

Why the PPP and Not the Programme?

The asymmetry of scrutiny in sustainability, gender and environment — and the 3% of public capital spending it falls on

Working paper, 27 August 2026. Austral series #26. All figures in §4 are reproducible from `calcular_cifras.py` ; their construction and limits are documented in `METODOLOGIA.md` , and the caveats that matter are carried in the text rather than buried in notes.

Cite as: Duarte Arancibia, David, *Why the PPP and Not the Programme?* (15 August 2026). SSRN: ssrn.com/abstract=7389958 · DOI: [10.2139/ssrn.7389958](https://doi.org/10.2139/ssrn.7389958). Open access under CC BY.

Abstract

Public-private partnerships are held to sustainability, gender and environmental standards that are not applied to the public investment programme they are drawn from. This paper asks why, and finds that the asymmetry does not track performance — it tracks contractibility. The standard is applied where there is a counterparty to bind, not where the problem is largest.

Three findings, each documented rather than asserted. **First**, magnitude: by the IMF's Investment and Capital Stock Dataset, cross-checked against the World Bank's PPI database, the modality carrying this architecture is about **3% of public capital spending** in the countries that report it. **Second**, the comparator: the general appraisal rulebooks of seven jurisdictions — Chile, Peru, Colombia, Mexico, the United Kingdom, France and the European Union — were counted against two chapters of the 2026 PPP Certification Guide and then, because word frequency cannot distinguish rhetoric from obligation, every relevant passage was **coded for normative force** on a six-point scale from absent to required-with-a-method. Gender is absent as an appraisal requirement in all seven; the PPP curriculum names it ninety times and supplies no procedure, parameter or threshold either. On climate the ranking inverts: the strongest treatments are public investment manuals, not the PPP curriculum. **Third**, the symmetric test: holding the publisher fixed, the World Bank's *Public Investment Management Reference Guide* and the IMF's *PIMA Handbook* — 526 pages from the two institutions that jointly built the PPP certification

— name gender twice and require it nowhere, while the Bank's guide contains no substantive treatment of climate at all. The asymmetry is authored, not a lag in national capacity.

The paper argues the standard should be identical across modalities, with only the *timing and mechanism of specification* differing, and takes the UK Green Book as the working model of a mandate with a method attached to spending rather than to an instrument. It closes on the same pattern recurring three levels down — in national rules, in the international guidance that shapes them, and in the survey instruments that would be needed to measure what either is missing.

Keywords: public-private partnerships · public investment management · project appraisal · cost-benefit analysis · gender · climate · safeguards versus standards · PIMA · Green Book · distributional weighting · fiscal policy

1. The question nobody asks

On 25 February 2026 the World Bank, with APMG International and the multilateral development bank community, published a revised edition of the *PPP Certification Program Guide* — the body of knowledge behind the international PPP certification, and the closest thing the field has to a global curriculum. The revision gives, in the publishers' own summary, "greater emphasis on sustainability and resilience throughout the PPP lifecycle, including risk assessments," and an "enhanced focus on gender considerations" — alongside improved text on the fiscal risk of PPPs. Around it sits the rest of the architecture: development agencies offering free courses on designing "sustainable and responsible" partnerships aligned to the Sustainable Development Goals; multilateral lenders attaching environmental and social performance standards to what they finance; a United Nations centre codifying a "people-first" PPP standard around five outcomes, with benchmarks, indicators, weights and a scoring method.

Every one of these is a reasonable thing to do. Taken together they amount to a substantial architecture of sustainability obligations, built with care, and built almost entirely on **one procurement modality**.

The question this paper asks is the one the architecture does not: **why this modality and not the programme?**

Two answers are possible. The first is that public–private partnerships are more likely than direct public provision to be socially irresponsible, environmentally damaging or blind to gender — and therefore need correcting. The second is that direct public provision already possesses these attributes, and what is delegated to a private party does not.

Neither is ever stated, and neither has been examined. This paper examines them.

It should be said at the outset what the paper does not argue. It does not argue that the sustainability agenda is misguided, that PPPs are unfairly maligned, or that private provision performs better than public. It argues something narrower and, we think, harder to dismiss: **the asymmetry of scrutiny does not track performance. It tracks contractibility**. The standard is applied where there is a counterparty to bind, not where the problem is largest — and the modality it is applied to is, by the best available measure, about **3% of public capital spending**.

And the asymmetry turns out not to be an accident of national capacity. §5 holds the publisher fixed and counts: the same two institutions that built the PPP certification also publish the reference guidance on managing *all* public investment. In two chapters of the PPP curriculum, gender is named ninety times. In five hundred and twenty-six pages of their public investment guidance, it is named twice. Counting, however, is not enough to settle what that means, and §5 does not stop there: every passage was coded for normative force. The result qualifies the claim in one direction and sharpens it in another — the curriculum's ninety mentions contain no method either, while on climate the strongest treatments in this study are not in the PPP curriculum at all.

The paper ends somewhere it did not intend to. The same pattern turns up at three scales — in national appraisal rules, in the international guidance that shapes them, and in the survey instruments that would be needed to measure what either is missing — and at the second and third the corroboration comes from literatures that were not looking for it. §8 closes on that.

2. What does justify treating a PPP differently

The case for a differential treatment is real, and it deserves to be made properly before it is qualified. Four arguments carry weight, and the fourth is the one this paper found hardest to answer.

Contractual rigidity. A PPP fixes design, service standards and the payment mechanism for twenty or thirty years. Whatever is not written at signature is effectively excluded for the life of the contract, and adding it later requires a renegotiation — at a moment when competitive tension is gone and the counterparty knows it. Direct public provision has no equivalent lock-in: if a government decides in 2030 that its works must meet a new accessibility standard, it instructs and the standard applies. **This is the strongest argument, and it is sufficient on its own to justify treating the two modalities differently.**

Residual control over what the contract does not say. Incomplete contracts are the normal case, not the pathological one. Where a PPP contract is silent, the decision falls to a private party whose objective function is not the public one. Where public provision is silent, the decision falls to an official bound by public-law duties. The *default* differs, and that difference is not rhetorical.

