

# PROTOCOL FOR SYSTEMATIC REVIEW ON

## *evidence of the impact on investment rates of reforms to improve the enforcement of contracts*

Main title	The impact on investment rates of reforms to the enforcement of contracts.
Sub title	A systematic review of the evidence.
Review group	
Section	
Authors IN ORDER OF CREDIT (Please include first and surnames, institutions. Include titles - Dr, Prof - if you want them to be used.)	<p>Andrés Rius Instituto de Economía, Universidad de la República, Uruguay</p> <p>Diego Aboal Centro de Investigaciones Económicas (CINVE), Uruguay</p> <p>Nelson Noya Centro de Investigaciones Económicas (CINVE), Uruguay</p>
EPPI-Centre reference number	[To be completed by EPPI-Centre]
Month/year of publication	[To be completed by EPPI-Centre]
This report should be cited as...	Protocol for a systematic review of the impact on investment rates of reforms to the enforcement of contracts.
Contact details (address, phone number, email)	<p>CINVE Uruguay 1242 11100 Montevideo Uruguay</p> <p>Phone: +598-2900-3051</p> <p>Emails: <i>Rius:</i> <a href="mailto:arius1965@gmail.com">arius1965@gmail.com</a> <i>Aboal:</i> <a href="mailto:aboal@cinve.org.uy">aboal@cinve.org.uy</a> <i>Noya:</i> <a href="mailto:nelsonn@cinve.org.uy">nelsonn@cinve.org.uy</a></p>
Institutional base	CINVE (Centro de Investigaciones Económicas) and Instituto de Economía, Universidad de la República; Uruguay
Review Group (with institutions)	<ul style="list-style-type: none"> <li>- Rius, Andrés (Instituto de Economía, Universidad de la República, Uruguay)</li> <li>- Aboal, Diego (CINVE, Uruguay)</li> <li>- Noya, Nelson (CINVE, Uruguay)</li> </ul>
Advisory group (with institutions)	<p>Angeriz, Alvaro (Queen Mary, University of London, UK)</p> <p><i>NB: We applied for methodological support and peer review from International Initiative for Impact Evaluation (3ie)</i></p>
Conflicts of interest (if any)	None.
Acknowledgements	

# 1. Background

## 1.1 Aims and rationale for review

The key role of capital accumulation in promoting economic development has been almost a truism in Economics since the classical economists (Smith, Marx, Ricardo). Investment was one of the obvious ways to promote economic growth in the basic “modern” growth models (Solow, 1956), and the specific circumstances of underdeveloped economies in this regard have been more systematically explored since development economics became a recognizable sub-discipline (see, e.g., Rostow, 1960; Hirschman, 1958; and various contributions in Meier and Seers, 1984).

Even if development is conceptualized to be broader than economic growth, the growth of incomes and wealth are generally recognized to be powerful instruments to expand opportunities and reduce deprivation (e.g., Sen, 1999). For most development agencies, poverty reduction is the primary concern. Since 2000, the broader Millennium Development Goals have guided the efforts of those agencies, as well as those of most of the 180+ governments that adopted them. It is for these reasons that many development agencies are interested in strategies to promote sustained growth, and they have been promoting regulatory and policy reforms to attain that (White, 2008; and next section).

Growth is generally assumed to depend on sustaining high rates of investment, not exclusively but fundamentally by the local private sector (DFID, 2009). The investment-growth assumption is straightforward for many practitioners and agencies, and supported by some evidence (Levine and Renelt, 1992; Hausman, Pritchett, and Rodrik, 2005), but it has also been challenged (e.g., Dollar and Easterly, 1999; Devarajan, Easterly and Pack, 2001). The latter group of scholars would claim that investment is at least partly endogenous—i.e., growth promotes investment—and that, particularly in low income countries, both low growth and low investment can be symptoms of other underlying factors.

Among the factors that could determine growth and/or investment performance, institutions have become more prominent in the scholarly literature in recent years (North, 1990; Rodrik, 2000; Acemoglu and Johnson, 2005). The literature identifies a host of growth-and investment- relevant institutions, and their effects on either growth or investment can be direct or through some indirect channel. Salient among them are the institutions that protect investors from expropriation and those that determine how contracts are enforced. Douglass North argues that with increased specialization, larger numbers of trading partners, and geographic dislocation of transactions, more complex contracts became necessary, and therefore the institutions that reduce uncertainty about their execution became more crucial (North, 1990).

This systematic review focuses on the evidence about one specific causal mechanism: from better enforcement of contracts to higher rates of capital accumulation, either directly or indirectly (e.g., through availability of financing). The review does not assume necessarily that investment *strictly* causes growth (though it is justified, to some degree, by the assumption that facilitating investment will somehow benefit the growth process), nor does it examine all the linkages from quality of institutions to investment.

The rationale for the review does not rest exclusively on the (still debated) link from investment to growth, but also on the fact that donors and governments do invest resources and political capital in improving the business environment, and in particular in seeking to improve the enforcement of contracts (see 1.3 below). While some of these efforts could be simply justified on grounds of promoting the rule of law, the underlying assumption for many of those reform efforts is that investments will be unleashed by them, so analyzing systematically the evidence in favour of that assumption may eventually help in deciding what priority those reforms should have.

## 1.2 Definitional and conceptual issues

The New Institutional Economics (NIE) has highlighted the role of institutions in shaping economic agents' decisions and, ultimately, shaping also aggregate economic performance. It is then apposite to look to one of its founders for guidance on key definitional issues. Douglass North (1990) provides a conceptual framework that has been largely followed by many or most of the studies that will be reviewed.<sup>1</sup>

According to North, economic exchanges inevitably involve transaction costs and asymmetries of information, and it is to make these manageable (and the fundamental exchange viable) that economic agents devise institutions. *Institutions* are "the humanly devised constraints that shape human interaction" (North, 1990, p. 3; all the following page citations refer to the same volume). They can be formal or informal, and their main difference is that *formal institutions* are written ones. In fact, other than their written character, formal institutions only differ in degree from informal ones (p. 46), and they emerge as the increasing complexity of society raises the benefits from the formalization of constraints.

"Formal rules include political (and judicial) rules, economic rules, and *contracts*" (p. 47). This distinction is important: the review will focus on enforcement of contracts (on which, more below), and will exclude the enforcement of general economic rules. While the latter are typically associated with some enforcement mechanism, to the extent that they represent unilateral impositions of the State and affect a broader population of agents that are expected to abide by them, we will not include them in our study (*contracts* is therefore reserved for mostly bilateral and voluntary agreements). As with regards to *informal institutions*, they can be "(1) extensions, elaborations, and modifications of formal rules, (2) socially sanctioned norms of behaviour, and (3) internally enforced standards of conduct." (p. 40). They may also stipulate enforcement mechanisms, as we will see.

The whole theory of institutions and economic performance advocated by North rests on the adoption of new institutional arrangements as the agents' response to the increasing complexity of economic transactions. In pre-modern societies, transactions were essentially personalized exchanges among "neighbours", and production and trade was in small scales. Reputation and the risk of isolation from a community could function effectively to prevent or address opportunism. Gradually, impersonal exchanges among more distant parties became more frequent and economically significant, which led to the emergence of informal institutions with more explicit enforcement arrangements (these would include, for example, the ostracism of those who violated agreements, stipulated in unwritten codes of commercial conduct; p. 43).

With complex contracts that contain many hard-to-measure attributes about exchanged goods and services, and that are plagued by information asymmetries; and with the expansion of the reach of trade and the chances that transactions may never be repeated between the same two parties, it became necessary to devise third-party enforcement. In fact, it would be more appropriate to say that in modern societies the three forms of exchanges (and enforcement arrangements) co-exist, and even archaic and seemingly dysfunctional informal rules can have major impacts, as demonstrated by the evidence that the same formal rules imposed on different societies produce different outcomes (p. 36)

Our review will focus on one set of impacts of different forms of *enforcement of contracts*.

"Enforcement poses no problem when it is in the interests of the other party to live up to agreements. But without institutional constraints, self-interested behaviour

---

<sup>1</sup> North's and the NIE's framework are not free of conceptual problems (for some difficulties with North's, see Field, 2006; for an appraisal of the NIE see Rutherford, 1994). However, it is appropriate to borrow basic definitions from the original source, given that these definitions are widely used, and the framework provides a basic benchmark to which complications or refinements can be compared.

will foreclose complex exchange, because of the uncertainty that the other party will find it in his or her interest to live up to the agreement.” (p. 33)

Enforcement can come from societal sanctions, from second-party retaliation or from a coercive third party (typically, the State), and the long-range economic history of the world shows each of these forms prevailing at some set of space-time coordinates. We do not exclude informal (multilateral, bilateral, or third party) enforcement from our review, but anticipate that most of the empirical literature will focus on formal, third-party (state) enforcement.

As noted by Acemoglu and Johnson (2005), the NIE has persuaded many economists and political scientists that institutions are a primary determinant of economic performance. However, in much of the literature there has been a tendency to conflate a variety of economic institutions in a “cluster” that presumably defines a favourable business environment (on “business environment”, see next section). In their work, Acemoglu and Johnson distinguish *contracting institutions*, which are the institutions supporting private contracts, from *property rights* institutions, which are the institutions constraining government and elite expropriation.

