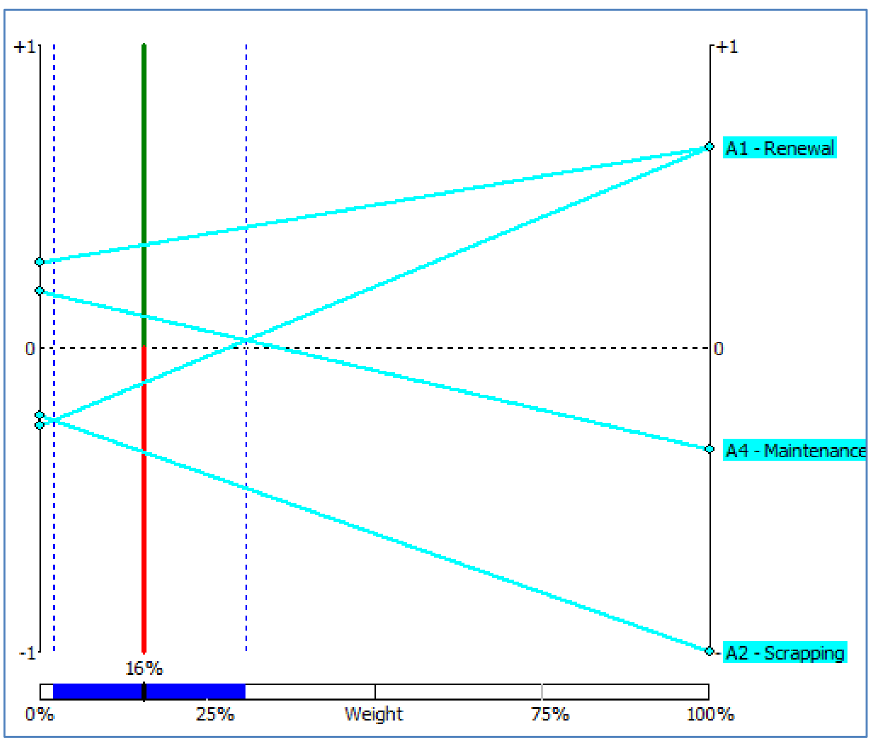


Figure A4.4. Criterion: Enforcement; WSI = [1.98%; 30.63%]



[Note: A3 is represented by the uppermost line]

Figure A4.5. Criterion: Implementation; WSI = [8.07%; 30.08%]

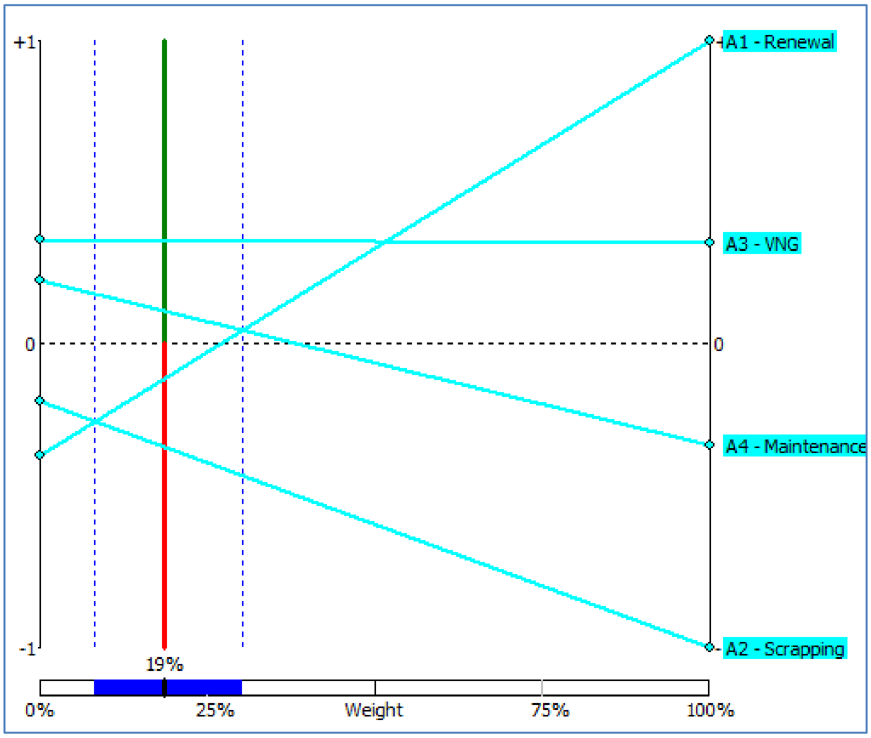


Figure A4.6. Criterion: Cost; WSI = [0.00%; 20.66%]