The incentive to shave what is not measured. Hart, Shleifer and Vishny's argument on the proper scope of government applies here directly: a fixed-price, long-duration contract creates strong incentives to reduce cost, including through quality reductions on dimensions that are not contracted and therefore not paid for. Accessibility beyond the legal minimum, inclusive design, environmental performance above compliance — these are precisely the non-contracted dimensions. Specify and price them, or they degrade.

And the enforcement problem is empirical, not just theoretical. The most uncomfortable evidence for this paper's thesis comes from the historical record on water services. Troesken and Masten's work on US municipal water, surveyed by Fabre and Straub, suggests that local governments with distributional and social concerns "might have found it easier to run water services themselves than to enforce specific objectives through contractual arrangements with private operators." The same survey concludes more generally that enforcing "wide-ranging distributional and social objectives that are not always aligned with productivity-enhancing and profit-maximizing choices through long-term contracts with private partners" has proved difficult, and that this difficulty partly explains why many local governments keep such services in house. There is also a distributional channel specific to the modality: PPPs may raise user prices and subsidies and so exclude the poorest consumers, whose demand is relatively elastic (Fay et al.).

That is a real argument for asking more of a PPP *ex ante*, and this paper accepts it. If social objectives are harder to enforce through a contract than through a hierarchy, then the specification has to be tighter precisely because the enforcement is weaker.

Note what these four arguments establish and what they do not. They establish that a PPP requires **more ex ante specification** than direct provision, because the opportunity to specify closes at signature and does not reopen. They do not establish that a PPP should be held to a **higher standard**. That is a different claim, and it is the one the architecture actually makes.

3. What does not justify it

Contractibility is not performance. The sustainability standard is applied to PPPs because there is a contract to write it into and a counterparty to enforce it against. It is not applied to a road built by force account because there is no one to bind — not because the road performs better. Nobody audits the gender impact of directly executed public works. The asymmetry measures where obligations are cheap to impose, and it is being read as though it measured where problems are concentrated.

The implicit premise has no evidence behind it, and the discipline's own survey shows it. The definitive review of what is known about PPP impacts — Fabre and Straub's 2023 survey in the *Journal of Economic Literature*, covering infrastructure, health and education across roughly a hundred pages — **does not contain the word "gender" once.** "Equity" appears three times, "distributional" twice, "inequality" twice. This is not a criticism of that survey, which reviews the literature that exists. It is a statement about the gap between what the curriculum now requires practitioners to assess and what the research has established about it.

The comparison problem is sharper still. The survey notes, of one of the better-identified studies it reviews, that "the counterfactual against which the program is tested is not 'public schools' but rather 'no PPRS school in the village'," so "it is then difficult to know whether the PPRS schools are more effective... than public schools would have been." That is the field's characteristic evidential shape: PPPs are usually compared against *nothing*, rarely against *public provision*. The survey's own verdict is that "the jury is still very much out regarding the efficiency gains provided by PPPs, and more good-quality studies with convincing identification strategies are needed."

So the honest position is not that public provision performs worse on these dimensions. It is that **almost nobody has measured either side, and nobody has measured them against each other.** A standard justified by an unexamined comparison is a standard justified by an assumption.

Much of what looks like a norm is a condition of finance, and the international lawyers got there first. Dann and Riegner, writing in the *Leiden Journal of International Law*, describe the mechanism precisely: safeguards operate as ex ante conditionality on the funding application, and are then incorporated into a loan agreement which is a treaty, so that borrowers "implement projects under two sets of norms: their own domestic law, and international obligations towards the Bank." Where national law falls short, the borrower "faces the choice of forgoing the loan or raising its national standards." They read this as the Bank's transformation from lender into lawmaker — operating, in their phrase, through financial leverage rather than formal hierarchy. This paper reached the same conclusion from the arithmetic of a word count and a set of appraisal manuals, which is some comfort that it is not an artefact of either.

What follows from it for this paper's purpose is the ratchet. The dual-norm structure raises standards **only where the loan is.** A budget-financed road in the same country, built by the same ministry, is

outside both norms — outside the Bank's because it is not Bank-financed, and outside any national equivalent because, as §5 documents, most national rulebooks do not contain one.

And the frameworks say so themselves, which §5 confirms. A substantial share of the sustainability requirements that attach to PPPs in developing countries are the environmental and social standards of the institutions lending to them. That is a fact about who is financing the project, not a claim about the modality — and the frameworks themselves say so: the World Bank's applies to what the Bank finances and expressly not to what it does not, whoever executes the work. A budget-financed programme in the same country faces no equivalent screen, and would face the identical one if a multilateral financed it.

Distinguishing the safeguard from the standard is half of this paper's work, and it is routinely not done — including, often, by the ministries that inherit the requirement and assume it is a norm.

4. How large is the PPP channel, really

If the sustainability architecture is built on PPPs, it is worth knowing what fraction of public capital formation PPPs represent. The question turns out to be harder than it looks, and the difficulty is itself informative.

The figures

Using the IMF's Investment and Capital Stock Dataset (2021 vintage), which reports public, private and PPP investment for 194 countries, over 2010–2019:

PPP investment as a share of world public investment	3.3%
Median across the 118 countries with any PPP investment	4.6% (interquartile range 1.3–9.7%)
PPP investment as a share of GDP, median country	0.18%
Advanced economies · emerging markets · low-income	2.1% · 6.4% · 4.7%

The world figure is the ratio of summed flows: US\$2,021bn of PPP investment against US\$61,847bn of public investment, in constant 2017 international dollars. In Latin America, where the PPP discourse is loudest, the country ratios are Brazil 31.1%, Peru 26.5%, Colombia 20.9%, Mexico 9.1%, Argentina 3.2%.

What these numbers are, precisely

These are numbers that will be cited and doubted, so the construction matters more than the result.

The IMF's PPP series is **not a direct measurement**. Its own manual states the construction: project commitments at financial close, taken from the European PPP Expertise Centre for European countries and from the World Bank's Private Participation in Infrastructure database for low- and middle-income countries, with divestitures, rentals and merchant projects excluded, and each commitment spread linearly over five years. The manual is candid about why: annualised PPP investment data are unavailable for most countries, and "there is no consistent standard framework to classify PPPs as public or private."

Four qualifications travel with the figures, and none of them is a footnote.

The numerator is not "public investment delivered as a PPP." It is total project value, private finance included. The ratio measures the relative size of two channels, not the share of public spending routed through one of them.