Like these authors, we also note that there is much overlap between the two types (inadequate enforcement of private contracts could result in some form of expropriation, and constraints on governments’ expropriatory powers could be contained in seemingly voluntary contracts). In this review, we adopt a definition of relevant *contracts* that includes private contracts (as in Acemoglu and Johnson, 2005) but also the bilateral, mostly voluntary agreements that can be established between states and private parties (such as, e.g., when a private company gets a concession through public bidding, and agrees to certain terms that either the state or the private party could be tempted to violate, thus requiring third-party enforcement). We exclude from our review, in turn, general economic rules imposed unilaterally by the state, as these lack the voluntary aspect.

Our review focuses on the impact of reforms affecting the enforcement of contracts on rates of investment. We define *investment* as the accumulation of productive assets. These can be tangible (such as buildings, equipment or permanent plantations) or intangible (such as productive methods, or commercial patents). We are interested in investments that enhance productive capacities and are either made by domestic or foreign agents; we are not interested in foreign “direct investment” that simply acquires existing companies (or parts of them) without adding to the stock of productive assets (although the distinction may sometimes be hard to make in practice).

There is a range of possible reforms that may directly or indirectly impact on the enforcement of contracts. Typical donor-funded reform programs, for example, tend to tackle simultaneously a number of “problems” in the laws and their enforcement (for example, rectifying identified flaws in the letter of commercial or civil laws, creating non-judicial arbitration mechanisms, facilitating access by aggrieved parties to the judicial system, reducing various costs of litigation, strengthening the capacities of the courts and judges, etc.). All these have some bearing on the speed and effectiveness of contract enforcement, and more broadly on “the rule of law” (see, e.g., World Bank, 2004). Moreover, other policies not directly connected to the contents or enforcement of written laws, may directly impact on *contract enforcement*. Woodruf (1998), for example, identifies a more or less direct effect of trade liberalization on informal contract enforcement; substantial investments in information technologies in the judiciary -i.e., public investments—may reduce litigation times and procedural mistakes, and therefore improve enforceability; etc.

This wide variety of possible “interventions” has important consequences for our review:

- (a) There may be very few, if any, studies of a reform (“intervention”) that removed swiftly a single problem, leaving everything else unchanged and enhancing the quality of contract enforcement;
- (b) When/if we find them, there may not be anything similar to a control group, as the reform is likely to have affected a whole country (this has to do with the “public good” nature of justice by the courts and the formal equality of citizens before the

law), although some “reforms” may have been deployed gradually, thus generating within-country variety of “exposure” over time; and other countries may serve as “controls”.

- (c) More frequently, interventions will be complex sets of simultaneous changes, taking place over more or less extended periods of time.
- (d) “Reforms” can be expected to be reflected in bigger or smaller “jumps” in some of the usual international indicators for quality of contract enforcement, but may be indistinguishable from other, unintended, unplanned, uncontrolled, and/or endogenous changes in institutions. Our typical quantitative study will treat all variations in enforcement recorded by the usual indicators similarly, regardless of whether actual “reform processes” are behind them.
- (e) The “interventions” will likely reflect very differently in the usual international indicators (see section 1.4), and the existence and size of a variation in some of them may be affected by the nature, support, ideology, politics or other spurious aspects of the intervention, as seen by stakeholders or key informants, more than by its practical/objective effects on enforceability.

The weak enforcement of contracts has been found (analytically and empirically) to impact on investment through a number of channels. First, it could most directly influence the uncertainty surrounding an investment project, or some of its critical activities or dimensions, and therefore influence investors’ decisions by increasing the project’s costs or reducing its expected returns (note that we are not considering here the risk of expropriation of assets created by the investment, which would pertain to the effects of “property rights institutions”). Second, weak enforcement could inhibit lending, or otherwise influence financial markets in a way that hinders investment (Acemoglu and Johnson, 2005). In this line, some authors have found analytical grounds for the idea that “limited enforceability” not only affects the level of firms’ investments but it also increases its “sensitivity to the arrival of new technologies and generates greater macroeconomic volatility” (Cooley, Marimon and Quadrini, 2004). To the extent that aggregate (output) volatility influences investment (a simple accelerator model could show this), there is here another causal channel from enforcement of contracts to rates/levels of capital accumulation. Others have argued that imperfect enforcement influences, through financial contracts, the size distribution and overall heterogeneity of firms, which could reflect on the level or rate of investment (Monge-Naranjo, 2009).

Our systematic review adopts an approach to synthesis that borrows from “realist synthesis” literature a concern with uncovering patterns of context-mechanism-outcome (Pawson et al, 2004). This means that we are not only interested in discerning whether enforceability reforms have been followed by increases in investment (and of what size), but also in the mechanisms that may cause changes in enforceability to influence investment decisions and aggregate investment rates. Therefore, in our review we will record systematically the theoretically plausible causal explanations stated or suggested by the empirical literature. In other words, the “channels” discussed in the previous paragraph (and see also Appendix 1) are just an illustration of early findings in this search rather than a complete listing of all that seem relevant. Our protocol leaves us open to finding other theoretical or empirical arguments for alternative causal chains, or refinements in main channels already identified (such as, e.g., the *financing* channel), and to recording the nature of the empirical evidence that supports them.

### 1.3 Policy and practice background<sup>2</sup>

Donors and developing country governments are publicly committed to, and concerned with promoting high(er) and (more) sustained growth as a means to reduce poverty. In the aid field, this has translated in the last decade or so into a series of efforts to support “private

---

<sup>2</sup> This section is based to a large extent on White (2008).

sector development" (or PSD). While donors and multilateral institutions have long recognized the role of the private sector in development, and strived to address the internal constraints to enterprise development and investment, the focus of their efforts has changed over time. In recent years, there has been a greater concern with the "enabling environment" for the formation, consolidation and growth of private enterprises.

Business environment reforms refer to "interventions aimed at reforming laws and regulations that set the 'rules of the game' for businesses and reduce unnecessary obstacles to doing business" (White, 2008, p. 18). Among other programs or activities, they include specifically those interventions targeted at improving the policy, legal and regulatory framework and legal and institutional reforms that have to do with strengthening "the rule of law", reforming contract and commercial dispute mechanisms, and protecting property rights. Donors and UN agencies surveyed by Habib and White (2007) ranked these interventions at the top of their efforts to promote growth and the development of a vibrant private sector. DFID (2009) has stated that "working with developing country governments and international agencies to create an environment that supports the growth of business" (p. 5) is one of the three pillars of its PSD strategy.

These interventions are almost universally expected to promote investment and growth, but the specific causal chains that connect concrete interventions with expected results are hardly ever explicated, and the evidence is frequently taken casually.<sup>3</sup> Monitoring and evaluation mechanisms put in place by donors and their local partners have helped develop gradually more sophisticated understanding of the effects of reforms, but impact evaluation still poses a significant challenge (White, 2008, p. 22). Among donors' concerns, the first deals with the extent to which the interventions of development agencies actually contribute to economic growth and poverty reduction. The second deals with programme performance and the causal links between outputs, outcomes and impacts. The interest in the impact on investment of reforms to strengthen contract enforceability can be understood in this context, as addressing primarily the latter concern.

The *Doing Business* initiative spearheaded by the World Bank has made valuable contributions to both measure the progress of business environment reforms, and at the same time provide the means to investigate if those reforms have the intended impacts.<sup>4</sup> Even though it has been noted that the dissemination and use often made of the Doing Business indicators leave much room for improvement (World Bank IEG, 2008), they make it possible to trace the impact of institutional reforms at the country level. Moreover, the efforts by this group to develop more and better indicators of the enforcement of contracts (e.g., Djankov et al, 2003) has contributed to the program of understanding the effects of institutions on economic performance, and confirms the practical significance of this review.

## 1.4 Research background

Research on the effects of institutions on economic performance has grown very rapidly since the early 1990s. Theoretical developments (e.g., North's work) have prompted the generation of indicators and proxies for use in empirical analysis (see, e.g., Knack and Keefer, 1995; Kauffman et al., 2004). At the same time, proliferation of databases that include measures of institutional configurations have also stimulated empirical investigation

---

<sup>3</sup> For example, DFID (2009) presents a graphic illustration of 2002-2006 averages of private investment as a share of total investment against real GDP growth, to support the claim that private investment promotes growth. Leaving aside the choice of indicators and the issues of causality, the pictured regression line does not seem to have a very good fit or to reflect a strong effect. The picture (borrowed from the World Bank) is barely discussed at all in the section of the PSD Strategy document that is supposed to present "The evidence".

<sup>4</sup> The program's stated objective is to advance the World Bank Group's private sector development agenda in four ways: motivate reforms through country benchmarking; inform the design of reforms; enrich international initiatives on development effectiveness; and inform theory.

of new or suitably modified research questions, but also “data-driven” research that has not shed much light on the causal chains or the robustness of meaningful theoretical hypotheses (Aaron, 2000; Keefer, 2004; Williams and Siddique, 2008).

