On the Determinants and Consequences of Informal Contracting

Ricard Gil Giorgio Zanarone

Johns Hopkins Carey Business School CUNEF

July 2015

Abstract

As documented by Macauley (1963) and others, informal contracts are pervasive in modern economies. Yet, systematic empirical evidence on them is still limited. In this paper, we provide a framework to investigate the determinants and consequences of informal contracts. First, we present a model that organizes key predictions from the theoretical literature. Next, we discuss selected empirical works that shed light on the model’s empirical relevance. Finally, we discuss strategies for testing theoretical predictions for which conclusive evidence is still missing, as well as unexplored research opportunities offered by available studies and data.

Keywords: Contracts, Enforcement, Empirical evidence, Testability.

JEL codes: D23, L24, M21.

* We thank Pol Antras, Benito Arruñada, Jordi Blanes-i-Vidal, Alessandro Bonatti, Albert Choi, Oscar Contreras, Gillian Hadfield, Bob Gibbons, Fernando Gomez, Desmond Lo, Kevin J. Murphy, Lorenzo Sacconi, Orie Shelef, Patrick Warren, Michael Whinston, and audiences at MIT, Pompeu Fabra University, ISNIE and the AEA Meetings for useful comments and suggestions. This study received financial support from the Spanish Ministry of the Economy and Competitiveness through grant ECO2014-57131-R.
1. Introduction

Thanks to scholars like Coase, Williamson, Hart, Holmstrom, and to the empirical works inspired by their theories, the importance of formal contracts as vehicles of commitment and exchange is now well understood in economics.¹ At the same time, economists, sociologists and legal scholars alike have documented the existence and pervasiveness of informal contracts—that is, contracts that are enforced by parties, markets or communities through the threat of terminating trade, rather than by courts.² For instance, managers often rely on “hand-shake” agreements to support their deals (Macauley 1963); formally centralized companies and governments motivate employees (Foss 2003) and bureaucrats (Xu 2011) by informally promising to delegate authority; large corporations such as General Motors rely on informal, internally enforced routines to manage their workers and suppliers (Helper and Henderson 2014); and long-distance traders enter commercial contracts even in the absence of reliable courts because they are afraid of being ostracized from the market, as under the medieval Law Merchant (Milgrom et al. 1990).³

Inspired by these and other works, a rich theoretical literature has emerged in economics, investigating the conditions that make informal contracts feasible (Bull 1987; MacLeod and Malcomson 1989; Williamson 1991), their dynamic patterns (Levin 2002,

---

¹ For comprehensive reviews of the empirical evidence on contracts see, for instance, Lafontaine and Slade (1997, 2013).
² The economics literature often refers to contracts that are not enforced by courts as “relational”. However, some economists, and many legal scholars, define relational contracts more broadly as those that govern long-term relationships, and include in their definition contracts that are enforced by courts based on non-standard legal rules (e.g., McNeil 1978; Williamson 1979; Schwartz 1992). To avoid confusion, we refer to contracts that are not enforced by courts as “informal”, rather than “relational”.
³ Ostracism by local communities has also been shown to play an important role in protecting public goods from free-riding (Ostrom 1990). See Dixit (2004) for a comprehensive review of the literature on community enforcement.
03; Ray 2002; Halac 2012; Li and Matouschek 2013), and the way formal contracts help sustain and enforce informal ones by reducing the parties’ temptation to renege on their promises (Klein and Murphy 1988, 1997; Klein 1995, 2000; Baker et al. 1994, 2001, 02, 11; Battigalli and Maggi 2008; Kvaløy and Olsen 2009; Zanarone 2013).4

Despite the ubiquitousness of informal contracts, documented by the case studies, there is scarce systematic evidence on whether existing economic theories correctly predict their determinants and consequences.5 In a way, this is not surprising given that, by definition, informal contracts are difficult to measure. However, direct data on the use of formal and informal contracts is sometimes available because organizations keep internal records of, and optimize upon, these two contractual forms (Gil 2013). Moreover, the existing theories may predict patterns that are unique to informal contracting environments, and hence can be used to develop indirect empirical tests that may be performed using more conventional data on contracts and outcomes. Building on these observations, in this paper we provide a framework to empirically investigate informal contracts, in the hope that it may both shed light on the relevance of existing theories and guide future empirical research. To the best of our knowledge, this is the first such study in the literature.

In the first part of the paper, we present a simple model that illustrates how informal contracts can be optimally used, and combined with formal ones, to solve a variety of organizational problems, from moral hazard (e.g., MacLeod and Malcomson 1989; Baker et al. 1994; Levin 2003) to holdup (e.g., Baker et al. 2002; Halonen 2002) and ex post

---

4 Up-to-date reviews of the theoretical literature on informal contracts can be found in MacLeod (2007) and Malcomson (2013).

5 For a discussion of early evidence and case studies consistent with the presence of informal contracts, see Shelanski and Klein (1995).
adaptation (e.g., Williamson 1991; Baker et al. 2011). For reasons of space, we focus on informal contracts that are optimally stationary—that is, that do not change over time. We discuss dynamic informal contracts, whose theoretical study is still at its outset, in a companion paper (Gil and Zanarone 2015).

The central idea in our model is that informal contracts trade off enforcement security, which is stronger under court-enforcement, in exchange for savings in the cost of formalizing contracts in ways that courts can understand. The optimal contract—that is, the contract that maximizes total performance net of contracting costs—may be thus purely formal, purely informal, or it could contain a mix of formal and informal terms. The model generates three sets of testable predictions. The first set relates to factors that drive the use of purely informal contracts, as opposed to purely formal, or formal and informal in combination—namely, the parties’ intertemporal discount rate, their fallback option after the relationship breaks down, and their degree of incentive misalignment. The second set of predictions relates to the interaction between formal and informal contracts—that is, how the formal contractual provisions and outcomes that we expect to observe when informal contracts are being used, differ from those we would expect in the absence of informal contracts. The third set of predictions highlights how informal contracts become more effective when the parties are embedded in networks, which we model, following Levin (2002), as multilateral contracts where non-performance by one party triggers termination of his relationship with all the counterparts, rather than with the victim of non-performance alone.

In the second part of the paper, we discuss recent empirical works that may be interpreted as tests of our theoretical predictions. By doing this, we are not aiming to
review the empirical literature on informal contracting, but to draw inferences on how to measure and identify key variables that may be replicated by future studies. Hence, we deliberately focus on selected works that 1) offer systematic econometric evidence with 2) strong identification and 3) accurate measurement of variables, and that 4) use promising datasets. Because we focus on works that satisfy all or most of the above criteria, we do not discuss papers referring to purely anecdotal evidence and case studies. The empirical works that we discuss vary from CEO compensation (Gillan et al. 2009) and movie exhibition (Gil 2013; Barron et al. 2015) to procurement contracts (Johnson et al. 2002; Corts and Singh 2004; Camuffo et al. 2007; Shi and Susarla 2011; Gil and Marion 2013), franchising (Zanarone 2009, 2013), international trade (Macchiavello and Morjaria 2015; Antras and Foley 2014), public transportation (Jackson and Snyder 2011), and the relationships between elected politicians and interest groups (Shelef 2013), among others. Overall, we find combined support for most theoretical predictions from the model. Nevertheless we notice that a few predictions, mainly on the interaction between formal and informal contracts, are still awaiting clear supportive evidence, perhaps due to their demanding empirical requirements. We present and discuss these untested predictions as a clear opportunity for future empirical research.

In the final part of the paper, we discuss empirical strategies that may allow researchers to explore the untested theoretical predictions, and we comment on further research opportunities offered by existing data. In particular, we discuss three recent empirical studies on contracts (Forbes and Lederman 2009; DeVaro et al. 2015; Bolton et al. 2014) whose data, alone or in combination with data from other studies, could be used to further test predictions from our illustrative model.
The rest of the paper is organized as follows. Section 2 presents our illustrative model and derives testable implications. Section 3 draws empirical lessons on how to test the model’s predictions from selected empirical works on informal contracts. Section 4 goes back to theory, extending the analysis of informal contracting to transactions backed by business and social networks. Section 5 draws lessons from selected empirical works on informal contracts in networks. Section 6 discusses opportunities for future research. Section 7 concludes.

2. An illustrative model

2.1. Model setup

There are two parties, denoted as A and B. Both A and B live forever, and discount next-period payoffs by the factor $1/(1+r)$. Time evolves in discrete periods.