The denominator is all public investment, not infrastructure. Against public investment *in infrastructure*, the fraction would be higher — how much higher we cannot say, because no comparable cross-country series of public infrastructure investment exists. **These figures are therefore a floor**, and the honest way to read them is as a lower bound.

Coverage stops at the sources. Because the underlying data are EPEC (Europe) and PPI (low- and middle-income), there is no PPP series for high-income countries outside Europe. The United States, Canada, Australia, Japan, Korea — and **Chile**, which has run one of the longest-standing concession programmes in the world — simply do not appear. This is a limitation of the global measurement infrastructure, not of the calculation, and it is worth pausing on: *the dataset with which the world measures the size of its PPP programmes does not cover several of the oldest.*

The extremes are denominator artefacts. A handful of low-income countries show PPP investment approaching or exceeding their entire public investment, because their public investment is very small and one large project dominates. Use the median and the interquartile range; never the maximum.

The literature says the same thing, in passing

The finding is not novel, only unquantified in most places where it matters. Fabre and Straub open the conclusion of their survey with it: "While PPPs still represent a **relatively small share of the overall worldwide investments** in physical and social infrastructure, they enjoy a high popularity in policy and government circles and are currently widespread across places and sectors." The observation is made as an aside, in a survey about something else, and then not pursued. This paper's contribution on this point is to put a number on it and to draw the consequence.

What it means

Even read generously — as a floor, on the modality's most favourable sample — the PPP channel is a small fraction of public capital formation. In advanced economies, where most of the sustainable-PPP literature is written, it is around one-fiftieth. The sustainability architecture is being built on the part of the capital programme where obligations can be written into a contract, and the remainder — the overwhelming majority of public investment — is not audited through the same lens.

5. What is asked of the PPP, and what of the programme

The mapping this section needs is a side-by-side one: for each modality, what sustainability and fiscal requirements apply, who sets them, who verifies them, at what point in the cycle, and what follows from non-compliance. Its most consequential row can now be filled, because the curriculum was revised while this paper was being written — and reading it changed what this section says.

What the curriculum actually contains

An earlier draft of this paper expected to find that the 2026 revision had strengthened sustainability while leaving the fiscal side thin. That expectation was wrong, and the correction is worth stating plainly because it is the more useful finding.

The 2026 *PPP Certification Program Guide* is **thorough on both**. Its framework chapter runs to 115 pages and mentions contingent liabilities seventy-two times, fiscal commitments thirty-five times and affordability thirty-four; it sets out a control framework for identifying and quantifying fiscal commitments, and it names the IMF–World Bank PFRAM model as the tool for the job. Its appraisal chapter runs to 133 pages, carries a dedicated box on the project fiscal risk matrix, another summarising contingent liabilities per project, another on typical affordability tests — and, alongside them, boxes on collecting gender-disaggregated data, on integrating gender into social impact assessment, and on the social baseline. Its own foreword records that the 2026 edition "improves the text around the fiscal risk of PPPs" as well as adding the social and climate material.

The publishers' summary line — that value for money was updated with "qualitative approaches, and broader definitions beyond fiscal terms" — is narrower than it sounds. It describes a change to the *value-for-money* method, not a weakening of fiscal risk assessment, and conflating the two would have been an error.

Which makes the asymmetry sharper, not weaker

The finding that survives is better than the one we went looking for.

The PPP side of the ledger is **extraordinarily well developed**. A practitioner is trained to screen the project, appraise it, test affordability against a fiscal envelope, quantify contingent liabilities, run a fiscal risk matrix, build a social baseline, collect gender-disaggregated data, assess climate resilience, and manage the contract after financial close. Nearly 250 pages, across two chapters, of accumulated international practice — built collaboratively by the World Bank, a certification body and the multilateral development bank community, taught worldwide, examined at three levels.

Now ask the corresponding question of the other 96.7% of public capital spending. **What is the equivalent?**

For a road built by force account, in most jurisdictions: a social appraisal methodology that in many cases is decades old, no gender screen, no climate resilience assessment, no contingent liability register, no affordability envelope distinct from the annual budget line, no certification, no examination, and no international body of knowledge. The state that would require a concessionaire to produce gender-disaggregated baseline data requires nothing comparable of itself.

That is the paper's claim, and the quality of the PPP curriculum is the evidence for it rather than against it. **The problem is not that too much is asked of the PPP. It is that this is the only place the asking happens.**

The comparator, documented: five systems

The paragraph above was an assertion. Here it is as a count, across seven public investment systems on three continents — four Latin American, two European high-income, and one supranational.

Each document is the rulebook that governs the appraisal of **every** project seeking public funds in its jurisdiction, of every delivery mode and sector. Set beside them the two chapters of the PPP curriculum covering the same ground.

Term	PPP Guide 2026 ch. 2+4 ≈248 pp	Chile SNI 79 pp	Peru 44 pp	Colombia 124 pp	Mexico 14 pp	UK Green Book 88 pp	France 212 pp	EU 98 pp
gender	90	0	0	3*	0	0	0	0
women	—	0	0	0	0	0	0	1
equity / equality	—	0	0	4	0	11	2	4
distributional analysis	—	0	0	0	0	53	12	1
climate	87	2*	0	4	0	19	33	55
sustainability	43	0	10	6	1	0*	6	53
citizen participation	—	0	0	0	0	0	3	5
indigenous peoples	—	1	0	3	0	0	0	0
accessibility / disability	—	1	0	3	0	0	1	20

Three readings, in descending order of confidence.

Gender is absent from public investment appraisal in every one of the seven. This is the robust finding, and it survives inspection. Colombia's three hits are not a counter-example: all three are a demographic data field in the beneficiary characterisation — populations are "classified as masculine or feminine," alongside age bracket, ethnicity and victim-of-conflict status. Colombia asks you to *count* beneficiaries by sex. No system in the table asks you to *assess* a project's differential impact on women, or to design against it. The PPP curriculum uses the word ninety times.

Climate has reached the rulebooks, unevenly, and mostly without a method. The EU Vademecum is the outlier at 55, with France at 33 and the UK at 19 behind it — France's figure resting on the *valeur tutélaire du carbone*, a shadow price the state publishes and updates. Chile's two mentions are both in the statement of statutory purpose — the SNI allocates resources "ensuring the reduction of the adverse effects of climate change," citing the ministry's governing statute — and then the term appears nowhere

in the seventy-nine pages of rules, procedures, thresholds and required documents that follow. Peru's directive does not mention it at all. Chile has a mandate without a method; Peru has neither.

