On the Determinants and Consequences of Informal Contracting*

Ricard Gil Giorgio Zanarone

Johns Hopkins Carey Business School CUNEF

June 2017

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 contracting. First, we present an illustrative model that organizes key predictions from the theoretical literature. Next, we discuss selected empirical works that shed light on the model's relevance and testability. We conclude by discussing strategies for testing theoretical predictions for which conclusive evidence is still missing.

Keywords: Informal Contracts, Relational Contracts, Enforcement, 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 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); firms include in their employees' compensation both formal bonuses based on objective performance evaluation and informal ones based on subjective evaluation (Fast and Berg, 1975); formally centralized companies and governments motivate employees (Foss 2003) and bureaucrats (Xu 2011) by informally promising to delegate authority; and large corporations such as General Motors rely on informal, internally enforced routines to manage their workers and suppliers (Helper and Henderson 2014).

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), their dynamic patterns (Levin 2002, 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

[1] For comprehensive reviews of the empirical evidence on formal 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".

(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).[3]

Despite the ubiquitousness of informal contracts as documented by case studies, there is scarce systematic evidence on whether existing economic theories correctly predict their determinants and consequences.[4] 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 economic model that illustrates how informal contracts can be optimally combined with formal contracts to provide incentives, and whose predictions can be easily generalized to contractual problems other than moral hazard, such as adaptation to change (e.g., Williamson 1991; Baker et al. 2011). The model generates three testable predictions. The first is that formal contracts are more

[3] Up-to-date reviews of the theoretical literature on informal contracts can be found in MacLeod (2007) and Malcomson (2013). See Gil and Zanarone (2017) for an integrated review of the theoretical and empirical literatures that combines contributions from law and economics.

[4] For a discussion of early evidence and case studies consistent with the presence of informal contracts, see Shelanski and Klein (1995).

likely to be used when self-enforcement of informal agreements is imperfect—that is, when a low discount factor or an attractive outside option limit the parties' interest in preserving their ongoing collaboration. The second, and related, prediction is that a high discount factor, or an unattractive outside option, increase the parties' ability to generate surplus under an informal contract. The third prediction is that when informal contracts are available, parties will not perform as dictated by formal contract terms even though they may sign such formal contracts to facilitate self-enforcement.

In the second part of the paper, we discuss recent empirical works testing the theoretical predictions from the model. By doing this, we are not aiming to comprehensively review the empirical literature on informal contracting, but rather 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 offer systematic econometric evidence based on strong identification and accurate measurement of variables and/or promising datasets. Accordingly, we do not discuss purely anecdotal evidence and case studies.

The empirical works we discuss span a variety of industries and contractual settings, including CEO compensation (Gillan et al. 2009), movie exhibition (Gil 2013; Barron et al. 2017), procurement (Johnson et al. 2002; Corts and Singh 2004; Gil and Marion 2013; Chassang and Ortner 2017), supply chains (Macchiavello and Miquel-Florensa 2017), international trade (Macchiavello and Morjaria 2015), transportation (Jackson and Snyder 2011; Gil et al. 2017), and politics (Shelef 2013).

Overall, we find combined support for most theoretical predictions from the model, and we observe that significant progress has been made over time at measuring key

determinants of informal contracting such as the parties' discount factor and fallback option. At the same time we notice that a few theoretical questions that naturally emerge from our basic model—mainly, whether formal and informal contracts are substitutes or complements and how the presence of informal contracts affects the design of formal ones—are still awaiting clear supportive evidence. In the final part of the paper, we briefly discuss empirical strategies that may allow researchers to investigate these untested theoretical questions. We also notice that more complex models of informal contracting than the one analyzed here, such as those in Halac (2012) and Li and Matoushek (2013), generate additional predictions on the evolution of contractual relationships over time that future empirical researchers may attempt to test. We discuss some possible strategies for testing those dynamic informal contracting models in a companion paper (Gil and Zanarone 2016).

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 contracting. Section 4 discusses opportunities for future research. Section 5 concludes.

2. An illustrative model

In this section, we adapt the agency model in Gibbons et al. (2012) to derive testable implications that will help us organize the empirical studies on informal contracting discussed in section 3.[5] We also use the model as a basis for our discussion of future research opportunities in section 4.

[5] Other related models are Holmstrom and Milgrom (1991), Baker (1992, 2002), and Gibbons (2005).

2.1. Model setup

There are a principal and an agent, denoted as P and A, who are both risk-neutral, live forever, and discount next-period payoffs by the common factor $\delta \in [0,1]$. Time evolves in discrete periods.

At the beginning of any given period, P offers A an incentive contract (hereafter simply a *contract*), $w(y,p) \equiv s + bp + By$, where s is a fixed salary, b and B are outcome-contingent bonuses, and p and y are outcomes to be defined momentarily. If A rejects the offer, P and A take their outside options, receiving payoffs $\underline{\pi} \geq 0$ and $\underline{u} \geq 0$, respectively, such that the joint outside option is $\underline{v} \equiv \underline{\pi} + \underline{u}$. If A accepts the offer, he simultaneously chooses a *productive action*, $a_1 \in \mathbb{R}^+$, and an *unproductive action*, $a_2 \in \mathbb{R}^+$, at cost $C(a_1, a_2) \equiv \frac{k}{2}(a_1^2 + a_2^2)$. After A chooses the actions, the two outcomes, $y \in \{0,1\}$ and $p \in \{0,1\}$, are realized and jointly observed by P and A. Outcome y is the *output* that A generates for P, which is only affected by the productive action: $y = 1$ with probability a_1, and $y = 0$ with probability $1 - a_1$. Outcome p is A's *measured performance*, which may be affected by both the productive and the unproductive action: $p = 1$ with probability $a_1 cos\theta + a_2 sin\theta$, and $p = 0$ with probability $1 - a_1 cos\theta - a_2 sin\theta$. After the outcomes are realized the contract is enforced, such that P receives payoff $y - w(y,p)$, while A receives payoff $w(y,p) - C(a_1, a_2)$.

Two types of transaction costs drive the model's results: 1) potential *misalignment* between measured performance, p, and output, y; and 2) *imperfect enforceability* of the incentive contract, $w(y,p)$. We briefly discuss both frictions below.

Misalignment between p and y is measured by the angle $\theta \in [0,90°]$. At $\theta = 0$, $cos\theta = 1$, and $sin\theta = 0$. As a result, p and y are perfectly aligned, in the sense that their realizations follow the same probability distribution. As θ grows, $cos\theta$ decreases while $sin\theta$ increases. Consequently, p and y become progressively misaligned, as the probability distribution of measured performance, p, depends more on the unproductive task, a_2, than on the productive task, a_1. At the limit, at $\theta = 90°$, $cos\theta = 0$ and $sin\theta = 1$, in which case p and y are completely orthogonal: the probability distribution of y only depends on the productive task, a_1, whereas the probability distribution of p only depends on the unproductive task, a_2.

Regarding enforceability, we assume that while a court can observe A's measured performance, p, it cannot observe the output, y. This implies that only the part of A's variable compensation that depends on measured performance, bp, can be part of a *formal contract*. P may nevertheless agree with A on a bonus contingent on the observable-but-non-verifiable output y, $B > 0$. However, because a court cannot enforce such *informal contract*, P's promise to pay $B > 0$ must be self-enforcing.

Given the above definitions and assumptions, we analyze below the three cases where P incentivizes A through a formal contract (section 2.3), an informal contract (section 2.4), or a combination of the two (section 2.5).