As mentioned above, the enforcement of contracts can be “private” (Cungu et al., 2008; Hamish et al., 2000). In these cases, it tends to be informal and will be affected by reforms to the formal enforcement mechanisms, but also by other policy reforms (as, e.g., in Woodruff, 1998). That said, countries will have some form of third-party, formal enforcement mechanism, and that is why quality of contract enforcement is usually taken to be an attribute of nation-states or sub-national jurisdictions (Djankov et al., 2003; Acemoglu and Johnson, 2005). For that reason, the quantitative operationalization of quality or effectiveness of enforcement has largely been done by creating indicators that rate such quality for pairs of country-time coordinates. At the same time, the nature of the problem itself leaves limited room for experimental designs in the implementation of reforms: principles of territoriality of the law and equal treatment of all citizens (in addition to the typical complexity of legal reforms) make this an unfriendly territory for randomized control trials and other quasi-experimental designs.

Three broad approaches have been followed to generate cross-country and longitudinal evidence:

- Indirect measures such as the use of “contract-intensive” money: The relative use of currency in comparison with contract-intensive money is taken as an indicator of inadequate/weak contract enforcement (Clague et al., 1999)
- Experts’ assessment: effectiveness, efficiency and/or fairness of the formal enforcement mechanism is assessed by practitioners and other key informants and conveyed and aggregated through surveys (Knack and Keefer, 1995; but also La Porta et al. 1997, and Berkowitz, 2003; Staats et al., 2000)
- Quantification of time and pecuniary costs to enforce standard contracts: Legal experts are not asked about their opinions but to estimate the time and financial costs incurred by a private party to enforce some rather simple economic contracts (e.g., collect a bounced check, or evict a delinquent tenant; Djankov et al. 2003), or more complex lending contracts (Djankov et al. 2008)

A significant body of research has focused on exploiting these data sets and the variations over space and time of institutional factors and economic outcomes. We expect to find several cross-section, panel or longitudinal analysis, for countries or sub-national jurisdictions as the units of analysis. These studies are the ones that might yield estimates of quantitative effects. Li and Resnik (2003) are a good example of the conventional approach. One caution, however, is that the meaning of “reform” in these studies is not the customary one, of a deliberate change in formal institutions, but the studies actually examine more generally the effects of all recordable *variations* in certain attributes of institutional arrangements.

However, there are various reasons to include other research designs, even if they are less amenable to offering quantitative estimates of aggregate effects on investment. In fact, some authors have argued that the time series or panel approach is already showing diminishing returns, and that other sources of variation in the quality of institutions need to be exploited to understand their effects (Pande and Udry, 2005). This is so because, to have extensive international and time coverage, available indicators of institutional arrangements are mostly based on expert judgement or just the variation in formal procedures as stipulated in the law, so they may not really measure the variations of interest (i.e., they have validity problems; see Pande and Udry, 2005; Shirley, 2008). Moreover, institutions being partly endogenous to the process of development, the causal analysis of their effects based on time series requires ever more sophisticated and scarce “instruments” (that is, variables that are correlated with the institutional quality indicator but can logically be assumed to be exogenous to the economic outcome being “explained”; see Pande and Udry, 2005; Rodrik, 2005; Rehme, 2007). These two problems are hard to circumvent and provide the justification for alternative approaches that may have other limitations but are better equipped to discern causality in various contexts.

Thus, we will review and expect to find valuable evidence about the existence and sign of effects from national or sector studies that exploit natural experiments (e.g., a policy change generating conditions for a before-after analysis, or affecting one industry in a particularly interesting manner; see, e.g., Woodruff 1998), intra-national variation in institutions (e.g., variations determined by federal arrangements or from the slow deployment of an institutional innovation to various sub-national jurisdictions), and differential firm or individual exposure to alternative institutional arrangements (e.g., varied presence of institutions across regions that expose otherwise similar firms to varied institutional environments; Visaria, 2005, cited by Pande and Udry, 2005).

Our synthesis will examine and summarize differently the contribution to our knowledge from each of the previously summarized two broad strands of research designs, and we expect that each will contribute something valuable from different positions in the trade-off between institutional detail and context vs. external validity and quantification.

## 1.5 Objectives

The objectives of the review are to provide a synthesis of the evidence about the impact of policies to enhance contract enforcement on investment. Investment is defined as physical capital accumulation or intangible productive asset accumulation. We will not focus exclusively on the issue of what kind of policy interventions work, but we will ask the questions of why some policies work and others do not, in which contexts and by which mechanisms.

## 2. Methods used in the review

### 2.1 User involvement and composition of the review team

#### 2.1.1 Approach and rationale

This systematic review is undertaken in response to a DFID's grant, awarded through a call for proposals on pre-defined themes. One of the broad research questions was "What is the evidence of the impact on investment rates of implementing the following investment climate reforms: starting a business, protecting investors and enforcing contracts?", and DFID indicated that questions could be partitioned into some of its components for the purpose of submitting an application (to make the review manageable, we applied to address the last of the three sets of reforms identified in the question). As mentioned before, DFID specifically identifies creating an environment that favours the development of businesses as a key objective of its Private Sector Development strategy.

It is therefore safe to assume that informing policy and practice (rather than contributing to scholarly literature or academic teaching) is the main purpose of the whole initiative, and therefore of this review.<sup>5</sup> We derived from such "revealed preference" of the funding donor a decision to undertake a *narrative synthesis* approach, which seeks "to contribute to policy-makers' and practitioners' 'sense-making' -the way they understand and interpret the situations they encounter and the interventions they deploy" (Popay et al., 2006).

The user has been interacting with the review team through comments on the draft protocol. We expect to continue this interaction through the review and commentary of intermediate and final outputs. Moreover, our application proposed to present the draft

---

<sup>5</sup> The call invited applications to participate in "a cutting - edge pilot to increase the use of evidence in policy and contribute directly to shaping international development policy and practice".



report in London, at a seminar at DFID's Headquarters and/or at the School of Business and Management, Queen Mary College, London (where Angeriz, the external advisor, is a Lecturer); and to undertake joint proactive dissemination, with DFID's teams, of a policy brief to target audiences and at convenient venues/events. These could be opportunities for receiving feedback that could result in revisions to the final report before further dissemination.

### *2.1.2 Division of labour within the review team*

Andrés Rius will be responsible for coordinating the study. With Diego Aboal and Nelson Noya, they will carry out most of the activities of the review process. In particular, they will define search criteria, do first and second screening based on inclusion criteria (through abstracts and/or full texts), define keywording strategy, do in-depth reviews and quality assessments, and lead the synthesis and reporting.

Taking into account their research background, Andrés Rius and Diego Aboal will contribute their knowledge of the institutional economics and political economy literature; meanwhile Nelson Noya will provide inputs from the literature of the determinants of investment, both micro and macroeconomic.

María Eugenia Jung will contribute her librarian skill to carry out the operational aspects of the searching and coding processes. Sebastián Fleitas will provide general research assistance. Alvaro Angeriz will act as internal referee.

## 2.2 Identifying and describing studies

### 2.2.1 Defining relevant studies: inclusion and exclusion criteria

Review question: What is the evidence of the impact on investment rates of reforms to improve the enforcement of contracts?

<b>Inclusion criteria</b>
<ol style="list-style-type: none"><li>1. <u>Relevance to the general question</u>: We will include studies that address empirically the causal chain from quality of contract enforcement to levels of investment. A complete causal chain (i.e., going from empirical indication of changes in enforceability to variation in investment) must be empirically assessed in the study for it to be included. (See Appendix 1) While this includes both studies that enable some attribution of impact and others that can only detect correlation, the two will be analyzed separately (the latter group will be treated as a separate population, and will be examined mainly to contribute to the discussion of contexts-mechanisms-outcomes).</li><li>2. <u>Accessibility</u>: There is sufficient information available to allow screening, or it is possible to retrieve the full-text</li><li>3. <u>Languages</u>: English (Spanish, French and Portuguese, provided studies are abstracted in the indexed databases and key websites).</li><li>4. <u>Publication date</u>: on or after 1990;</li><li>5. <u>Temporal coverage</u>: studies must document changes in enforcement (and investment) occurred in the 20th century.</li><li>6. <u>Geographic coverage</u>: all of the world (despite the focus of the review on reforms in developing countries, it is considered that studies of developed countries, or of countries that have not undertaken reforms, may still shed light on causal mechanisms, or serve comparative purposes)</li><li>7. <u>Units of analysis</u>: individual investors, individual firms, industries, regions or countries;</li><li>8. <u>Type of contracts</u>: (i) among private parties (including labour contracts), (ii) between private parties and government agencies or branches.</li><li>9. <u>Type of investment</u>: those affecting gross capital formation; foreign direct investment (FDI), included if effects on "green field" FDI or capital accumulation by foreign subsidiaries (as opposed to "mergers and acquisitions") can be discerned.</li></ol>
<b>Exclusion criteria</b>
<ol style="list-style-type: none"><li>1. General studies of quality of institutions and growth, if they do not allow to discern the marginal effects of enforceability on investment. Either the quality of institutions variables do not distinguish quality of contract enforcement, and/or investment is not assumed at least partly explained by it (or the assumption not tested with empirical evidence).</li><li>2. Full text not available/accessible.</li><li>3. Language other than English, Spanish, French or Portuguese.</li><li>4. Published before 1990.</li><li>5. Historical studies of pre-XXth century institutions and investment.</li><li>6. Studies focused on the enforcement of general economic rules; studies focusing on regulations (unless the effect of regulation on "voluntary contract" enforcement,</li></ol>

and of this on investment, can be discerned empirically).