In any given period, A chooses an action vector $\mathbf{a}$ from the feasible set $H$, which produces payoff $\pi_B(\mathbf{a})$ for B and $\pi_A(\mathbf{a})$ for A. A contract implementing action plan $\mathbf{a}$ is a fixed payment $s$, plus a payment schedule $b(\cdot)$, which may vary with A’s actions, such that, if A accepts the contract, $\mathbf{a}$ is his preferred action plan—that is, $\pi_A(\mathbf{a}) + b(\mathbf{a}) > \pi_A(\mathbf{a'}) + b(\mathbf{a'})$ for any $\mathbf{a'} \neq \mathbf{a}$. If B and A do not enter a contractual relationship, B obtains payoff $\pi_B$ and A obtains payoff $\pi_A$, yielding surplus $S = \pi_B + \pi_A$. 
We assume B and A jointly observe either the chosen actions or the realized payoffs \( \pi_B(a) \) and \( \pi_A(a) \). We also assume B and A jointly observe monetary payments.\(^6\)

**Relation to existing models**

The above formulation captures a broad set of contracting situations. For instance, B may be A’s employer, \( a \) the unobservable effort spent by A on various tasks, \( s \) the base salary that B pays A, and \( b(\cdot) \) a bonus contingent on some performance measures that depend on A’s actions, as in agency models of incentive contracts (e.g., Holmstrom and Milgrom 1991; Baker 1992, 2002).

Alternatively, A may be a seller, B a buyer, \( a \) a vector of observable specific investments affecting the quality of the good or service that A supplies to B, \( s \) the good’s pre-agreed price, and \( b(\cdot) \) a price adjustment that B and A expect to negotiate ex post given the initial allocation of assets and decision rights, as in property rights and holdup models (e.g., Grossman and Hart 1986; Hart and Moore 1988, 1990). If the initial contract involves a reallocation of asset ownership, \( s \) may also include a per period fee covering the costs of such reallocation, and \( b(\cdot) \) may include the stream of asset returns and payoff rights that are transferred from B to A (or from A to B, if \( b(\cdot) \) is negative) after the reallocation, as in incentive-system models of firm boundaries (e.g., Holmstrom and Milgrom 1991, 1994; Gibbons 2005).

Finally, \( a \) may be a set of observable managerial, administrative, or political decisions that B wants A to implement (within an organization, or as part of a supply or distribution

---

\(^6\) Since none of the empirical papers we discuss here relates to uncertainty, we use a deterministic model to economize on notation. All of our results, and in particular those on informal contracts, survive in a non-deterministic environment provided that the principal and the agent are risk-neutral (Levin 2003).
network, a strategic alliance, and the like), and $b(\cdot)$ a bonus or penalty contingent on such decisions, as in adaptation models (e.g., Simon 1951; Williamson 1991; Baker et al. 2008, 2011; Hart and Holmstrom 2010). Notice that sometimes B may want A to make no decisions at all, as in models of delegation (Aghion and Tirole 1997; Baker et al. 1999) where the boss (A) promises not to interfere with the autonomy of his subordinate (B).

**Formal and informal contracts**

The key feature of our model is that B and A may contract either formally or informally. A formal contract is enforced by the threat of court-imposed damages, whereas an informal contract is enforced by the threat of terminating the relationship. Since courts are backed by the State’s coercive power, we assume a formal contract is automatically enforced—that is, B has no option to withdraw the contingent payment once A has performed, whereas B has such option under an informal contract.\(^7\) On the other hand, since a formal contract is enforced by courts, it must depend on variables that courts can verify. In contrast, an informal contract can directly depend on A’s actions.\(^8\)

To capture this tradeoff between enforcement and contracting costs (e.g., Battigalli and Maggi 2008, Kvaløy and Olsen 2009), we define the cost of formally contracting

\(^7\) In reality, enforcement of courts’ decisions by the State requires that the State’s rulers be willing to incur enforcement costs, and so it must be itself self-enforcing (Aldashev and Zanarone 2014). However, if a ruler fails to enforce, he compromises his multilateral relationship with all citizens, whereas if two parties fail to comply with their contract, they only compromise their own relationship, which is narrower in scope. Hence, given a court-verifiable contract, State enforcement may be more credible than informal second-party enforcement.

\(^8\) Baker and Choi (2015) study the interaction between formal and informal contracts when verification by the parties is also costly. Gauza and Gomez (2015), building on the survey data in Bozovic and Hadfield (2015), show how in that case the parties may want to use legal or formalistic language even in the absence of court enforcement.
action plan $\mathbf{a}$, $f(\mathbf{a})$, as the minimum expenditure necessary to make a contract that implements plan $\mathbf{a}$ court-verifiable. We assume for simplicity that given plan $\mathbf{a}$, the formal contract that implements it at the minimum cost $f(\mathbf{a})$ is unique, and we define it as $b^r(\mathbf{a})$. Without loss of generality, we also assume that $f(\cdot)$ is paid by $B$ at the time of contracting. We interpret $f(\cdot)$ as the cost—in terms of time and lawyers’ fees—necessary to draft the contract in such a way that a court will correctly interpret $B$’s and $A$’s obligations and enforce them as specified.\footnote{For a detailed discussion of the costs of specifying contracts in order to make them court-verifiable, see Battigalli and Maggi (2002), and Schwartz and Scott (2003). An alternative interpretation of the contracting cost $f(\mathbf{a})$ is as the cost of proving breach in court. In that case, $f(\mathbf{a})$ will not be incurred unless breach arises in equilibrium. However, $f(\mathbf{a})$ will still shape $A$’s equilibrium actions under formal contracting by imposing a credibility constraint that $B$ and $A$ be willing to pay $f(\mathbf{a})$ in case of breach. For the purposes of this paper, the predictions from such a model are qualitatively equivalent to those presented here.}

The costs of formal contracting are (mostly implicitly) present in the literature, where key variables are assumed to be “non-contractible”. For instance, in models of formal incentive contracts the cost of paying a bonus that is directly contingent on the agent’s actions is infinite, and so is the cost of paying a bonus based on the realized payoffs. As a result, the second-best incentive contract tends to be distortionary—for instance, because it is based on a contractible but imperfect performance measure that induces the agent to spend an inefficient amount of effort in some states (Baker 1992; Baker \textit{et al.} 1994), or to misallocate effort across tasks (Holmstrom and Milgrom 1991; Baker 2002). In holdup and property rights models, it is assumed that formally contracting on the seller’s (or the buyer’s) specific investments is infinitely costly, and that contracting on the outcomes of such investments is infinitely costly before the investments are made, and costless afterwards. Moreover, decision rights about the investments’ outcomes are costlessly
contractible. Hence, the only way a formal contract can affect specific investments is by allocating asset ownership and decision rights, which will influence ex post bargaining between the buyer and the seller about the investments’ outcomes, and hence their ex ante investment choices.

2.2. Purely formal contracts

Under a purely formal contract, in every period B contracts with A the action plan $a^F$, which solves:

$$\max_{a \in H} \pi_B(a) - b^F(a) - s - f(a),$$

subject to B’s and A’s participation constraints:

$$\pi_B(a) - b^F(a) - s - f(a) \geq \pi_B,$$  \hspace{1cm} (1)

$$\pi_A(a) + b^F(a) + s \geq \pi_A.$$  \hspace{1cm} (2)

Substituting for the lowest fixed payment $s$ consistent with (2), we can write:

$$a^F \equiv \arg \max_{a \in H} \left\{ \pi_B(a) + \pi_A(a) - f(a) - \pi_A \right\},$$

subject to

$$\pi_B(a) + \pi_A(a) - f(a) \geq S.$$  \hspace{1cm} (3)

Two things are noteworthy. First, in the absence of contracting costs, A’s action plan under the optimal formal contract coincides with the first best action plan. Second, the joint surplus under an optimal formal contract, given by $S^F \equiv \pi_B(a^F) + \pi_A(a^F) - f(a^F)$, is non-increasing in the contracting cost $f(\cdot)$. This is consistent with the literature: when
writing a formal contract that implements first best actions is too costly, the parties may agree on a second best contract, which is less costly to write but also generates less efficient actions.

2.3. Purely informal contracts

Under a purely informal contract, no contracting costs are incurred because courts are not involved and B perfectly observes either A’s actions (as in agency models) or their outcomes (as in property rights and adaptation models). Hence, B and A can agree on a simple contract whereby B pays a discretionary bonus \( b'(a) \) if A takes the promised action plan \( a \), and zero otherwise.\(^{10}\)

At the same time, under an informal contract B and A do not benefit from the State’s commitment to enforce, so their future losses from terminating the relationship must be sufficiently large to deter non-compliance. As standard in the literature, we model an informal contract as a trigger-strategy equilibrium of the infinitely repeated game: B and A keep honoring the informal contract until either party breaches and, in case of breach, they revert to the optimal equilibrium without informal contracts forever after. Given the absence of liquidity constraints, the optimal informal contract is stationary, in the sense that A’s actions and B’s payments are the same in every period (MacLeod and Malcomson 1989; Levin 2003).\(^{11}\)

\(^{10}\) In practice, incentives in informal contract do not always come from discretionary bonus payments. For instance, employment contracts and franchising agreements often rely on formally contracted quasi-rents, such as efficiency wages and guaranteed revenues, which the employer/franchisor can severe by terminating the relationship if he detects non-compliance (Shapiro and Stiglitz 1984; Klein 1980; Klein and Murphy 1988; Klein 1995; Lafontaine and Raynaud 2000). Evidence consistent with the existence of quasi-rents in franchising agreements is provided by Kaufmann and Lafontaine (1994) and, more recently, Kosova and Sertsios (2012).