And the United Kingdom is the case that breaks the pattern — which is the most useful thing in the table.

What the Green Book does that the others do not

A word count would report the Green Book as having zero mentions of "gender" and zero of "sustainability," and would be badly wrong about what it requires. It uses its own vocabulary, and behind that vocabulary sits a method.

Chapter 7 is titled *Distributional Analysis and Place-Based Analysis*. Its subsections are equivalisation, distributional weighting, and place-based analysis; welfare weights appear seven times, and practitioners are told they "must consider economic transfers in distributional analysis" and quantify which groups gain and lose. That is not an aspiration. It is an arithmetic procedure with published parameters.

Gender is there too, in statutory form rather than as a keyword. The Green Book instructs practitioners to use distributional analysis to show the effects of a proposal on "those groups with the protected characteristics defined in the Equality Act 2010," and cross-refers to the Public Sector Equality Duty and to the Equality and Human Rights Commission's guidance. Sex is one of the nine protected characteristics. The obligation is broader than gender, not narrower.

Nor is climate rhetorical: greenhouse gases appear sixteen times, carbon values three, net zero three, natural capital six — with the valuation carried in the supplementary guidance rather than the main text.

So the Green Book is a public investment appraisal manual that carries distributional and environmental obligations, with methods, applying to all central government spending irrespective of delivery mode or financier. **The gap documented in the other systems is therefore not inevitable, and not a property of cost-benefit analysis.** It is a choice about what the appraisal rules are asked to contain — and §7 takes the Green Book as the model rather than the PPP contract.

The symmetric test: the same institution, the other product line

The counts so far invite an easy explanation — that national systems lag, and that the international guidance is ahead of them and pulling them along. The way to test that is to hold the publisher fixed.

The World Bank and the IMF jointly built the PPP certification programme. Both also publish a reference guide on the *other* modality — the World Bank's **Public Investment Management Reference Guide** (2020) and the IMF's **PIMA Handbook** (2022). Same genre, same audience of finance and planning ministries, same houses, overlapping years. Between them, 526 pages and 225,000 words on how a country should appraise and select **all** of its public investment.

Term	PPP Guide 2026 ch. 2+4 ≈248 pp	IMF PIMA Handbook 2022 247 pp	WB PIM Ref. Guide 2020 279 pp
gender	90	1	1
women	—	0	0
climate	87	11* → 4	4* → 0
indigenous peoples	—	0	0
accessibility / disability	—	3	0
citizen participation	—	0	4
sustainability	43	68* → ~0	76* → ~4
distributional analysis	—	1	15

The asterisked cells are the whole argument, so they are worth spelling out.

Gender appears once in each, and both times it is rhetoric rather than requirement. The World Bank's single instance is a subordinate clause listing what a social impact assessment looks at — impacts "on poverty, on unemployment, on gender equality, and on minorities." The IMF's single instance is in the opening case for why infrastructure matters: "with broad-based access, public infrastructure supports income and gender equality." One describes a tool; the other justifies the subject. Neither asks an appraiser, a planning ministry or an assessor to do anything at all. Across 526 pages, **gender is named twice and required zero times.**

Climate, in the World Bank guide, appears four times and not once about climate. Two are idioms — the "fiscal climate" and the "business climate." One is the name of a Canadian agency in a list of data sources. The fourth is a footnote in a table of other countries' discount rates, recording that the Netherlands uses 4.0% for climate change effects. The Bank's reference guide on managing public investment contains **no substantive treatment of climate change whatsoever.**

The IMF Handbook is more revealing, because it explains itself. Of its four uses of "climate change," one is a footnote stating that the Fund "is currently piloting a PIMA climate change module, which will assess countries' ability to systematically reflect climate change considerations in their public investment." That is the primary source confirming what the architecture implies: climate did not enter the fifteen institutions of the assessment. It was built alongside them, as a separate optional module. The arithmetic of the separation is stark — a 29-page Climate-PIMA country report uses the word "climate" 410 times, because that is the only place the subject lives. **Gender never got a module at all.**

And "sustainability," at 68 and 76, is a false friend in both. In the IMF Handbook the dominant phrase is the title of Institution 5 — *Planning Sustainable Levels of Public Investment* — joined by debt and fiscal sustainability; environmental sustainability is essentially absent. In the World Bank guide it is "affordable and sustainable," "financially sustainable," "sustainable operation of the asset," "long-run fiscal sustainability," with roughly four of seventy-six environmental or social. In this literature *sustainability* means the government can keep paying for it. It is the same trap as Chile's three hundred uses of "social," and as reading "equity" in a Treasury manual as social equity.

So the explanation that national systems merely lag does not survive. **The asymmetry is authored.** The same two institutions, in the same years, for the same governments, wrote one body of guidance that names gender ninety times in two chapters and another that names it twice in five hundred pages. This is not a gap between rich and poor countries, or between advanced and developing practice. It is a gap between two product lines.

Which also relocates the responsibility. A ministry that adopts the PPP curriculum and the PIM reference guide as written will hold its concessionaires to a standard it never applies to its own works — not through hypocrisy, but because that is what the two manuals on its shelf say.

From frequency to force

A word count cannot distinguish ninety mentions of rhetoric from one that binds, and a critic is entitled to say that one mention is enough if it creates an enforceable duty. The objection is correct, and the counts above cannot answer it. So they were treated as an index — a way of locating where to read — and every passage in which any document touches gender or climate was then read and coded for normative force.

This is not an improvised procedure. Measuring legal texts by systematic coding is an established method — legal epidemiology, or policy surveillance, developed in public health law research and set out by Anderson, Tremper, Thomas and Wagenaar. Its defining discipline is that coding rests on **the apparent features of the text**, eschewing interpretation, so that the data are replicable by someone who disagrees with the analyst. That is the standard this section tries to meet, and §"What this count can and cannot support" states plainly where it falls short of it.

The scale asks four questions of each passage: is there an actor; is there a deontic verb; is there a procedure — a formula, a published parameter, a shadow price, a threshold, a required annex; and does non-compliance block anything.