2.2. The first best

Before analyzing formal and informal contracts, we characterize as a benchmark the actions that maximize the joint surplus that P and A may produce if they enter a

relationship—that is, the difference between P's expected profit and A's cost of taking the actions:

$$max_{a_1,a_2}\left\{a_1 - \frac{k}{2}(a_1^2 + a_1^2)\right\}. \quad (1)$$

The solution is given by $a_1^* \equiv \frac{1}{k}$ and $a_2^* \equiv 0$, and generates joint surplus $V^* \equiv \frac{1}{2k}$. We assume that $V^* > \underline{v}$, so that it is efficient for P and A to enter a relationship, and actions a_1^* and a_2^* achieve the first best.

2.3. Limit case 1: purely formal contracts

Suppose that at the beginning of each period P is constrained to offer a purely formal contract to A: $B = 0$, and $w(y,p) = w(p) \equiv s + bp$, to A. If A accepts the contract, she chooses the two actions to solve:

$$max_{a_1,a_2}\left\{s + b(a_1 cos\theta + a_2 sin\theta) - \frac{k}{2}(a_1^2 + a_1^2)\right\}. \quad (2)$$

The solution is given by $a_1^F(b) \equiv \frac{b}{k}cos\theta$, and $a_2^F(b) \equiv \frac{b}{k}sin\theta$. Clearly, an increase in the formal bonus b increases A's incentive to produce, but also her incentive to engage in the unproductive activity, and the more so the greater the misalignment parameter θ. Hence, the overall effect of the bonus on P's profit is ambiguous.

Anticipating A's responses to the formal contract, P chooses s and b, ex ante, to maximize profits, subject to A's participation constraint. Thus, P offers the minimum fixed salary that induces A to accept the contract:

$$s^F \equiv \underline{u} + C\left(a_1^F(b), a_2^F(b)\right) - b(a_1^F(b)cos\theta + a_2^F(b)sin\theta). \quad (3)$$

After substituting s^F into (2), P's problem simplifies to that of choosing the formal bonus, b, to solve:

$$max_b\{a_1^F(b) - C(a_1^F(b), a_2^F(b)) - \underline{u}\}. \quad (4)$$

The solution is given by $b^F \equiv cos\theta$, which generates joint surplus $V^F \equiv \frac{cos^2\theta}{2k}$. In the absence of transaction costs—that is, if $\theta = 0$ and hence the verifiable performance measure p is perfectly aligned with the output y—the optimal formal contract has strong incentives, $b^F = 1$, and induces A to take the first best actions. As θ grows, $b^F < 1$, and the optimal formal contract generates lower productive effort, and higher unproductive effort, than in the first best. At the limit, when y and p are orthogonal, $b^F = 0$, and the optimal formal contract solely consists of a fixed salary, $s^F \equiv \underline{u}$, so that A takes no actions. In that case, P is better off not offering a contract to A and taking his outside option.

2.4. Limit case 2: purely informal contracts

Suppose now that at the beginning of each period, P is constrained to offer a purely informal output-contingent contract to A: $b = 0$, and $w(y,p) = w(y) \equiv s + By$. If A accepts the contract, she chooses the two actions to solve:

$$max_{a_1,a_2}\left\{s + Ba_1 - \frac{k}{2}(a_1^2 + a_1^2)\right\}. \quad (5)$$

The solution is given by $a_1^I(B) \equiv \frac{B}{k}$, and $a_2^I(B) \equiv 0$. In this case, A has no incentive to take the unproductive action, a_2, because she is evaluated on output, which solely depends on a_1. Thus, P could achieve the first best by offering a high informal bonus, $B = 1$. However, an informal incentive contract cannot be enforced in court. Thus, the

contract must be designed in such a way that P has an incentive to pay the bonus once A's productive action is sunk, or else the bonus promise would not be credible to A.

As standard in the literature, we model the informal contract as a subgame perfect equilibrium of the infinitely repeated game, such that on the equilibrium path, A takes action $a_1^I(B)$, and P pays bonus B, in every period, and if P fails to pay, A terminates the relationship forever after.[6] We assume that if the relationship terminates, P and A take their outside options in all subsequent periods, as in Levin (2003). In section 4, on future work, we discuss additional testable predictions that may be obtained by relaxing this assumption.

After imposing the minimum fixed salary consistent with A's participation constraint, $s^I \equiv \underline{u} + C(a_1^I(B), 0) - Ba_1^I(B)$, P's problem simplifies to choosing the informal bonus B to maximize profits, subject to the constraint that the bonus be credible (hereafter, enforcement constraint):

$$max_B\{[V^I(B) \equiv a_1^I(B) - C(a_1^I(B), 0)] - \underline{u}\},$$

subject to

$$\frac{\delta}{1-\delta}\left[V^I(B) - \underline{v}\right] \geq B. \qquad (6)$$

Let $\underline{\delta}(\underline{v}) \equiv \frac{2k}{1+2k(1-\underline{v})}$, increasing in the fallback option $\underline{v}$, be the critical discount factor such that (6) holds with equality at $B = 1$. Then, for $\delta \geq \underline{\delta}(\underline{v})$, the enforcement constraint is slack, the optimal bonus is $B = 1$, and the informal contract achieves the

[6] Given the absence of liquidity constraints, restricting attention to stationary informal contracts where equilibrium behavior is the same in every period is without loss of generality (MacLeod and Malcomson 1989; Levin 2003).

first best. In contrast, for $\delta < \underline{\delta}(\underline{v})$, the enforcement constraint is binding, the optimal informal bonus is smaller than one, and the first best cannot be achieved. In that case, the optimal informal bonus is given by the largest real positive bonus such that (6) holds with equality.

It follows from (6) that given the lack of court enforcement, joint surplus under an informal contract (weakly) increases in the value of future interactions between P and A, as measured by δ, and (weakly) decreases in the fallback option following termination of the relationship, $\underline{v}$.

2.5. General case: Informal contracts backed by formal contracts

Having studied the two special cases where P is constrained to use either purely formal or purely informal contracts, we now turn to the more general case where P may combine formal and informal contracts: $w(y,p) \equiv s + bp + By$.

If A accepts the contract offered by P, she chooses the two actions to solve:

$$max_{a_1,a_2} \left\{ s + b(a_1 cos\theta + a_2 sin\theta) + Ba_1 - \frac{k}{2}(a_1^2 + a_1^2) \right\}. \qquad (7)$$

The solution is given by $a_1^{IF}(b,B) \equiv \frac{bcos\theta + B}{k}$, and $a_2^{IF}(b,B) \equiv a_2^F(b) = \frac{bsin\theta}{k}$. Since the formal part of A's compensation only depends on measured performance, p, A chooses the unproductive action as he would under a purely formal contract. However, A's choice of productive action is now affected by both the formal and the informal bonuses, b and B.

Anticipating A's response, for a given level of the informal bonus B, P chooses the formal bonus b to maximize profits:

$$max_b\{a_1^{IF}(b,B) - C\left(a_1^{IF}(b,B), a_2^F(b)\right) - \underline{u}\}. \quad (8)$$

After substituting for A's response functions and solving P's program, the optimal formal contract, as a function of the informal contract, is given by:

$$b(B) \equiv (1-B)cos\theta. \quad (9)$$

If output and measured performance are orthogonal ($\theta = 90°$), or if $B = 1$, so that the first best can be achieved by conditioning A's pay on output alone, P does not benefit from including a distortionary formal term in the contract: $b(1) = (1-B)cos(90°) = 0$. Conversely, if $\theta \in [0, 90°)$ and $B < 1$, P benefits from including a formal term so long as the marginal increase in expected output due to A's enhanced productive effort outweighs the marginal compensation cost: $b(B) > 0$. The smaller θ, the smaller A's unproductive effort and hence the smaller the marginal increase in compensation cost. Therefore, $b(B)$ decreases in θ.