\(^{11}\) The testable predictions derived here would still hold in a model with liquidity constraints, where optimal informal contracts may be non-stationary (e.g., Ray 2002; Li and Matouschek 2013). A non-stationary
Let $S^{sp} \equiv \max \{S^F, S\}$ be the parties’ fallback option—that is, their joint surplus in the optimal post-breach equilibrium (either an optimal formal contract or termination)—and let $\pi^B$ and $\pi^A$ be B’s and A’s individual fallback payoffs, so that $\pi^B + \pi^A = S^{sp}$. Finally, let $a^0$ be the “opportunistic” action plan, which maximizes A’s gross payoff $\pi_A(a)$, and hence will be chosen by A if he reneges on the informal contract. Then, in each period, B contracts with A the action plan $a^1$, which solves:

$$\max_{a \in H} \pi_A(a) - b^I(a) - s,$$

subject to B’s and A’s participation constraints

$$\pi_B(a) - b^I(a) - s \geq \pi^B$$ and

$$\pi_A(a) + b^I(a) + s \geq \pi^A,$$ (4)

and to A’s and B’s incentive constraints, which insure that A be willing to take the promised actions, and B be willing to pay the promised bonus:

$$\pi_A(a) + b^I(a) + \frac{1}{r} \left[ \pi_A(a) + b^I(a) + s \right] \geq \pi_A(a^0) + \frac{1}{r} \pi^A$$ for A, and

$$-b^I(a) + \frac{1}{r} \left[ \pi_A(a) - b^I(a) - s \right] \geq \frac{1}{r} \pi^A$$ for B. (6)

Substituting for the lowest fixed payment $s$ consistent with (5) and the lowest bonus $b^I(a)$ consistent with (6), constraint (4) becomes $\pi_B(a) + \pi_A(a) \geq S^{sp}$, whereas constraint (7) becomes:

---

model would generate additional predictions on the dynamics of informal contract terms. We abstract from those because no empirical evidence on contract dynamics has been developed yet.
\[ \pi_n(a) + \pi_A(a) \geq S^{sp} + r\left[\pi_A(a') - \pi_A(a)\right]. \]  

Since (8) is more restrictive, it is both necessary and sufficient for a purely informal contract to be self-enforcing, so that:

\[ a^I \equiv \arg\max_{a \in H} \left\{\pi_n(a) + \pi_A(a) - \pi_A(a')\right\}, \text{ subject to } (8). \]

Condition (8) implies that the joint surplus under a purely informal contract, \( S^I \equiv \pi_n(a^I) + \pi_A(a^I) \), is non-increasing in \( r \). When \( r \) is close to zero, A’s actions under an informal contract coincide with the first best actions, whereas when \( r \) is large enough, informal contracts cannot generate more efficient actions than formal ones. Moreover, \( S^I \) is non-decreasing in \( f(\cdot) \), the cost of formal contracting. That is because, when \( S^{sp} = S^F \), so that B and A revert to formal contracting if B breaches the informal contract, an increase in \( f(\cdot) \) reduces B’s post-breach payoff, thereby making informal contracts easier to enforce.

### 2.4. Informal contracts backed by formal contracts

While we have so far analyzed formal and informal contracts as alternative governance modes, in practice they are often used together. For instance, Lincoln Electric notoriously uses both formally contracted piece rates and discretionary bonuses to incentivize workers (Fast and Berg 1975; Baker et al. 1994). As discussed in several papers by Klein and coauthors, and by Baker, Gibbons and Murphy, one reason for using formal and informal contracts together is that formal contracts may be used as credible
threats within an overarching informal agreement in order to reduce the parties’ temptation to breach.\textsuperscript{12} In this section, we use our framework to broadly state this point.

When formal and informal contracts are used together, B informally contracts with A the equilibrium action plan $a^I+F$, and formally contracts the off-the-equilibrium action plan $a^F$. In any given period, if B and A honor the informal contract, the formal one is not enforced, and the game is repeated in the following period. If either party reneges on the informal contract, the formal contract is enforced in the current period, and B and A revert to the spot equilibrium forever after. Let action plan $a \in H \subseteq H$ and payment $b^F(a)$ be a formal contract (i) that B is willing to enforce in case A reneges on the informal contract ($\pi_B(a) - b^F(a) \geq \pi_B(a^0)$) and (ii) that reduces A’s temptation to renege ($\pi_A(a) + b^F(a) \leq \pi_A(a^0)$). Then, $a^{I+F}$ and $a^F$ solve:

$$\max_{a \in H} \pi_B(a) - b^I(a) - s - f(a),$$

subject to the participation constraints:

$$\pi_B(a) - b^I(a) - s - f(a) \geq \pi_B^{sp}, \quad \text{and}$$

$$\pi_A(a) + b^I(a) + s \geq \pi_A^{sp}, \quad (9)$$

and to the incentive constraints:

$$\pi_A(a) + b^I(a) + \frac{1}{r} \left[ \pi_A(a) + b^I(a) + s \right] \geq \left[ \pi_A(a) + b^F(a) \right] + \frac{1}{r} \pi_A^{sp}, \quad \text{and}$$

$$-b^I(a) + \frac{1}{r} \left[ \pi_B(a) - b^I(a) - s - f(a) \right] \geq \max \left\{ 0, \pi_B(a) - \pi_B(a) - b^F(a) \right\} + \frac{1}{r} \pi_B^{sp}. \quad (12)$$

To understand the right-hand side of (12), notice that if B decides to renege after A has taken the informally agreed actions, he can either withdraw the informal bonus and receive payoff $\pi_a(a)$ from the informal actions, or withdraw the informal bonus and enforce the formal contract, thus receiving payoff $\pi_b(a) - b^f(a)$. Substituting for the minimum fixed payment $s$ consistent with (10) and the minimum informal bonus consistent with (11), we have:

$$\left( a^{1-f}, a^f \right) \equiv \arg \max_{a^{1-f}, a^f} \left\{ \pi_b(a) + \pi_a(a) - f(a) - \pi_a^{sp} \right\}, \text{ subject to }$$

$$\pi_b(a) + \pi_a(a) - f(a) \geq S^{sp} + r \left[ \pi_a(a) - \pi_a(a) + \max \left\{ b^f(a), \pi_b(a) - \pi_b(a) \right\} \right]. \quad (13)$$

The joint per period surplus under a mix of formal and informal contractual provisions is thus $S^{1-f} \equiv \pi_n \left( a^{1-f} \right) + \pi_a \left( a^{1-f} \right) - f \left( a^f \right)$.}

Notice that while in the above formulation the cost of writing an off-the-equilibrium formal contract must be incurred every period, all of the results that follow would continue to hold if we allowed for long-term formal contracts whose cost is incurred by B only once, at the beginning of the relationship. In that case, the cost of formal contracting would no longer affect the long-term value of the relationship (the left-hand side of condition 13), but it would still affect B’s choice at the outset of whether to write a formal contract, and of what kind. Moreover the actions and payments in the formal contract would continue to affect B’s and A’s joint temptation to renege on the informal contract (the right-hand side of 13). Hence, the tradeoff between enforcement and contracting costs at the root of our model would exist even in the presence of long-term formal contracts.
2.5. First set of testable predictions: the choice of enforcement regime

A first set of testable implications from our model regards the choice of enforcement regime—that is, the extent to which B and A use formal or informal contracts to govern their transactions. First, an increase in the intertemporal discount rate \( r \) and in the post-termination surplus \( S \) does not favor informal contracts. For a given \( f(\cdot) \), the surplus under pure formal contracting, \( S^F \), does not depend on \( r \) and \( S \) (constraint 3), whereas the surpluses under both pure and mixed informal contracting, \( S^I \) and \( S^{I+F} \), weakly decrease in \( r \) and \( S \) (constraints 8 and 13). Notice that, when \( S \) grows too large, no informal contracts can be enforced, so B and A do not enter a relationship. Conditional on B and A being in a contractual relationship, though, both an increase in the intertemporal discount rate \( r \) and in the post-termination payoff \( S \) disfavor informal contracts relative to formal ones.

