Coding Scheme and Instructions for Capture in World Bank Projects Database

Matthew S. Winters[1]
University of Illinois at Urbana-Champaign
mwinters@illinois.edu

For each project in the dataset, there are two main questions to be answered:

- At whom is the project targeted?
- Are there signs of capture in the project?

Capture is most easily thought of as misallocation of or a failure to account for funds. For this reason, the targeting coding is based on where the money is intended to go; and capture codings are based on (a) how transparently the project's resources were allocated and managed; and (b) how well the use of money was documented.

Keeping this in mind, capacity or education issues (i.e., unfamiliarity with or requiring training on how the World Bank works) and non-intentional or situational delays, even for procurement or audit reasons, for examples, would not alone *prima facie* count as instances of capture.

Coding Instructions for Question 1 – Targeting

In general, we can distinguish between three main types of foreign aid lending:

1. **project lending** goes to support specific investment project (e.g. building roads, rehabilitating classrooms, funding job-training programs);
2. **programmatic lending** goes to support broad reform agendas and therefore calls for the government to complete specific reforms (e.g. of macroeconomic policy, of labor policy, of social welfare policy) in exchange for the money; and
3. **technical assistance** lending goes to build the capacities of government bureaucrats in the recipient countries or help them with planning and management.

The first coding goal is to figure out what type of aid project we are dealing with and who the intended beneficiaries (or ultimate end-users) of the project are (if ultimate end-users exist). In general, the question to ask is "to where or to whom is the money directed?" Coders should consider total project budget allocation, not just the Bank portion of the funding, in determining the target coding.

There are 16 possible categories – nine of which are for project lending; and seven of which are for programmatic or technical assistance lending.

Project lending – involves specific investments or service delivery

1. **Single City** – all or nearly all project works will be completed in a single city
2. **Two or More Cities** – all or nearly all project works will be completed in a set of cities; if this set encompasses more than 50 percent of cities within one province, coding (3) might be more appropriate
3. **Single Province, State, Department or Region** – all or nearly all project works will be completed in a single sub-division of the country
4. **Two of More Provinces, States, Departments or Regions** – all or nearly all project works will be completed in a set of sub-divisions of the country. Projects aimed at some defined number of districts (often that must

[1] Thanks to Amanda Cronkhite for valuable research assistance in helping to create this file and the coding instructions.

meet some predefined selection criteria) should be coded as regional if the number of chosen districts is less than half the country. If this set encompasses more than 50 percent of the provinces of the country (by number or by population), coding (9) might be more appropriate

5. **Urban Sector** – a nationwide project but benefits are limited to urbanized areas
6. **Rural Sector** – a nationwide project but benefits are limited to rural areas
7. **Social Group** – whether geographically-concentrated or nationwide, the project goes to an identifiable social group that is not defined by geography but rather by some ascriptive personal characteristic (examples: university students, a certain ethnic group, ex-combatants, researchers); this coding should not be used for very large groups, such as women or children, which should be coded on geography
8. **Business/Industry** – the project primarily benefits one or more companies in particular, even if there are then spillover benefits for a region. Deliverables from the project would be considered private goods or private benefits for the company or industry. (Projects with deliverables likely to be classified as public goods for a community, such as expanded electricity transmission or improved water quality, would be more appropriately coded regionally.) Small and medium enterprises are included here, as SME projects are necessarily aimed at businesses, usually entrepreneurs.
9. **Nationwide**

For the dataset, you must choose ***only one*** category, even though a project may have multiple components, targeted at different sets of beneficiaries. When this is the case, choose the category to which the majority of project funding will go. If the majority of project funding does not go to a single category, then choose the modal category.

Note: Projects that fund research can be difficult to classify. If the project funds agricultural extension services (which implies research from which farmers can be expected to benefit in the near-term) this should be coded geographically or for the rural sector, as appropriate. On the other hand, if the project funds specialized scientific research within institutions of higher education, then, although there will be eventual general benefits to that research, the appropriate coding for the near-term beneficiaries is the social group of university researchers. If unsure, as always, it is preferable to mark the case for later review than to code inaccurately.