0 · Absent	does not appear
1 · Purposive	appears in a statement of purpose or opening rationale; no actor, no action
2 · Descriptive	describes a tool or lists a consideration; requires nothing of anyone
3 · Discretionary	instructs with a weak verb — <i>should, where relevant, when so required</i>
4 · Required, no method	actor and strong verb, but no procedure or parameter to make it auditable
5 · Required, with method	obligation plus an operable procedure: a formula, a published parameter, a threshold

Coded, the picture changes — in both directions, and one of the changes corrects an earlier reading in this paper.

Document	gender	climate
PPP Guide 2026, ch. 2+4	3	4
Chile · SNI	0	1
Peru · Invierte.pe	0	0
Colombia · MGA	2	3
Mexico · Lineamientos	0	0
UK · Green Book	4	5
France · Guide	0	5
EU · Vademecum	1	5
IMF · PIMA Handbook	1	1
World Bank · PIM Reference Guide	2	0

Nobody has a method for gender — including the PPP curriculum. This is the finding the frequency count concealed. The PPP Guide names gender 102 times across the two chapters, and in not one of those passages is there a procedure, a parameter or a threshold. Five use an obligating verb, and the strongest of them is an alignment duty — the legal framework "must align with the country's... policies... including those related to social inclusion and gender equality" — or sits inside a description

of social impact assessment good practice adapted from IFC, alongside a box on collecting sex-disaggregated data. That is real content and more than anyone else offers. It is not a method. So the gap between the curriculum and the appraisal manuals is a gap between **abundant aspirational language and silence**, not between rigour and silence.

That matters for what follows from this paper. The recommendation cannot be to port the PPP Guide's gender material into the public investment manual, because there is no method there to port. On gender the honest statement is that infrastructure appraisal, in every document examined here, lacks one. Whether that is a permanent condition or a neglected one is a separate question, and §8 offers a worked existence proof — with the author's interest in it declared — that it is the second.

On climate the ranking inverts, and this corrects the framing earlier in this section. The three strongest treatments are not in the PPP curriculum. They are public investment manuals. The Green Book carries an arithmetic procedure — change in fuel quantity multiplied by the emissions factor, converted to tCO₂e — with a proportionality rule and an instruction that where practitioners doubt whether assessment is required they should lean towards conducting it. France publishes *valeurs tutélaires*: a table of shadow prices in euros per tonne of CO₂, which is precisely the parameter the others lack. The EU Vademecum carries obligation or procedure in 46 of its 103 climate passages. The PPP Guide codes 4 — climate risk assessed, allocated and mitigated, with the cost of government support estimated, but no shadow price of its own.

So on climate the public investment side leads wherever it exists at all. An earlier draft of this section implied the reverse. The coding does not support that, and the corrected statement is narrower and true: **the hole is not "public investment versus PPPs." It is a specific set of systems — and, uncomfortably, the international guidance itself.**

Which is where the coding does the most damage. The IMF's *PIMA Handbook* codes 1 and 1. The World Bank's *PIM Reference Guide* codes 2 and 0. On both themes, the two international reference guides on public investment management score at or below Chile, Peru and Mexico — and far below the national manuals of the United Kingdom, France and the European Union, which needed the guidance least. The institutions that publish the world's most developed body of knowledge on sustainability in one modality publish, for the other, guidance that a British or French appraiser would find thinner than their own domestic manual.

And the United Kingdom is the reason the whole exercise was necessary. Its gender count is zero and its coded force is 4 — the second-highest in the table. The obligation is statutory rather than lexical: the Public Sector Equality Duty "requires decision makers to have due regard" to the Equality Act 2010, and the Green Book instructs practitioners to use distributional analysis to show effects on those with protected characteristics. The verb at appraisal level is *should*, which is why it codes 4 and not 5. But the method beside it is fully specified, and is worth quoting because it is the only formula of its kind found in any document here: the welfare weight of a group is the ratio of national to group median equivalised income raised to the power 1.3, with equivalisation scores given for each household

member, and — the hard requirement — practitioners **must** present weighted estimates alongside unweighted ones so that decision makers can see what the weighting did.

That is what a mandate with a method looks like. Zero mentions of gender, and more enforceable content than ninety.

What this count can and cannot support

Seven documents are more than the existence proof this section previously offered, but they are still seven, and they are not of equal weight. Three caveats travel with the table.

Word frequency is a crude instrument, and the UK row is the proof: two of its most important requirements are invisible to it. That is why the counts are not the argument. They are an index of where to read, and the coding above is what they were built for. Every non-obvious cell was read in context before being reported, and where the raw count misleads — Colombia's gender field, Chile's statutory-purpose climate, the UK's zeroes, the PPP Guide's methodless ninety — the text says so.

Mexico's row in particular should be read with care. Its *Lineamientos* run to fourteen pages and eight thousand words — a binding regulation published in the *Diario Oficial*, not a training document, but far the shortest text here. A row of zeros in fourteen pages is weaker evidence than a row of zeros in two hundred and seventy-nine. Mexico's methodological content sits in a companion CEPEP guide that could not be obtained.

The coding has a single coder and no intercoder reliability statistic. In legal epidemiology that is a named deficiency, not a general caveat: the method's own guidance is that two coders work independently and their agreement is measured, precisely because a scale like this one cannot avoid interpretation entirely — deciding whether a passage states a procedure or merely gestures at one is a judgement. The extraction is scripted and the passages are reproducible, so a second coder can be added without redoing the work. Until then the coded table should be read as one careful reader's reading.

Second, these are the *general* rulebooks. Sector methodologies, environmental impact assessment law and procurement rules sit alongside them and carry obligations of their own; Chile's rail methodology, counted earlier, is one such document and it too records zero for gender. What the table measures is what the system requires of every project as a condition of entering the investment programme — which is the relevant comparison, because that is exactly what the PPP curriculum requires of every PPP.

Third, the sample is not random. It is, however, unfavourable to the claim: Chile and Colombia run two of the region's most established systems, the EU document is the product of a decade of methodological work, and the UK and France are the acknowledged references — the French guide is the longest document in the table. If the gap appears here, it is unlikely to be narrower elsewhere.

The last row: safeguards or standards?