Second, the equilibrium surplus under pure informal contracts is weakly decreasing in \( \pi_a(a^v) \), A’s payoff from reneging (constraint 8). Hence, an increase in \( \pi_a(a^v) \) favors the use of formal contracts as a means to reduce the parties’ reneging temptation.

Finally, for a given discount rate \( r \), there is a non-positive relationship between the cost of formal contracts, \( f(\cdot) \), and their use when the fallback option following reneging is termination (\( S > S^F \)), whereas the relationship is ambiguous when the fallback option is continuation under pure formal contracting (\( S < S^F \)). When \( S > S^F \), an increase in \( f(\cdot) \) reduces \( S^{I+F} \) relative to \( S^I \), and hence disfavors formal contracting, because it increases the cost of using formal contracts to mitigate the reneging temptation. When \( S < S^F \), an
increase in \( f(\cdot) \) has the additional effect of making the fallback option after reneging less attractive, and hence it has an ambiguous effect on the choice between formal and informal contracts. In other words, as shown by Baker et al. (1994) for the case of incentive pay, formal and informal contracts can be both substitutes and complements.

We summarize our predictions on the choice between formal and informal contracts in the following proposition.

**Proposition 1:** The likelihood of using formal contracts (i) is weakly increasing in the intertemporal discount rate and in the post-termination surplus, and (ii) is weakly increasing in A’s payoff from choosing opportunistic actions. Moreover, (iii) the relationship between the cost of formal contracts and their use is non-positive when the fallback option following breach of an informal contract is termination, and ambiguous when the fallback option is continuation under pure formal contracting.

Notice that Proposition 1 predicts that, as the discount rate decreases, the surplus under pure informal contracting increases so that, for low enough discount rates, it is efficient to use informal contracts alone, without formal ones. This does not necessarily imply, as argued at times, that when formal contracts are used, their complexity (in our model, that may be measured by the contracting costs incurred in equilibrium) monotonically decreases in the discount rate. To obtain sharp predictions on the relationship between contractual complexity and the discount rate, the model should be tailored to match the idiosyncrasies of the particular contracting environment under study.\(^{13}\)

---

\(^{13}\) Some of the existing evidence suggests such a modeling effort may be worthwhile. For instance, in a study of technology development contracts, Ryall and Sampson (2009) find that contracts between long-
2.6. Second Set of Testable predictions: Contracts and outcomes under a given enforcement regime

A second set of testable predictions from our model regards the optimal actions and contract terms within a given enforcement regime (purely formal, informal, or mixed). When the predictions differ depending on whether informal contracts are used or not, they allow us to indirectly test for the presence of informal contracts and their interaction with formal ones.

First, constraints (8) and (13) imply that, when informal contracts are used, the equilibrium surplus is weakly decreasing in the discount rate $r$ and in the post-termination surplus $S$. Formally, as $r$ and $S$ increase, $S^I$ and $S^{I+F}$ tend to collapse into $S^{sp} = \max \{S, S^F\}$, the surplus available in the absence of informal contracting.

Second, it follows from constraints (3) and (13) that formal contracts may take different forms depending on whether they are used alone, or to support informal contracts. For instance, a formal contract $a$ that produces little payoff for $B$ ($\pi_B(a) \approx 0$) but a low payoff for $A$ ($\pi_A(a) < \pi_A(a^*)$) is unlikely to be optimal under pure formal contracting (equation 3). However, the threat to enforce such a contract may reduce the informal payment that $B$ must promise to induce $A$ to perform, and hence $B$’s temptation not to pay (equation 13).\textsuperscript{14} This implies that one strategy to investigate the interaction between formal and informal contracts is to look for seemingly suboptimal formal contracts that reduce the parties’ reneging temptation.

\textsuperscript{14} See Iossa and Spagnolo (2011) for a model in this spirit.
Third, and related, constraint (13) implies that observed actions and payments, as well as the way actions respond to payments, may differ from formally contracted ones when formal contracts are used to facilitate informal contracts but are not enforced in equilibrium.

Fourth, (3) and (13) imply that formal contracts depend on the discount rate $r$ if, and only if they are used to facilitate informal ones. Hence, empirical studies that aim to identify the role of formal contracts (dictating behavior, as in section 2.2, or facilitating superior informal contracts, as in section 2.4) should test for joint variation between formal contracts and $r$.

Notice that our general model cannot predict the specific patterns of divergence between observed formal contracts and contracts that would be optimal if enforced, the patterns of divergence between formal contracts and observed behavior, and the patterns of co-variation between formal contracts and the discount rate. Hence, empirical studies that aim to test for the interaction between formal and informal contracts by looking for such systematic patterns should first develop a specific model, tailored to the environment under study, that generates immediately testable hypotheses.

Finally, by the same argument as in Proposition 1 (iii), when the fallback option following breakdown of an informal contract is termination ($S > S^f$), an increase in the cost of formal contracting, $f(\cdot)$, cannot increase the equilibrium surplus under informal contracting. On the other hand, when the fallback option is continuation under pure formal contracting ($S < S^f$), an increase in $f(\cdot)$ has an ambiguous effect on the equilibrium surplus, as on one hand, it raises the cost of using formal contracts to reduce
the reneging temptation (a surplus-reducing effect), while on the other hand it reduces the fallback option after reneging (a surplus-increasing effect).

We summarize our second set of predictions in the following proposition.

**Proposition 2:** When informal contracts are used, (i) realized performance is weakly decreasing in the discount rate and in the post-termination surplus, (ii) observed formal contracts may be suboptimal if enforced, (iii) observed actions and payments may differ from formally contracted ones, and (iv) formal contract terms may vary systematically with the discount rate. Finally, (v) increases in the cost of formal contracting weakly reduce performance when the fallback option after breach of an informal contract is termination, whereas they have an ambiguous effect on performance when the fallback option is continuation under pure formal contracting.

Note that our predictions in (i) and (iv) of Proposition 2 above, on the relationship between contract terms, surplus and the discount rate, are sensitive to the static technology assumed in our model. If actions in one period affected future payoffs, surplus and contract terms under both formal and informal contracting may depend on the discount rate. In that case, one could no longer infer that informal contracts are being used from the simple observation that surplus decreases in the discount rate or that formal contract terms vary systematically with the discount rate. Hence, some evidence that the production technology is static would be necessary to fully support that conclusion.  

---

15 The surplus under formal contracts may also depend on the discount rate if courts work imperfectly, and hence enforce contracts with delay. This is not the case, though, if one interprets the discount rate as probability of termination due to exogenous events, because court-imposed damages will be paid to the victim of breach irrespective of whether she is still in a relationship with the breaching party when the court rules. Indeed, all of the empirical papers discussed here measure the discount rate as termination probability.
However, even in the presence of a dynamic technology, an increase in the discount rate would cause a greater surplus reduction under informal contracts than under formal ones, because it would have the additional effect of tightening the self-enforcement constraints (8) and (13). Hence, the prediction from part (i) of Proposition 1, which relates the discount rate to the choice between formal and informal contracts, seems robust to a non-static production technology.\textsuperscript{16} We will return on these caveats in the next section, where we compare the two sets of predictions from our theoretical model to the evidence in the informal contracting literature.

3. Inferences from selected empirical studies on informal contracting

In this section, we evaluate the empirical relevance of our illustrative model of informal contracting through the lens of recent empirical works. In the process, we hope to draw lessons on how to measure key variables from the model that may provide guidance for future empirical studies.

\textsuperscript{16} This may not be the case if liquidity constraints are present. For instance, in their model of apprenticeships, Garicano and Rayo (2013) assume a dynamic technology (an expert permanently transfers knowledge to a novice, thus affecting the latter’s future productivity) and liquidity constraints (the novice cannot compensate the expert upfront). They show that, while total surplus under an efficient informal contract on knowledge transfer decreases in the discount rate, as predicted by our model, surplus under the contract offered by the expert when he has bargaining power increases in the discount rate.
3.1. Choice between formal and informal enforcement

H1: High discount rate and post-termination surplus increases the use of formal contracts

Ideally, to test this first hypothesis we would need measures for the following variables. First, we would need a dependent variable indicating whether a formal contract exists. Second, we would need exogenous measures for the intertemporal discount rate and the parties’ payoffs outside the relationship, where by exogenous we mean that they must be uncorrelated with the quality of the match between the two contractual parties (Ackerberg and Botticini, 2002). If the measures of discount rate and post-termination payoffs are exogenous, a simple correlation will be informative about the causal relationship between the use of formal contracts and the discount rate regardless of whether we use cross-sectional or longitudinal variation in the corresponding data analysis.