Non-project lending – involves reform and/or capacity building

10. **National Government Ministry Project** – The project provides capacity-building training or planning assistance for one or more ministries of the national government or other national-level governmental entities.
11. **National Government Budgetary Support** – The project provides general budget support to the recipient's national government. Reforms *may* be included in the loan description, but the goal of the loan is explicitly to support the *regular functioning of government* rather than to achieve structural reforms (less structural, more money). Such support is often, but not exclusively, given in times of transition or crisis. There may be a general description of the types of things the loan is expected to finance, though few specific deliverables will be included in pre-financing documentation.
12. **National Adjustment Lending** – The project provides funds to the recipient's national government with the *expectation that significant reforms will be undertaken*. Usually these reforms will be macroeconomic in character, but they might refer to specific economic sectors. Include under this coding sectoral-adjustment loans that seek reform only in one sector. World Bank documentation for these projects is usually just a PID (Project Information Document), not a PAD (Project Appraisal Document) or a SAR (Staff Appraisal Report).
13. **Sub-National Adjustment Lending** – The project provides funds to a sub-national government with the expectation that significant reforms will be undertaken at that level of government.
14. [Does not exist]
15. **Multiple Borrowers** – The project lists multiple countries as borrowers.

16. **Debt Relief** – The project explicitly (and using the majority of its resources) funds the government paying off external debts.
17. **Sub-National Ministry Project** – The project provides capacity-building training or planning assistance to one or more sub-national government entities , such as state or municipal governments.

Note: there should be less of a tendency within the categories of programmatic and technical assistance lending to want to classify a project into multiple categories. If, however, a project involves both programmatic aid and technical assistance, for instance, code it based on the function of the majority or modal amount of overall funding (as above). Similarly, if a particular lending operation appears to have both programmatic and investment project components, choose its ultimate coding based on the majority or modal amount. In all cases, *make note of the lack of certainty* in the coding in the "summary" notes in the Word document.

Method for determining project classification

There are numerous documents on which the target coding might be done:

- The preferred document for coding is the Staff Appraisal Report (SAR) or Project Appraisal Document (PAD). (The PAD is the successor to the SAR.)
- For Global Environment Fund projects, there will be a similar document available called the GEF Project Document.
- A Project Information Document (PID)
  - The PID contains less information than the SAR or PAD, often only stating the broad goals of the project and not describing the specific on-the-ground goods and services to be funded by the project money.
  - For programmatic lending, however, often only the PID will be available – there will not be a SAR or a PAD.
- Technical Annex (TA)
  - TAs generally appear for emergency or quickly arranged projects.
- Other available project documents – such as Project Agreements or Credit Agreements – tend not to contain specific information about the project as it will be implemented.

In some cases, none of these pre-project documents will be available. In this case, you can code for targeting based on the Implementation Completion Report (ICR). However, use this only as *a last resort* and read *only* the background before doing the target coding. The targeting coding should be based on pre-implementation information given that the ultimate project implementation might be changed, reduced or canceled. (Accordingly if information about project change is noted in the background, coders should nonetheless consciously do the target coding on the original project.) If any document other than the SAR or PAD is used for coding, please state the same in the summary and later in the Excel notes (Column G).

Coding steps for project classification:

1. Data will be entered into both an Excel spreadsheet and a Word document.
2. ***Excel:*** Always *enable the macro* when opening the Excel spreadsheet – it keeps track of the date on which entries were most recently edited.
3. ***Word:*** In the Microsoft Word document, copy the project template and use the hyperlinks to reach the appropriate country listing.
4. ***Word:*** Projects are ordered alphabetically by project name within each country. Begin a record for the project to be coded by noting down the Project ID, Country and Project Name (from the information in the Excel spreadsheet).
   a. If the project name on the World Bank page and the name in the spreadsheet differ, use the name on the World Bank page for the Word document (and for alphabetizing purposes) and correct the spreadsheet to reflect the proper name. If the name difference is significant, make a note of this.