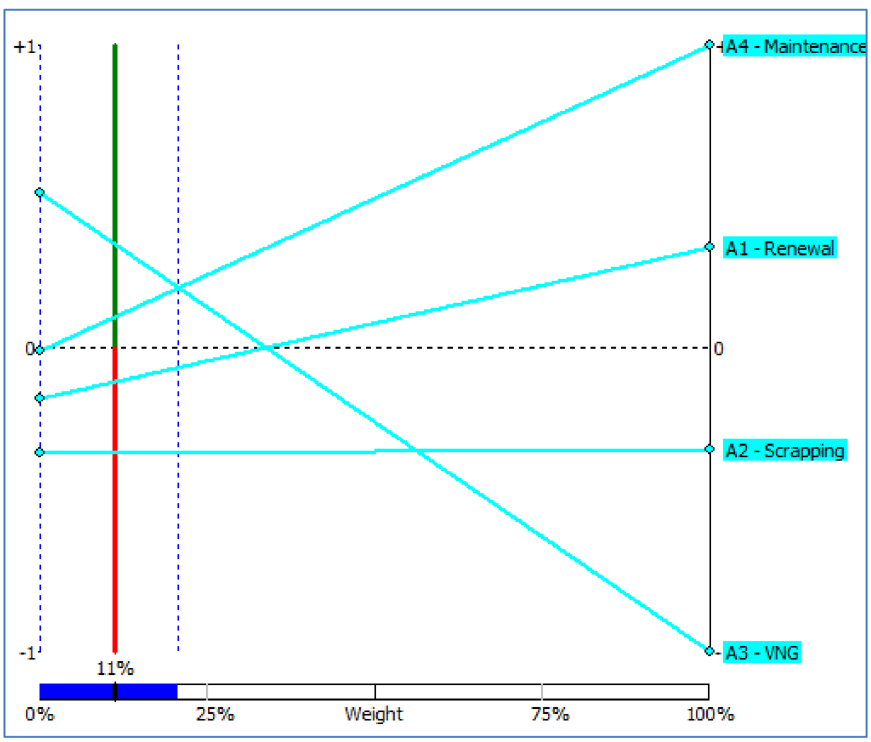


Figure A4.7. Criterion: Incentives; WSI = [0.00%; 100.00%]