Gillan et al. (2009) study the choice between explicit and implicit employment agreements (EAs) for CEOs in S&P 500 firms. They find that the use of explicit EAs (measured by an indicator for whether the firm’s SEC filings report an explicit agreement) increases in the CEO’s perceived uncertainty about the firm’s future prospects (measured by sales volatility, the rate at which firms in the industry change ownership and control, and an indicator for whether the CEO is new to the firm). They also find that the use of explicit EAs increases in the industry’s homogeneity (measured by the extent to which variations in industry stock returns are explained by an equally weighted historical industry index). Note that an interesting, and potentially replicable feature of Gillan et al. (2009) is that it exploits a property of the legal environment—
namely, the obligation to disclose all formal dimensions of CEO employment agreements to the SEC—to measure whether the compensation terms in such agreements are formal or informal.

Both of their results are consistent with hypothesis H1. In particular, when the firm’s prospects are uncertain (in terms of our model, the intertemporal discount rate is high), promising a purely discretionary compensation to the CEO is not credible, because there’s a high chance that the CEO-firm relationship will soon end, and hence the compensation promise will not be honored. Similarly, discretionary compensation will not be credible when the firm belongs to a homogeneous industry, and hence can more easily replace the CEO, because in that case the firm’s fallback option after reneging on the compensation promise will be too attractive. An explicit contract will mitigate these credibility problems, either by reducing the firm’s temptation to renege on the CEO or, in the case of severe uncertainty/extreme industry homogeneity, by fully substituting discretionary compensation with formally contracted one.

H2: Agent’s opportunism increases the use of formal contracts

To test this second hypothesis, in addition to variation in the use of formal and informal contracting, we need variation in the agent’s opportunism—formally, her opportunity cost of honoring an agreement.

Gil (2013) exploits a data set of movie exhibition contracts where 22 distributors place their movies on the screens of one Spanish exhibitor. These distributors may use formal contracts that weekly specify revenue-sharing terms. Alternatively, they may use informal contracts (no formal written terms are in place). If a formal contract is in place,
the distributor may pay an additional “kickback” to the exhibitor after her movie has been pulled out of the theater. Like Gillan et al. (2009), Gil (2013) suggests a strategy for measuring the use of informal contracts—namely, looking at internal company records. Assuming the practice of keeping records on the formal or informal nature of contract terms is not unique to the movie industry, the measurement strategy in Gil (2013) is potentially replicable.

The paper shows that movies that did well during their US release (which occurs a few months earlier than the Spanish release) are more likely to use a formal contract than movies that were not released in the US, or were released but did not perform well. This result is consistent with hypothesis H2. When the agreement is completely informal, the exhibitor, who collects revenues upfront, is tempted to renege, the more so the larger the movie’s revenues. To mitigate the exhibitor’s reneging hazard, movies with high expected revenues are governed by formal contracts, which prevent the exhibitor from retaining revenues at the cost of increasing the distributor’s temptation to renege on kickbacks.

\[ H3: \text{the relation between the costs of formal contracts and the use of purely informal ones is positive when the fallback option is termination, and ambiguous otherwise}\]

To this day we are not aware of any empirical paper that provides evidence regarding H3. Note that in order to test this hypothesis we would need, first, to be certain that in the empirical setting under study the fallback option is termination and, second, we would need good measures of both the cost of formal contracts and the use of informal ones. These are admittedly demanding requirements in terms of data availability. We comment
on H3 and other untested hypotheses later in the paper when we discuss opportunities for future empirical work.

3.2. Contracts and outcomes under a given enforcement regime

H4: Realized performance decreases in the discount rate and in the post-termination surplus

To test this hypothesis, we need clean measures of the discount rate (see above in H1) and measures of realized performance that are relevant to the potential contractual relationship. Several recent papers provide supporting evidence for H4.

Gil and Marion (2013) test how relationships between contractors and subcontractors impact their bids for procurement contracts. Using data for 10 years of Caltrans (California Department of Transportation) auctions, they are able to account for how often contractors have worked in the past with listed subcontractors in their bids, and how the future value of relationships may influence their bids. To measure the future value of relationships, they use the number and value of projects auctioned over the next calendar year. Since projects are offered when they are needed and not as a response to a particularly good match between contractor and subcontractor, this measure of the potential future value of the relationship is not contaminated by endogenous considerations of how well contractors and subcontractors are matched or were matched in the past. The results show that, consistent with our hypothesis H4, the higher the potential future value of the relationship, the lower the posted bids. This result still holds after controlling for past interactions between contractors and subcontractors—in fact,
coefficients of the interaction term indicate that past interactions matter only if future contracts are at stake.

Another paper that offers consistent evidence with H4 is Antras and Foley (2014). They investigate the formal financing contracts that are used in international trade and the enforcing mechanisms that are available to exporters. They use data from one large exporter of frozen chicken, who distributes to importers located in over 140 countries. Formal contracts used by the exporter vary on how exposed the exporter is to importer’s default. Cash-in-advance contracts (CIA) and letters of credit protect the exporter the most, while open account terms protect it the least. Among other things, the authors find that: (1) as the number and value of accumulated transactions grows within an exporter-importer dyad, the likelihood of using open account terms increases; (2) during the recent crisis the exporter was more likely, all else equal, to use CIA contracts for importers that did not transact with her before the crisis; and finally, 3) customers that purchased a larger share of goods on open account terms before the crisis are more likely to remain active after the crisis.\(^{17}\) This last result suggests that, once the importers’ temptation not to pay becomes more binding due to the crisis, weaker relationships—that is, those governed by cash-in-advance terms that formally protect the exporter—are less likely to survive. This finding is consistent with hypothesis H4 in our model, whereby performance (in this case, survival) is stronger when the discounted value of the relationship is higher.\(^{18}\)

\(^{17}\) The paper also states (p. 27) that switching contract form over time within a given buyer-seller dyad is rare. Although our model suggests that switching contractual form should take place, we may be omitting important switching costs that are specific to international trade in this industry.

\(^{18}\) An alternative interpretation, proposed by the authors, is that the exporter is less informed on the creditworthiness of cash-in-advance buyers, so he is more likely to terminate them after the crisis, when the
Johnson et al. (2002) study how informal relationships affected supplier trade credit in transitioning countries that belonged to the former Soviet. Using survey data, and controlling for country and industry fixed effects, they find that a supplier’s willingness to give trade credit to his customers increased in the customer’s cost of searching for alternative suppliers (negatively measured by the availability of similar suppliers at short distance). This result is consistent with hypothesis H4 because customers that cannot readily replace their supplier stand to lose more if their supplier terminates them following non-payment. This makes informal trade credit agreements more likely to be self-enforcing, and hence increase the supplier’s willingness to lend.

Shelef (2013) develops a model of relational influence buying where elected politicians favor interest groups in exchange for campaign contributions, and shows that equilibrium contributions decrease in the parties’ discount rate (consistent with our hypothesis H4). Using data from elections and campaign contributions for the US Congress and State legislatures, he tests this prediction and finds that a candidate’s campaign contributions decrease in two alternative measures for her time horizon/discount rate: (1) age and (2) whether the candidate is running for a last election term as mandated by State law. As argued by Shelef (2013), two caveats should be borne in mind in interpreting these results. First, the evidence comes from purely longitudinal, within-candidate variation in campaign contributions and his measures of discount rates, so one cannot fully control for the potentially confounding effect of the candidates’ experience. Second, the production technology of politicians is likely to be dynamic, in the sense that it may take them multiple terms in office to “deliver” a promised favor.

buyers’ cash constraints become more binding. This interpretation is consistent with the fact that, before the crisis, cash-in-advance contracts were used for buyers who had not transacted with the exporter in the past.
Hence, the observed positive correlation between candidates’ time horizons and campaign contributions may be partially due to low discounting of future returns, rather than weak informal contracting.

Cohen et al. (2013) provide evidence that NBA coaches that are more likely to be terminated are less likely to play their rookies (i.e., novice players) in any given game. In most cases, NBA coaches are responsible for winning games and develop players. Since the former is observable but the latter is not, coaches under high risk of termination will try to outperform in the success measure that is easily observable and therefore give priority to players with experience. Their findings are consistent with H4, in that a coach’s termination risk can be interpreted as a measure of her intertemporal discount rate. This evidence differs from that in other papers discussed above in that not only the match between coach and team is endogenous and could be driving the nature of the informal agreement, but also different teams may use different informal agreements with their coaches that could be driven by unobserved heterogeneity that the paper does not account for. In other words, because not all coaches are under the same informal contract with their teams, it is empirically challenging to disentangle whether the empirical finding in Cohen et al (2013) is due to different informal agreements or variation in the coaches’ discount rates.