Finally, given $b(B)$, P chooses the informal bonus B to maximize profits, subject to the self-enforcement constraint:

$$max_B \left\{\left[V^{IF}(B) \equiv a_1^{IF}(b(B),B) - C\left(a_1^{IF}(b(B),B), a_2^F\left(b(B)\right)\right)\right] - \underline{u}\right\},$$

subject to:

$$\frac{\delta}{1-\delta}\left[V^{IF}(B) - \underline{v}\right] \geq B. \quad (10)$$

On the one hand, increasing the informal bonus B allows P to elicit productive effort at lower compensation cost, as P can rely less on the distortionary contract term b. This increases P's payoff as well as the right-hand side the self-enforcement constraint, (10).

On the other hand, raising B increases P's temptation to renege on the informal part of the agreement—that is, the left-hand side of the self-enforcement constraint.

At $B = 1$, self-enforcement constraint (10) is the same as constraint (6). Thus, if $\delta \geq \underline{\delta}(\underline{v})$, a purely informal contract achieves the first best: $B^{IF} = 1$, and $b^{IF} = 0$. If $\delta < \underline{\delta}(\underline{v})$ the optimal informal bonus, B^{IF}, is smaller than one, and is given by the largest real positive value of B such that (10) holds with equality.

2.6. Testable predictions

A number of testable predictions follow from the model developed in the previous sections. First, whether the contract between P and A includes formal terms or not depends on δ, the discount factor, and on $\underline{v}$, the fallback option. More specifically, it follows directly from (9) that $b^{IF} = 0$ for $\delta \geq \underline{\delta}(\underline{v})$, whereas $b^{IF} \geq 0$ for $\delta < \underline{\delta}(\underline{v})$, with the inequality holding strictly for $\theta < 90°$. To turn this prediction into an empirical hypothesis, let ε be a zero-mean random component of the discount rate (with δ being the fixed, deterministic component), which is realized before P and A start their relationship and is unobservable to the econometrician. Then, assuming a potentially useful but imperfect performance measure, $\theta \in (0{,}90°)$, the probability that a formal term is included in the contract is $Pr\left(\varepsilon + \delta < \underline{\delta}(\underline{v})\right) = Pr\left(\varepsilon < \underline{\delta}(\underline{v}) - \delta\right)$, decreasing in δ and increasing in $\underline{v}$.

H1*: The probability that a formal term is included in the contract between P and A decreases in the discount factor,* δ*, and increases in the fallback option following termination of the relationship,* $\underline{v}$*.*

A second, and related, prediction of the model, which follows directly from the self-enforcement condition, (10), is that the equilibrium joint surplus, $V^{IF}(B^{IF})$, (weakly) increases in δ, and (weakly) decreases in $\underline{v}$.

H2*: The joint surplus from the contractual relationship, $V^{IF}(B^{IF})$, (weakly) increases in the discount factor, δ, and (weakly) decreases in the fallback option, $\underline{v}$.*

Finally, the model predicts that when informal contracting is available, A's response to formal contract terms differs from the response that would be observed if P and A "played by the book", that is, if only formal contracts, but not informal ones, could be enforced. To see this point, suppose P and A agree on formal bonus $\tilde{b} \in (0,1)$. If $\delta \approx 0$, condition (10) implies that no informal bonus can be enforced. Consequently, A's productive action is given by $a_1^{IF}(\tilde{b}, 0) = \frac{\tilde{b}cos\theta}{k}$. Suppose now that δ grows large enough. Then, condition (10) implies that holding $\tilde{b}$ constant, it is optimal for P to offer a positive informal bonus: $\tilde{B} > 0$. As a result, A's productive action becomes $a_1^{IF}(\tilde{b}, \tilde{B}) = \frac{\tilde{b}cos\theta+\tilde{B}}{k} > a_1^{IF}(\tilde{b}, 0)$.

H3*: Under a given formal contract, A's productive action when informal contracting is available differs in general from the action that A would take in the absence of informal contracting.*

Note that hypothesis H2 is sensitive to the static technology assumed in our model. If actions in one period affected future payoffs, surplus 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. Hence, some evidence that the production technology is static would be necessary to fully support that conclusion.[7]

2.7. Beyond agency: reinterpreting the model

We have derived our testable predictions from an illustrative *moral hazard* model where actions are privately observed, and contracts specify incentive pay. Another important contractual problem is *adaptation*—the difficulty of tailoring economic decisions to a changing environment (e.g., Williamson, 1991). In adaptation models, state-contingent actions are observable but formally non-contractible, and may be affected by a richer set of contract terms than incentive pay, including decision rights (Simon, 1951; Hart and Moore, 2008; Baker et al., 2011), asset ownership (Hart and Holmstrom, 2010), and pricing terms (Bajari and Tadelis, 2001; Hart and Moore, 2008).

It would be simple to obtain testable predictions similar to H1, H2, and H3, from an adaptation model (see, for instance, the "relational adaptation" section of Baker et al., 2011). Thus, while we do not formally model adaptation due to space limits, we encourage readers to think of the predictions from our illustrative model as applying more broadly to both moral hazard and adaptation settings, and to bear that in mind when reading our discussion of the empirical evidence in the next section.

[7] 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.

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 aim to draw lessons on how to measure key variables from the model that may provide guidance for future empirical studies.

3.1. Measuring variation in the discounted value of relationships

The present discounted value (PDV) of contractual relationships—that is, the left-hand side of self-enforcing condition (10) in our model—plays an important role in driving the three testable hypotheses, yet it is typically hard to observe. Empirical studies attempting to measure variation in the PDV have followed two different approaches. A first approach uses proxy variables for the discount factor δ that may vary either over time within a relationship (Gil and Marion 2013, Shelef 2013) or across relationships (Corts and Singh 2004, Gillan et al. 2009). A second approach estimates lower bounds for the PDV by measuring the parties' present temptation to renege on an informal contractual agreement—that is, the right-hand side of the self-enforcement condition in the model (Macchiavello and Morjaria, 2015; Barron et al., 2017; Gil et al., 2017).

Let us discuss next the different proxy measures used in the literature according to the two approaches described above. We discuss how these measures, as well as measures for the other parameters in our illustrative model, have been used to test hypotheses H1, H2, and H3, in sections 3.2 and beyond.

Measuring the discount factor

Earlier empirical studies of informal contracting have used cross-sectional variations in past interactions between the parties as a proxy for the discount factor. For instance, Corts and Singh (2004) use the number of projects on which a given oil company and a driller have worked together in the previous six months as a proxy for whether the two firms expect a long-term relationship, and hence are likely to interact in the future. A similar measure is used by Kalnins and Mayer (2004) in the context of IT outsourcing. Relatedly, Gillan et al. (2009) use an indicator for whether a CEO has been hired from outside the firm as a proxy for the expected duration of the employment relationship.

An obvious drawback of using past interactions as a proxy for the likelihood of future interactions is that the former may also be correlated with unobservable determinants of the quality of the match between parties such as learning, experience solving problems together and accumulated information on performance. Even in the absence of enforcement concerns, more experienced or informed partners may be less in need of formal contractual documents to coordinate their activities and may achieve superior performance. Thus, observed correlations between past interaction and the use of formal contracts or performance may be driven by endogeneity.

