Transactional Governance Structures: New Cross-Country Data and an Application to the Effect of Uncertainty*

Peter Murrell† Nona Karalashvili‡ David C. Francis††

October 18, 2022

Abstract: To what extent are personal trust, mutual interests, and third parties important in enforcing agreements to trade? How do firms combine these to form transactional governance structures? This paper answers these questions in a whole-economy, cross-country setting that considers a full spectrum of transactional-governance strategies. The data collection requires a new survey question answerable in any context. The question is applied in six South American countries using representative samples, with the resultant survey weights facilitating a whole-economy analysis. Without imposing an a priori model, latent class analysis estimates meaningful governance structures. Bilateralism is always used. Law is never used alone. Bilateralism and formal institutions are rarely substitutes. Within country, inter-regional variation in governance is greater than inter-country variation. The usefulness of the data is shown by testing one element of Williamson's discriminating-alignment agenda: greater uncertainty in the transactional environment increases the involvement of third parties.

Keywords: governance, transactions, law, bilateralism, transaction-costs, uncertainty

JEL Classifications: L14, C81, D23, P5, M21

* This is a thoroughly revised version, with new content added, of the paper "Mapping the Landscape of Transactions: The Governance of Business Relations in Latin America". For helpful comments and suggestions, we would like to thank Mohammad Amin, Daniela Behr, Adrian Gonzalez, Peter Grajzl, Gillian Hadfield, Mary Hallward-Driemeier, Kathryn Hendley, Asif Islam, Ethan Kaplan, Aart Kraay, Stephanie Lanza, Norman Loayza, Jorge Luis Rodriguez Meza, Károly Mike, Helena Schweiger, and Jeroen Vermunt. Financial support from the Knowledge for Change Program (KCP) of the World Bank is gratefully acknowledged. The views expressed here are those of the authors, and do not reflect those of the World Bank, its Executive Directors, or the countries they represent.

† University of Maryland, pmurrell@umd.edu
‡ The World Bank, nkaralashvili@worldbank.org
†† The World Bank, dfrancis@worldbank.org

I. Introduction

To what extent do firms rely on each others' trustworthiness and mutual interests or invoke the assistance of various third parties when ensuring the fulfillment of their agreements to buy and sell goods and services? Which combinations of these approaches do firms use? How do answers to these questions vary across countries? Williamson (1979) refers to such combinations as transactional governance structures. In the 40 years since this concept was introduced, no study has produced economy-wide, cross-country data on a comprehensive set of approaches to supporting transactions, showing which governance structures are most commonly found effective by firms. This is surprising given the consensus that firm performance and levels of development are associated with the effectiveness of arrangements for enforcing agreements (Williamson 1985; North 1990; Moran and Ghoshal 1999, Greif 2001, World Bank 2002, 2017).

The aim of this paper is to fill this gap in the literature. We rely on data that are generated from the responses of individual firms to a survey question that is designed to be interpretable by a respondent from any firm in any sector in any country. The data reflect the experiences of 3,430 firms in six South American economies. This is a small number of somewhat similar countries, but the methodology is sufficiently general that it could be applied in any country, sector, or region in the world. By using representative samples and survey weights, we are able to characterize the aggregate importance of different governance structures in a very large part of economic activity.[1]

To our knowledge our paper is the first to present and analyze data that provide such a comprehensive picture of the effectiveness of the different approaches that firms use to support the fulfillment of agreements. While there are many datasets on particular elements of governance, such as the strength of contractual institutions, our data uniquely capture four elements: reflecting

[1] The samples are designed to be nationally representative of non-agricultural, non-extractive, formally registered, private firms with at least five employees. See Section II.3 for more details.

the responses of individual firms about their own experiences of transactional governance; providing a (nearly) whole-economy perspective; facilitating comparisons between different countries; and enabling the examination of a comprehensive set of strategies employed in transactional governance.[2] To emphasize the potential value of our dataset for further research, we provide an example of its use by examining a core element of transaction-cost theories: the effect of uncertainty on the choice of governance structures.

A (transactional) governance structure is the "institutional framework within which the integrity of a transaction is decided" (Williamson 1979: 240), a coordinated combination of different mechanisms that together constitute a strategy for encouraging the fulfillment of agreements to transact.[3] To date, there exists no generally agreed upon, encompassing theoretical framework that predicts which particular mechanisms appear often, and which can be safely ignored. Therefore, the first step in our quest is to specify a comprehensive list of such mechanisms. As described in detail in Section II, in formulating this list, we rely on three very common elements of the economics literature. First, the analysis of inter-firm agreements invariably begins with opportunism and its detrimental effects, thereby leading to a focus on enforcement.[4] Second, the mechanisms of enforcement are often designated by referring to the specific agents involved in enforcement. Third, the pertinent agents can be arranged on a spectrum, from the narrowest to the broadest, as in Greif's (1997) differentiation between first-, second-, and third-party enforcement.

[2] There are many high-quality studies that examine pieces of the picture. However, to provide background on how difficult it is to pursue all these elements, we examined those studies labelled as key studies by Cao and Lumineau (2015). Of the 49 studies, only six covered more than one country. Of these six, three covered three or fewer countries and five used small, fairly narrow samples of firms. One study (Yang et al. 2012) covered 969 firms in 17 countries, but confined its results to industrial sectors, did not use sampling weights to derive an economy-wide picture, and focused on a limited number of transactional strategies.
[3] We confine our attention to those aspects of governance relating to transactions in goods and services. Thus, a governance structure in the sense used in this paper, excludes many activities that fall under the umbrella of corporate governance.
[4] Williamson's pithy definition of opportunism was that it was "self-interest seeking with guile" (1975).

Using this agent-based approach and focusing on enforcement, we construct a survey question about the effectiveness of six generic agents involved in enforcement (individuals, organizations, etc.) that cover the spectrum from intimate first-party personal trust to remote third parties, such as the legal system. This set of six agents is exhaustive in the sense that all firms would be able to relate to the survey question no matter which approach they take to the enforcement of agreements. Given the centrality of the quality of data collection to this paper's contribution, Section II contains extensive discussion of our approach to question wording, provides information on the rigor of survey design, and describes elements of survey implementation that are relevant to making judgments on the validity of the resultant dataset. The formulation, justification, and implementation of this question is the first contribution of this paper.

The six agents specified in the survey question are not necessarily governance structures themselves but rather inputs into governance structures, which combine the six in varying ways. In building an evocative picture, the next challenge then is finding which individual mechanisms are combined by firms to produce coherent governance structures. In meeting this challenge, one cannot rely on a definitive theory of how combinations are chosen because none is available (Mike and Kiss 2019, Greif 1997). Nevertheless, the patterns of governance structures must be latent within our data, and therefore we can use an exploratory statistical technique to make them manifest. Latent class analysis (LCA) is particularly suitable for this task (Mike and Kiss 2019).[5] LCA models the governance structures as the categories, or classes, of a categorical latent variable, which is to be estimated. The estimated governance structures are then purely data-driven, not reliant on any a priori conception of those governance structures that are actually used in practice.

[5] LCA would be labeled as an unsupervised machine-learning approach had it not been developed before the age of big data. Collins and Lanza (2010) provide an intuitive introduction to LCA; Masyn (2013) presents a precise description of the decisions to be made in practical implementations; Vermunt and Magidson (2016) provide a comprehensive technical introduction to the statistical theories and methods included in the software used to produce the results for this paper.

Given the identification of a number of distinct governance structures (or classes), LCA produces estimates of the posterior probability that each firm belongs to each class.

Thus, a second contribution of this paper is the discovery and estimation of governance structures in the six-country dataset using LCA and also the substantive interpretation of the governance structures captured by these estimates.[6] As we describe in Section IV, we find that pure bilateralism (i.e., a strategy combining only first-party personal trust and second-party mutual interest without invoking any third-party agent) is the most common governance structure and that all governance structures embody bilateral enforcement mechanisms. Hence, none of the estimated governance structures correspond to purely arm's length transactions, where firms rely only on impersonal mechanisms and formal institutions to support their transactions. A corollary is that bilateralism and formal institutions are rarely substitutes, and for many firms, they are complements. Additionally, while much attention has been devoted in the literature to various unpaid, third-party, mechanisms of supporting agreements, such as social clubs and culturally defined groups, our results suggest they are not very important in the enforcement of transactions.