5. ***Excel/Web/Acrobat:*** Click on the hyperlink in column A (far left side) of the Excel spreadsheet to get to the project page and open the initial project document (the SAR or PAD, or PID or TA if one of the first two is not available) in Adobe Acrobat.
   a. If for any reason that link is missing or is not functional, click on the ICR link (column X). Then click "show more" and you will see the hyperlinked project ID that will get you to the project page, which shows all available project documents.
      i. If for any reason those links do not work, search using the project ID number in the "projects" database, which can be reached from the World Bank homepage by clicking on "Projects."
      ii. If it is still not possible to quickly locate the document, note the case for extended review later and continue on to the next project.
   b. If any document other than the SAR or PAD ws used for coding, fill in the notes (Column G).
   c. If you have any doubt that the project in the spreadsheet is the project mentioned on the World Bank page, include it among cases for extended review later. *Do not code the project* in the presence of such doubt.
6. ***Web:*** The method for reading the summary at the beginning of the SAR is:
   a. Read the summary page or project chart at the very beginning of the document
   b. Read the section titled "The Project"
   c. If necessary, if supplied at the end of the SAR, look at the project map to determine the intended reach of the project. (Google Maps may also be useful.)
   d. If necessary, review the annexes for details on budget allocations

   To avoid priming effects and to ensure consistency among coders, you should *not* read the background or other sections, unless the two indicated sections reference them.
7. ***Coding:*** What is you preliminary coding of the project's target?
8. ***Word:*** Record a brief description of the project in the Word document (with page numbers if copying and pasting exact language). For projects in which the target decision was made on budget allocation, note the allocation of funds.
9. ***Coding:*** Use the **search** function in Adobe Acrobat to find all instances of:
   target
   benefic

   Do you see any reason to change your coding of the project's target?
10. ***Word:*** Record your final coding of the project's target in the Microsoft Word file, adding any relevant information to the summary.
11. ***Excel:*** When you are certain you have the correct target coding, record the same in the Excel spreadsheet (column U). If the money passes through a non-government or quasi-government entity, mark column T.

Nearly all project documents are searchable PDFs, but you occasionally may come across a document that is not searchable. For obvious reasons, spend additional time on such documents to be sure of your coding.

Note: for non-government or quasi-government borrowers, such as onlending to the private sector or through a development bank, coders should note the same (especially the intermediary entity) in the summary and be prepared, later, to fill in the binary coding in the Excel spreadsheet (column T) for a non-government borrower. Coders should continue, then, to review the project and base the target coding on the beneficiaries' geographic or group distinctions, resulting ultimately in (almost always) an investment project coding.

Note: for programmatic and technical assistance lending, relatively less detail is needed; for project lending, however, the description of the project is more important

If the case is an investment project, proceed to the next section. Otherwise, delete capture-related lines from the Word document and move on to the next project.

Coding Instructions for Question 2 – Corruption/Capture in the Project

This step is *only for project lending* (codes 1 – 9 above). If a case is an instance of programmatic or technical assistance lending, you do *not* need to provide any additional information.

The goal here is to produce a binary coding (0/1) for *whether or not there are indicators* that the project may have been subject to capture. In addition, you will include a binary coding for *what indicators* you found in your review of the project documentation.