Cognizant of this problem, other studies have attempted to directly measure variations in the likelihood and value of future interactions over time. For instance, Gil and Marion (2013) use the number and value of projects auctioned by Caltrans (California Department of Transportation) over the next calendar year as a proxy for the likelihood and expected value of future interactions between contractors and subcontractors bidding for Caltrans procurement projects. In a study of relational contracting in US Congress and

State legislature campaigns, Shelef (2013) uses a candidate's age, and an dummy for whether the candidate is running for a last election term as mandated by State law, as alternative proxies for the likelihood of future interactions between the candidate and the interest groups contributing to her campaign.

The "lower bound" approach to measuring PDV

Macchiavello and Morjaria (2015) analyze a rich dataset of contracts between Kenyan rose exporters and buyers. Importantly, formal enforcement of those contract is prevented by the lack of effective courts in Kenya. They compute the exporters' temptation to renege on their informal agreements with buyers as the stipulated quantity times the spot market auction price of roses. Self-enforcement implies that the present discounted value of a given buyer-exporter dyad must be at least as large as the reneging temptation. Thus, while Macchiavello and Morjaria (2015) cannot measure the discounted value of future relationships for each buyer-exporter dyad, they can use their estimates of the reneging temptation as lower bounds for such value. To validate their interpretation of the reneging temptations as lower bounds for continuation value, they provide evidence that following an unforeseen increase in the spot market price, informal contracts between exporters and buyers proportionally reduce the stipulated quantity, such that the self-enforcement condition is restored.

Beyond the specific context analyzed by Macchiavello and Morjaria (2015), the lower bound approach relies on the fact that under an informal contractual relationship, high-temptation contracts must be assigned to partners with high PDV. Reneging temptations, unlike future interactions, can often be measured based on market and contractual data (such as the spot market for roses used by Macchiavello and Morjaria, 2015). Thus, this

indirect approach to measuring PDVs is replicable and potentially attractive. The main challenge one faces when using the lower bound approach to test the hypotheses from our illustrative model is the endogenous selection of "good" partners into high-temptation contracts, which may correlate with performance and the use of formal versus informal contracting. This challenge should be borne in mind when assessing the evidence discussed in the next sections.

3.2. Testing H1: low discount factor and high post-termination surplus increase the likelihood of using formal contracts

Ideally, to test this first theoretical prediction we would need a dependent variable indicating whether a formal contract exists, and measures of the intertemporal discount factor and the parties' payoffs outside the relationship that are exogenous—that is, uncorrelated with the quality of the match between the contracting parties (Ackerberg and Botticini, 2002).

We have discussed in section 3.1 the challenges empirical researchers face when trying to measure the PDV of contractual relationships. Measuring the formal or informal nature of contracts is also challenging. One strategy, followed by Gil (2013) in his study of movie exhibition agreements, is to look at internal company records, as companies have a clear interest in keeping track of whether their contractual agreements are legally enforceable. Another strategy, followed by Gillan et al. (2009), is to exploit legally mandated disclosure rules. Since Gillan et al. (2009) provide an empirical test of hypothesis H1, we discuss their measurement strategy in greater detail below.

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 the results in Gillan et al. (2009) are consistent with hypothesis H1. In particular, when the firm's prospects are uncertain (in terms of our model, when δ is high), promising a purely discretionary compensation to the CEO is not credible, because there is 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 ($\underline{v}$ in the model) 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.

While the results in Gillan et al. (2009) are consistent with H1, it must be reckoned that their discount factor measures are endogenous. The new CEO indicator, like any variable measuring past interactions (see discussion in section 3.1), may be associated to low experience and learning, whereas sales volatility and industry-wide stability in ownership and control may be positively correlated with firm/industry complexity and risk. In turn, unobserved low experience and high complexity and risk may be positively correlated with the use of more complex formal contracts for coordination purposes.

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,[8] 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.[9] They interpret this result as evidence that informal self-enforcing agreements and formal incentive contracts (i.e., turnkey contracts) are substitutes, rather than complements.

[8] 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.

[9] Similar results are obtained by Kalnins and Mayer (2004) in a study of IT services procurement contracts.

The negative association between repeated interactions and formal incentive contracts is consistent with hypothesis H1. 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 have a higher discount factor, and hence are less likely to need formal contract to replace or complement informal ones. As explained in section 3.1, however, their proxy for the discount factor is problematic because it may correlate with determinants of the goodness of fit between drillers and oil companies that may be the true drivers of contractual form.

DeVaro et al. (2017) present a model of CEO incentive pay where current shocks in firm performance are a good predictor of future shocks. In their model the CEO's effort only affects current output and shocks do not affect the CEO's effort productivity (that is, they reflect pure "luck"). Under such conditions, a purely formal incentive contract would pay equal bonuses to the CEO following a positive or a negative shock. In contrast, and consistent with our H1 hypothesis, informal bonuses should increase—potentially up to the first best level—following a positive shock, as the latter correlates with future positive shocks, and hence implies a greater PDV from the relationship. For the same reason, the difference between the informal bonuses following a positive and a negative shock should increase in state persistence.

While the theoretical predictions in DeVaro et al. (2017) 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 measures of firm state and state persistence. To construct these measures, they apply the Hodrick-Prescott filter to the time series of each firm's annual sales to disentangle its cyclical

component from the long term trend component. Then, they use the cyclical component as their state proxy, and the correlation between the cyclical sales component and its one-year lagged value as their firm-level proxy for state persistence. They find that, consistent with their model and with H1 in our model, firms pay larger bonuses following a good state, and the more so the greater the degree of state persistence.

A potential caveat when interpreting these results is that the authors cannot empirically verify two key assumptions in their model—namely, that CEO effort does not affect future output and shocks do not affect effort productivity. If both assumptions were violated, formal bonuses, not only informal ones, may optimally increase in the state and persistence variables. One way to address this concern, and perhaps an opportunity for future research, would be to complement the data on bonus levels and firm performance in DeVaro et al. (2017) 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 the theoretical prediction from DeVaro et al. (2017)—that is, that the correlation between bonuses and the state and state persistence variables be stronger in the subsample of firms that use informal contracts.

3.3. Testing H2: Realized performance increases in the discount factor and decreases in the post-termination surplus

To test this hypothesis, we need relationship-level measures of realized surplus and performance, the discount factor/PDV (see the discussion above in section 3.1), and the post-termination payoffs. Several recent papers provide supporting evidence for H2.

Performance and the discount factor

Gil and Marion (2013) tests 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 the subcontractors listed in their bids, and how the PDV of relationships may influence their bids. To proxy for the PDV, they use the number and value of projects auctioned over the next calendar year. This is an exogenous proxy insofar as projects are offered when they are needed, and not as a response to a particularly good match between contractor and subcontractor. The results show that consistent with our hypothesis H2, the higher the number and value of forthcoming projects—and hence the larger the PDV of the contractor-subcontractors relationship—the lower the bids posted by contractors, where low bids reflect more efficient contractor-subcontractor teams. 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.