A third contribution is to provide illustrative examples of the use of the data derived from LCA. We do so in Sections V and VI. In Section V, we present descriptive statistics on the variation in the prevalence of different governance structures across countries and regions. These are descriptive in the sense that they do not isolate ceteris paribus causal effects of single variables. We do, however, use techniques previously developed for LCA that produce consistent estimates

[6] No existing studies obtain a cross-country, economy-wide picture on how these mechanisms are combined to form governance structures (Greif 1997, 2005, Mike and Kiss 2019). To be sure, there are high quality studies that analyze particular combinations. See, for example, Cao and Lumineau (2015) on contractual and relational governance, which is the focus of many studies in the business economics literature. Hendley et al. (2000), Hendley and Murrell (2003), and Mike and Kiss (2019) are the only country-wide studies of which we are aware that use datasets to which the current paper's methods could be applicable. These datasets are all on single countries. The World Bank's Regional Project on Enterprise Development (RPED) collected cross-country comparable data on, inter alia, firms' attempts to solve transactional problems in seven African countries beginning in the early 1990's (see Fafchamps 2004). However, the questions used in the RPED surveys are not in a form that would make the responses suitable for estimation of governance structures. Unlike the current paper, none of the studies cited in this footnote use a sampling design and survey weights so that estimates are representative of a whole economy.

of the descriptive parameters. We show, for example, that inter-regional variation in the prevalence of governance structures that use legal institutions is larger than cross-country variation. In the countries analyzed, this is unexpected for two reasons: because institutional rules relevant to transactions are set at the national level and because there is significant cross-country variation in the strength of legal institutions. This is a puzzle that is surely in need of further investigation, providing just one example of how our approach to data collection can stimulate new avenues for research.

Section VI provides a specific illustration of how that research might proceed by examining a hypothesis that has been a central feature of theories of transactional governance at least since Williamson (1975, 1985) observed that exogenous shocks could enable opportunism. We examine how the presence of uncertainty affects the use of purely bilateral governance. LCA provides the dependent variable, the estimated posterior probability that a firm uses pure bilateralism.[7] This variable indirectly measures the non-involvement of third-parties due to our finding that all estimated classes embody bilateralism, with varying degrees of third-party support. Our analysis thus examines if firms are more likely to value third-party mechanisms as a complement of bilateralism when there is uncertainty.

The explanatory variable of chief interest is the degree of uncertainty in the conditions surrounding the exchange between the two parties. In both fixed-effects OLS and IV regressions, we find that transactions that are more affected by uncertainty use pure bilateralism less and are more likely to use governance structures that add third parties to bilateralism. Our interpretation of this result is that reference to a third party provides increased adaptability and a gap-filling role when firms face an uncertain environment (Williamson 1975, 1985). The adaptability-facilitating

[7] Section VI focuses solely on firms' relations with their suppliers.

role of third parties is suggested by complementarity between bilateralism and third-party agents that we found in the process of estimating governance structures. A neutral referee is especially useful as an addition to bilateral governance when shocks require adjustments that create room for opportunism. Hence, the role of third parties increases as uncertainty increases: the incidence of pure bilateralism decreases.

These illustrative examples of the use of our data suggest a fourth contribution of our research. We produce a derived dataset of the estimated posterior probabilities that each firm adopts each of the estimated governance structures. This dataset is available for use by other researchers to advance their own research agendas by combining our methodology and their own data.[8]

II. The Conceptual Framework, the Question, the Surveys and Raw Responses

We spend much longer than is usual in addressing issues of question construction, wording, and survey implementation, since the relevance and validity of the data obtained from our questions is central to this paper. Nevertheless, the problems we faced, and tradeoffs we encountered are typical of any study that aims to collect data that is not recorded in the accounts of firms. To put the process of question design and implementation into context, it is worth remembering that in the seminal paper in the current line of inquiry, Macaulay (1963) was forced to remark that "…to a great extent, existing knowledge has been inadequate to permit more rigorous procedures—as yet one cannot formulate many precise questions to be asked a systematically selected sample…Much time has been spent fishing for relevant questions…" Despite the years since the publication of Macaulay's paper and the subsequent recognition that it had raised fundamental issues, researchers still face a notable challenge in constructing questions

[8] The concluding section of the paper provides details on how to access the relevant datasets and tools.

that address the problems of data collection that he so clearly articulated. We have endeavored to implement a question that covers a comprehensive set of agents involved in transactional governance, with the constraint that it must be addressed to economy-wide samples of firms from different countries.

II.1 The Conceptual Framework

Three elements in the economics literature underpin the conceptual basis of the survey question that lies at the core of this paper. The first is that the analysis of inter-firm agreements almost invariably begins with Williamson's (1975, 1985) emphasis on opportunism and its possible detrimental effects on the attainment of the overall objectives of the firms that are undertaking those agreements. Hence, our approach can be thought of as emphasizing enforcement in its broadest sense.[9] We thus follow the prevailing approach in the literature, which places enforcement at the center of the analysis of agreements (e.g., MacLeod 2007).[10] Agreements can also be coordination devices, whereby the details can serve a purely informational role to coordinate the actions of firms (e.g., Hadfield and Bozovic 2016). While we do not focus specifically on this aspect of agreements, coordination and enforcement can be intertwined, as firms must be able to both commit to the agreement and communicate this commitment to each other to achieve coordination (Greif, 2000).

Second, an important strand of the literature ultimately casts analysis in terms of what specific agents (individuals, firms, organizations, or institutions) are involved in enforcement.[11] We call

[9] Greif (1997) illustrates this broad emphasis on enforcement as follows (p. 247): "Since the agent can act opportunistically, the principal would not initiate an exchange (even if it were efficient) in the absence of a contract enforcement institution." See also, Greif (1994, 2005).
[10] See also Johnson, McMillan and Woodruff (2002), Dixit (2003, 2009), Hart and Moore (1988), Susarla et al. (2020), Zhou et al. (2008), Bernstein (2016).
[11] This approach is present when focusing on the actions of organizations, such as courts or government officials, that implement institutions. Further examples are provided by criminal groups, arbitration courts, social groups practicing ostracism, and business associations cutting off trade.

this an agent-based approach, in which broad strategies emerge as descriptive summaries of the types of agents chosen to aid enforcement. That approach can be seen most clearly in the foundational analytical approach to understanding inter-firm agreements, game theory, which begins with the naming of agents. Our survey question effectively does the same, by asking respondents to consider specific sets of agents (e.g., government officials).

An alternative approach would begin with strategies, not agents. The most popular form of this approach in the literature contrasts contractual and relational transactional strategies, with the origins of the former often interpreted as lying within economics and the latter reflecting sociological theory.[12] Measures of the use of these two strategies usually aggregate responses to multiple questions on various behaviors and beliefs that characterize the formation and implementation of inter-firm agreements.[13] Strategy-based approaches have been successfully applied across multiple contexts, especially using smaller, bespoke samples that reflect narrower ranges of cultures, sectors, or activities.[14] Such a focus facilitates the use of highly nuanced survey questions, which is necessary in understanding strategies and their use. We chose the agent-based approach since the terms designating distinct types of agents will likely have a more consistent meaning than terms corresponding to the nuances of broad strategies when implemented via

[12] The definitions of these two strategies vary across different studies, often significantly. For example, contrast the narrow approach on whether the use of social sanctions is implied in relational strategies used by Gil et al. (2019) with the broader approach of Johnson, McMillan, and Woodruff (2002). For definitions, see, e.g., Cao and Lumineau (2015) and Abdi (2017). Contractual and relational transactional strategies can be used together or separately. Argyres et al. (2020) argues that a contractual approach might be embedded in a relational approach to produce more powerful inter-firm cooperation. For the roles of economics and sociology in conceptualizing these strategies, see, e.g., Poppo and Zenger (2002), Greif (2005).

[13] For example, Poppo and Zenger (2002) measure relational governance by combining responses to three questions on the level of collaborative relationship, on the extent of shared goals, and the reliability of the partner to keep promises. Zhou and Xu (2012) measure it by aggregating information on whether the exchange is frequent and whether the parties keep each other informed about changes. Homburg et al. (2009) include multiple questions on trust, along with multiple questions on the extent to which the relationship is governed by written contracts.