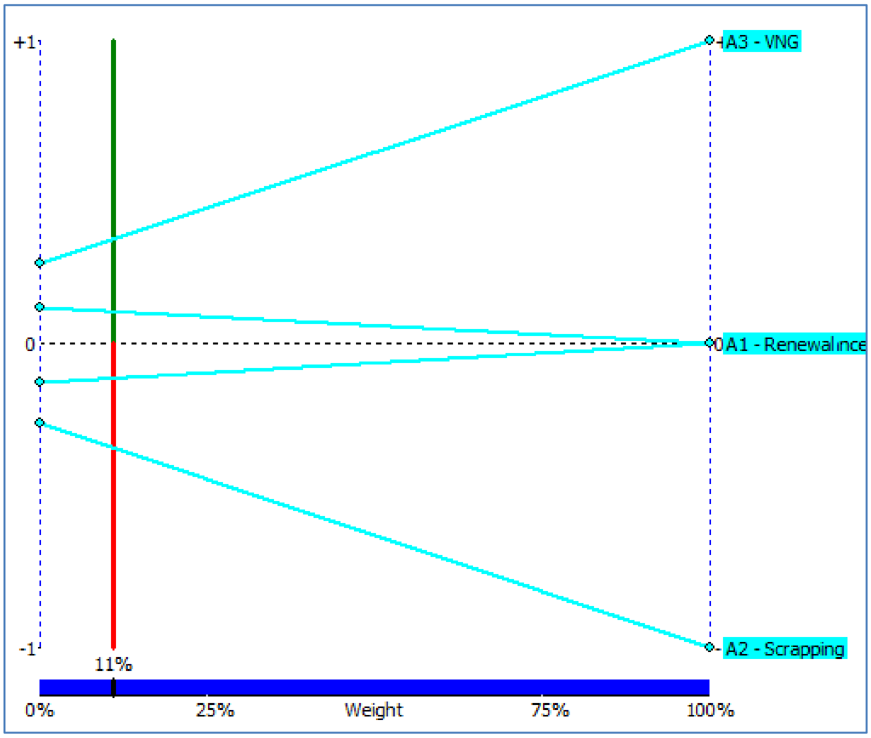


Table A4.1. Comparison of WSI with a plus/minus 50% of the weight interval

Criterion	w_j^-	$w_j - 0.5w_j$	$w_j + 0.5w_j$	w_j^+
BoD	0.0920	0.0955	0.2865	0.3100
Enforce.	0.0198	0.0776	0.2327	0.3063
Implem.	0.0807	0.0933	0.2798	0.3008

ANNEX 5. Effect of normalization in the COSIMA approach on the case study

In the baseline MCA problem for the application of the COSIMA approach (Barfod et al., 2011) to the MALC case study, all criteria except "implementation" share the same scale (1 to 5) and preference structure (better is more, and there is no indifference threshold), and are unit-less. By contrast, the "implementation" criterion is measured in years and its preference structure is reversed (less is better). For the purposes of the translation into "Assessment currency", I transformed its preference structure into a maximization one by subtracting the implementation time from an arbitrary maximum of 10 years, but I did not transform the values of "implementation" into a 1 to 5 scale. Below are the results of such normalization, for contrast with the chosen path. We can transform the values to a 1 to 5 unit-less scale through a basic normalization:

$$X_{is} = 1 + 4 \times \left(\frac{X_i - X_{min}}{X_{max} - X_{min}} \right) \qquad \text{Eq. A5.1}$$

Where:
X_{is} is the scaled value of the original number
X_i is the original number
X_{min} is the minimum value of X
X_{max} is the maximum value of X

The transformed values of the modified "implementation" criterion are listed in table A5.1 below:

Table A5.1. Scaling of "implementation" criterion

Action	Implementation (years)	Time left from a 10 year horizon	Equivalent on a 1 to 5 scale
A1 renewal	5.77	4.23	5.00
A2 scrapping	6.63	3.37	1.00
A3 VNG	6.36	3.64	2.26
A4 maintenance	6.52	3.48	1.51

We can now calculate the new weights proportionally to the previous so as not to change the order, by dividing the individual weight of each remaining criterion by the sum of weights of the remaining ones. The resulting MCA problem is summarized in table A5.2.

Table A5.2. MCA problem with all criteria in the same scale and preference function

Criterion	Inclusion	Pub.op.	Enforce.	Implem.	Incentives
Preference fn.	Usual	Usual	Usual	Usual	Usual
P, Q	N/A	N/A	N/A	N/A	N/A
Units	N/A	N/A	N/A	N/A	N/A
Optimization	MAX	MAX	MAX	MAX	MAX
Weights	0.1955	0.1730	0.2157	0.2472	0.1685
A1 renewal	1.80	1.60	3.80	5.00	2.40
A2 scrapping	2.60	3.80	2.40	1.00	2.00
A3 VNG	4.20	4.40	3.80	2.26	4.80
A4 maintenance	2.40	3.00	3.00	1.51	2.40

The Promethee rankings I and II are listed in table A5.3 below.

Table A5.3. PROMETHEE preference flows for "combination of outputs" approach

Ranking	Action	Phi	Phi+	Phi-
1	**A3 VNG**	0.7633	0.8457	0.0824
2	**A1 renewal**	0.0225	0.4472	0.4247
3	**A4 maintenance**	-0.2772	0.3333	0.6105
4	**A2 scrapping**	-0.5086	0.2457	0.7543

While the values of the PROMETHEE flows change in each action when compared with the MCA where "implementation" is not normalized, this does not change the prioritization of actions. However, given the change in the values of "implementation", the normalization could still have an effect in the COSIMA TRR. To check this possibility, we can use the same methodology laid out in the main text to calculate the $VF_j(Y_{jk})$ term for each action considered, and the resulting Total Rate of Return (TRR). The results are listed in table A5.4.