Two main caveats should be borne in mind when interpreting the findings in Gil and Marion (2013). First, the paper cannot capture variations in the PDV among different subcontractors as changes in future projects are assumed to affect all contractors and subcontractors in a district in the same way. Second, their data is limited to Caltrans contracts and therefore cannot quantify differences in relationship value between contractors and subcontractors in procurement contracts at the city and county level (non-federal projects). These differences may drive the distribution in the quality of the match between contractor and subcontractor in a way that is correlated with the arrival of

Caltrans project to a district, and therefore undermine the assumption that changes in the PDV of contractor-subcontractor relationships within a Caltrans district are exogenous.

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 increase in the parties' discount factor (consistent with our hypothesis H2). 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 inverse measures for her time horizon/discount factor: (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 variations in campaign contributions and the age and last term variables, so one cannot fully control for the potentially confounding effect of the candidates' experience and skill depreciation over time. 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. 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.

Finally, Gil et al. (2017) study how US major airlines rely on informal agreements to induce their outsourced regional partners to exchange slots and adapt flight schedules in the face of adverse weather. They present a model where a regional airline's maximum temptation to renege on its relational adaptation contract with the major is the amount of

slots the regional would need to supply if landing slots were simultaneously reduced on all the routes in the outsourced regional network. Building on previous work by Forbes and Lederman (2009), they use aggregate bad weather conditions across routes in the network as a proxy for the maximum reneging temptation. Finally, they rely on the "lower bound" approach discussed above to argue that under an informal adaptation agreement between the major and the regional, the maximum reneging temptation must be at least as large as the PDV of the major-regional relationship. Consistent with hypothesis H2, Gil et al. (2017) show that following the 2008 profitability shock caused by the financial crisis, routes in bad-weather/high-PDV partnerships were more likely to stay outsourced to the same regional partner, and hence to keep benefiting from the low labor costs achievable under outsourcing.

Consistent with the lower bound approach, Gil et al. (2017) relies on the endogenous selection of "strong" relationships to high-temptation routes and uses the variation in network-level weather as a proxy for variation in PDV across partnerships. Just as in any other paper using the lower bound approach, there will be measurement error and this may bias (albeit towards zero) their estimates. On a different note, the paper assumes that the shock was unanticipated, and that conditional on pre-shock observables, all in-route collaborations active before the shock would be equally likely to survive if the financial crisis had not occurred. This assumption may be violated if major airlines started their most recent relationships anticipating their termination. To attenuate this concern, they measure relationship strength two years before the 2008 shock and check whether these relationships are still ongoing two years after the shock. Similarly, even though Gil et al. (2017) run a placebo test between 2003 and 2006, their empirical exercise lacks a control

group as the shock applies industry-wide. Therefore, there is little evidence of what may have been the course of action of major airlines in the absence of the 2008 shock.

Performance and the fallback option

In a study of taxi transportation in New York City, Jackson and Schneider (2011) use cross-sectional evidence to show that drivers who lease their taxi from owners from the same country are less likely to have accidents, be convicted for driving violations, or receive summonses. This result is consistent with hypothesis H2 because drivers who lease from owners belonging to the same socio-ethnic network are potentially subject to harsher punishments (from reduced access to community service and assistance to outright ostracism) if they renege on the driving behavior informally agreed with the owners.

To control for the endogenous selection of drivers into social networks, Jackson and Schneider (2011) perform number of imperfect and yet informative additional tests. First, they show that drivers who switch to leasing from same-country owners between 2005 and 2007 improve their driving behavior. Second, they aggregate data at the network level and show that networks with higher percentages of in-network drivers present lower levels of driving violations. Finally, they instrument for in-network driving using the number of same-country taxi owners living in the driver's neighborhood, and find consistent results. Notice that the estimated effects of social networks on performance in the three latter empirical tests are 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.

The paper's main limitation is that it ultimately lacks a clear source of exogenous variation in the drivers' decision to lease from an owner from the same social network. On the one hand, improvements in a driver's performance after leasing from a same-network owner may be driven by unobserved variables that correlate with both the leasing decision and performance, including changes in ability, preferences or intangible compensation. On the other hand, the instrument used in the 2SLS approach uses variation in the number of taxi owners from the driver's social network within distance of the driver's address, but where drivers choose to live is endogenous and may be correlated with other factors that influence (see list above) the decision of driving a taxi within her social network.

Chassang and Ortner (2017) investigate cartel agreements among contractors in procurement auctions in Japan. Their paper argues that price wars are a channel through which cartels punish members' deviations. Therefore, procurement auctions that constrain the pricing strategies of participants through legally imposed minimum bids limit self-enforcement of cartel agreements. This implies (perhaps counter-intuitively) that in the presence of informal collusive agreements, minimum price policies may *lower* prices instead of increasing them. Interpreted through the lenses of our illustrative model, a legally imposed minimum pricing policy is equivalent to a permanent increase in the fallback option of cartel members following a deviation (that is, an increase in $\underline{v}$), which according to hypothesis H2 should lead to a decrease in cartel performance—that is, a decrease in prices.

Chassang and Ortner (2017) test their theoretical prediction using a difference-in-difference methodology where the treatments are minimum procurement option bidding

policies introduced in six Japanese cities. They find that consistent with H2, the distribution of bids in the city of Tsuchiura prior to the policy change first-order stochastically dominates the distribution of bids after the change. Their results on the remaining treated cities are not as strong as their results in Tsuchiura.

While Chassang and Ortner (2017) carefully motivate the exogeneity of Tsuchiura as treatment city, a potential concern—almost unescapable in a non-experimental setting—is that cities changing their procurement auction rules may derive unobserved benefits from such changes that could be correlated with the bidders' anticipated change in behavior.

3.4. Testing H3: Agents respond differently to formal contract terms in the presence of informal contracting

In order to test this hypothesis, one needs to observe formal contracts, parties' responses to such contracts, and whether informal contracting is available or not. A natural indicator of the latter, which could be measured using the techniques discussed in the previous sections, is the parties' discount factor.

Recent empirical research in the movie industry provides promising results in support of H3. Gil (2013) explores how movie distributors adjust revenue sharing terms in their contracts with a large movie exhibitor in Spain. He documents that the exhibitor's actual revenue shares are systematically different from those stipulated in its formal agreements with the distributors. Enriching the dataset used by Gil (2013), Barron et al. (2017) provide evidence that the revenue share "discounts" granted by distributors after a movie's has ended its run at the exhibitor's theaters informally reward the exhibitor for

efficiently allocating movies to screens in a manner that is consistent with H3. In particular, they show that the exhibitor is more likely to keep showing high-risk movies (i.e., movies whose continuation would be least profitable for the exhibitor given their formally contracted revenue share and box office turnouts) when the present discounted value of its relationship with the distributors is higher. Following the "lower bound" approach discussed above, Barron et al. (2017) measure the value of a given exhibitor-distributor relationship through the distributor's maximum reneging temptation—that is, the maximum discount ever granted by that distributor to the exhibitor. Consistent with H3, this result provides evidence that when informal contracting is available (that is, when distributors' relationship with the exhibitor has high present discounted value), the exhibitor responds to informal incentives (the discounts) rather than formal ones (the formally contracted revenue share) in choosing its movie-continuation and screen allocation decisions. Conversely, when informal contracting is not available, the exhibitor responds to formal incentives—that is, it discontinues movies that are unprofitable given the formally contracted revenue share.