[14] See e.g., Homburg et al. (2009) who collect data on contracts in the US and Germany in the chemical mechanical and electrical industries. To examine the effect of national culture on the comparative use of contractual and relational governance, Handley (2015) collects data from firms that are large, US-based, and engaged in outsourcing of business processes. To analyze the effect of culture on the relationship between governance and of opportunism, Handley and Angst (2015) survey service providers that are located across countries but all having customers that are based in the US.

economy-wide and cross-country surveys. We regard the agent-based and strategy-based approaches as true alternatives since they could not be combined in a single survey question of the type used in this paper, for reasons we clarify below.

The third element of the economics literature that underpins the conceptual basis of our survey question is the view that the pertinent agents can be identified and arranged on a spectrum, from the narrowest to the broadest. Although the language varies enormously across papers, this view is common, although often implicit and not comprehensive in scope (see, e.g., Dixit 2003, World Bank 2002, Greif 1997). To clarify this point and connect it to the structure of our survey question, we use Greif's (1997, 2005) formulations since the notion of a spectrum of agents is an important feature of his highly influential work.

Greif (1997) differentiates between first-, second-, and third-party enforcement (see also Dixit 2009), any combination of which can be relevant to any specific agreement. First-party enforcement reflects the personal characteristics of agents themselves, characteristics that are often relevant to all of the agent's actions, not only those undertaken in implementing a particular agreement. Problems arise less frequently when that agent is trustworthy, trusting, honest, and has absorbed common social norms. The rest of society can provide background to these purely first-party motives, since many moral values arise as the result of social interactions (Granovetter, 1985).[15] However, these motives are purely first-party because enforcement occurs without any reference to second or third parties: guilt (as opposed to shame) does not imply direct social interaction (Greif 1994, 1997, Argyres et al. 2020). Problem-solving occurs simply because of the "… the tendency of individuals to derive utility from acting according to their values (first-party

[15] Poppo et al. (2016) describe this as follows: "relational trust arises from social relationships when there are strong beliefs about the goodwill, honesty, and good-faith efforts of others, which mitigate risk by aligning core values" (p. 2). See also Granovetter (1985), Gulati (1995), Zhou et al. (2008), Poppo and Zenger (2002), Susarla et al. (2020).

enforcement)" (Greif 1997: 250).[16] The agent enforcing the agreement is the intrinsic person and the mechanism is personal trust.[17]

Second-party enforcement reflects calculative trust, the prime example being the repeated-game-folk-theorem equilibrium, whereby each of the parties has a direct long-term financial interest in fulfilling the agreement (e.g., MacLeod 2007). The information used for second-party enforcement is mostly forward looking, in which the personal identities of the two parties can be irrelevant, and the focus is upon the continuation of a profitable business arrangement.[18] In second-party enforcement, we do not have social beings, but rather agents as profit-seeking computing machines.[19]

Third-party enforcement involves implicit or explicit reliance on additional economic agents to prevent or to resolve problems that might arise in the implementation of the agreement. In this paper, we take 'third parties' to include all agents other than the two that are directly party to the agreement. Such third parties include courts, social or economic groups, arbitrators, and government. They might be explicitly mentioned in the agreement (e.g., arbitrators) or be implicit (e.g., criminal groups). Potential sanctions include the loss of valuable future transactions with third parties. Specific actions by third parties might never occur: merely referring to their existence and powers might lead to cooperative adjustments that prevent problems arising or lead to obvious

[16] See also Greif (1994), Argyres et al. (2020) and Guiso et al. (2009). Ellickson (1997) comments on Greif (1997) as follows "[a] first party system involves internalized norms against defection that individuals enforce against themselves." (page 268)
[17] First-party enforcement is referred to as morality in Dixit (2003) and Mike and Kiss (2019). Greif (1997) uses terms such as general morality or personal trust to describe the same concept. Other terminology used to describe this type of enforcement includes individual or “internalized” norms, such as honesty (World Bank, 2002, Bohnet et al. 2001). Also see, e.g., Di Domenico and Ryan (2017) for neuroscientific evidence of such intrinsic enforcement.
[18] The forward-looking calculation in this context refers to calculative assessment of "whether it pays to cooperate" (Poppo and Zhou 2016, p. 3) which is absent in the case of the personal trust. In MacLeod's terms (2007), the existing stock of reputation capital held by the agents might be relevant to the calculations, but no consideration needs to be made concerning how this capital was built. The repeated-game-folk-theorem equilibrium could involve statistical analysis of past data when agents estimate parameters that they can try to deduce from past behavior.
[19] While the distinction between 'personal trust' and 'calculative trust' is important and has been acknowledged in the literature for many years, interest in it has risen recently, with attempts at empirical analysis (Williamson 1993, 1996, Baker et al. 2002, Poppo et al. Zhou 2016, Argyres et al. 2020, Susarla et al. 2020).

resolutions when problems do arise. But sometimes there are court cases, social ostracisms, or acts of violence.

We now turn to the wording of the survey question and the tradeoffs faced in its formulation. These primarily arise from practical matters of survey implementation, particularly using terms that are clear to a heterogeneous group of respondents, and constraints on question length.

II.2 The Question

We begin by stating the question, as implemented. First, the interviewer reads the following aloud:

> *When making agreements with [suppliers][customers], please indicate to what degree each of the following is effective in resolving or preventing problems.*[20]

Then, the respondent is presented with a 'show card' that contained a list of six 'agents' in the sense discussed above.[21] These are: personal relationship and trust; mutual interest in maintaining business relationship without involving others; paid private dispute resolution; assistance of government officials; intervention of other third-parties (excluding paid, private dispute resolution and government officials); and legal system. The interviewer then reads aloud the Likert scale of possible responses, also displayed on the show card, which remains visible the whole time the respondent is being prompted to give individual answers for each of the six agents:

Not at all	Slightly	Moderately	Very much	Extremely

[20] We chose to ask about agreements with suppliers and customers separately because firms might employ very different strategies when managing upstream relations than downstream ones.

[21] Interviewers were instructed to read questions as written on the survey instrument and present the 'show card'. Then, on respondent inquiries, interviewers could consult a written manual to explain the meaning of questions. The manual contained the following interpretation of the opening statement of the question: "Resolving or preventing problems in agreements incudes negotiations, how agreements are finalized, and how other available means are implemented to resolve problems when they appear. Each one of the following six mechanisms includes not only their use but also the promise or threat of their use when a problem emerges."

Respondents rate each of the six items individually. Because respondents could continually see the content of all six sub-questions when responding on any particular one, they could implicitly understand the underlying logic of the question.[22]

II.2.1 The Specification of the Six Agents

The six agents in the question are arrayed on the first-party to third-party spectrum.[23] The first sub-question, 'personal relationship and trust' aims at evoking unilateral personal trust. It begins with 'personal' to indicate motives that lie deep within the agent. It uses 'relationship' and 'trust' to steer the respondent away from any hint of pecuniary motives.

The next sub-question, 'mutual interest in maintaining business relationship without involving others', refers to calculative trust. It places 'mutual' at the fore, bringing into play the second party to the agreement, but not any other agents, and mentions 'business relationship' in order to move the respondent's attention to more worldly motives. The use of 'business' primes the idea of pecuniary benefits. The term 'maintaining' evokes a repeated game (e.g., McMillan and Woodruff 1999), while the 'without involving others' forces attention only on the interaction with the other party to this agreement. Note that maintaining a business relationship might also implicate aspects of reputation since breach of an agreement may reduce the reputational capital of the breacher (e.g., MacLeod 2007).

The next four sub-questions refer to different kinds of third parties. Each is worded to steer the respondent away from personal trust or mutuality. 'Paid private dispute resolution' refers to arbitration, perhaps by business associations, mediation, and other similar fora. These are not

[22] The exact wording of the questions in Spanish and English is included in Appendix A.1, together with the show card. Nuances raised in the translation process are discussed in Appendix A.2. The stages of survey implementation that are relevant to the question design are described in Appendix A.3, with the relevant sections of the interviewer manual provided in Appendix A.4. Appendix A.6 and A.7, respectively, include information on fieldwork dates and the total number of observations available for each country; and summary statistics on rates of "Don't Know" responses, which are negligible.