7. FDI focused exclusively on mergers and acquisitions (M&A)
8. Magazine and newspaper articles, editorials, letters, comments/opinion articles

### 2.2.2 Identification of potential studies: Search strategy

We will undertake a comprehensive search of the following databases, sites and through the relevant search engines.

- Econlit
- RePEc (Research Papers in Economics; [www.repec.org](http://www.repec.org))
- Scopus
- Social Sciences Citation Index (Web of Science)
- JSTOR
- Citeulike ([www.citeulike.org](http://www.citeulike.org))
- Academic Search Complete
- SciELO (Scientific Electronic Library Online)
- ProQuest Social Science Journals
- PAIS International
- HW Wilson Web Social Sciences Full Text
- Jolis Library Catalog (World Bank and IMF)
- Dialnet ([dialnet.unirioja.es](http://dialnet.unirioja.es))
- PRISMA (Publicaciones y Revistas Sociales y Humanísticas) [prisma.chadwyck.com](http://prisma.chadwyck.com)
- Latindex ([www.latindex.unam.mx](http://www.latindex.unam.mx))
- Theses databases:
  - EconLit's theses database from North American universities
  - [www.theses.com](http://www.theses.com) for the UK and Ireland
  - EThOS; British Library's database of digitized theses from UK
  - Canadian theses: [www.collectionscanada.gc.ca/thesescanada/index-e.html](http://www.collectionscanada.gc.ca/thesescanada/index-e.html)
  - Australasian Digital Thesis Program: <http://adt.caul.edu.au/>
  - Networked Digital Library of Theses and Dissertations (NDLTD) [www.ndltd.org/](http://www.ndltd.org/)
  - Dialnet's theses in Spanish and Portuguese
- Working paper series from major research organizations, possibly not covered by the previous
  - Global Development Network ([www.gdnet.org](http://www.gdnet.org))
  - SSRN (Social Science Research Network) eLibrary ([www.ssrn.com](http://www.ssrn.com))
- Scholarly search engines
  - [scholar.google.com](http://scholar.google.com)
- Grey literature databases
  - European System for Information on Grey Literature (SIGL) ([www.kb.nl/infolev/eagle/frames.htm](http://www.kb.nl/infolev/eagle/frames.htm))

These databases provide acceptable coverage of the relevant sources, and access to title and abstract of potential studies, regardless of nature of the publication (journal article, working paper, thesis) or language of the full text.

Searches will start with broad parameters derived from our research question and inclusion/exclusion criteria. These searches will be gradually narrowed under guidance from the three reviewers. All searches will be stored to ensure replicability.

The *EPPI-Reviewer* software will be used throughout the review to record searches, manage references, generate reports, record decisions, analyse data and report results.

The parameters of the first search will consist of the Boolean union of type 1 and type 2 keywords (see table of keywords below). For some words, especially the word reform we will use some common variants of it as well. For example: reform, reforming, reformed.

Also in order to broaden the search we will use general search terms as contract\* AND enforce\* AND invest\*. Additionally when databases allow it we will search for words with close proximity to each other or within the sentence. To allow replicability the exact search terms used in the searches in the different databases will be recorded and reported in the final report.

The search terms will be translated into the other languages employing the usual equivalents as illustrated, e.g., by the various language versions of the World Bank's World Development Report 2002, *Building Institutions for Markets*. Slight variations in usage/translation will be identified and added to the terms list by inviting suggestions from one native speaker economist with thematic expertise, for each of French and Portuguese (the team is fully capable of identifying relevant keywords for Spanish searches).

"Snowballing searching" will complement these strategies. When empirical core studies or review studies will be identified, we will trace back the citations contained, and will trace forward the publications that cite those key outputs. Some of the above cited electronic databases are useful for this task: JSTOR, Web of Science, Repec, SSRN.

"Hand searching" of some of the most relevant academic journals of the area will also complement the strategy. We will include:

- European Journal of Law and Economic
- International Review of Law and Economics
- Journal for Institutional Innovation, Development and Transition
- Journal of Comparative Economics
- Journal of Institutional and Theoretical Economics
- Journal of Institutional Economics
- Journal of Institutional Research (South-East Asia)
- Journal of Law and Economics
- Journal of Law, Economics and Policy
- Review of Economics and Institutions
- Review of Law and Economics
- The World Bank Legal Review
- Revista de Economía Institucional (Universidad Externado de Colombia)

Additionally, we will search key websites of programs, networks or organizations that show up frequently and with relevant materials in our searches. We will include the following:

- Doing Business Project (The World Bank Group, [www.doingbusiness.org](http://www.doingbusiness.org));
- Multilateral organizations: World Bank, Inter American Development Bank, African Development Bank, Asian Development Bank, Organization for Economic Cooperation and Development, United Nations Commission for Europe, United Nations Commission for Latin America and Caribbean, UN Commission for Africa, and United Nations Economic and Social Commission for Asia and the Pacific.
- Bilateral donor organizations: DFID; US AID; CIDA; AusAid; SIDA; NORAD; GTZ, etc.

Research institutes, foundations and networks: IRIS Centre at the University of Maryland; International Society for New Institutional Economics ([www.isnie.org](http://www.isnie.org)); European School on New Institutional Economics ([esnie.u-paris10.fr](http://esnie.u-paris10.fr)); The Ronald Coase Institute ([coase.org](http://coase.org)); Law and Economics Centre, George Mason University.

Keywords	
Type 1 Specific policy interventions and reforms	Type 2 Outcomes
enforceability of contracts	Investment
contract enforcement	capital investment
contracting institutions	capital accumulation
enforcement of contracts	capital formation
enforcement costs	capital stock
contracting institutions	R&D investment
contractual practices	R&D expenditure
contract non-performance	fixed assets
enforceability of agreements	machinery and equipment
contractual arrangements	Infrastructure
dispute resolution systems	growth
contractual unreliability	development
contractual reliability	business start up
contract hold-ups	
contract enforceability	
contract unenforceability	
judicial quality	
court enforcement	
contract intensity	
contractibility	
third part enforcement	
state enforcement	
judicial enforcement	
informal enforcement	
relational contracting	
formal enforcement	
limited enforcement	
limited enforceability	
ability to enforce contracts	
contract intensive money	
legal formalism	
business climate reform	
business environment reform	
judicial reform	
reform of the judiciary	
investment climate reform	
private sector development	

The time period for the searches will be 01-01-1990 to the date of performing the first search (it will be specified in the systematic review report to ensure replicability).

Identified and “included” studies will be classified and coded (see Appendix 2), to allow reviewers to move forward with the review.

#### *2.2.3 Screening studies: applying inclusion and exclusion criteria*

Initially, studies will be screened using pre-defined inclusion and exclusion criteria. Two reviewers will apply the criteria independently to the results of the searches, based first on titles and abstracts only. Studies will then be (i) excluded, (ii) included, or (iii) marked as “pending” if the reviewer is unsure about their inclusion. The two independent reviews will be compared and contradictory judgements or “pendings” will be temporarily “included”, and moved to the next phase of review of full texts.

Once full texts have been retrieved, two reviewers will independently apply inclusion and exclusion criteria, based on quick assessments of the full texts. Disagreements between the two reviewers will be adjudicated by the third reviewer. A record will be kept in *EPPI Reviewer* of all decisions, of criteria that had been used, and of those studies that were eventually excluded.

#### *2.2.4 Characterising included studies*

Once the review team is satisfied that the search and identification are not yielding significant new additions to the population of “included” studies, the reviewers will fill out the descriptive portion of the coding and appraisal form. The form is presented in Appendix 2. Each reviewer will be assigned a sub-set of the references, and all the forms will be doubly-reviewed for possible mistakes.

We will use the information extracted using the coding and appraisal form presented in Appendix 2 to describe the population of studies, and provide some key general statistics for included studies (for example, geographic coverage, type of methods used, types of reforms or institutional variations examined). In particular, we will characterize the two sub-populations of studies that enable some causal inference from those that use method designs that cannot support such inference, and essentially test for the presence of correlation.

#### *2.2.5 Identifying and describing studies: quality assurance process*

The following mechanisms are part of the quality assurance methods:

- Search strategy controlled by senior reviewers
- Moderation exercises undertaken by senior reviewers, with a sample of papers screened and coded (depending on the stage of the review) by all members of review team. Results compared and discussed until consensus is reached.
- Screening (title/abstract and full text) conducted by more than one reviewer for each reference found
- Double-screening of dubious cases (looking at full text) and disagreements discussed by review team.