As signs of capture, you are looking for the following things:

- Direct mention of corruption (= "corruption")
  - Examples include: bribes, kickbacks, fraud or stealing
- Negative descriptions of financial management procedures (= "financial")
  - Examples include: failure to maintain accounting books; non-compliance with Bank reporting procedures; irregularities with disbursements
- Negative descriptions of procurement practices (such as incomplete or incorrect documentation or improper bidding procedures) (= "procurement")
  - Examples include: non-delivery of procured goods (unless funds repaid); awarding of contracts without bidding or some other transparent process (unless the document mentions a specific, usually an emergency or sole provider, reason for the same)
- Negative descriptions of audit outcomes (such as missing, late or unfinished audits) (= "audit")
  - Examples include: qualified audits; audits never delivered
  - Late audits themselves are not necessarily indicative of capture. Repeated mention of late audits – in other words what might be described as *chronic* disinterest in or an *unwillingness* to submit audits – merits a capture coding, with the coder noting the citations and, if possible, years in which the audits were not delivered.
  - Audits not delivered before the printing of the ICR should be coded as capture, as Bank policy suggests the staff would have, by then, repeatedly attempted to get any missing audits.
- Descriptions of political interference (= "political")
  - Examples include: political (non-transparent or unequal) targeting in the selection of beneficiaries (or, conversely, the exclusion of an opposition or other non-supporting group); interference in hiring/firing decisions; threats to staff salaries or job security if they do not comply with Government wishes
  - Note: mere government haggling or delays caused by the same would not alone qualify for a 1 coding in this category. Administration-opposition disagreement is quite common in most countries. This category is more about interference in the way in which project outputs are delivered.
- Other specific signs of capture or diversion of funds or interference not covered by the above categories.

Where you find evidence of these things, you will record them *under the appropriate indicator*.

What is the threshold for capture?

For some categories, thresholds will be obvious: political interference in hiring decisions or qualified audits. For others, such as repeated problems with procurement procedures despite training, coders may be less sure. Any explicit mention of a capture event – even incidentally or in passing or if it is the only such mention in the ICR – should generate a positive coding for that category. Absent an identifiable indicator of capture (that should always be quoted and cited), coders should default to a no capture coding.

Mitigating circumstances

Instances of impropriety should be listed as mitigated if the government takes immediate action to remedy the situation, such as dismissing an official indicated for fraud or remedying a procurement issue. In that case, in the Excel spreadsheet, the action should be noted as having happened and then having been mitigated, resulting in zero capture. Example:

Corruption	Financial	Procurement	Audit	Political	Other		Mitigating		Capture		Bureaucracy
0	1	0	0	0	0		1		0		0

Note: in cases with more than one indicator for capture (for example, qualified audits and a procurement problem), coders should *not* note a 1 for mitigating even though an issue was resolved. If only one issue was resolved though and one remains, the capture problems overall were not mitigated. (If both problems were mitigated, a 1 coding would be appropriate; but that would be extremely rare.) Otherwise, the positive capture coding would remain, based on the other remaining (unresolved) indicator.

The following should *not* be coded as capture:

- Capacity or competence issues (should be noted in the Word document with an open bullet)
  - The Bank works in many countries in which the education standards, for example in international accounting or civil engineering, may not be at internationally ideal levels. Often, the Bank seeks out a citizen educated abroad as a lead consultant in such a case. Day-to-day workings of the project will nonetheless still be executed almost entirely by those educated domestically and, therefore, training is often needed.
  - Additionally, interlocutors with whom the project works (such as suppliers or government officials) may face the same capacity issues, particularly with World Bank documentation requirements.
  - The Bank sees it as important to provide domestic training for project sustainability (and overall growth); and the need for the same (or delays as a direct result of that need) should not be counted against a project.
- The government's failure to supply counterpart funding.
  - Budget shortfalls are not the same as capture, even if it is a significant project problem.
  - Note: if a report mentions that the government *chose* not to fund a portion of the project, that section of the report should be read carefully for corruption and/or political interference markers.
- Excessive bureaucracy or delay (= "bureaucracy")
  - A significant portion of ICRs mention bureaucracy or significant delay. This is not alone a problem, but rather an unfortunate reality in large institutions.
  - However, the Bank will often indicate avoidable or excessive delays, such as, for example, those caused by different ministries' wanting different documentation.
    - It may be helpful to think of this category with its analogous, everyday realities: having to get four documents to renew one's drivers' license or having the rules change every year to renew your vehicle registration – that kind of annoying, ultimately costly bureaucracy for which the costs outweigh the benefits, for most people.
    - Rewriting tax laws or resubmitting project information to three different ministers in three years is, in some ways, a bigger version of that problem.
  - When problems are alleged to have been intentional or when problems continued despite Bank attempts to address the delay, code 1 for the category.
  - For delays related to capacity issues, problems on the part of Bank staff and contractors, delays corrected with training or delays addressed promptly and efficiently by the government, code 0.
- Significant project change is not in and of itself pejorative, although it may be linked to problems with the project. In the event that there is a significant difference between the project as described in the PAD and