[23] A number of the points made in this subsection were included in the interviewer's manual, which determines interviewer explanations. See Appendix A.4.

named in the question because specific names would vary between countries, and the use of them might involve country- or language-specific priming. This general term also gives parties a means of indicating the effectiveness of criminal groups while not revealing sensitive information (Hendley, Murrell, and Ryterman 2000).

We include the 'assistance of government officials' for a variety of reasons. Even in developed countries, government officials play a role: when a market is regulated, when local officials use zoning, inspections, and fines to influence the outcomes of disputes, or when criminal-justice officials suspect fraud.[24] Hendley and Murrell (2003) included government because they were studying a country with a communist past, and the same logic would surely apply when there has been a history of statist policies.[25]

The 'intervention of other third-parties (excluding paid, private dispute resolution and government officials)' can include unpaid involvement, or threat thereof, by business associations, by officials of third-party firms with which the parties have other relationships, citizens prominent locally, or social groups, such as churches, political parties, and clubs, even family members who are not party to the agreement, etc. This is a broad, catch-all category but that is intentional: respondents can invoke the effect of agents who do not fit easily into the other categories.[26] When constructing this sub-question, there was the decision on whether to use the word 'intervention' in the question. Ultimately this decision rested on the phrasing of the question as a whole, because mention of 'third-parties' alone without some type of action would be much too general. Without the words 'dispute resolution', 'assistance' and 'intervention' in the third, fourth, and fifth sub-

[24] Williamson (1991) and Greif (1997) mention regulations in this context, Greif (1997) and Greif (2005) mention bureaucratic procedures, and World Bank (2002) mentions fines.
[25] We are agnostic on whether the effectiveness of government officials in resolving or preventing problems in agreements is positive or negative from the perspective of economic development in general. Our purpose is to solely understand the extent of their role.
[26] For the role of these kinds of third parties in governance of agreements see, e.g., Greif (2005), World Bank (2002), Ellickson (1997), Greif (1994, 1997, 2000), Johnson, McMillan and Woodruff (2002).

questions, there was the possibility that respondents might invoke too many aspects of transactions in their responses, such as individual norm formation, partner selection, or informational, or coordination aspects of entering agreements. Certainly, this is one example of the trade-offs inherent in the design of short survey questions: we chose to elicit responses that are indicative of any concrete effect of the transactional agents, rather than casting a broader brush to mere association with the parties to the agreement.

The 'legal system' is the last item. Because it is the farthest removed of all six sub-questions from moral, personal, or relationship considerations, it is placed at the end, to evoke a spectrum of options from personal to remote. The phrasing steers respondents away from thinking that any role for the law would imply choosing a high rating: most agreements will mention the law in some way even if it never has any effect on behavior; all actions are under the shadow of some law. The use of the term 'system' is aimed at evoking the idea of a real organization in the respondents' minds, without using the word 'courts' to avoid the assumption that lawsuits are a necessary condition for giving this option a high rating.

There are two words that we studiously avoided including in our questions even though they describe important concepts widely used in the pertinent literature—'reputation' and 'networks'. These words refer to broad strategies or tools that agents use but are not the entry points of analysis in the agent-based approach that we use. Moreover, these words can be rather ambiguous, describing powerful and versatile strategies that can vary considerably across contexts and agents in the ways that they function. Reputation can, for example, have a societal value, implying third-party involvements (Greif 2000, Williamson 1991). A reputation for a willingness to commit violence is at the heart of the enforcement by criminal groups. Government, social groups, and the legal system all use reputation as their currency. Reputation can also invoke the probability of

certain behaviors that one agent assigns to other agents (World Bank 2017, Macchiavello and Morjaria 2015, Bernstein 2016), percolating into second-party enforcement in which reputation sometimes lies at the center of analysis (MacLeod, 2007). People can value their reputation simply because that is the lens through which they view their intrinsic selves and they preserve their reputation because that is what they value, implying aspects of first-party enforcement. Simply put, reputational considerations can be a tool of all the agents invoked in our question and our use of the term would loosen, rather than tighten the wording of the question.

Similarly, networks can be invoked in many different senses. The most common definition of a network is a set of agents with some characteristics in common, each agent potentially having a connection with some other agents. Under this definition, a network might play many kinds of roles in transactions. Networks can be used to operationalize reputation in enforcing agreements (e.g., World Bank 2002, Bernstein 2016, Argyres et al. 2020). They also transmit information about law, people, markets, goods, and technologies that serve alternative purposes, such as finding suppliers or forming norms. These additional elements are important, but they are too far removed from opportunism and, our focus: the central task of enforcement.

Avoiding the terms 'reputation' and 'networks' shows why we did not attempt to combine the agent-based and strategy-based approaches in a single survey question. Elements of reputation and networks are crucial in relational strategies. But, as we note above, reputation can be used in some way by each of the six agents we focus upon. Networks have a similarly broad scope of application in relational strategies. To use these terms to find which of the six agents are involved in relational, rather than contractual, agreements would lead to a much too complex question.

We now turn to the decisions made in choosing the precise wording used in the question.

II.2.2 Question Wording

In conducting economy-wide, cross-country research using only one composite survey question, it is inevitable that survey administration would piggyback on existing efforts.[27] Our questions were asked within a regular cycle of the World Bank Enterprise Surveys (WBES, described in more detail in the next subsection). The form and phrasing of the questions were therefore framed to fit the approach of the WBES, an important element of which is asking busy top managers or owners of firms a large set of questions on a collection of heterogenous topics. This means that questions must be terse and understood in a natural way by businesspeople, leading to budget constraints on words and concepts embodied in questions.[28]

A first decision was to use 'agreements', and not 'contracts', since contracts are only one type of agreement. Then one must decide whether to ask firms about their agreements with transactional partners in general, or to ask about one specific partner, since the transaction-cost and contract-theory literatures suggest that relationships may differ across different partners. Focusing on a specific partner is conceptually and practically complicated. First, the question must suggest to respondents which specific partner they should focus upon.[29] A possible solution is asking respondents to choose a partner when answering questions, as when Mike and Kiss (2019)

[27] For example, the construction of an economy-wide sampling frame for any particular developing country is often a challenging, burdensome and expensive task, which requires considerable resources and expertise.

[28] Terseness in wording has benefits for the whole survey. Longer questions cause problems because they place a greater cognitive burden on respondents (Yan and Tourangeau 2008). This affects respondent comprehension (Holbrook et al. 2006), adding complexity without altering the distribution of responses (Shaeffer et al. 2005), leading to respondent satisficing and decreased measurement reliability (Alwin and Beattie 2016). It also leads to a potential bias as respondents with more cognitive resources will answer complex questions differently than those with fewer (Oberski 2016). Contrast our terse questions with those of Hendley and Murrell (2003), on which they are based, but which faced fewer constraints on question length as their survey covered only one country, did not use a representative sample, and could fully focus on issues of transactional governance and the law.

[29] This approach originated in McMillan and Woodruff (1999) who asked respondents to reflect on their firm's first and most recently added customers and suppliers. However, McMillan and Woodruff (1999) were interested in exploring the determinants of trade credit, for which a focus on highly specific transactions is necessary. Note that a case-study approach can be applied as well. For example, Mahapatra et al. (2010), chose the organizations whose interactions they studied in great detail. However, this approach would be impossible to undertake on the scale required for the current investigation, which is economy-wide in multiple countries.

suggested choosing a 'typical' supplier and buyer.[30] But this solution invites selection bias: Mike and Kiss (2019) found that respondents equated typical business partners with those having longstanding ties, implying that the resultant dataset characterized relationships in which familiarity between partners was a strong component. Conceivably, such selection bias could be avoided by working with each firm to generate a list of its partners and then choosing one randomly. But this would increase the length and complexity of questions and the survey implementation overall. Asking firms about their agreements with transactional partners in general avoids these types of conceptual and practical complications, reflecting the plausible assumption that the approach taken by a firm will transcend one specific relationship and be based on a firm's overall comparative advantage when choosing a specific government structure.

Moreover, the WBES surveys establishments (which we have called 'firms', for convenience). The WBES is meant to broadly capture firms' experiences, and so it includes questions on a variety of topics, which are generally pitched at the level of overall operations. This provides a fitting context in which to embed a question about agreements with transactional partners in general. Importantly, establishments usually produce a narrow range of products with a uniform technology. Hence, our approach is analogous to that taken in many studies where establishments are viewed as a single entity with a single technology, one example being the estimation of establishment-level production functions. We therefore use a level of aggregation that is common in the economics and business literatures by posing questions about the establishments' transactional partners in general. Finally, our question was preceded in the WBES questionnaire by questions that specifically asked respondents to focus on the whole set of transactional partners.