## **2.3 Methods for synthesis**

### *2.3.1 Assessing quality of studies*

The review team will use a critical appraisal approach based on a multi-dimensional concept of quality in research. This approach covers quality of reporting, methodological rigour, conceptual depth and breath, and relevance.

The quality assessment will be done with two purposes: first, to exclude studies that clearly do not meet minimum professional/academic standards (e.g., omit sources of data, report mere opinions, obviously inadequate choice of econometric method); second, to generate quality ratings to qualify synthesis results.

We develop an encoding tool consisting of a series of questions and a checklist (see Appendix 2, from 4.1 to 4.23) that is inspired by different sources: a non-published document recommended by one of the reviewers (Dickson, personal communication), the EPPI-Centre Data Extraction and Coding Tool for Education Studies V2 "Coding Studies and Extracting Data for a Review"; Cochrane Qualitative Research Methods Group, 2010, (draft of chapter 6); and Spencer, Ritchie, Lewis and Dillon (2003). Many of the contents and concepts of such sources overlap, and others differ in some degree.

Many of the contents and concepts of such sources overlap, and others differ in some degree. Moreover, these sources are primarily designed to deal with qualitative studies, while our review will also include quantitative studies.

With regard to the quantitative studies, we consider useful to clarify some criteria in evaluating quality.

First, it is important to note that the macro social and complex nature of the "intervention" to be analyzed (i.e., reforms that affect the quality of contract enforcement), generally prevents the existence of studies based on experimental designs. Judicial reforms are typically deployed simultaneously within a national jurisdiction, and it is hard if not impossible to exclude citizens from their effects. For this reason, we expect that nearly all studies will be based on observational data.

Second, we will consider some standard criteria of research quality appraisal in quantitative studies. Studies that used only statistical correlation analysis without any attempt to establish causality will be rated as low quality. Other classic criteria are sample size and design, consideration of omitted variables, treatment of heteroskedasticity, and, when time series are used, proper analysis of stationarity conditions.

Even after assigning studies to two separate sub-populations, depending on whether or not they provide support for some kind of causal inference and attribution of impact, the problem still remains of assessing quality among those that do. We will consider with particular care the analysis of causality and the treatment of endogeneity/exogeneity of independent variables. Even though there is not a "golden standard" to treat these problems with observational data, econometric theory has developed many techniques: instrumental variables estimations, two or three stage ordinary least squares, simultaneous testing of systematic policy reaction functions (implicitly -e.g., reduced form specifications, in vector autoregressions- or explicitly -structural specification, etc.), Granger-exogeneity tests in time series, some Bayesian techniques, etc. (Stock and Watson, 2006, Chapter 9). When instrumental variables are used, particular care must be taken to judge the adequacy of instruments.

Third, the appropriate use of control or confounding variables, based on the most generally accepted theories of the determinants of investment, has to be considered.

We are aware of the existence of a strong debate coming from the empirical growth literature, about the validity of cross country regressions, and panel data analysis for countries. This debate has extended, in particular, to the question of interaction between growth and institutions (Rodrik 2005). While we do not intend to cover exhaustively this literature in our review, some of its key references will be useful to evaluate the strengths and weakness of the research reviewed using regression analysis with similar kind of data (Rodrik 2005; Clemens & Bazzi 2009; Rehme 2007; Deaton 2010).

Following Gough (2007) and the example of Tripney et al (2009), we will apply a "weight of evidence" approach to rate quality of studies. Pairs of review team members working

independently will fill out those sections of the coding tool. The third reviewer will arbitrate differences of judgement, when the first two reviewers disagree. In any case, the three reviewers will discuss and attempt to reach a consensus.

### *2.3.2 Overall approach to and process of synthesis*

We will undertake a narrative synthesis (Popay et al. 2006; Mays, Pope and Popay, 2005), slightly augmented with the focus on context-mechanism-outcome (CMO) characteristic of the realist synthesis (Pawson, 1989; Pawson and Tilley, 1997 and Pawson, 2006). Besides synthesizing the findings regarding the direction and size of effects (of contract enforcement on investment), we will examine the studies seeking to identify various plausible causal mechanisms, and how they interact with context to determine outcomes.

The reasons why we choose these narrative methods of synthesis are:

- the complex and multiple embedded nature of the social “intervention” to study, i.e., reforms of judiciary and other reforms to improve contract enforcement. We want to understand the contexts and mechanisms that are part of the explanation of the outcomes of a given intervention or complex, multi-intervention reforms.
- the realist synthesis is a way of evaluating the external validity of the studies (van der Knaap et al. 2006). It pushes us to identify, for each intervention, the mechanisms that underlie the effectiveness of the reform and the conditions under which they seem to work. Both ingredients will be key to evaluate if and when the findings can be generalized to other areas (countries, regions, etc.), times, societies, etc.

The expected small number of relevant studies, and the heterogeneity in study designs, make it unlikely that we will be able to undertake a standard meta-analysis (Petticrew and Roberts, 2006). We do not, however, rule out that possibility, if sufficient comparable studies turn out from the searches.

We will follow the four steps that Popay et al. (2006) identify as key in a narrative synthesis: (1) develop a theory of how and why the “intervention” works, (2) carry out a preliminary synthesis, (3) explore the relationships within and between studies and (4) assess the robustness of the synthesis. Through the four stages we will seek to identify the mechanisms that are under way in the policy action as well as describe the specific contexts in which this action may be successful or not. The mechanisms “are the engines of explanation” (Pawson, 2006). Nevertheless, the mechanism is only part of the explanation of an outcome (given an intervention): the same mechanism in action can deliver different outcomes in different contexts. Context “is mechanism’s partner concept in the realist understanding of causality” (Pawson, 2006 p.24). The context refers to all the contextual constraints in which the mechanisms are supposed to work.

Even though some of the information that we will be extracting with the coding and extraction tool will give us some indication about the contexts and mechanisms implicit in each intervention/study, we will need to dive deeper in each of the selected studies to identify them (that is, go back to full text of key articles, and discuss among three reviewers). As pointed out by van der Knaap et al. (2008) it is not an easy task to try to infer from the different studies the contexts and mechanisms at work in any given intervention because there can be many missing links. The population of “correlation only” studies (see first inclusion criteria on p. 10) will also be examined to this effect, but making a clear distinction in its use, not to blurr but to complement sensibly the conclusions drawn from the population of “impact-detecting” studies.



### *2.3.2.1 Selection of studies for synthesis (if not all studies that are included in the synthesis)*

Not applicable. There will be no further filtering of studies (i.e., application of further exclusion criteria) at this stage.

### *2.3.2.2 Selection of outcome data for synthesis*

The outcome relevant for this synthesis is levels/rates of investment. To the extent possible given the primary research, we will attempt to measure and discuss the impact of reforms to contract enforcement on different types of investments: fixed assets in general, machinery and equipment, R&D expenditure, infrastructure.

For the quantitative studies we will tabulate the sign, the size, and the significance of the coefficients measuring the impact of contract enforcement on investment.

See below for some additional ways proposed for combining and synthesising this information.

### *2.3.2.3 Process used to combine/ synthesise data*

Given the expected scarcity and heterogeneity of studies, it does not seem likely that we will be able to use meta-analysis to synthesise the information. In any case, we would not rule out the use of meta-analysis as a method of synthesis if some conditions apply. In particular, our use of the statistical aggregation of studies through meta-analysis will depend on:

- i. the number of impact studies and their quality,
- ii. homogeneity in the measurement of the independent variable,
- iii. homogeneity in outcome measures.

Even in the case that these condition do not apply and therefore meta-analysis cannot be used, the information contained in the different studies can and will be summarized in other ways.

For the group of studies containing a quantitative measure of impact, some frequently used graphical devises will be drawn:

- i. Funnel graphs (a scatter graph of the estimated impact coefficient versus the inverse of the estimate's standard deviation)
- ii. Forest plots (point estimate of impact coefficient versus 95% confidence interval for each study).

To generate these graphic analyses, the studies will be grouped according to homogeneous populations, and measures of dependent and independent variables.

We also plan to build the following tables:

- i. information on the study setting, main characteristics of the intervention or the measure of the independent variable, the study methods, the units of analysis (countries, sub-national jurisdictions, firms, etc.), and the study findings (narrative summary of findings).
- ii. results of the critical appraisal for each study.
- iii. Weight of Evidence ratings for each study (individual components A to C, and aggregate rating D).

We conduct our narrative synthesis with the focus on CMO as "themes". According to Popay et al. (2006 p. 18):

*"Where results are presented in the form of themes or concepts, as is the case in qualitative research or some surveys, studies focusing on similar topics may have*

*conceptual overlaps, even if these are not apparent from the way the results are reported. Alternatively, apparently similar concepts in different studies may actually be referring to different phenomena. In this context a process of 'translation' of primary themes or concepts reported across studies can be used to explore similarities and/or differences between different studies.*

...