Besides the evidence it provides, Barron et al. (2017) suggests a broader methodological point that may be useful to future researchers interested in testing for the interaction between formal and informal contracting. It is easier to measure responses to formal and informal contract terms in *adaptation contexts*—i.e., those where parties need to agree on observable actions such as showing a movie, exchanging a landing slot, refurbishing a store, and the like—relative to moral hazard settings, where the actions, such as employees' efforts, are typically unobservable. Despite their encouraging results, an important caveat remains in that there is no exogenous variation in the PDV between

the 22 distributors supplying movies and the sample exhibitor. Instead, relying on their better more granular data than Gil (2013), Barron et al. (2017) exploit variation in opportunity costs across movie reels within a theater and week using movie release and continuation decisions at the theater and week level. They also split the sample in half to estimate a cross-sectional variation in lower bounds of PDV, and while their results are robust to the methodology used in their calculation of PDV, their observed variation may also be driven by differences in unobservable factors (to the econometrician) that explain the distribution in distributor-exhibitor match quality.

Finally, Macchiavello and Miquel-Florensa (2017) study contracts between coffee buyers and coffee mills in Costa Rica. They are able to observe the terms in their formal contracts and the actual realizations of coffee exchanges between buyers and mills. Their evidence shows that when good weather unexpectedly increases coffee production, mills are able to allocate the additional unforeseen output to vertically integrated buyers and "relational" buyers—defined as buyers that have traded with a given mill for at least four consecutive harvest seasons—despite the fact that buyers are not required to purchase it under their formal contract with the mills. Conversely, mills are not able to allocate the extra output to spot market, non-repeat buyers. Consistent with H3, this evidence indicates that buyers respond to the formal contract when they are in a spot market relationship with the mill, whereas they *do not* respond to the formal contract when their interaction with the mill is repeated.

The results in Macchiavello and Miquel-Florensa (2017) are subject to two potential caveats. First, in their empirical exercise spot and relational transactions are defined ex-ante by the observed volume of past trade, and not by measures of the discount factor or

PDV. Thus, the interpretation of their results suffers from the weaknesses of using past interactions as a proxy for future ones, as discussed in section 3.1. Second, the authors cannot rule out that mills may persuade coffee buyers to purchase the unforeseen extra output by granting unobserved price discounts. If that is the case, what Macchiavello and Miquel-Florensa (2017) classify as "relational" transactions may in fact be transactions between buyers and sellers who can more easily bargain with each.

4. Untested predictions and research opportunities

Table 1 below provides a snapshot of the relation between the theoretical predictions from our model and the empirical papers that we have discussed. Using this table as a reference, we provide guidance on future research opportunities in this final section. First, we highlight theoretical predictions that remain untested, and second, we describe how data sets used by existing studies may be exploited to further test for informal contracting.

[TABLE 1 HERE]

4.1. Formal and informal contracts: substitutes or complements?

Theory

In a seminal paper, Baker et al. (1994) assume that if a principal reneges on the informal bonus promised to his agent, the parties may revert to purely formal contracting rather than separating as assumed in our illustrative model, so long as doing so generates greater surplus. Baker et al. (1994) show that given this assumption on off-the-

equilibrium behavior, improvements in formal contracting may increase or decrease surplus—that is, formal and informal contracts may be substitutes or complements[10].

To understand this point in the context of our illustrative model, consider a reduction in the misalignment parameter θ. This reduction in θ increases the surplus that can be achieved when formal and informal contracts are combined, V^{IF}, because it enables P to rely more on the formal bonus, and less on the informal one, for incentive provision, thereby relaxing the self-enforcement constraint, (10).

In a departure from our model, suppose now that if P fails to pay the informal bonus, P and A revert to whichever type of arm's-length relationship—separation or pure formal contracting—maximizes their joint surplus, as in Baker et al. (1994). In that case, P's fallback option following non-payment is $max\{V^F, \underline{v}\}$ rather than $\underline{v}$. If $V^F > \underline{v}$, an improvement in formal contracting (a reduction in θ) has the additional effect of increasing P's fallback option, thus potentially hampering enforcement of the informal bonus. Therefore, the overall effect of a reduction in θ on surplus is ambiguous: formal and informal contracts may be either substitutes or complements, depending on whether $V^F \gtreqless \underline{v}$ and on the model's parameters.

Evidence and testability

The analysis in Baker et al. (1994) is compelling, and yet we could not find any empirical study providing convincing related evidence. Thus, testing for the complementarity and substitutability between formal and informal contracts constitutes a clear opportunity for future empirical research.

[10] See Baker et al. (2001, 2002, 2011) for related models.

Any test of Baker et al. (1994) will face two major challenges—namely, finding accurate and exogenous measures for 1) the quality of formal contracting (θ), and 2) whether the parties' fallback option is separation or spot market contracting ($max\{V^F, \underline{v}\}$). A useful starting point to think of how to address these challenges is the cross-country survey by Johnson et al. (2002). They find that in European countries from the former Soviet block, trade credit increased when courts were believed to enforce contracts. However, this efficiency-enhancing effect of formal contracting was weaker when clients and suppliers were locked into a bilateral monopoly relationship. While not constituting a direct test, this evidence seems consistent with the predictions of Baker et al. (1994). Under bilateral monopoly the fallback option following breakdown of an informal client-supplier relationship is more likely to be spot market contracting than separation (that is, $V^F > \underline{v}$). In such a setting, improvements in formal contracting (the presence of reliable courts) may not always facilitate trade credit, as they may hamper the enforceability of informal contracts.

For the purpose of testing Baker et al. (1994), the attractive feature of Johnson et al. (2002) is that court availability is a plausibly exogenous measure of formal contracting quality. The potential drawback is that Johnson et al. (2002) lack an exogenous measure of the fallback option. Indeed, in interpreting their results we implicitly made a strong assumption that cannot be verified in the data—namely, that locked-in pairs of suppliers and customers are less likely to separate following breach of an informal agreement, and vice versa.

Based on these considerations, we would advise interested empirical researchers to further explore court quality as a measure of formal contracting quality, while searching

for exogenous sources of variation in the parties' fallback option. A potential strategy to achieve this second goal would be to exploit exogenous changes in contract law that may affect the feasibility of separation as a fallback option, and hence the extent to which parties are "stuck" with each other following breakdown of their informal collaboration. An example is given by laws protecting franchisees from unfair termination. Since both court quality (Berkowitz and Clay, 2006; Choi et al., 2008) and the regulation of contract termination (Brickley et al., 1991) have been found to vary both across countries and industries and within countries, the empirical strategy we are suggesting may exploit variation across countries, within a country over time, or even within a country and year but across different industries.

4.2. Informal contracting affects the design of formal contract terms

Theory

An implication of much of the literature on the interaction between formal and informal contracting is that the latter affects the former: "although Macaulay and others are correct in noting that many business relationships are self-enforced, transactors are not indifferent regarding the contract terms they choose to govern their self-enforcing relationships" (Klein, 2000, p. 68). Consistent with this observation, numerous models emphasize how formal contracts in the presence of informal contracting differ from those the parties would sign if informal contracting were unavailable (e.g., Battigalli and Maggi, 2008; Kvaloy and Olsen, 2010; Iossa and Spagnolo, 2013; Zanarone, 2013).

This point can be easily seen in our illustrative model from section 2. Suppose that θ is neither too large nor too small, so that a purely formal contract is preferable to no

contract at all ($V^F > \underline{v}$) but does not achieve the first best. If $\delta \approx 0$, no informal bonus can be enforced, so the optimal formal bonus is given by $b^F \equiv cos\theta$, with $0 < b^F < 1$. Suppose now that δ grows to an intermediate value, such that the optimal informal bonus is positive but limited ($0 < B^{IF} < 1$). Then, the optimal formal bonus shrinks to $b^{IF} = b(B^{IF}) = (1 - B^{IF})cos\theta < b^F$ —that is, optimal formal bonuses with and without informal contracting are in general different.