[30] The questions could be made even more granular by asking about a specific transaction, though the strategies that firms employ to ensure fulfillment of agreements in a transaction likely transcends the particular transaction at hand (e.g., Parmigiani and Mitchell 2010), diminishing the value of such granularity.

Therefore, respondents begin to answer our question on supplier (or customer) relationships having been primed to think about the circumstances under which they engaged with suppliers (or customers) in general.[31]

Next, in framing a question, one must decide which aspect of agent performance to ask respondents to assess.[32] We chose to focus on 'resolving or preventing problems in agreements' because this was a circumscribed objective that all respondents would quickly understand.[33] It directly resonates with the first element of our conceptual framework that emphasizes enforcement in its broadest sense and implicitly invokes notions of good overall transactional performance, including through the deterrence of opportunism.[34] Importantly, we do not constrain ourselves to enforcement in the sense of specific performance of the terms of an agreement, since uncertainty invariably means that efficiency would necessitate adjustments in those terms during implementation. Nor is our focus on actual disputes: the use of 'preventing problems' prompts respondents to include information about cooperative adjustments that are made under the shadow of enforcement without any real dispute explicitly arising. In this sense, our approach might be better summarized by the term 'safeguarding', than the more common 'enforcement'.[35]

In combining an inquiry about the resolution and prevention of problems, we have introduced unfortunate 'multibarreled' phrasing, where a question inquires about two or more concepts that

[31] Our question on customer relations was identical to the one for suppliers, except for the obvious substitution of 'customers' for 'suppliers'.
[32] In the literature reviewed by Cao and Lumineau (2015), respondents were asked a series of questions about different properties tailored to different transactional mechanisms. While this approach is suitable in many areas of research, our overall objective dictated that all transactional mechanisms be treated in exactly the same manner. Thus, for example, we could not model our approach on the methodologies of any of the voluminous number of papers reviewed by Cao and Lumineau (2015).
[33] Therefore, we do not inquire about the selection of partners, which occurs before agreements are made. See, for example, Dekker (2008) on how the amount of effort invested in partner selection affects the choice of governance structures.
[34] The deterrence effect of enforcement is difficult to quantify directly for multiple reasons, including due to respondents' potential guesswork as to why problems did not arise. Given our conceptual framework's focus on enforcement in its broadest sense, we consider it crucial to attempt to measure the deterrence aspect of enforcement.
[35] See Zhou and Poppo (2010), Carson, Madhok and Wu (2006), Abdi (2017), Devarakonda, McCann, Reuer (2018).

are not clearly separated.[36] The relative focus on either one of these two aspects might vary non-randomly between different types of respondents, introducing bias. This is a shortcoming but one that would be prohibitively costly to avoid. In an ideal theoretical scenario, we would have asked questions about resolution and prevention separately. In practice however, there are at least two important complications. First, the two properties will be difficult for respondents to disentangle. A mechanism that prevents complications from evolving into problems can also be viewed as a way to resolve issues. A mechanism that resolves problems, perhaps with the explicit threat of sanctions, can also prevent problems from even arising because of the deterrence effect. Indeed, 'preventing' can be thought of as being produced by the shadow of 'resolving'. Second, respondents might view two separate but very similar questions (on preventing and resolving) as repetitive, which reduces the overall quality of a survey.[37] Third, piggybacking on the regular cycle of the WBES meant a strict constraint on the total number of questions that we could ask. Given these conceptual and practical considerations, we chose to ask about both resolution and prevention in one question.

A further issue revolved around the phrasing of the term to connote success when resolving and preventing problems. We chose to use a single, simple term 'effective'.[38] We did not ask respondents to focus upon whether agreements work in an efficient way because this is too demanding a criterion and probably one that is less familiar for business managers than economists. In using this word, we avoided weighted terms such as importance, usefulness, or

[36] Survey methodologists suggest avoiding the use of such combinations of terms, to allow respondents to focus on one aspect of an inquiry without making a choice of which aspect to focus upon (Fowler and Cosenza 2008).
[37] Survey methodologists warn against such repetitive design since it contributes to respondents' survey fatigue and affects their responses through the rest of the questionnaire, and even their overall participation.
[38] In the papers reviewed by Cao and Lumineau (2015), equivalent notions are satisfaction with the exchange relationship and performance of the partner in the exchange. These notions are entirely analogous to our 'effective', although measured in a variety of different ways in the papers reviewed. Poppo and Zhou (2016) refer to the notion of "harms" to exchange performance, which too is analogous to our 'effective'.

frequency of either use or occurrence, even as these elements, and perhaps more, enter in determining effectiveness.[39] The terms "important" or "useful" carry the meaning of value and can be prone to more subjective interpretation than the term "effective". Frequency of use or occurrence does not necessarily indicate effectiveness since, for example, a legal system with high level of effectiveness works through its shadow rather than through explicit use of the courts.[40]

Final decisions on the form of questions reflected information gained in the field at various stages of survey implementation, during translations and scripting, while training local surveyors, and from a 'cognitive audit' of the questions. These stages are described in detail in Appendix A.3, and Appendix A.7 reports the (negligible) rates of "Don't Know" responses.[41]

II.3 The Surveys and Raw Responses

Our question was posed to owners and top managers of a total of 3,430 firms in Argentina, Bolivia, Ecuador, Paraguay, Peru, and Uruguay, as part of a regular cycle of the WBES in 2017 and 2018.[42] The WBES samples of firms are designed to be nationally representative of non-

[39] Hendley and Murrell (2003), whose survey question is adapted in this paper, mention all these terms but also provide detailed clarifications to respondents, which was not possible in the context of the WBES.

[40] It is conceivable that respondents could confuse the term 'effective' with other terms we avoided, such as importance, usefulness, or frequency of use or occurrence. However, as noted below, various stages of survey implementation included a 'cognitive audit' of our survey question, which is described in more detail in Appendix A.3. Interviewees did interpret the term 'effective' as referring to a combination of both the extent of reliance and productiveness of the enforcement agent: the lowest ranking on the scale would be chosen if either the enforcement agent was not relied on, including due to lack of knowledge about that possibility for enforcement, or if it had been relied upon and it was ineffective.

[41] Low rates of "don't know" responses may be suggestive of good comprehension of questions by respondents (Krosnick 1991). The questions used in this paper had frequencies of don't knows that were comparable to those questions that had proven satisfactory by their staying power through many cycles of the WBES surveys. Krosnick and Presser (2010) note that frequent selection of a middle point on a scale (our 'moderately') is evidence of satisficing behavior on the part of respondents. There is no evidence in our data that that the mid-point is used frequently. Given the negligible rates of "don't know" responses, we omit from our analysis any observation that has at least one "don't know" as a response to any of the set of six sub-questions.

[42] While these countries are quite similar from a global perspective, there are important differences across them in terms of economic development, quality of institutions, natural resource endowments, etc. Table A.1 in Appendix A.5 lists some standard statistics on these six countries, such as GDP per capita, measures of rule of law, trust and fractionalization, together with regional and global comparisons. See Appendix A.6 for detailed timing of each survey.

agricultural, non-extractive, formally registered, private firms with at least five employees.[43] The survey used a stratified design with simple random sampling.[44]

Table 1 displays the raw response percentages for the six sub-parts of each of the two identical questions, one on suppliers and one on customers. Respondents tend to regard first- and second-party enforcement (i.e., trust and mutual interest) as most effective, while regarding others as less effective.