We conclude this section on the link between discount rate and performance with a discussion of an innovative and potentially replicable strategy to estimate the discounted future value of informal relationships developed by Macchiavello and Morjaria (2015). Using a rich dataset on Kenyan rose exporters, Macchiavello and Morjaria (2015) simultaneously observe that the relationships between rose exporters and buyers are
never governed by formal contracts, and record spot auction prices for roses. Hence, although they cannot exogenously measure the discounted value of future relationships for each buyer-exporter dyad, they can estimate lower bounds of their value by computing the largest amount of money that the exporters forgo when honoring their informal agreements with buyers instead of selling on the spot market. Notice that the exporters’ efforts to deliver roses in one period are unlikely to affect future periods’ deliveries (the production technology is likely to be static), which suggests that the observed relationship between discounted value and performance may be attributed to the presence of informal contracting.

Because (given the self-enforcement constraint) the temptation to renege on an informal contract is a lower bound for its present discounted value, the Macchiavello-Morjaria approach could be replicated in other studies, and used to test predictions from informal contracting models such H4 and H1 above. A potential caveat is that a positive correlation between reneging temptation and performance may be partially caused by the fact that high-temptation deals are assigned to high-reputation partners. Hence, controlling for a partner’s type would be important in a test of H4 that estimates the discounted value of relationships using the Macchiavello-Morjaria approach.

\textit{H5: Observed formal contracts would be suboptimal if enforced}

In order to test this hypothesis, we would need to compare the observed formal contracts with those we would expect to observe in the absence of informal contracting. Here we comment on a few papers that present evidence consistent with hypothesis H5.
Zanarone (2009) studies how vertical restraints in Italian car dealership contracts changed after a 2002 EU regulation prohibited manufacturers to assign dealers to exclusive territories. Among other results, Zanarone (2009) finds that, while contracts before the legal change mostly relied on quantity floors to contain dealers’ double marginalization, contracts after the legal change contained a mix of both quantity floors and price ceilings.

This result is consistent with hypothesis H5 because if dealership contracts were purely formal, retail prices should decrease once intrabrand competition is liberalized. Hence, price ceilings should be less necessary after the legal change. Based on a series of field interviews, Zanarone (2009) argues that after the liberalization manufacturers planned to keep enforcing exclusive territories through informal agreements. If such informal agreements were in place and all dealers honored them, quantity floors would be sufficient to prevent double marginalization. However, if an individual dealer decided to deviate from the informal agreement, the new law would allow him to “game” the quantity floor by selling in other dealers’ territories, while restricting car sales and overpricing in his own territory. Including a price ceiling in the formal contract prevents such gaming, and thus reduces the dealers’ present gain from reneging on the informal exclusive territory agreement.19

Result (2) in the aforementioned paper by Antras and Foley (2014)—whereby during the recent crisis the exporter was more likely, all else equal, to use cash-in-advance contracts for importers that did not transact with him before—is also consistent with hypothesis H5. A model of purely formal contracts would predict that, anticipating a

---

19 This argument is explicitly modeled in a companion working paper (Zanarone 2007, accessible at gzanarone.cunef.edu).
more uncertain future and/or tighter cash constraints, the exporter would use CIA contract terms after the crisis for all similarly situated customers (i.e., importers located in countries with similar court quality, and characterized by similar sales volume, transaction type, etc.), irrespective of whether those importers have or have not transacted with him before. Instead, the result in Antras and Foley (2014) suggests that, consistent with the presence of informal contracts, efficient agreements involving trade credit (that is, open account terms) were used for customers who had a relationship with the exporter, while CIA terms were used for new customers.20

\[ H_6: \text{Observed outcomes are inconsistent with formally contracted ones} \]

In order to test this hypothesis, one needs to observe formal contracts and outcomes. Absent informal contracts, these two variables should coincide, as parties “play by the book”. Hence, a systematic discrepancy between formal contracts and outcomes indicates that formal contracts are not being enforced, suggesting that informal agreements may be in place. In addition, one should observe that the patterns of discrepancy be consistent with a specialized model of the informal contractual relationship under study.

As described above, Gil (2013) explores how movie distributors and exhibitors informally adjust revenue sharing terms in movie exhibition contracts. Consistent with H6, Gil (2013) finds that adjustments are not contracted formally—that is, realized revenue sharing systematically differs from the terms in the formal revenue-sharing agreements. Filson et al. (2005) argue that such adjustments occur informally to save on

---

20 One may argue that past interactions provide the exporter with “soft” information on the importer’s liquidity that is unobservable to the econometrician. In that case, as argued before, the result in Antras and Foley (2014) may also be consistent with a model where contracts are purely formal but customers have private information on their own creditworthiness.
contracting and negotiation costs. As a complementary argument, Barron et al. (2015) suggest that such adjustments occur to affect the allocation of movies to screens.

In particular, Barron et al. (2015) match Gil's (2013) contract data with detailed weekly box-office data by theater and movie and show that renegotiations of pricing contracts are related to decisions about whether to continue show a movie and, if so, whether to show it in prime time. Thus, Barron et al. (2015) are consistent with H6 in a broad sense: not only they show that observed payments differ from formally contracted ones, but also that observed actions respond to informal payments whereas they do not respond to formally contracted payments.

A methodological lesson from Barron et al. (2015) is that adaptation contracts—that is, contractual relationships whose goal is to agree on observable actions such as showing a movie—provide an easier setting for testing for the interaction between formal and informal contracts than agency settings, where the actions, such as employees’ effort, are typically unobservable.

Like Barron et al. (2015), Zanarone (2013) studies informal adaptation. Specifically, he studies theoretically how franchise chains adapt local service standards when the law requires that differences in such standards across franchisees be objective and verifiable. He shows that as a result of informal agreements, franchisors may dictate softer standards to small franchisees even when the formal contract gives them authority to impose uniform standards—that is, when flexible adaptation is not guaranteed by the formal contract. Using data on the formal allocation of authority over service standards in automobile distribution agreements for 19 car brands in Italy, and on the ex post adaptation of service standards for 10 car brands in one year, Zanarone (2013) finds that,
consistent with the theory, when the formal franchise contract assigns to car manufacturers the authority to modify standards, manufacturers set milder standards for smaller dealers ex post, even when the standards pertain to aesthetic features of dealers’ outlets and hence are unrelated to sales volumes.\footnote{Zanarone (2013) also provides anecdotal evidence that some car manufacturers selectively tolerate non-compliance by dealers when they have formal authority, and use discretionary or non-monetary payments to reward compliance when they lack formal authority.} While this evidence is descriptive due to the paucity of data, it suggests a potentially replicable strategy for comparing formal contracts and outcomes, and hence for testing H6, in future works—namely, to look at whether a franchisor’s ex post decision on the characteristics and service standards of franchisees’ outlets, which are recorded in operating manuals, are or not specified in the formal franchise contract.

Finally, Camuffo et al. (2007) study the vertical relations between two A/C manufacturers in Italy and a sample of long-term suppliers. They observe that formal contracts between manufacturers and suppliers rarely include risk-sharing clauses, and yet changes in supplier costs are not followed by equivalent changes in their profits, which is consistent with informally agreed risk-sharing.

\textit{H7: Formal contracts vary systematically with the discount rate}

To test this hypothesis, we need variations in observed formal contracts and, as in the case of hypotheses H1 and H4, variation in the contracting parties’ intertemporal discount rate. Moreover, and in the spirit of what discussed for H5 and H6, we also need to rely on specific theories that predict how formal contracts and the discount rate should co-vary in the relationship under study.
Corts and Singh (2004) study the choice between turnkey contracts (akin to fixed-price) and day-rate contracts (akin to cost-plus) in offshore oil drilling. Turnkey contracts provide drillers with stronger incentives to cut costs than day-rate contracts, but are also more rigid, and hence costlier to renegotiate when project specifications need to be changed. Using a sample of 1476 drilling projects, and an instrumental variable approach to control for the endogenous choice of drillers, Corts and Singh (2004) find that, all else equal, projects are less likely to be governed by a turnkey contract when the oil company and the driller have worked together in the past. They interpret this result as evidence that informal self-enforcing agreements and formal incentive contracts (i.e., turnkey contracts) are substitutes, rather than complements.

The negative association between repeated interactions and turnkey contracts is consistent with hypothesis H7. While Corts and Singh (2004) cannot directly measure the discount rate, insofar as past interactions also imply a higher likelihood of future interactions, oil company and driller pairs that have interacted in the past can more easily enforce informal contracts (that is, they have a lower discount rate). The fact that repeated interactions favor turnkey over day-rate contracts would then suggest that drillers are more tempted to renegotiate aggressively than to inflate costs, so that day-rate is the contract form that minimizes the driller’s aggregate reneging temptation.