*Thematic analysis, a common technique used in the analysis of qualitative data in primary research, can be used to identify systematically the main, recurrent and/or most important (based on the review question) themes and/or concepts across multiple studies"*

Then, we proceed by assimilating the general notions of recurrent themes or concepts to our more focused CMO questions. There are some items in the coding tool that will allow us to extract data for this task from the primary studies, but others may probably emerge during the iterative process of detecting CMOs across all studies. This inductive manner of extracting data is typical in thematic analysis for narrative synthesis (Popay et al. 2006, p. 18).

Vote counting of studies will be used as a preliminary descriptive tool. It must be qualified taking into account the critical appraisal of the studies and the heterogeneity of the research designs. Therefore, vote counting among homogeneous studies will be a first valid approach to synthesis.

Graphs and tables, though illustrative by themselves, are inputs for the synthesis and not the synthesis. They can be used to conduct the aggregation of results by the following steps:

- i. Organization, grouping the studies into categories according to different contexts, mechanisms and outcomes;
- ii. analyzing the findings within each of the categories;
- iii. synthesizing the findings across all included studies.

To organize and obtain a synthesis across all studies, we first elaborate "a benchmark theory of change" against which to compare the different CMO findings of the studies. This is a non linear process, so some iterations in the comparison process across all studies are needed. This comparison will progressively lead to a preliminary synthesis. Then as, Popay et al 2006 indicate:

*"As patterns across study results begin to emerge from preliminary attempts at a synthesis reviewers should begin to subject these to rigorous interrogation in order to:*

- Identify any factors that might explain differences in direction and size of effect across the included studies or in the type of facilitators and/or barriers to successful implementation*
- To understand how and why interventions have or do not have an effect or why particular barriers and/or enablers to implementation operate."*

In answering the questions of what works, for whom and in what circumstances it is useful to consider "moderator variables", i.e., variables which can be expected to moderate the effects of the independent variable. An analysis of the moderator variables can be guided by question like;

- "• What are the moderators that the authors of the primary studies identify?*
- What are the contributing factors that appear to recur across the studies even if they have not been explicitly identified by authors as moderators?*
- How much difference do the likely moderators appear to make to the study results?*
- What possible relationships are there among the moderators?"*

We will use Weight of Evidence rating to summarize the reviewers critical appraisal and to examine robustness of findings. We will explore if the results of the synthesis change in a meaningful way if we progressively include studies from top to bottom of this scale (particularly for components A and summary rating D of the WoE ratings).

As we expect that nearly all the studies that quantify the impact will fall in the lowest category of component B of the WoE rating (due to their non experimental nature) we plan to use our specific evaluation of quality about methodological issues (items 4.7 to 4.17 of Coding tool) to make a sensibility analysis. We will not, however, mix the sub-population of "correlation only" studies with the focal population of those that enable some attribution of impact (i.e., "impact-detecting" studies).

## 2.4 Deriving conclusions and implications

After completing our synthesis, we will examine the findings from three complementary angles, all of various degrees of interest to specific audiences.

First, we will summarize the findings regarding the existence of evidence supporting the hypothesis that better enforcement of contracts has a positive effect on aggregate gross capital formation (i.e., rates of investment). The summary will discuss the robustness of the evidence; that is, how confident can we be -based on quality of data, internal validity of studies, external validity, etc.—that the findings are not an accident or a result of biased searches, analyses or syntheses. The conclusions will be of relevance to donor agencies and others that support reforms (judicial, administrative, etc.) to strengthen contract enforcement, particularly in developing countries. They will also expose the possible need to improve study designs, data (and/or the design and deployment of reforms, so as to enable their more rigorous analysis), thus illuminating an applied research agenda for research funders and the scientific/analyst community.

Second, we will summarize the findings regarding the contexts in which enforcement reforms appear to have been more/less likely to have the hypothesized effect, the mechanisms revealed by the review that appear to explain when reforms do/do not have the expected effect, and how those mechanisms interact with specific contextual conditions to produce (or fail to produce) the hypothesized outcomes. We will seek to answer the questions of what kinds of policies work and which ones do not; why some policies work and others do not in certain settings; in what circumstances or conditions do some policies work and others do not. This analysis will also highlight more specific gaps in the empirical literature, pointing to narrower research questions that scholars might want to address (and donors might want to help address). They will also suggested entry points for more theoretically inclined analysts, to either adapt, transform, or develop new analytical frameworks.

Third, the results on effects, and on the contexts-mechanisms-outcomes will illuminate the broader contours of the policy space for reforms aimed at enhancing the contribution of the private sector to development. The findings will be looked at from that vantage point, and the implications spelled out. Depending on the findings about context-mechanism-outcomes, we may be able to reflect on weak spots in the implementation of broader reform agendas; and assumptions that might need revision (e.g., regarding causal chains or the behaviour of key actors, etc.).

## References

- Aaron, J. 2000. "Growth and institutions: A review of the Evidence", *World Bank Research Observer*, 15(1): 99-135.
- Acemoglu, D., and S. Johnson. 2005. "Unbundling institutions", *Journal of Political Economy*, 113 (5): 949-995.
- Ardagna, S.; and Lusardi, A. 2009. "Heterogeneity in the Effect of Regulation on Entrepreneurship and Entry Size". *National Bureau of Economic Research Working Paper Series No. 15510* (November).
- Berkowitz, D., Pistor, K; and Richard, J.-F. 2003. Economic Development, Legality, and the Transplant Effect. *European Economic Review* 47 (1): 165-195.
- Clague, C., Keefer, P., Knack, S. and Olson, M. 1999. "Contract-Intensive Money: Contract Enforcement, Property Rights, and Economic Performance". *Journal of Economic Growth*, 4 (2):185-211.
- Clemens, M.A., and Bazzi, S. 2009. *Blunt Instruments: On Establishing the Causes of Economic Growth*. 171. Working Papers. Washington DC, USA: Center for Global Development, May 20.
- Cooley, T, Marimon, R. and Quadrini, V. 2004 "Aggregate Consequences of Limited Contract Enforceability", *Journal of Political Economy*, 112(4): 817-847.
- Cungu, A., Gow, H., Swinnen, J.F.M. and Vranken, L. 2008. "Investment With Weak Contract Enforcement: Evidence from Hungary during Transition", *European Review of Agricultural Economics*, 35: 75-91.
- Deaton, A.S. 2010. Instruments, Randomization, and Learning about Development. *Journal of Economic Literature* 48: 424-455.
- Devarajan, S., Easterly, W.R. and Pack, H. 2001. "Is Investment in Africa Too Low or Too High?", Policy Research Working Paper 2519. Washington: The World Bank.
- DFID. 2009. *Private sector development strategy. Prosperity for All: Making Markets Work*, London, UK: Department for International Development.
- Djankov, S., La Porta, R., Lopez-de-Silanes, F. and Schleifer, A. 2003. "Courts", *Quarterly Journal of Economics* 118(2): 453-517.
- Djankov, S.; R. La Porta; F. Lopez-de-Silanes and A. Shleifer (2008) "Debt Enforcement Around the World". *Journal of Political Economy* 116(6): 1105-1149.
- Dollar, D. and Easterly, W.R. 1999. "The Search for the Key: Aid, Investment and Policies in Africa", *Journal of African Economies* 8(4):546-77.
- EPPI-Centre. 2007. *EPPI-Centre Methods for Conducting Systematic Reviews*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London. March.
- Field, A. 2006. "North, Douglass (b. 1920)", in David A. Clark (ed.), *The Elgar Companion to Development Studies*, 423-426. Cheltenham: Edward Elgar Publishing.
- Habib, M. and White, S. 2007. *UNDP Private Sector Benchmark Study*. New York, UNDP.
- Hamish R., D. Streeter and J. Swinnen (2000) "How Private Contract Enforcement Mechanisms Can Succeed Where Public Institutions Fail: The Case of Juhocukor a.s.", *Agricultural Economics* 23 (2000):253-265.
- Hausmann, R., Pritchett, L. and Rodrik, D. 2005. "Growth Accelerations" *Journal of Economic Growth* 10, no. 4: 303-329.
- Hirschman, A.O. 1958. *The Strategy of Economic Development*. New Haven, USA: Yale University Press.
- Kaufmann, D., Kraay, A. and Mastruzzi, M. 2004. *Governance Matters IV: Governance Indicators for 1996-2004*, Washington, DC, USA: The World Bank.
- Keefer, P. 2004. "A Review of The Political Economy of Governance: From Property Rights to Voice", Washington, DC, USA: The World Bank.
- Knack, S.; and Keefer, P. 1995. "Institutions and Economic Performance: Cross-Country Tests Using Alternative Institutional Measures". *Economics and Politics*, 7: 207-227.
- Levine, R. and Renelt, D. 1992. "A Sensitivity Analysis of Cross-Country Growth Regressions", *American Economic Review*, 82(4): 942-63.