The remaining test was whether the multilateral environmental and social requirements attach to sovereign-financed works as they do to project-financed ones. If a lender applies the same screen to a budget-financed road as to a concession, then these are standards about infrastructure and the asymmetry is national. If it does not, they are safeguards about money.

The answer is that they are **conditions of the financing, and they say so**.

The World Bank's Environmental and Social Framework applies to all Investment Project Financing operations — which are executed by the borrower, that is, by the government itself. Its ten standards cover risk assessment, labour, pollution, community health and safety, land acquisition, biodiversity, indigenous peoples, cultural heritage and stakeholder engagement. A ministry building a road with a Bank loan carries every one of them.

And the framework is explicit about where it stops: activities implemented by the borrower that are **not financed by the Bank** are not subject to it.

So the identical road, built by the identical ministry, to the identical design, carries ten environmental and social standards if a multilateral finances it and none if the budget does.

What the three findings amount to

Read together, the counts and this last test point somewhere slightly different from where this paper began, and more precisely.

The asymmetry is not fundamentally between PPPs and direct provision. It is between **investment where someone has leverage to impose conditions and investment where nobody does**. A PPP has a counterparty; a multilateral loan has a lender; a domestically financed public works programme has neither. The obligations cluster at the points of leverage — which is why the PPP curriculum runs to hundreds of examined pages, why the lender's framework runs to ten standards, and why the national rulebook governing the whole portfolio contains the word "gender" zero times.

That is a more precise claim than the one this paper set out to make, and it subsumes it. **The standard follows the leverage, not the spending.**

6. What the asymmetry costs

It raises the price of the modality asked to close the gap. Every requirement that attaches to the PPP and not to its alternative is a cost differential in favour of direct provision, applied to the instrument being promoted precisely because budgets are constrained.

It converts a fiscal decision into a design standard. Requiring a concessionaire to deliver gender or environmental outcomes means paying for them, in the tariff or in the subsidy. That is a legitimate choice, but it is a choice about spending, and it should be made as one. A state that could not afford those attributes under direct provision does not become able to afford them by contracting for them.

It holds the private party to a standard the state does not hold itself to. This is uncomfortable and worth stating plainly: the sustainability agenda is enforced at the point where the state is the principal and someone else is the agent, and relaxed at the point where the state is both.

And the aggregate effect is bounded by construction. Whatever the architecture achieves, it achieves on roughly 3% of public capital spending. That is not an argument against doing it. It is an argument against believing it is a strategy for sustainable infrastructure.

7. The correct formulation: same standard, different moment of specification

The resolution is not to lower the bar for PPPs. It is to state the rule correctly.

The standard should be the same for both modalities. What differs is when and how it must be specified. A PPP requires the requirement to be written at signature, because the contract will not reopen cheaply; direct provision allows it to be imposed and revised administratively. That is a difference in the *timing and mechanism* of specification, not in the *level* of the obligation.

And the same logic extends past the modality, once §5's last finding is admitted. If the obligation is a property of the *investment* — if it matters that a road be accessible, or that its impacts be assessed for the people who live along it — then it cannot be a property of the *contract* or of the *lender*. A requirement that attaches when a multilateral finances the road and lapses when the treasury does is not a standard about roads. It is a covenant about money, and it should be named as one.

Stated that way, three things follow. Sustainability requirements belong in the **investment programme's** rules, applying to every project regardless of delivery mode **and regardless of who financed it**. This is not a hypothetical design: it is what the UK Green Book does, and §5 documents it. A chapter of the appraisal manual, with an arithmetic procedure, published parameters, and a cross-reference to the statute that makes the duty binding — attaching to the spending rather than to the instrument. The question for every other system in the table is not whether such a thing can be written, but why it has not been. The PPP-specific guidance becomes what it should always have been — **a specification discipline**: here is what must be fixed at signature, because after signature it cannot be fixed. And where a requirement is genuinely a lender's safeguard rather than a public standard, it should be labelled as such, so that the government knows which of its obligations survive a change of financier.

8. Implications

For the certification curriculum — the primary target, and the timing is now known. The revision has already happened: the Guide was published in February 2026 and the syllabi and training materials across all three certification levels are being brought into line with it. This paper is therefore not an input to that revision. It is a reading of a completed one, and an input to two things that remain open — the training materials still being written, and whatever cycle comes next.

The claim is not that the curriculum is deficient — §5 found the opposite, and says so. It is that the curriculum is the most developed instrument of its kind in public investment, and it governs a small and shrinking share of it. The proposal that follows is not to add to the Guide. It is to ask, in its own framing chapter, the question the Guide is uniquely placed to raise: **which of these requirements are properties of the partnership, and which are properties of the investment — and if the latter, why do they stop at the contract boundary?**

On gender, a correction to what this paper can ask for. The coding in §5 found no method anywhere — not in the appraisal manuals, and not in the PPP curriculum either, whose 102 passages contain no procedure, parameter or threshold. So the recommendation cannot be to transplant. Infrastructure appraisal lacks a method for gender, and the modality with the most words about it has not produced one. That makes it a research task nobody has completed, rather than a diffusion task somebody is neglecting.

An existence proof, and a declared interest

Disclosure: the author's firm builds appraisal software, and what follows describes a module in it. It is included because a paper that says a gap cannot be closed should be held to producing evidence when it can be, and because the result cuts against the paper's own earlier expectation. Readers should discount it accordingly; the method is described in full so that it can be reproduced without the software.

The gap is closable, and the machinery required is already standard.

Distributional weighting prices a benefit by the income of whoever receives it, because the marginal utility of income falls as income rises. Transport and access projects, though, deliver most of their benefit in **hours**, and every appraisal manual examined here values those hours at a single price regardless of who receives them. Discretionary time — what remains after paid work, unpaid care and personal care — has the same property income has: an hour returned to someone with very little of it is worth more. And it is measured. National time-use surveys report unpaid work by sex; labour statistics report paid hours and employment rates by sex. The weight is then the same algebra already used for income:

$$w_g = (\bar{T}/T_g)^{\eta_t}$$

applied to the time-denominated benefits only, with the incidence of the time savings across groups supplied by the analyst.