III. Latent Class Analysis: Data-Generating Process and Diagnostics

This paper uses the data collected in response to our question in order to undertake a cross-country characterization of transactional governance structures. Much analysis could be conducted using the types of information in Table 1. However, a fuller understanding is achieved by analyzing response patterns across all six sub-questions in combination: the governance of transactional relations likely involves the coordinated use of different agents. As noted by Cannon et al. (2000: 184), "exchange is best understood as embedded in a complex matrix of economic, social, and political structures and…the governance of exchange relations more often relies on combinations of market, social, authority-based mechanisms than on any one category exclusively." However, "theory alone cannot predict which system of societal organization—the integrated individualist system based on second-party enforcement, formal organizations, general morality, and legal

[43] The survey universe comprises the following sectors as per ISIC Rev. 3.1: the entire section D covering manufacturing (two-digit codes 15 through 37), section F covering construction (code 45), section G covering wholesale and retail trade as well as repair of motor vehicles, motorcycles and personal and household goods (codes 50 through 52), section H covering hotels and restaurants (code 55), and section I covering transport, storage and communications (codes 60 through 64). Formal registration usually refers to registration with tax agencies or in general firm registries. The universe includes partially privately owned firms and excludes fully state-owned firms. The sources of the sampling frames for these WBES's were: for Argentina – Dun & Bradstreet 2016, for Bolivia – Economic Census, updated by Encuestas y Estudios consulting group 2016, for Ecuador – Superintendencia de Compañías Valores y Seguros del Ecuador 2016, for Paraguay – Directorio General de Empresas y Establecimientos 2015, for Peru – the combination of Top 10mil 2011, Registro Mype Callao 2010, Registro Mype 2012 and SUNAT (Hacienda) 2011, and for Uruguay – Instituto Nacional de Estadística 2015.

[44] Full details of the methodology can be found at http://www.enterprisesurveys.org/methodology. Strata were based on firm size, geographical location, and economic sector. The data includes sampling weights. All results are obtained using these weights and thus refer to the entire population of pertinent establishments in the six countries.

enforceability, or the segregated collectivist system based on personal trust and third-party enforcement—is more efficient in promoting exchange." Greif (1997: 253) The main goal of this paper is to fill this gap, not with theory, but rather by using induction from data.

Many studies have explored, theoretically or empirically, whether various types of enforcement substitute or complement each other (e.g., McMillan and Woodruff 2001, Poppo and Zenger 2002), often concluding that context is important (e.g., Cao and Lumineau 2015). But there is a daunting number of possibilities when combining different types of enforcement. With our six sub-questions, each with five categories of responses, there are 15,625 possible response patterns (5^6). The actual responses do include far fewer combinations than this. But even so, for relations with suppliers, we observe 711 distinct combinations and for relations with customers 631.[45] This confirms the theoretical prediction of the coexistence of many different combinations of transactional governance approaches rather than a single one that works across all contexts (Greif 1997, Mike and Kiss 2019). Given that responses are spread over hundreds of distinct patterns, statistical methods are necessary to gain the types of insights that arise from parsimonious summaries of the data. These insights and more are provided using LCA.

We introduce our statistical methodology in fine detail, because LCA is not commonly used to study firm behavior.[46] It is assumed that there are different types, or classes, of firms; but, it is unknown to which class any firm belongs. LCA leverages patterns in the responses to a multi-part survey question like ours and estimates the probability that each firm lies in each class. Parallels with factor analysis are immediate, with the key difference that factor analysis estimates

[45] Appendix A.8 includes tables listing the most common reported response patterns.
[46] LCA has been applied in a number of areas of business and economics, but its use is not common in any area, with the arguable exception of marketing. (See, e.g., Kamakura and Russell (1989) for an early empirical application and Jedidi et al. (1997) for a methodological discussion.) The few existing examples of the application of LCA to data on producers tend to focus on niche activities. Grewal et al. (2006) focus on individual software developers, Samuelsson and Davidsson (2009) study nascent entrepreneurs, Ebers and Oerlemans (2016) use a small sample of firms in the German construction industry, and Mike and Kiss (2019) study a sample of privately owned small and medium sized enterprises in Hungary.

continuous, cardinal latent variables, while LCA estimates a nominal latent variable, the categories of which correspond to classes. When the estimated classes are meaningful and recognizable in the context of the existing literature, the analysis is especially illuminating. In terminology that is becoming increasingly standard, LCA is an unsupervised, machine-learning classification procedure.

III.1 The Simplest LCA Data-Generating Process

In the most basic data-generating process (DGP) envisioned by LCA, the set of responses of firms of a similar type, or class, forms an unobserved probabilistic contingency table.[47] Since there are different classes of firms, a complete data set comprises mixtures of contingency tables, with unobserved mixing proportions. LCA separates the mixture into its constituent parts by estimating the contingency table of each class and the class of each firm.

Each firm, i, reports on the effectiveness of K enforcement agents by choosing, for each K, one of R points on a Likert scale.[48] The response vector, $Y_i = (y_{i1}, \ldots, y_{iK})$, is observed. In our data, $K = 6$ and $R = 5$.[49] In its operations, firm i chooses one of several governance structures (i.e., combinations of agents). The choice of governance structure places the firm in one category of a nominal and latent variable, c, with C values. c_i denotes the value of c for firm i: $c_i = j$ if firm i chooses governance class j. c_i is to be estimated for each firm.

Denote by $\theta_{kr|c}$ the probability that a firm in latent class c chooses answer r concerning agent k. Denote by π_c the probability that a firm is in latent class c. Then the most basic DGP assumes that the probability of observing a specific vector of responses, Y_i, for firm i is:

---

[47] Appendix B.1 provides an intuitive introduction to LCA, emphasizing the contingency-table data-structure and providing a simple example framed in this paper's context.

[48] Consistent with our conceptualization of transactional governance, as detailed in Section II, we refer to enforcement agents, or simply agents, rather than mechanisms or strategies, terms that fit better with the strategy-based approach. Of course, the terms 'agent' and 'enforcement' are to be understood in the broad sense in which we have defined them in Section II.

[49] We observe separate Y_i's for relations with customers and relations with suppliers, but since we keep the analysis of each type of relation entirely separate, we use only one Y_i in specifying the DGP.

$$P(Y_i) = \sum_{c=1}^{C} \pi_c \prod_{k=1}^{K} \prod_{r=1}^{R} [\theta_{kr|c}]^{I(y_{ik}=r)} \quad (1)$$

where $I(y_{ik} = r)$ is an indicator function equal to 1 if $y_{ik} = r$, and 0 otherwise. The $\theta_{kr|c}$ and π_c are to be estimated. By multiplying the $\theta_{kr|c}$'s to obtain the probability of observing a contingency table of each class c, the DGP of (1) assumes that, conditional on class, c, the response of a single firm to any sub-question k, y_{ik}, is independent of the response to any other sub-part m, y_{im} for all $k \neq m$ and all i. This is the local independence assumption in LCA, which is invoked in the overwhelming majority of LCA applications. The term 'local' indicates that independence is conditional on class c, which is very important here. Local independence is the assumption that the correlation across sub-questions of a firm's responses is only due to economic fundamentals, which are fully reflected in class membership. To the extent that such correlations arise for reasons other than the fundamentals—for example, because of cognitive problems in distinguishing between sub-questions—then local independence is violated.[50] If such a tendency for pairs of responses to partially contain the same extraneous information is not considered, then LCA will give correlated responses too much weight (Vermunt and Magidson, 2002: 95). The DGP should be modified to take this into account: the analogy to clustering in OLS is immediate.

III.2 Relaxing Full Local Independence

Exceptions to assuming local independence between all *K* sub-questions can be built into the DGP by replacing the simple multiplication of probabilities (the $\theta_{kr|c}$'s) in (1) with a specification of the joint distribution of responses. Split the *K* agents into *H* subsets. Assume that the responses

[50] For example, when somewhat similar enforcement agents are probed in succession, respondents might not exert the cognitive effort to distinguish responses between these agents: the errors in consecutive sub-questions are possibly correlated This phenomenon is often referred to as "nondifferentiation of responses" (e.g. Krosnick and Alwin, 1988). For example, in the piloting of the survey, we found that respondents had some difficulty in separating the notions of mutual interest and personal trust. This is not surprising. One of the takeaways from Poppo et al. (2016) is that there have been few efforts in the past to distinguish empirically between these two different types of trust.

about any two agents within the same subset are correlated even conditional on the latent class of the firm (c_i). Let Y_{ih} be the vector of firm i's responses on the agents in the h^{th} subset, with each Y_{ih} being a sub-vector of Y_i. Y_{ih}, $h=1,...H$, is observed. Denote by $f(Y_{ih}|c)$ the pdf of Y_{ih} given c. Then the probability of observing a specific response vector, Y_i, for firm i is:

$$P(Y_i) = \sum_{c=1}^{C} \pi_c \prod_{h=1}^{H} f(Y_{ih}|c) \tag{2}$$

Estimates of $f(\cdot \mid \cdot)$ and the π_c are then obtained by maximizing the following likelihood:[51]

$$\sum_i w_i \, P(Y_i) = \sum_i w_i \sum_{c=1}^{C} \pi_c \prod_{h=1}^{H} f(Y_{ih}|c) \tag{3}$$

where the w_i denote sampling weights. Use of the sampling weights implies that our estimates are representative of the entire universe of firms covered by the WBES.