Table A5.4. Change in VF and TRR with various values of a and of COSIMA score per unit of difference in a criterion

Value of α	Items	A1 renewal	A2 scrapping	A3 VNG	A4 maint.
N/A	CBA cost (ND) M. S/	8,131	5,122	388	1,280
N/A	CBA benefit (ND) M. S/	1,159	8,804	1,325	10,888
0.25	VF with SP=10 M. aS/	7.72	5.67	9.42	6.03
	100 M. aS/	77.22	56.70	94.22	60.34
	1000 M. aS/	772.25	566.97	942.21	603.41
	TRR with SP=10 M. aS/	0.14	1.72	3.43	8.51
	100 M. aS/	0.15	1.73	3.65	8.55
	1000 M. aS/	0.24	1.83	5.83	8.97
0.75	VF with SP=10 M. aS/	23.17	17.01	28.27	18.10
	100 M. aS/	231.67	170.09	282.66	181.02
	1000 M. aS/	2,316.74	1,700.90	2,826.64	1,810.24
	TRR with SP=10 M. aS/	0.15	1.72	3.48	8.52
	100 M. aS/	0.17	1.75	4.14	8.65
	1000 M. aS/	0.43	2.05	10.68	9.92

The differences with the TRRs in the main text are marked in red; below is the percentage change from the TRRs in the main text:

Table A5.5. Percentage change in TRRs due to the normalization of the "implementation" criterion

Original	Normalized	% change
5.80	5.83	0.52
4.13	4.14	0.24
10.57	10.68	1.04
8.64	8.65	0.12
9.91	9.92	0.10

The differences are minor and most importantly, they do not change the prioritization of the actions.

ANNEX 6. Isolating the effect of BCR in Scenario 2

To analyze which actions acquire a higher ranking by taking into account their BCR performance, using Scenario 2 directly is problematic, because in it are mixed the effect of both the addition of one criterion (BCR) and the relative performance of the actions within that new criterion BCR. Instead, we would have to compare Scenario 2 with a counterfactual scenario (we will call it Scenario 2CF, see table A6.1 below) where all remains equal, but where the values of the BCRs are identical in a BCR-CF criterion.

Table A6.1. MCA Scenario 2CF problem, with modified criteria and weights, and BCR criterion of Scenario 2 for comparison

Criterion	Inclusion	Pub.op.	Enforce.	Implem.	Incentives	BCR-CF	~~BCR~~
Preference fn.	Usual	Usual	Usual	Usual	Usual	Usual	~~Usual~~
P, Q	N/A	N/A	N/A	N/A	N/A	N/A	~~N/A~~
Units	N/A	N/A	N/A	Years	N/A	N/A	~~N/A~~
Optimization	MAX	MAX	MAX	MIN	MAX	MAX	~~MAX~~
Weights	0.1348	0.1124	0.1551	0.1865	0.1079	0.3034	~~0.3034~~
A1 renewal	1.80	1.60	3.80	5.77	2.40	1.00	~~0.14~~
A2 scrapping	2.60	3.80	2.40	6.63	2.00	1.00	~~1.72~~
A3 VNG	4.20	4.40	3.80	6.36	4.80	1.00	~~3.41~~
A4 maintenance	2.40	3.00	3.00	6.52	2.40	1.00	~~8.50~~

Red: worst performing; Green: best performing

In this way, we would visualize what actions benefit and which ones "suffer" (i.e. their net preference flow diminishes) from their BCR performance only (table A6.2.):

Table A6.2. Net preference flows in Scenario 2 and a counterfactual (all BCR=1)

Action	Scenario 2-CF	Scenario 2	Difference
A3 VNG	0.5206	0.6217	0.1011
A1 renewal	0.0427	-0.2607	-0.3034
A4 maintenance	-0.1962	0.1071	0.3034
A2 scrapping	-0.3671	-0.4682	-0.1011