Run over ECLAC time-use data and ILO labour statistics for fifteen Latin American countries, this produces results that are material and, importantly, **heterogeneous**. The gap in weekly discretionary hours between men and women runs from 12.3 in Ecuador and 11.0 in Peru, through 6.0 in Chile, to 0.9 in Argentina and approximately zero in Brazil. On a stylised Chilean project, the same investment moves from -6.7% to $+9.0\%$ of its economic NPV depending on whether its time savings fall on a male-dominated commuting corridor or on care-related trips. That is a sixteen-point range which no appraisal in any of the seven systems in §5 would currently see.

Three things must be said alongside that number, and the third is the reason this subsection exists.

The elasticity is not estimated — for time. For income it is, and well: Acland and Greenberg's meta-analysis of 158 studies and 1,711 estimates puts the elasticity of marginal utility of income at 1.61, with a 95% interval of 1.18 to 2.05, and recommends sensitivity testing at 1.2 and 2.0. That literature tells us what the parameter looks like *if the analogy between income and time holds*. It says nothing about whether it holds. So the honest position is that the functional form is borrowed from a well-estimated parameter for a different good, and the range should be reported rather than a point: on Peru the result runs from $+16.0\%$ at an elasticity of 1.0 to $+36.6\%$ at 2.0.

The variable does not exhaust the question. Brazil scores approximately zero, and it does not follow that Brazil has no gender asymmetry in infrastructure — it follows that this variable does not find it. Unpaid work may be less substitutable, more fragmented, and constrain labour-market participation in ways that hours alone do not capture.

And the method is blocked on data the sector does not collect — which is the paper's own thesis, one level down. The weight needs one input the analyst must supply: how the project's time savings are distributed across groups. Sánchez de Madariaga's work on the *mobility of care* shows why that input is hard to obtain and why the difficulty is not accidental. Care-related travel is not a category in conventional transport surveys: it is dispersed across escorting, shopping, leisure, strolling and visits, while employment is presented as a single large heading. Short pedestrian trips — under a kilometre or fifteen minutes, which is much of what care travel is — are, in her account, deliberately omitted from many surveys because they are not considered relevant to infrastructure policy. Trip-chaining, where women's care travel actually lives, is still not routinely recorded.

So the instrument that would measure the incidence was designed around commuting, and the reason given for excluding the rest is that it does not matter for infrastructure decisions. **The measurement follows what is already valued, exactly as the standard follows the leverage.** A gender weight is therefore not blocked on econometrics. It is blocked on a survey category, and adding that category is cheap.

And the first version of this was wrong, in the direction that flattered nobody. An early build guessed men's paid hours in Chile at 38 per week and concluded that total work time was equal between the sexes and that the whole approach therefore yielded no gender content. The measured per-capita figure is 27.2, because men's hours must be multiplied by the male employment rate. With measured data, women's total work week in Chile is about six hours longer. What nearly destroyed the result was not omitting the parameter but inventing it — which is the same failure, at the level of a single number, that §5 documents at the level of whole rulebooks. The lesson generalises: **a method is not the formula. It is the formula plus the discipline about where its inputs come from.**

For the public investment guidance — the target this paper did not expect to find. The World Bank's *Public Investment Management Reference Guide* and the IMF's *PIMA Handbook* are, between them, the international statement of how a country should appraise and select its capital spending. They are careful, useful documents, and on the dimensions this paper examines they are close to silent — and the coding is harsher than the count. On gender and climate the Handbook codes 1 and 1, the Reference Guide 2 and 0: at or below Chile, Peru and Mexico, and far below the domestic manuals of the United Kingdom, France and the European Union, which needed the guidance least. The Fund's own footnote records that climate entered as a **separate pilot module** rather than as a revision of the fifteen institutions. Whatever the merits of that sequencing, its effect is that the guidance shaping the other ninety-seven per cent of capital spending carries none of what the guidance shaping the three per cent carries. That is a fixable asymmetry, and it is fixable by the same authors.

For the lenders. Distinguish the safeguard from the standard, in writing. A borrower is entitled to know which requirements are conditions of your money and which are claims about good practice.

For finance ministries. If the agenda matters, apply it to the investment programme. Applying it to 3% of capital spending and describing the result as a sustainable infrastructure policy is a category error — and one this paper's own arithmetic makes hard to sustain. The model already exists and is not hypothetical: a chapter of the appraisal manual with an arithmetic procedure, published parameters and a cross-reference to the binding statute, attached to the spending rather than to the instrument. The UK Green Book has had one for years, and France publishes the parameter — a shadow price in euros per tonne of CO2 — that most of the others lack.

The same result, three levels down

This paper set out to test one asymmetry and found the same shape three times, at three different scales, and only the first was the one it went looking for.

In the rulebooks. Sustainability obligations attach where there is a counterparty to bind. Seven appraisal systems and two international guides, coded for normative force, put gender at 0 to 2 almost everywhere and climate at 5 only where a shadow price exists. The standard follows the leverage, not the spending.

In the guidance that writes the rulebooks. The same two institutions that built the PPP curriculum publish the reference guidance on public investment management, and it carries almost none of what

the curriculum carries. This was not a hypothesis; it was found by holding the publisher fixed and counting. And it was reached independently, from an entirely different direction: Dann and Riegner, reading loan agreements as treaties, describe the Bank operating through financial leverage rather than formal hierarchy, and borrowers living under two sets of norms — one of which stops precisely where the money stops.

In the instrument that would measure the gap. The weight described above needs to know who receives a project's time savings. That number cannot be obtained from conventional transport surveys, because care travel is not a category in them and short walking trips are dropped — and the reason given, in Sánchez de Madariaga's account, is that they are not considered relevant to infrastructure policy. The data collection encodes the same judgement the appraisal rules encode. **What is not valued is not counted, and what is not counted cannot later be valued.**

The recursion is what makes the problem stubborn, and it is also where the leverage is. A requirement written into an appraisal manual with a method attached would break the loop at the first level. A survey category for care travel would break it at the third, and cost almost nothing. Neither requires anyone to agree that PPPs are held to too high a standard — which is the argument this paper opened with, and the least important thing in it.

For research. Two objects are missing. The first and most useful is a comparable cross-country series of **public investment in infrastructure**: it would convert this paper's floor into an estimate, and would let the same question be asked of every other claim built on the share of infrastructure delivered through partnership. The second is cheaper and this paper only started it — a systematic reading of national appraisal rulebooks against the obligations the PPP curriculum imposes, coded for normative force rather than counted. Nine documents are coded here on two themes. The extraction is mechanical and the scale is stated; what it needs is coverage, and a second coder.