Evidence and testability

In order to investigate empirically the effect of informal contracting on formal contract design, one could follow three approaches. The first approach, longitudinal, would be to study how a given formal contract adjusts to an exogenous shock to the feasibility of informal contracting—for instance, a change in δ, the intertemporal discount factor. The second and related approach, cross-sectional, would be to study how differences in the feasibility of informal contracting affect formal contracts in otherwise similar relationships. The third approach would be to rely on structural estimation to construct counter-factual formal contract terms in the absence of self-enforcement.

While we are not aware of empirical studies following the first two approaches, our discussion in section 3 points out recent progress at identifying variations in δ, both across sections (Macchiavello and Morjaria, 2015; Gil et al., 2017; Barron et al., 2017) and over time (Gil and Marion, 2013). At the same time, variations in formal contract terms have been extensively documented both between firms (see the surveys in Lafontaine and Slade, 1997, 2013) and within (see the surveys in Prendergast, 1999, and Lazear and Oyer, 2013). Thus, a clear and exciting opportunity for empirical researchers is to combine existing datasets on formal contract terms in franchising, procurement, and

employment, with estimates of the discounted value of informal relationships in the spirit of the empirical studies discussed in section 3.

So far as the structural approach is concerned, it has not been yet exploited in the literature on informal contracting. However, a few case studies on vertical restraints have documented formal contracting patterns that appear suboptimal in the absence of informal contracting, but potentially optimal when informal contracts are feasible (e.g., Klein and Murphy, 1988; Zanarone, 2009). We see the results from these case studies as an encouragement towards the development of a structural approach, so we briefly discuss them below.

Klein and Murphy (1988) document the widespread use of minimum resale price maintenance in contexts where free-riding on presale services is not a serious concern, such as fashion retailing. They argue that the purpose of price floors in this context is to create quasi-rents that retailers would lose if they reneged on informal obligations on store maintenance or customer management. Zanarone (2009) shows that counterintuitively, provisions against double marginalization *increased* in Italian car dealership contracts after EU law banned exclusive territories, switching from quantity floors to a mix of quantity floors and price ceilings. Relying on field interviews, Zanarone (2009) argues that after the legal change, manufacturers continued to enforce exclusive territories through informal agreements. A deviating dealer could "game" the quantity floor by selling in other dealers' territories, while charging monopoly prices in

his own territory. By preventing such gaming, price ceilings reduced the dealers' gain from reneging on the informal exclusivity agreement.[11]

5. Concluding remarks

In this paper we have investigated 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 have derived testable predictions from a parsimonious model where formal and informal contracts are optimally combined, and we have discussed selected empirical works to shed light on the model's relevance. Three major conclusions emerge from this exercise. First, the evidence supports the predictions of economic models of informal contracting across a broad spectrum of collaborative relationships, industries, and institutional settings. Second, significant progress has been made at measuring key determinants of informal contracting, such as the present discounted value of a given relationship and the parties' fallback option following its breakdown. Third, there are important theoretical questions, such as whether formal and informal contracts are substitutes or complements, or how the existence of informal agreements affects the design of formal contracts, on which systematic evidence is still missing. These questions constitute important opportunities for future empirical research.

We conclude this essay by highlighting a broad methodological point that emerges from our discussion of selected empirical studies on informal contracting. Future researchers in this field should search for industries, institutional environments, firms, or types of contractual relationships, where informal contracts are either observed (as in

[11] This argument is explicitly modeled in a companion working paper (Zanarone 2007, accessible at gzanarone.cunef.edu).

Gillan et al. 09, Gil, 2013, and Barron et al., 2017) or are likely to be important due to observed imperfections in formal contracting. The latter is typically true in developing or transition countries with limited court enforcement (as in Macchiavello and Morjaria, 2015, 2016, and Johnson et al., 2002), and in collaborations where desired performance is too complex or subject to too frequent change to be formally contracted (as in the case of flight schedule adaptation in the airline industry studied by Gil et al., 2017).

We hope our paper will push empirical researchers to use data from these industries and settings, sometimes publicly available, to further advance our understanding of informal contracts and their interaction with formal ones.

References

Ackerberg, D. and M. Botticini (2002), "Endogenous Matching and the Empirical Determinants of Contract Form," *Journal of Political Economy* 110: 564-591.

Baker, G. (1992), "Incentive Contracts and Performance Measurement," *Journal of Political Economy* 100: 598-614.

Baker, G. (2002), "Distortion and Risk in Optimal Incentive Contracts," *Journal of Human Resources* 37: 728-751.

Baker, G., R. Gibbons, and K. J. Murphy (1994), "Subjective Performance Measures in Optimal Incentive Contracts," *Quarterly Journal of Economics* 109: 1125-56.

Baker, G., R. Gibbons, and K. J. Murphy (2001), "Bringing the Market Inside the Firm?" *American Economic Review* 92: 212-218.

Baker, G., R. Gibbons, and K. J. Murphy (2002), "Relational Contracts And The Theory Of The Firm," *Quarterly Journal of Economics* 117: 39-84.

Baker, G., R. Gibbons, and K. J. Murphy (2011), "Relational Adaptation" MIT working paper.

Barron, D., R. Gibbons, R. Gil, and K. J. Murphy (2017), "Reel Authority: Relational Renegotiation in the Movie Industry," mimeograph.

Battigalli, P., and G. Maggi (2008), "Costly Contracting in a Long-Term Relationship," *RAND Journal of Economics* 39: 352-377.

Berkowitz, D., and K. Clay (2006), "The Effect of Judicial Independence on Courts: Evidence from the American States," *Journal of Legal Studies* 35: 399-440.

Brickley, J., F. Dark, and M. Weisbach (1991), "The Economic Effect of Franchise Termination Law," Journal *of Law and Economics* 34: 101-132.

Bull, C. (1987), "The Existence of Self-enforcing Implicit Contracts," *Quarterly Journal of Economics* 102: 147-159.

Chassang, S., and J. Ortner (2017), "Collusion in auctions with Constrained Bids: Theory and Evidence from Public Procurement," Princeton University Working Paper.

Choi, S., G. Gulati, and E. Posner (2008), "Professionals or Politicians: The Uncertain Empirical Case for an Elected rather than an Appointed Judiciary," *Journal of Law, Economics and Organization* 26: 290-336.

Corts, K. and J. Singh, (2004) "The Effect of Repeated Interaction on Contract Choice: Evidence from Offshore Drilling," *Journal of Law, Economics and Organization* 20: 230-260.

DeVaro, J., J. Kim and N. Vikander (2017), "Non-Performance Pay and Relational Contracting: Evidence from CEO Compensation," *Economic Journal* (forthcoming).

Ellickson, R. (1991), *Order without Law: How Neighbors Settle Disputes*, Cambridge, MA: Harvard University Press.

Fast, N., and N. Berg (1975), "The Lincoln Electric Company," Harvard Business School Case #376-028.

Forbes, S., and M. Lederman (2009), "Adaptation and Vertical Integration in the Airline Industry," *American Economic Review* 99: 1831-1849.

Foss, N. (2003), "Selective Intervention and Internal Hybrids: Interpreting and Learning from the Rise and Decline of the Oticon Spaghetti Organization," *Organization Science* 14: 331-349.

Gibbons, R., (2005) "Four Formal(izable) Theories of the Firm?" *Journal of Economic Behavior and Organization* 58: 202-247.