- Li, Q., and Resnick, A. 2003. Reversal of Fortunes: Democratic Institutions and Foreign Direct Investment Inflows to Developing Countries. *International Organization* 57 (1): 175-211.
- Mays, N., Pope; C. and Popay; J. 2005. Systematically Reviewing Qualitative and Quantitative Evidence to Inform Management and Policy-Making in the Health Field. *Journal of Health Services Research & Policy* 10, no. Supplement 1 (2): 6-20.
- Meier, G. M and Seers, D., eds. 1984. *Pioneers in Development*. New York: Oxford University Press.
- Monge-Naranjo, A. 2009. "Entrepreneurship and Firm Heterogeneity with Limited Enforcement." *Annals of Finance* 5, no. 3 (June 1): 465-494.
- North, D. 1990. *Institutions, Institutional Change and Economic Performance*. Cambridge: Cambridge University Press.
- Pande, R. and Udry, C.R. 2005. *Institutions and Development: A View from Below*. Center Discussion Papers. New Haven, USA: Economic Growth Center, Yale University, November.
- Pawson, R. 1989. *A Measure for Measures: A Manifesto for Empirical Sociology*. London: Routledge.
- Pawson, R. 2006. *Evidence-Based Policy. A Realist Perspective*. London: Sage.
- Pawson, R. and Tilley, N. 1997. *Realist Evaluation*. London, UK: Sage Publications Ltd.
- Pawson, R., Greenhalgh, T., Harvey, G. and Walshe K. 2004. "Realist Synthesis. An Introduction". ESRC Research Methods Programme. University of Manchester, RMP Methods Paper 2/2004.
- Petticrew, M., and Roberts, H. 2006. *Systematic Reviews in the Social Sciences: A Practical Guide*. Oxford, UK: Blackwell Publishing Ltd.
- Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M., Britten, N., Roen, K. and Duffy, S. 2006. *Guidance on the Conduct of Narrative Synthesis in Systematic Reviews. A Product from the ESRC Methods Programme*. London, UK: ESRC Methods Programme.
- Rehme, G. 2007. *Endogenous Policy and Cross-Country Growth Empirics*. Darmstadt Discussion Papers in Economics. Darmstadt, Germany: Darmstadt Technical University, Department of Business Administration, Economics and Law, Institute of Economics.
- Rodrik, D. 2000. "Institutions for High Quality Growth: What They Are and How to Acquire Them." *Studies in Comparative International Development* 35(3): 3-31.
- Rodrik, D. 2005. *Why We Learn Nothing from Regressing Economic Growth on Policies*. Mimeo, Harvard University. Boston, MA, USA.
- Rostow, W.W. 1960. *The Stages of Economic Growth*. Cambridge: Cambridge University Press.
- Rutherford, M. 1994. *Institutions in Economics: The Old and the New Institutionalism*. Cambridge: Cambridge University Press.
- Sen, A. K. (1999) *Development as Freedom*. New York: Anchor Books.
- Shirley, M. M. 2005. *Institutions and Development*. Cheltenham, UK and Northampton, MA: Edward Elgar.
- Solow, R. 1956. "A Contribution to the Theory of Economic Growth", *Quarterly Journal of Economics*, 70:65-94.
- Spencer, L., Ritchie, J., Lewis, J. and Dillon, L. 2003. *Quality in Qualitative Evaluation: A Framework for Assessing Research Evidence*. London, UK: Government Chief Social Researcher's Office.
- Staats, J., Bowler, S. and Hiskey, J. 2000. "Measuring Judicial Performance", *Latin American Politics and Society* 47(4): 77-106.
- Stock, J.H. and Watson, M.W. 2006. *Introduction to Econometrics*. Boston: Addison Wesley.
- Visaria, S. 2005. *Legal Reform and Loan Repayment: The Microeconomic Impact of Debt Recovery Tribunals in India*. Ph.D. dissertation. New York, USA: Columbia University.
- White, S. 2008. "A conceptual framework to guide research on private sector development in developing countries", IDRC Working Papers on Globalization, Growth and Poverty, Number 6, Ottawa, Canada: IDRC.
- Williams, A. and Siddique, A. 2008. "The Use (and Abuse) of Governance Indicators in Economics: A Review", *Economics of Governance* 9:131-175.
- Woodruf, C. 1998. "Contract Enforcement and Trade Liberalization in Mexico's Footwear Industry", *World Development* 26(6)979-991.

World Bank IEG. 2008. "An Independent Evaluation of the Doing Business Indicators", Washington: World Bank, Independent Evaluation Group.

World Bank. 2001. *World Development Report 2002: Building Institutions for Markets*. Washington DC, USA: The World Bank.

## Appendix 1

### Examples of causal chains related to the research question, and interpretation of inclusion criteria

Causal chain:		Include?	Reference			
<b>Contracting insts. and laws</b> (legal formalism, proced. complexity, number of proceds.)	→	Forms of financing (equity vs debt)	→	<b>Investment</b>	<i>Yes</i>	Acemoglu and Johnson (2005)
<b>Contract enforcement</b> (number of proceds.)	→	Determinants of Entrepreneurship	→	Size of new businesses	<i>No</i>	Ardagna and Lusardi (2009)
<b>Legality</b> <sup>6</sup>	→	GNP growth			<i>No</i>	Berkowitz, Pistor and Richard (2003)
<b>Contract intensive money</b>	→	Gains from trade, ecs of scale, specialization	→	<b>Investment</b> → Capital stock	<i>Yes</i>	Clague, Keefer, Knack, Olson (1999)
<b>Contract intensive money</b>	→	Ability of firms to raise capital	→	<b>Investment</b>	<i>Yes</i>	Clague, Keefer, Knack, Olson (1999)
<b>Weak enforcement</b>	→	Sensitivity to new technologies	→	Macroeconomic volatility → <b>Investment</b>	<i>Yes</i>	Cooley, Marimon and Quadrini (2004)

<sup>6</sup> “Legality” is “the effectiveness of institutions that enforce the law” (Berkowitz, Pistor and Richard, 2000)

## Appendix 2

### Coding tool - Systematic Review on The Impact of Contract Enforcement on Investment

- 0 Unique identifier code
- 1 General descriptive information
  - 1.1. Author(s)
  - 1.2. Date (month, year)
  - 1.3. Publication status
    - Published or forthcoming in refereed journal or book
    - Published or forthcoming in non refereed journal or book
    - In press
    - Unpublished
    - Unknown
  - 1.4 Type of publication
    - Article
    - Chapter in a book
    - Working paper
    - Conference presentation
    - Government or institutional report
    - Mimeo
  - 1.5 Source
    - Electronic database
    - Handsearch
    - Citation
    - Website
    - Unknown
  - 1.6 Language
    - English
    - French
    - Portuguese
    - Spanish
- 2 Study Aims and Rationale
  - 2.1 What are the broad aims of the study? Please write in authors' description if there is one. Elaborate if necessary, but indicate which aspects are reviewers' interpretation. Other, more specific questions about the research questions and hypotheses are asked later.
    - Explicitly stated (please specify)
    - Implicit (please specify)
    - Not stated/unclear (please specify)



- 2.2 **Topic focus/foci of the study** (tick as many as necessary)
- Contract enforcement and investment
  - Contract enforcement and economic growth
  - Institutions and investment
  - Institutions and economic growth
  - Finance and growth
  - Finance and investment
  - Contract enforcement and international trade
  - Contract enforcement and finance
  - Contract enforcement and vertical or horizontal integration
  - Contract enforcement and substitutes or complements (social capital, trust, etc.)
  - Reform of the judiciary
  - Investment climate reform
  - Private sector development
  - Formal and informal enforcement
  - Other (specify)
- 2.3 **What are the study research questions and/or hypotheses?**  
 Research questions or hypotheses operationalise the aims of the study. Please write in authors' description if there is one. Elaborate if necessary, but indicate which aspects are reviewers' interpretation.
- Explicitly stated (please specify)
  - Implicit (please specify)
  - Not stated/ unclear (please specify)
- 2.4 **Which theoretical/conceptual framework/tradition does the study adopt?**
- Not stated
  - Not explicitly stated (Write in, as worded by the reviewer)
  - Stated (Write in, as stated by the authors)
- 2.5 **Is this report linked to one or more other reports (included in the searching results) in such a way that they also report the same study?**
- Linked (specify)
  - Not linked

- 2.6 **Was the study informed by, or linked to, an existing body of empirical and/or theoretical research?** Please write in authors' description if there is one. Elaborate if necessary, but indicate which aspects are reviewers' interpretation.
- Explicitly stated (please specify)
- Implicit (please specify)
- Not stated/unclear (please specify)
- 2.7 **Please provide details of any user involvement in the study**
- Details
- 2.8 **Do authors report how the study was funded?**
- Explicitly stated (please specify)
- Implicit (please specify)
- Not stated/unclear (please specify)
- 2.9 **When was the study carried out?** If the authors give a year, or range of years, then put that in. If not, give a 'not later than' date by looking for a date of first submission to the journal, or for clues like the publication dates of other reports from the study.
- Explicitly stated (please specify )
- Implicit (please specify)
- Not stated/unclear (please specify)
- 2.10 **Place where the study was reported** (country or countries of authors' affiliation)
- Explicitly stated (please specify )
- Not stated/unclear (please specify)
- 2.11 **Geographic scope**
- Individual country (specify)
- Group of countries (specify)
- Individual region inside a country (specify)
- Group of regions inside a country (specify)
- 2.12 **Urban/Rural setting**
- Urban
- Rural
- Both
- Not clear