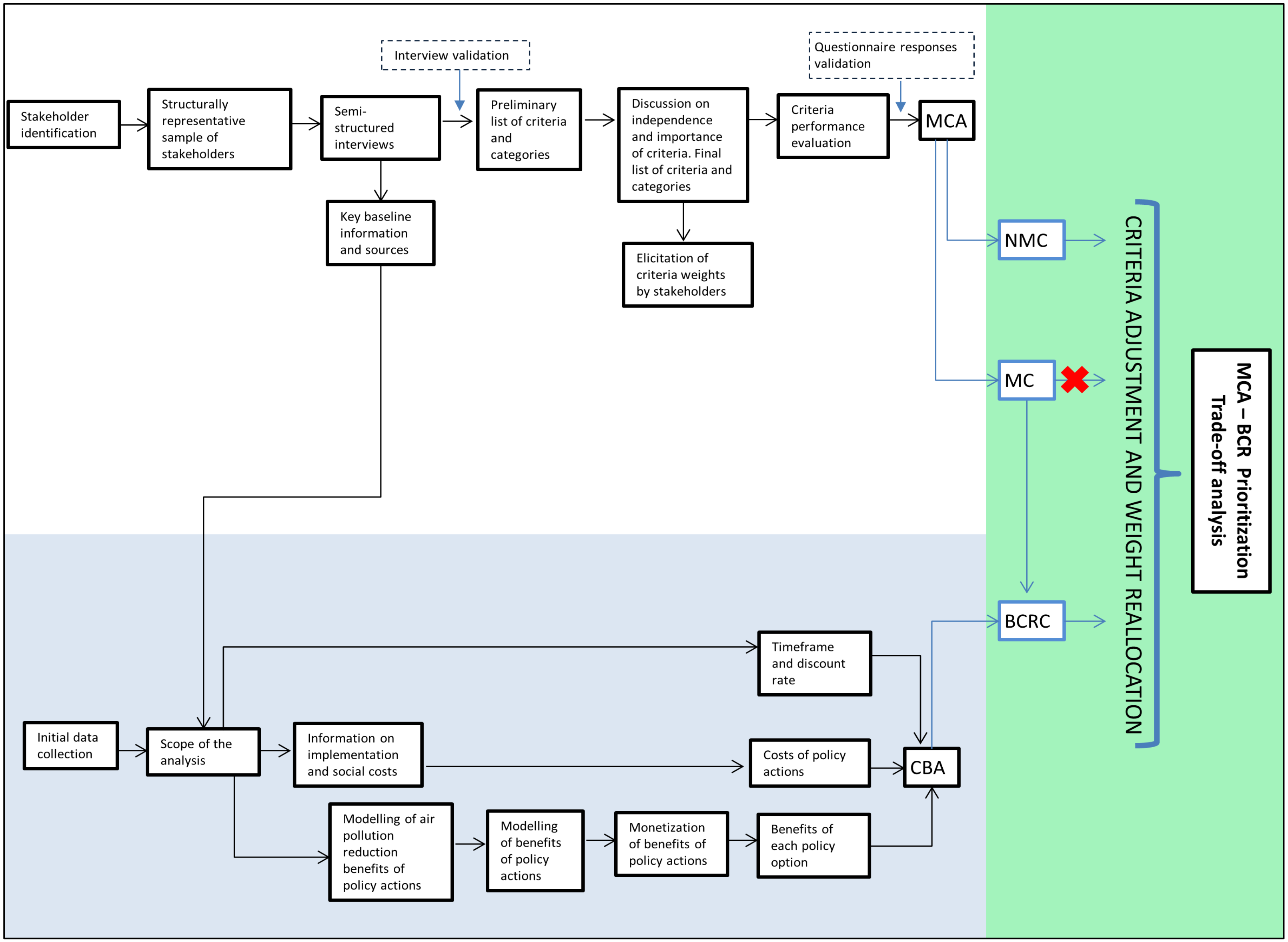
We can observe that A1 (worst BCR) loses the maximum possible amount of preference flow, equivalent to the full weight of the BCR criterion, to A4 (best performing), whereas A2 loses to A3 a third of that amount. Most importantly, the ranking of the alternatives changes as well (see table A6.3. below):

Table A6.3. Variation in ranking due to different BCR performance of actions

Ranking	Scenario 2-CF	Scenario 2.0
1	A3	A3
2	A1	A4
3	A4	A1
4	A2	A2

Because of their BCR performance, action A1 drops from 2nd to 3rd position, whereas action A4 rises from 3rd to 2nd. In other words, we can clearly visualize the effect of the incorporation of the CBA element in the final ranking of actions evaluated. In an example with several actions, the effects would likely be more subtle and less clear, but with an equally valid conclusion: the incorporation of the relative BCR performance of the actions can help visualize misalignments between the preferences of the decision-makers and those of society as measured by a societal CBA.

ANNEX 7. Integration of CBA and MCA through MCA-BCR approach



This page was intentionally left blank