LCA also facilitates studying which factors are associated with the class to which a firm belongs. This is Williamson's (1991) discriminating-alignment research agenda. Then, the π_c in (1), (2), and (3) are viewed as functions of observed firm characteristics. However, it is possible to estimate the classes themselves before delving into the specification of and estimation of those functions. Appendix B.2 outlines the details, presenting the logic of our empirical approach in first estimating the classes (or governance structures) using (3) and then relating the classes to characteristics of firms and their environments. Results from the former appear in Section IV and from the latter in Sections V and VI.

III.3 Selecting a Specific LCA Model and Evaluating Its Properties

LCA is typical of latent-variable models in that implementation requires making detailed specification decisions before estimation of a specific model.[52] We follow standard model-

[51] We use the Latent GOLD software (Vermunt and Magidson 2016).
[52] Detailed background information can be found in Collins and Lanza (2010), Masyn (2013), and Vermunt and Magidson (2016).

selection steps, intentionally separating this process from the analysis and interpretation of findings. Thus, we see none of these decisions affecting the core substantive findings in any way, since their aim is simply to improve the quality of estimation.

In most practical applications of LCA, the choice of how many classes (*C*) to estimate is primary, with a relatively small number of options considered. In contrast, relaxing local independence enormously increases the number of possible options when specifying the DGP. This probably explains the fact that the vast majority of practical applications of LCA assume full local independence and do not, unlike this paper, question whether this assumption should be relaxed for some specific set of responses.

Settling on the details of model specification moves through three stages. First, statistical measures of model-fit are used as criteria to choose a very small set of satisfactory models. Second, the results from those models are evaluated using more subjective criteria. For example, parsimony is important to avoid over-fitting and to facilitate meaningful interpretation, which usually means the use of a simpler model (with fewer parameters). Third, the researcher examines class homogeneity and separability. Homogeneity is the notion that the members of a specific class exhibit similar characteristics or, equivalently, that there are certain configurations of responses typifying each class. Separability captures whether each class looks quite different from all other classes, or, equivalently, that there are certain configurations of responses that distinguish each class from the others. In all three stages, statistical measures are used as model-selection criteria.

To begin the model selection process, it was necessary to make three preliminary decisions. First, we chose to analyze the responses for suppliers and customers separately. We presume that firms might employ very different approaches in conducting upstream relations than downstream ones: after all, in dealing with suppliers, a firm's primary objective is timely delivery at an

appropriate level of product quality while, in contrast, downstream relations are primarily about getting paid by a satisfied customer. The large samples meant that sufficient statistical power could be generated in two separate analyses.[53]

The second decision was to choose the numbers of classes (*C*) to be considered. Invoking parsimony, we focused on 3-, 4-, 5-, and 6-class specifications. Robustness results outlined in Appendix C show that considering even more options would not have affected our analysis.

The third decision was on particular choices for the relaxation of local independence. Relaxing local independence results in a proliferation of design possibilities. The six agents can form 15 unique pairs, with 32,766 distinct combinations of these pairs possible.[54] We thus chose to focus on situations where theory and the observations from survey implementation pointed us. As mentioned previously, respondents might not clearly distinguish between trust (sub-question 1) and mutual interest (sub-question 2). In the piloting of the survey, we learned that sometimes individuals did not clearly distinguish the assistance of government officials (4) from the intervention of other third parties (5). As a result, it was natural to consider specifications that relaxed independence between the responses on sub-questions 1 and 2 and those on 4 and 5.

Considering these possibilities for the relaxation of local independence and the four possibilities for the number of classes, there were already 16 models to consider. With this starting point, we conducted an empirical exploration of whether there was a need to relax the local independence assumption on other pairs of answers to the sub-questions. To accomplish this, we estimated the 16 models and examined the size of bivariate residual correlations, a measure of the

[53] Compare Mike and Kiss (2019), which merges data from the two types of questions, and Hendley and Murrell (2003), which does not differentiate between upstream and downstream. Note that our data describe relations of firms with the full set of their suppliers and customers, and these sets likely differ, meaning that there is no reason to believe that a firm's strategy on the customer side is the same as on the supplier side. Moreover, it is easy to construct examples where two firms engaged in a transaction use different strategies.

[54] $32{,}766 = \sum_{k=1}^{14} \binom{15}{k}$. For identification, LCA needs at least one pair of sub-question responses to be locally independent; hence the 14.

marginal increase in the log-likelihood function that could be obtained by removing the local independence assumption for any specific pair (Vermunt and Magidson, 2016: 83-5). The pairs that had particularly large bivariate residual correlations were then added to the set of pairs that were candidates for non-imposition of local independence.[55] As a result, we chose to consider a total of 20 models for each of supplier relations and consumer relations.[56] The selection of the best model was based on several criteria, particularly the Bayesian information criterion, the consistent Akaike information criterion, and the approximate weight of evidence criterion.

Providing further details of the model selection process is necessary for completeness, but understanding these details adds little for those readers focused primarily on the paper's central substantive findings. Interested readers can consult Appendix C for details regarding the set of models considered, the statistical criteria used to compare the performance of the different models, the final choice of the preferred models, and robustness analysis.

IV. The Estimated Model: Class Characteristics

For each of supplier and customer relations separately, the process of model selection led us to estimate four latent classes. That is, we conclude that four governance structures adequately describe the choices that firms make when combining the six enforcement agents.

Assigning names to the classes is a crucial, substantive element of the analysis because important insights are generated only if LCA uncovers readily recognizable types of governance structures. Finding resonance between our estimates and existing ideas and concepts provides additional validation of the analysis.

---

[55] The exploratory results detailed in Appendix C suggested that answers to adjacent sub-questions were related. As already noted, this may be due to nondifferentiation of responses. This effect is known to be smaller in face-to-face surveys (Holbrook et al, 2003), which may be a reason that only adjacent responses were correlated in our data. These correlations might also reflect an "anchoring effect" (e.g. Furnham and Boo, 2011), where a subsequent response is biased towards a previously selected response.

[56] In particular, in addition to the full local independence and three correlation structures mentioned above, we allowed the correlations of 1-2, 2-3, 3-4, 3-5, 4-5 for supplier relations, and 1-2, 4-5, 4-6, 5-6 for customer relations, for each of the four possibilities of class size.

IV.1 Characteristics of Chosen Models

The naming of classes builds primarily on an examination of how the estimated response probabilities vary across classes. These are the $\hat{\theta}_{kr|c}$, the estimated probability of choosing response *r* for agent *k* if the firm is in class *c*.[57] Tables 2a and 2b list these estimated probabilities, with accompanying graphical depictions of the estimated governance classes. In the tables, classes are labeled in two ways, by a number and a name. The numbers are artefacts of the estimation process; the class names are evocative labels provided by us and justified below.

The probabilities are precisely estimated. (The tables report standard errors.) Most non-zero estimated probabilities do not lie in the 95% confidence intervals of either their vertical or horizontal neighbors. Thus, easily discerned differences in the figures are almost certainly statistically significant.

IV.2 Class Names: Characterizing the Classes

We believe the following names capture the characteristics of the estimated governance structures:

	Relations with Suppliers	Relations with Customers
class 1	Pure bilateralism	Pure bilateralism
class 2	Bilateralism with private support	Bilateralism with private support
class 3	Bilateralism with legal support	Bilateralism with weak support
class 4	Strong comprehensive governance	Weak comprehensive governance

The nature of class 1 for both upstream and downstream relations is transparent and is the same for both types of relations. Only trust and mutual interest (first- and second-party enforcement) are endorsed as effective. For firms that rely solely on trust and mutual interest and hardly use any third party, enforcement involves only the actions of the two parties to the agreement themselves.