as executed (as described in the ICR), make a note of this in column Y. Examples include project components being dropped or project funds being reallocated to other endeavors not specified in the original project. Any event of this sort should receive a 1 coding in column Y.

- World Bank errors, particularly in planning and project evaluation, that later resulted in project problems
  - "The Bank team also significantly overestimated the costs of some activities, which resulted in the cancellation of a portion of the credit"
    - In this case the cancellation is clearly not pejorative to the borrower.
  - Procurement problems that seem related to poor planning or inability to understand better beneficiaries' needs, for example those mentioned in the Bank Performance section

For the recording of information in the Word document, there are many instances in which the same note will fall into multiple categories. Example:

Corruption
Financial
o Mentions early capacity issues for financial management and procurement
Procurement
o Mentions early capacity issues for financial management and procurement
o "There were no significant procurement or disbursement issues that affected project performance" found in the mid-term review (12).
Audit
o "Project accounts were audited annually by independent auditors acceptable to the Bank; reports were clean with unqualified auditors' opinions" (14).
Political
Other
Mitigating
Bureaucracy
o "Overall, the Ministry team was responsive and efficient in resolving project issues as they emerged" (13).

Coding Steps for determination of capture:

The ICR link in the Excel spreadsheet should bring you directly to the ICR page, from which you will be able to access the document.

1. ***Web/Acrobat:*** Record the World Bank's codings of *Project Outcome*, *Borrower Performance* and *Bank Performance* listed on the first pages of the ICR.
   - Collapse "Highly Satisfactory" ratings into "Satisfactory" and "Highly Unsatisfactory" into "Unsatisfactory" to produce a binary coding.
   - If summary evaluations are not given, search the document for:
     satisfact
   - In this case, under implementation information, note that outcome measures were imputed from the document text. If there is any doubt about the coding, for example two satisfactory measures and two deficient notes, coders should leave the category blank, as missing data is preferable to inaccurate data.
2. ***Web/Acrobat/Word:*** The method for reviewing the ICR is:
   a. Read "Achievement of Objective and Outputs" (Section 4 in newer ICRs)
   b. Read "Major Factors Affecting Implementation and Outcome" (Section 5)
   c. Read "Bank and Borrower Performance" (Section 7)

   For each of the aforementioned sections, make note *in the Word document* of any of the behaviors described above (including the ICR page number as a citation) under the related category headings.
   - Actions indicating capture are noted with dark bullets

- Notation of positive or productive behavior, as well as Bank information of questionable but not clearly capture actions, are noted with open bullets.

d. In the case of substantial project changes, coders should the same under implementation information and, later, translate the same into the binary coding in Column Y in the Excel spreadsheet.

3. ***Web/Acrobat/Word:*** Using the search function, search the entire document for the following terms (where the dash indicates that you also care about derivatives of the root provided) and note of any of the behaviors described above that may have appeared in other sections:
   corrupt
   fraud
   procur
   misman
   audit
4. ***Word:*** Note any mitigating information *with particular care and detail* under that category heading.
5. ***Word:*** Note any relevant information about poor bureaucratic performance.
6. ***Word:*** If the information that you have collected falls into the categories described above and there is no mitigating information (e.g. the government taking action), then code as a case of capture/corruption (1). Otherwise code as not capture (0, the default coding).
7. ***Excel:*** Once you are certain of your coding, translate the same to the binary code in the Excel spreadsheet.