Gibbons, R., H. Li, and S. Venables (2012), "Notes on Relational Incentive Systems," http://web.mit.edu/rgibbons/www/Relational%20Incentive%20Systems%20copy.pdf.

Gil, R. (2013), "The Interplay of Formal and Relational Contracts: Evidence from Movies," *Journal of Law, Economics and Organization* 29: 681-710.

Gil, R., and J. Marion (2013), "Self-enforcing Agreements and Relational Contracts: Evidence from California Highway Procurement," *Journal of Law, Economics and Organization* 29: 239-277.

Gil, R., and G. Zanarone (2016), "New Frontiers in the Empirical Research on Informal Contracting," *Journal of Institutional and Theoretical Economics* 172: 390-407.

Gil, R., and G. Zanarone (2017), "Formal and Informal Contracting: Theory and Evidence," *Annual Review of Law and Social Science* 13 (forthcoming).

Gil, R., M. Kim, and G. Zanarone (2017), "The Value of Relational Adaptation in Outsourcing: Evidence from the 2008 Shock to the US Airline Industry," Johns Hopkins University, working paper.

Gillan, S., J. Hartzell, and R. Parrino (2009), "Explicit versus Implicit Contracts: Evidence from CEO Employment Agreements," *Journal of Finance* 64: 1629-1655.

Halac, M. (2012), "Relational Contracts and the Value of Relationships," *American Economic Review* 102: 750-779.

Halonen, M. (2002), "Reputation and the Allocation of Ownership," *Economic Journal* 112: 539-558.

Hart, O., and J. Moore, (2008), "Contracts as Reference Points," *Quarterly Journal of Economics* 123: 1-48.

Hart, O., and B. Holmstrom (2010), "A Theory of Firm Scope," *Quarterly Journal of Economics* 125: 483-513.

Helper, S., and R. Henderson (2014), "Management Practices, Relational Contracts, and the Decline of General Motors," *Journal of Economic Perspectives* 28: 49-72.

Holmstrom, B., and P. Milgrom (1991), "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design," *Journal of Law, Economics and Organization* 7: 24-52.

Iossa, E., and G. Spagnolo (2011), "Contracts as Threats: On the Rationale for Rewarding A While Hoping for B," CEPR Discussion Paper 8195.

Jackson, K., and H. Schneider (2011), "Do Social Connections Reduce Moral Hazard? Evidence from the New York City Taxi Industry," *American Economic Journal: Applied Economics* 3: 244-267.

Johnson, S., J. McMillan, and C. Woodruff (2002), "Courts and Relational Contracts," *Journal of Law, Economics, and Organization* 18: 221-77.

Kalnins, A., and K. Mayer (2004), "Relationships and Hybrid Contracts: An Analysis of Contract Choice in Information Technology," *Journal of Law, Economics, and Organization* 20: 207-29.

Klein, B. (1995), "The economics of franchise contracts," *Journal of Corporate Finance* 2: 9-37.

Klein, B. (2000), "The Role of Incomplete Contracts in Self-Enforcing Relationships," *Revue d'Économie Industrielle* 92: 67-80.

Klein, B. and K. M. Murphy (1988), "Vertical Restraints as Contract Enforcement Mechanisms," *Journal of Law and Economics* 31: 265-97.

Klein, B., and K. M. Murphy (1997), "Vertical Integration as a Self-Enforcing Contractual Arrangement," *American Economic Review* 87: 415-20.

Kvaløy, O., and T. Olsen (2009), "Endogenous Verifiability and Relational Contracting," *American Economic Review* 99: 2193-2208.

Lafontaine, F., and M. Slade (1997), "Retail Contracting: Theory and Practice," *Journal of Industrial Economics* 45: 1-25.

Lafontaine, F., and M. Slade (2013), "Inter-Firm Contracts: The Evidence," in Gibbons, R., and J. Roberts (editors), *The Handbook of Organizational Economics*, Princeton: Princeton University Press.

Levin, J. (2002), "Multilateral Contracting and the Employment Relationship," *Quarterly Journal of Economics* 117: 1075-1103.

Levin, J. (2003), "Relational Incentive Contracts," *American Economic Review* 93: 835-57.

Li, J., and N. Matouschek (2013), "Managing Conflicts in Relational Contracts," *American Economic Review* 103: 2328-2351.

Macaulay, S. (1963), "Non-Contractual Relations in Business: A Preliminary Study," *American Sociological Review* 28: 1-19.

Macchiavello, R., and P. Miquel-Florensa (2017), "Vertical Integration and Relational Contracts: Evidence from the Costa Rica Coffee Chain," University of Warwick Working Paper.

Macchiavello, R., and A. Morjaria (2015), "The Value of Relationships: Evidence from a Supply Shock to Kenyan Rose Exports," *American Economic Review* 105: 2911-2945.

Macchiavello, R., and A. Morjaria (2016), "Competition and Relational Contracts: Evidence from Rwanda's Coffee Mills," University of Warwick Working Paper.

MacLeod, B. (2007), "Reputations, Relationships, and Contract Enforcement," *Journal of Economic Literature* 45: 595-628.

MacLeod, B., and J. Malcomson (1989), "Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment," *Econometrica* 57: 447-480.

MacNeil, I. (1978), "Contracts: Adjustments of Long-Term Economic Relations under Classical, Neoclassical, and Relational Contract Law," *Northwestern University Law Review* 72: 854-906.

Malcomson, J. (2013), "Relational Incentive Contracts," In Gibbons, R., and J. Roberts (editors), *The Handbook of Organizational Economics*, Princeton: Princeton University Press.

Ray, D. (2002), "The Time Structure of Self-enforcing Agreements," *Econometrica* 70: 547-582.

Schwartz, A. (1992), "Relational Contracts in the Courts: An Analysis of Incomplete Agreements and Judicial Strategies," *Journal of Legal Studies* 21: 271-318.

Shelanski, H., and P. Klein (1995), "Empirical Research in Transaction Cost Economics: A Review and Assessment," *Journal of Law, Economics and Organization* 11: 335-361.

Shelef, O. (2013), "Relational Influence Buying," UC Berkeley Working Paper.

Williamson, O. (1979), "Transaction Cost Economics: The Governance of Contractual Relations," *Journal of Law and Economics* 22: 233-261.

Williamson, O. (1991), "Comparative Economic Organization: The Analysis of Discrete Structural Alternatives," *Administrative Science Quarterly* 36: 269-96.

Xu, C. (2011), "The Fundamental Institutions of China's Reforms and Development," *Journal of Economic Literature* 494: 1076-1151.

Zanarone, G. (2009), "Vertical Restraints and the Law: Evidence from Automobile Franchising," *Journal of Law and Economics* 52: 691-700.

Zanarone, G. (2013), "Contract Adaptation under Legal Constraints," *Journal of Law, Economics and Organization* 29: 799-834.

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

	Transaction Characteristics	
Variable to be explained	**Discount Rate/PDV & Fallback Option**	**Quality of formal contracts**
Choice Between Formal and Informal Contract	*H1*: Gillan et al. (2009) Corts and Singh (2004) De Varo et al. (2017)	Untested
Agent's actions/performance	*H2*: Jackson and Schneider (2011) Gil and Marion (2013) Shelef (2013) Chassang and Ortner (2017) Gil et al. (2017)	Untested
Agent's response to formal contract	*H3:* Barron et al. (2017) Macchiavello and Miquel-Florensa (2017)	
Design of formal contract	Untested	