Every figure in §4 is reproducible from `calcular_cifras.py` ; the counts in §5 from `contar_comparador.py` and the coded passages from `codificar_fuerza.py` , over the twelve source documents archived in `fuentes/` . The counts are an index, not the argument: every passage behind the coded table was read, and where a raw count misleads the text says so. What is still missing — Mexico's method manual, any Middle Eastern jurisdiction, and the coverage §8 asks for — is listed in `METODOLOGIA.md` §7.4 rather than left to be discovered.

References

- Acland, D. & Greenberg, D. H. (2023). The Elasticity of Marginal Utility of Income for Distributional Weighting and Social Discounting: A Meta-Analysis. *Journal of Benefit-Cost Analysis*, 14(2), 386–405.
- Anderson, E. D., Tremper, C., Thomas, S. & Wagenaar, A. C. (2012; rev. 2023). Measuring Statutory Law and Regulations for Empirical Research. In *Legal Epidemiology: Theory and Methods*, 2nd ed. Wiley.
- Dann, P. & Riegner, M. (2019). The World Bank's Environmental and Social Safeguards and the Evolution of Global Order. *Leiden Journal of International Law*, 32(3), 537–559.
- Dewatripont, M. & Legros, P. (2005). Public-private partnerships: contract design and risk transfer. *EIB Papers*.
- Engel, E., Fischer, R. & Galetovic, A. (2014). *The Economics of Public-Private Partnerships: A Basic Guide*. Cambridge University Press.
- Fabre, A. & Straub, S. (2023). The Impact of Public–Private Partnerships (PPPs) in Infrastructure, Health, and Education. *Journal of Economic Literature*, 61(2), 655–715. (Working-paper version in [fuentes/](#).)
- Fay, M. et al. (2018). Hitting the Trillion Mark: A Look at How Much Countries Are Spending on Infrastructure. World Bank Policy Research Working Paper.
- Grout, P. (2003). Public and private sector discount rates in public-private partnerships. *Economic Journal*.
- Hart, O. (2003). Incomplete Contracts and Public Ownership: Remarks, and an Application to Public-Private Partnerships. *Economic Journal*, 113(486).
- Hart, O., Shleifer, A. & Vishny, R. (1997). The Proper Scope of Government: Theory and an Application to Prisons. *Quarterly Journal of Economics*, 112(4).
- International Monetary Fund (2021). *Investment and Capital Stock Dataset (ICSD)* and accompanying Manual & FAQ: Estimating Public, Private, and PPP Capital Stocks. Fiscal Affairs Department. (Both in [fuentes/](#) ; see [METODOLOGIA.md](#).)
- Iossa, E. & Martimort, D. (2012, 2015). Risk allocation and the costs and benefits of public-private partnerships; The Simple Microeconomics of Public-Private Partnerships. *RAND Journal of Economics*; *Journal of Public Economic Theory*.
- Kappeler, A. & Nemoz, M. (2010). *Public-Private Partnerships in Europe — Before and During the Recent Financial Crisis*. EIB Economic and Financial Report. (Source of the five-year spreading convention used by the IMF ICSD.)

Martimort, D. & Pouyet, J. (2008). To build or not to build: Normative and positive theories of public-private partnerships. *International Journal of Industrial Organization*.

Masten, S. (2011); Troesken, W. Municipal water and the enforcement of social objectives. (Cited via Fabre & Straub 2023.)

Sánchez de Madariaga, I. (2013). Mobility of Care: Introducing New Concepts in Urban Transport. In I. Sánchez de Madariaga & M. Roberts (eds.), *Fair Shared Cities: The Impact of Gender Planning in Europe*. Ashgate.

Groom, B. & Maddison, D. (2019). New Estimates of the Elasticity of Marginal Utility for the UK. *Environmental and Resource Economics*. (Central estimate 1.5; cited via Acland & Greenberg's meta-analysis.)

International Labour Organization (2018). *Care Work and Care Jobs for the Future of Decent Work*.

ECLAC / CEPALSTAT. *Tiempo de trabajo no remunerado según sexo* (indicator 2640) and *según quintiles y sexo* (2646). Retrieved via the CEPALSTAT API.

Ministerio de Desarrollo Social y Familia, Chile (2024). *Normas, Instrucciones y Procedimientos del Sistema Nacional de Inversiones (NIP 2024)*. In [fuentes/](#).

Ministerio de Desarrollo Social / SECTRA, Chile (2016). *Metodología para la Evaluación Socioeconómica de Proyectos de Transporte Ferroviario*. Analysed in the companion paper, series #25.

Departamento Nacional de Planeación, Colombia (2023). *Metodología General Ajustada — documento conceptual*.

European Commission, DG REGIO, with JASPERS (2021). *Economic Appraisal Vademecum 2021–2027: General Principles and Sector Applications*.

France Stratégie, Direction générale du Trésor & SGPI (2023). *Guide de l'évaluation socioéconomique des investissements publics*, édition 2023.

HM Treasury (2026). *The Green Book: UK Government Guidance on Appraisal*.

Ministerio de Economía y Finanzas, Perú (2019, as amended). *Directiva General del Sistema Nacional de Programación Multianual y Gestión de Inversiones*, Directiva N° 001-2019-EF/63.01.

Kim, J.-H., Fallov, J. A. & Groom, S. (2020). *Public Investment Management Reference Guide*. World Bank.

International Monetary Fund (2022). *PIMA Handbook: Public Investment Management Assessment*, 1st edition.

International Monetary Fund. *Climate-Public Investment Management Assessment (C-PIMA)* — country reports, Infrastructure Governance portal.

Secretaría de Hacienda y Crédito Público, México (2013). *Lineamientos para la elaboración y presentación de los análisis costo y beneficio de los programas y proyectos de inversión*, Diario Oficial de la

Federación, 30 December.

World Bank (2016, in force from 2018). *Environmental and Social Framework* — ten Environmental and Social Standards, applicable to Investment Project Financing.

World Bank, APMG International & MDB community (2026). *PPP Certification Program Guide*, 2nd edition, published 25 February 2026. Chapters 2 and 4 in [fuentes/](#).

Citations marked "via" are second-hand.