Antras and Foley (2014), mentioned above, find that the frozen chicken exporter is more likely to use open account contract terms, as opposed to cash-in-advance terms,

---

22 To control for the endogeneity of driller’s choice, Corts and Singh (2004) use characteristics of drillers that had drilling rigs near a well’s location before the drilling contract was assigned as instruments for the characteristics of the contracted driller. Their identification strategy relies on the fact that rigs are costly to move, so that drilling projects are likely to be assigned to drillers who are already operating in the project’s location.

23 Similar results are obtained by Kalnins and Mayer (2004) in a study of IT services procurement contracts.
with his importers as the number and value of accumulated transactions grows, which is consistent with H7. If one interprets past interactions as a proxy for future ones, as in Corts and Singh (2004), this evidence suggests that contracts that are less likely to be enforced in court (i.e., open account contracts) are chosen only when the parties value their relationship.

Finally, Shi and Susarla (2011) study how reputation and the expectation of future interactions affect the choice between formal fixed-price and cost-plus contracts in IT outsourcing projects. They differ from Corts and Singh (2004) in that they attempt to identify and separate future interactions from past ones. Doing so is important, because even if past interactions predict future ones, their effect on contract choice may be at least in part due to reputation development and learning of a counterpart’s type, rather than informal enforcement. However, separate identification is not easy, and Shi and Susarla (2011) end up using contract renewal clauses as a proxy for future interactions, and arbitration and exclusivity clauses as instruments to address the endogeneity of renewal clauses. Despite the endogeneity of both proxy and instruments, their findings are qualitatively consistent with hypothesis H7 and evidence in Corts and Singh (2004). In particular, Shi and Susarla (2011) find that all else equal, clients are more likely to use cost-plus contracts when they expect their relationship with the vendor to last longer.

**H8: The relation between the costs of formal contracting and performance is negative when the fallback option is termination, and ambiguous otherwise**

Together with H3, we discuss more thoroughly hypothesis H8 later in this paper, in the context of potential opportunities for future research. Once again, the necessary conditions to test this hypothesis are rather demanding, as the researchers would need
exogenous measures for the costs of formal contracting and performance, as well as certainty on termination being the only fallback option available to contractual parties.

4. Back to theory: Informal contracting in networks

While we have focused so far on a bilateral contractual relationship, contracts are often multilateral. For instance, an employer typically has several employees, and a manufacturer deals with many suppliers and distributors. In a multilateral relationship, a party may have stronger incentives to honor an informal contract if her counterparts form a network—in the sense that they can credibly threaten to jointly terminate their informal contracts with the focal party after she breaches his informal contract with any one of them. To incorporate networks into our framework in the simplest possible way, we adapt Levin’s (2002) model of multilateral employment contracts. We focus on pure informal contracts, which are notation-wise simpler. The results immediately extend to the case where formal and informal contracts are used together.

Suppose B contracts with M counterparts, A₁, ..., Aₘ, and let \( \pi_{Bi}(a_i) \) and \( \pi_{Ai}(a_i) \) be the payoffs of B and counterpart Aᵢ under an informal contract, respectively. Also, let \( \pi_{Bi}^{sp} \) and \( \pi_{Ai}^{sp} \) be B’s and agent Aᵢ’s payoffs once their relationship terminates. Finally, let \( S_i(a_i) = \pi_{Bi}(a_i) + \pi_{Ai}(a_i) \) and \( S_i^{sp} = \pi_{Bi}^{sp} + \pi_{Ai}^{sp} \) be B’s and Aᵢ’s joint surpluses inside and outside the relationship, respectively. Given all the definitions and assumptions from the baseline model, when his counterparts form a network, B offers the action plan \( (a_1^n, ..., a_m^n) \), which solves

\[
\text{max} \sum_i \left[ S_i(a_i) - \pi_{Ai}^{sp} \right], \quad \text{subject to}

\sum_i S_i(a_i) \geq \sum_i S_i^{sp} + \tau \sum_i \left[ \pi_{Ai}(a_i^0) - \pi_{Ai}(a_i) \right].
\]
Consider now the case where there is no network, so that B enters purely bilateral informal contracts with each counterpart. In that case, B offers the action plan \( (a_1^b, \ldots, a_m^b) \), which solves
\[
\max_{(a_1, \ldots, a_m) \in \mathcal{A}} \sum_i \left[ S_i(a_i) - \pi_{ii}^{sp} \right],
\]
subject to
\[
S_i(a_i) \geq S_i^{sp} + r \left[ \pi_{ii}^a(a_i^b) - \pi_{ii}^a(a_i) \right] \quad \text{for every } i.
\] (15)

Clearly, the network condition (14) is necessary for the M bilateral conditions in (15), but not sufficient—that is, it could be that even though (14) holds, condition (15) does not hold for some of B’s counterparts. Hence, a given action plan can be more easily enforced in the presence of a network and, consequently, having a network is (weakly) efficient. Intuitively, this is the case because, if B breaches his informal contract with A_i and B’s counterparts form a network, B does not only lose his relationship with A_i, but also his relationship with all the other counterparts. Hence, B’s post-breach payoff is not \( S_i^{sp} + \sum_{j \neq i} S_j(a_j) \), as in the case of bilateral contracts, but \( \sum_i S_i^{sp} (a_i) \), which is smaller.

The above results are summarized by the following

**Proposition 3:** Suppose B has multiple counterparts. Then, informal contracts perform better when B’s counterparts form a network.

5. Inferences from selected empirical studies on informal contracting in networks

In this section, we discuss empirical evidence that relates to our model’s predictions on informal contracting in networks.
H9: Performance improves when the contracting parties belong to a network

There is a large literature on the impact of social and ethnic networks on the access to capital in developing economies (e.g., Banerjee and Duflo 2002; Fafchamps 2002). Here we focus on studies where the enforcement effect of networks is separated from their informational and reputational effects.

Jackson and Schneider (2011) use cross-sectional evidence to show that within-network drivers (i.e., those who lease their taxi from owners from the same country) are less likely to have accidents, be convicted for driving violations, or receive summonses. Furthermore, the authors also show that drivers who become “in-network” between 2005 and 2007 improve their driving behavior. Because selection into the network is a pervasive problem when estimating the effect of networks, the authors complement their cross-sectional evidence with a number of imperfect and yet informative empirical tests. On the one hand, they aggregate data at the network level and show that networks with higher percentages of in-network drivers present lower levels of driving violations. On the other hand, given the lack of exogenous variation, they instrument for in-network driving with the number of taxi owners from the same country living in the same neighborhood as the driver, and find consistent results. The regression coefficients in the three latter empirical strategies is three times as large as the one obtained in the original cross-sectional regression. Aside from concluding that social networks reduce moral hazard and improve performance, the significant change in magnitude is consistent with negative selection into the networks.

A second paper that provides evidence consistent with H9 is Kalnins and Chung (2007). They provide purely cross-sectional evidence that unbranded hotels in Texas are
more likely to survive when they are located close to branded hotels run by managers from the same ethnic network as the unbranded hotel’s manager. Because the cross-sectional evidence offered in this paper is at best suggestive, the authors complement their analysis with qualitative and anecdotal evidence from inside interviews with local hotel managers. Contrary to Jackson and Schneider (2011), this paper is silent about the endogeneity of entry, in-network selection, and hotel characteristics.24

6. Future work

Tables 1 and 2 below provide a snapshot of the relation between the theoretical predictions from our model and the empirical papers that we have discussed. Using these tables as reference, we provide guidance on future research opportunities in this final section. First, we discuss in greater detail the two predictions that remain untested, and second, we discuss how data sets previously used in other studies may be exploited to further test our illustrative model of informal contracts.

TABLE 1 HERE

TABLE 2 HERE

24 Gil and Hartman (2011) does not directly test hypothesis H9, but it does provide evidence on the impact of networks on outsourcing decisions in Korean drycleaners in Los Angeles. If they had provided evidence on quality (an accepted measure of performance in that industry) or timeliness, their findings would have spoken directly to H9.
6.1. Untested H3 and H8: Non-monotonic impact of formal contracts on informal ones

As discussed above, two hypotheses out of the nine presented above, H3 and H8, remain untested or without clear supporting evidence.