2.13	<b>Research design</b>	Quantitative studies Cross section regression Panel data regression Time series regression Others (specify)  Qualitative studies Historical Narrative Observation Survey Audit Action-based Case series Expert opinion Focus group Other (specify)
2.14	<b>Type of investment</b>	Fixed assets in general Machinery and equipment R&D expenditure Infrastructure Other (specify) Not stated
2.15	<b>Origin of investment</b>	Domestic Foreign General Group of countries (specify) Country (specify)  Both Not stated Not applied
2.16	<b>Unit(s) of analysis</b>	Country(ies). Specify Sub-national jurisdiction(s) Industry(ies). Specify (by use of 3 digit code of International Industrial Standard Classification) Firm(s) Family enterprises
2.17	<b>Which theoretical/conceptual framework/tradition does the study adopt?</b>	Not stated Not explicitly stated (Write in, as worded by the reviewer) Stated (Write in, as stated by the authors)
2.18	<b>Controls or confounding variables</b>	Details

- 3 Programme or Intervention description**
- 3.1 **If a programme or intervention is being studied, does it have a formal name?** Not applicable (no programme or intervention)  
Yes (please specify)  
No (please specify)  
Not stated/ unclear (please specify)
- 3.2 **Is the programme or intervention studied part of a broader set of interventions or reforms?** Not applicable (no programme or intervention)  
Yes (please specify)  
No  
Not stated/ unclear (please specify)
- 3.3 **Content of the intervention package.** Details  
Describe the intervention in detail, whenever possible copying the authors' description from the report.
- 3.4 **Aim(s) of the programme or intervention** Not applicable (no programme or intervention)  
Not stated  
Not explicitly stated (Write in, as worded by the reviewer)  
Stated (Write in, as stated by the authors)  
Not applicable (no programme or intervention)
- 3.5 **Causal Pathways/Theory of change.** Details  
Describe in detail authors description of/or evidence provided about casual pathways and/or theory of change
- 3.6 **Year intervention started** Not applicable (no programme or intervention)  
Details
- 3.7 **Duration of the intervention** (Choose the relevant category and write in the exact intervention length if specified in the report)  
When the intervention is ongoing, tick 'OTHER' and indicate the length of intervention as the length of the outcome assessment period  
Not stated  
Not applicable  
Unclear  
Other
- 3.8 **Time period of observation** (for intervention and observational studies)  
Beginning  
Month, year  
End  
Month, year

#### 4 CRITICAL APPRAISAL QUESTIONS

##### Quality of study - reporting

- 4.1 Is the context of the study adequately described? Consider your previous answers to questions on study aims and objectives
- Yes (please specify)
- No (please specify)
- 4.2 Are the aims of the study clearly reported?
- Yes (please specify)
- No (please specify)
- 4.3 Is there an adequate description of the sample used in the study and how the sample was identified?
- Yes (please specify)
- No (please specify)
- 4.4 Is there an adequate description of the methods used in the study to collect data?
- Yes (please specify)
- No (please specify)
- 4.5 Is there an adequate description of the methods of data analysis?
- Yes (please specify)
- No (please specify)
- 4.6 Do the authors avoid selective reporting bias? (e.g. do they report on all variables they aimed to study, as specified in their aims/research questions?)
- Yes (please specify)
- No (please specify)

##### Quality of the study - methods

- 4.7 Are there ethical concerns about the way the study was done? Consider consent, funding, privacy, etc.
- Yes, some concerns (please specify)
- No (please specify)
- 4.8 Were potential users of the research appropriately involved in the design or conduct of the study?
- Yes, a lot (please specify)
- Yes, a little (please specify)
- No (please specify)
- 4.9 Was the choice of research design appropriate for addressing the research question(s) posed?
- Yes (please specify)
- No (please specify)

- 4.10 Have sufficient attempts been made to establish the repeatability or reliability of data collection methods or tools? Yes, good (please specify)  
Yes, some attempt (please specify)  
No, none (please specify)
- 4.11 Have sufficient attempts been made to establish the validity or trustworthiness of data collection tools and methods? Yes, good (please specify)  
Yes, some attempt (please specify)  
No, none (please specify)
- 4.12 Have sufficient attempts been made to establish the repeatability or reliability of data analysis? Yes (please specify)  
No (please specify)
- 4.13 Have sufficient attempts been made to establish the validity or trustworthiness of data analysis? Yes, good (please specify)  
Yes, some attempt (please specify)  
No, none (please specify)
- 4.14 To what extent are the research design and methods employed able to rule out any other sources of error/bias which would lead to alternative explanations for the findings of the study? A lot (please specify)  
A little (please specify)  
Not at all (please specify)
- 4.15 How generalisable are the study results? Details
- 4.16 In light of the above, do the reviewers differ from the authors over the findings or conclusions of the study? Please state what any difference is. Not applicable (no difference in conclusions) Please state what any difference is.  
Yes (please specify)
- 4.17 What is the overall quality of the study? *(taking into account all the quality assessment issues)* High (quality)  
Medium (quality)  
Low (quality)

Relevance

4.18 **Can the study deliver inferences about the mechanisms at work in the review question?**  
 High  
 Medium  
 Low

4.19 **Can the study deliver inferences about the effectiveness of the intervention/independent variable of the review question?**  
 High  
 Medium  
 Low

Weight of evidence

4.20 **Weight of evidence A:** High trustworthiness  
 Taking account of all quality assessment issues, can the study findings be trusted in answering the study question(s)? How good is the execution of the study?  
 Medium trustworthiness

In some studies it is difficult to distinguish between the findings of the study and the conclusions. In those cases, please code the trustworthiness of these combined results/ conclusions.  
 Low trustworthiness

4.21 **Weight of evidence B:** High: RCTs  
 Appropriateness of research design for allowing causal inference  
 Medium: non-randomly allocated prospective evaluations

4.22 **Weight of evidence C:** Low: all other study designs  
 High  
 Medium  
 Low  
 Relevance of particular focus of the study for addressing the question, or sub-questions, of this specific systematic review

4.23	<b>Weight of evidence D: Overall weight of evidence.</b> Taking into account quality of execution, appropriateness of design and relevance of focus, what is the overall weight of evidence this study provides to answer the question of this specific systematic review?	High  Medium  Low
5	<b>Results and conclusions</b>	
5.1	<b>What are the results of the study as reported by the authors?</b>	Details
5.2	<b>Outcome: impact evaluation</b>	Quantitative studies Average impact coefficient (preferred specifications) Standard deviation of average impact Significance Upper limit of the 95% interval of confidence Lower limit of the 95% interval of confidence Qualitative studies Strong impact Low impact No impact Not applied
5.3	<b>Does the study find evidence of non linearities in effects?</b>	Not applied Not analysed Not clear Yes. Describe details. No. Describe details.
5.4	<b>Mechanism: what and how channels of intervention work to enhance the quality of contract enforcement</b>	Contract enforcement and finance (describe) Contract enforcement and international trade (describe) Contract enforcement and vertical or horizontal integration (describe) Other (specify and describe)
5.5	<b>Mechanism: what and how channels of intervention do NOT work to enhance the quality of contract enforcement</b>	Contract enforcement and finance (describe) Contract enforcement and international trade (describe) Contract enforcement and vertical or horizontal integration (describe) Other (specify and describe)
5.6	<b>Context</b>	Details



- 5.7 **Measurement of quality of contract enforcement used**
- Expert opinion
    - Ordinal
    - Cost and time of judicial procedures
    - Other (specify)
  - Business survey
    - Ordinal
    - Cost and time of judicial procedures
    - Other (specify)
  - Contract intensity money (CIM)
  - Informal enforcement
  - Others (specify)
- 5.8 **Measurement of investment**
- Level
  - Rate (investment/GDP)
  - Rate (flux/stock)
  - Stock (i.e. accumulated investment; level of asset)
  - Others (specify)
- 5.9 **Do the author(s) refer to other previous studies and refute or confirm their results?**
- No
  - Not clear
  - Yes, to confirm (included in our review; specify)
  - Yes, to confirm (not included in our review)
  - Yes, to refute (included in our review; specify)
  - Yes, to refute (not included in our review)
  - Yes, ambiguous results (included in our review; specify)
  - Yes, ambiguous (not included in our review)
- 5.10 **What do the author(s) conclude about the findings of the study?**
- 5.11 **What do the reviewer(s) conclude about the findings of the study? If different from the author(s) conclusions**
- 5.12 **Strength of findings**
- No clear conclusions can be drawn. Not significant
  - Results ambiguous, but there appears to be a trend.
  - Conclusions can probably be based on the results.
  - Results are clear and very likely to be true.
  - Results are unequivocal.
- 5.13 **Overall impression of the work**
- 5.14 **Any additional comments regarding the overall strengths and weaknesses of the work)**