H3: The relation between the costs of formal contracts and the use of purely informal ones is negative when the fallback option is termination, and ambiguous otherwise

H8: The relation between the costs of formal contracting and performance is negative when the fallback option is termination, and ambiguous otherwise

These predictions are important because they directly relate to the interaction and potential complementarity between formal and informal contracts, which is emphasized by prominent theoretical models (e.g., Klein 2000, Baker et al. 1994). While we could not find any paper providing evidence that is related to H3, under very stringent assumptions, Johnson et al. (2002) may provide suggestive evidence in favor of H8. In a nutshell, they find that in European countries from the former Soviet block, trade credit (performance) increased when courts were believed to enforce contracts. However, this relationship was weaker when customers were locked into a bilateral monopoly relationship with the supplier and, therefore, the fallback option was less likely to be termination.

An attractive feature of Johnson et al. (2002) is the use of court quality as an inverse measure of the costs of informal contracting. This measure of contracting costs potentially allows for broad empirical studies that may include multiple contractual relationships and industries. On the other hand, our interpretation of Johnson et al. (2002) in terms of hypothesis H8 requires two strong assumptions that cannot be verified in the
data—namely, that locked-in pairs of suppliers and customers are less likely to terminate their relationship following breach of an informal agreement, and that non-locked pairs are less likely to use formal contracts after termination of the informal relationship. To overcome this weakness, future research that attempts to test hypotheses H3 and H8 may exploit exogenous changes in contract law that affect the feasibility of termination as a fallback option, such as termination laws in franchising (e.g., Brickley et al. 1991). If H3 and H8 are correct, one would expect court quality to improve performance, and to increase the use of formal contracts, less after laws that restrict contract termination are passed. Since court quality and the regulation of contract termination have been found to vary across countries and industries and within countries, the test we are suggesting may exploit variation across countries, within a country over time, or even within a country and year but across different industries.25

An alternative strategy for testing H3 and H8 would be to measure the enforcement cost $f(\cdot)$ alone (for instance, by looking at variations in court quality as discussed above). If enforcement costs can be accurately measured, increases in $f(\cdot)$ reduce the surplus under pure formal contracting, $S^p$, but do not affect the post-termination surplus, $S$, allowing the researcher to identify variation in the fallback option. Because in our model, increases in $f(\cdot)$ raise the likelihood that the fallback option is termination—formally, that $S^{sp} = S$, the effect of an increase in enforcement costs, $f(\cdot)$, on performance under informal contracting (H8), and on the use of formal contracts (H3), should be non-monotonic at low enough levels of $f(\cdot)$, and negative at high levels of

\[\text{footnote}{25}\text{In the US, variations in court quality across States have been documented, among others, by Berkowitz and Clay (2006), and Choi et al. (2008).}\]
6.2. Unexploited opportunities from existing works

As mentioned in the introduction, the conventional wisdom is that there is limited empirical evidence on informal contracts due to a scarcity of data. In this section, we further challenge this view by briefly discussing three existing data sets that are potentially suitable for investigating informal contracts empirically, even though such tests have not been pursued by the authors. In our view, these data represent an opportunity for follow-up studies as well as a reminder that, contrary to the conventional wisdom, novel empirical evidence based on conventional data may be feasible.

Bolton et al. (2014) use the recent financial crisis as an exogenous shock in the provision of credit from Italian banks to firms. They identify banks in a given credit transaction as “relationship banks” (as opposed to “transactional banks”) when their headquarters are located in the same region as the headquarters of the firm they are lending to. Their evidence shows that the difference in interest rates before and after the crisis is large, but less so for relationship banks, presumably due to information differentials among the two types of bank. Using the data in Bolton et al. (2014), one could examine which relationships are more likely to end after the crisis. Similarly to Macchiavello and Morjaria (2015), the difference between the interest rates charged by relationship and transactional banks before the crisis may be interpreted as a lower bound for the discounted value of the credit relationship to firms, and therefore should be
positively correlated with firms’ survival after the crisis and other performance measures, as predicted by our model.\textsuperscript{26}

DeVaro et al. (2015) present a model of CEO compensation where current shocks in firm performance predict future shocks. According to their model, and consistent with H6 in our general model, CEOs’ formal bonuses should be the same during good and bad states, whereas informal bonuses should be larger during good states, because present good states correlate with future ones, and hence imply a greater discounted value from the relationship. For the same reason, the difference in bonuses during good and bad states should increase in state persistence. While the theoretical predictions in DeVaro et al. (2015) are sharp, they do not have the data required to directly test them — mainly, information on whether CEOs’ bonuses are formal or informal. As an indirect test, they regress bonus amounts on a state variable (growth versus no growth), a firm-level measure of state persistence, and other firm performance characteristics, and find that, consistent with their model, growing firms pay larger bonuses, and the more so the greater the degree of state persistence. A clear research opportunity here would be to complement the data on bonus levels and firm performance in DeVaro et al. (2015) with data on whether CEO bonuses are formally or informally contracted, such as those collected by Gillan et al. (2009). By doing so, it would be possible to directly test for the differential effect of shocks on formal and informal bonuses.

Finally, Forbes and Lederman (2009) study vertical integration in the US airline industry, and argue that managing flights outsourced to regional airlines requires

\textsuperscript{26} A potential caveat, similar to that discussed for Antras and Foley (2014), is that relationship banks may charge lower interest rates to firms they have learned to be more creditworthy, so low interest rates may proxy for superior creditworthiness rather than more valuable relationships. Indeed, Bolton et al. (2014) model “relationship banks” as those that have superior information on their clients.
especially intensive coordination with the national airline in routes characterized by bad weather conditions and hub connections. Consistent with their hypothesis, they find that national airlines are less likely to outsource flights in bad-weather routes and routes that involve a major hub. A potential research opportunity in this context and data (publicly available) is to study within the subsample of outsourced routes and flights whether valuable informal relationships with major airlines have an impact on the performance of regional airlines. This industry has not been short of exogenous changes in the tightness of reneging constraints lately, as national airlines have had to cut back on flights and presence in different airports in response to shocks in oil prices, demand, and the like. When making such decisions, major airlines presumably chose to cut outsourced flights over in-house flights and when doing so they cut those operated by regional airlines with which they had a weaker relationship. Consistent with this strategy and the methodology proposed by Macchiavello and Morjaria (2015), Gil et al. (2015) use weather conditions on a route to proxy for the reneging temptation of regional airlines, and hence for the informal relationship’s discounted value on that route. They find that outsourced bad weather routes are more likely to survive the 2011 9/11 and the 2008 Lehman Brothers’ financial crisis shocks.

In our view, the above discussion illustrates how more data than commonly thought are available to empirically investigate informal contracts. We believe that those or similar data, combined with specialized theories of the underlying contractual setting, can and should lead to novel empirical findings.
7. Concluding remarks

In this paper we investigate whether informal contracts behave as predicted by economic theory, and how further evidence on their determinants and consequences can be produced. To achieve these goals, we obtain testable predictions from a parsimonious model where formal and informal contracts are combined to trade off enforcement and contracting costs. We then discuss selected empirical works to shed light on the model’s empirical relevance. Overall, we find that the evidence is broadly consistent with the theoretical predictions. At the same time, we notice that the empirical research on informal contracts has substantial room to grow and presents plenty of opportunities. We have hopefully proved this point by showing how current tests of the theoretical predictions can be improved, which predictions are left untested, and how current data sets and papers could be exploited for that purpose. We hope our paper will foster future research that will test and feed our understanding of informal contracts and of their interaction with formal ones.

References


Table 1. Allocation of Evidence to Predictions: determinants of contract enforcement mode and outcomes

<table>
<thead>
<tr>
<th>Decision to be explained</th>
<th>Transaction characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Discount Rate/PDV &amp; Fallback Option</td>
</tr>
</tbody>
</table>

*Illustrative evidence on H8.
Table 2. Allocation of Evidence to Predictions: Formal Contract terms in the presence of informal contracts

<table>
<thead>
<tr>
<th>Hypotheses</th>
<th>Related Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_5$: Parties write formal contracts that would be suboptimal if enforced</td>
<td>Zanarone (2009)</td>
</tr>
<tr>
<td></td>
<td>Antras and Foley (2014)</td>
</tr>
<tr>
<td>$H_6$: Observed outcomes differ from formally contracted ones</td>
<td>Zanarone (2013)</td>
</tr>
<tr>
<td></td>
<td>Camuffo et al (2007)</td>
</tr>
<tr>
<td></td>
<td>Gil (2013)</td>
</tr>
<tr>
<td></td>
<td>Barron, Gibbons, Gil and Murphy (2015)</td>
</tr>
<tr>
<td>$H_7$: Formal Contracts vary systematically with the discount rate</td>
<td>Corts and Singh (2004)</td>
</tr>
<tr>
<td></td>
<td>Antras and Foley (2014)</td>
</tr>
<tr>
<td></td>
<td>Shi and Susarla (2011)</td>
</tr>
</tbody>
</table>