[57] The $\hat{\theta}_{kr|c}$ do not appear explicitly in the DGP in equation (2), which relaxes full local independence and which is the one we use. Therefore, $\hat{\theta}_{kr|c}$ should be interpreted here as the marginal probability that a firm in class *c* chooses answer *r* on question *k*.

We thus call both class 1's *pure bilateralism*. The use of 'pure' is emphasized as a contrast to the remaining classes, which differ primarily in what they add to first- and second-party enforcement. Given this, it is tempting to conclude that the classes are ordered from the narrowest to the broadest, implying that each subsequent class adds a component to the previous one. However, LCA does not impose the assumption of ordered categories: the classes do not necessarily lie along any single dimension.

In contrast to pure bilateralism where only two agents are found effective, all agents contribute significantly to both class 4's. For upstream relations, firms in this class find the legal system as effective as any other agent, with government officials and other third parties at least as important as each of the two bilateral agents. Because every single agent is rated as effective as in every other class, we use the label *strong comprehensive governance*. However, for downstream relations *weak comprehensive governance* is more appropriate for class 4, given that within this class all agents are less effective for customer relations than for supplier relations.

In class 2, first- and second-party enforcement are as important as in class 1, but all third-party agents also have a notable presence. Among the third parties, paid, private dispute resolution is the most important, followed by the legal system. This is consistent with the common practices of private third parties: arbitration is backed by the legal system; debt-collection firms harass with legal threats. We thus use the name *bilateralism with private support* for class 2, but only brevity precludes mentioning the additional role of the legal system.

There is a marked difference between upstream and downstream governance structures only for the class 3's, albeit with some similarities. For both supplier and customer relations, there is a contribution from first- and second-party enforcement, but this contribution is weaker than for all the other three classes within either upstream or downstream relations. Third-party agents have a

significant presence for supplier relations, with the legal system being the most important among the third parties, followed by paid private dispute resolution. That is, for supplier relations, class 3 differs from class 2 primarily in the relative emphasis on law and paid private dispute resolution. We thus use the name *bilateralism with legal support* for class 3 on the upstream side, but only brevity precludes mentioning the role of paid private dispute resolution.

While first- and second-party enforcement are weakly present in class 3 of customer relations, the contribution of the third parties is even weaker. Thus, we name this class *bilateralism with weak support*, recognizing that among all eight estimated latent classes, this is the governance structure where the aggregate effect of all 6 agents is rated lowest by respondents. Compared to other classes, the label 'ineffective governance' might also be appropriate.

IV.3 What Has Been Learned About Transactional Governance?

Table 3 presents estimates of the proportion of firms placed within each class (the $\hat{\pi}_c$ from LCA). Before discussing the prevalence of each estimated class, it is worth revisiting our conceptual framework and emphasizing that we have successfully accomplished our goal of using an exploratory statistical approach to uncover the governance structures that firms use for their transactions, while estimating the proportions of firms that use each of the structures. We know of no other study that has undertaken such empirical analysis in an economy-wide, cross-country framework, while considering a full spectrum of enforcement agents.[58] In setting and pursuing this

[58] As noted in sections I and II, we are aware of only four studies that are somewhat similar, Hendley et al. (2000), Hendley and Murrell (2003), Mike and Kiss (2019), and Fafchamps (2004). However, each of the first three cover only a single country, and the fourth uses data collected in a way that makes analysis of combinations of agents very difficult. The closest to our analysis is Mike and Kiss (2019), but it is on a single country and does not use a nationally representative sample of firms. They also collect responses on six mechanisms of enforcement, but use a strategy-based rather than an agent-based approach. Three of their sub-questions can be directly mapped to three of ours, namely first- and second-party enforcement, and the legal system. The remaining three are 'community norms', 'personal reputation', and 'impersonal market reputation' that do not have direct agent-based equivalents in our survey question. Nevertheless, their estimated three governance structures (using LCA) are somewhat similar to ours, with similar names: (i) their 'bilateral' is equivalent to our 'pure bilateral'; (ii) 'comprehensive' is equivalent to our 'strong' and 'weak' comprehensive; and (iii) 'third-party' involves the three strategies that do not have a clear agent-based equivalents but the estimated class seems closest to our 'bilateralism with legal support'.

goal, our methodology has reflected Greif's (1997: 253) concern that theory alone cannot predict which systems of the enforcement of agreements to exchange would be used in any particular society.

Looking at the prevalence of each estimated class, pure bilateralism dominates, accounting for nearly two-thirds of governance structures on the supplier side and more than half on the customer side. This result alone has significance. Any cursory reading of the literature would convince the reader that economists' priors on the importance of pure bilateralism are extremely diffuse.[59] For example, much attention in the literature is devoted to analysis of enforcement by third parties, such as social or merchant groups, or communities at large. While the inner workings of third-party enforcement are often investigated in detail, they less often analyzed as one of several enforcement options that firms may choose between. Moreover, the separation between first- and second-party enforcement often appears, implicitly, in the divide between different branches of the literature (e.g., in sociology and economics, respectively), not directly presented in the specific contexts of enforcement of business agreements between firms.

In general, firms rate their governance structures as less effective for customer relations than for supplier relations. (This follows from a quick visual comparison of the figures accompanying Tables 2a and 2b.) This characteristic is epitomized in the two class-4 names—*strong comprehensive governance* and *weak comprehensive governance*.

The estimation and naming of the classes not only reveal which governance structures are important, but also which are absent. All governance structures rely, at least in part, on bilateralism: no firm relies solely on a combination of third parties and formal institutions. This is

[59] One indication of the diffuseness of priors on this result is recent work on global value chains (GVCs). World Bank (2020) provides a "novel, relational conceptualization of GVCs" where "the identity of the agents participating in a GVC is crucial". Our identification of pure bilateralism as being the most important governance structure is completely consistent with that novel conceptualization.

also a finding of Mike and Kiss (2019) in the data from Hungary, leading the authors to state that "Law never stands alone", but it is inconsistent with many claims in the literature that characterize development as a process of escaping personalized interaction and moving to a rule-based, impersonalized set of interactions.[60] A very widely cited version of this view can be found in Peng (2003: 276), which claims that the most important transition for emerging economies is the process of moving "from a relationship-based, personalized transaction structure calling for a network-centered strategy to a rule-based, impersonal exchange regime". While our data do not capture the process of development, they do reflect countries at different levels of development, and no estimated governance structure involves only rule-based, impersonal transactions.

A corollary of this pattern is that bilateralism and the legal system should not be viewed as purely substitutes. In several of the classes they play complementary roles, and there is no obvious case where a move from one class to another will involve a significant decrease in bilateralism simultaneously with a large increase in the perceived effectiveness of the legal system.[61] This pattern goes against a view that long held a dominant position in the literature, which considered the use of formal legal arrangements for transactions as inconsistent with the use of personalized relationships based on trust: formality eroded trust. This view became less dominant after Poppo and Zenger's (2002) seminal contribution and is perhaps now a minority view (Cao and Lumineau 2015, and Poppo and Cheng 2017). By providing economy-wide and cross-country evidence, our results add to the literature that has argued that trust and formal contracts are most often complements.

[60] Mike and Kiss (2019) characterize this as the classical view, and give many references to its use.
[61] A move from pure bilateralism to the class 3's (bilateralism with legal or with weak support for suppliers and customers, respectively) does involve a decrease in bilateralism coupled with an increase in the use of the legal system, but these changes are rather weak (especially the change in bilateralism), suggesting partial substitution at best.

Paid private dispute resolution and the legal system are sometimes substitutes and sometimes complements. For example, for supplier relations, when moving from the pure bilateral class to any of the three other classes there is an increase in the importance of both paid private dispute resolution and the law. But, as indicated by their very names, a move from bilateralism with private support to bilateralism with legal support indicates substitutability.

Two further facets of the results contrast with extant emphases in the literature. First, the literature almost entirely neglects the role of government officials in supporting transactional governance. This role may entail the use of regulations, inspections, fines, or even aspects of criminal justice. The work by Hendley and Murrell (2003) is a rare exception to the lack of emphasis on the role of government in this context, though that paper's attention to this role is linked to the focus on the transition of Romania from socialism. Our data suggests that government officials do sometimes support the enforcement of agreements, even if that role is minor. In the *strong comprehensive governance* class of supplier relations, government officials are prominent.

Second, given the amount of attention paid in the literature to non-paid private third parties, especially networks, one would expect them to be more important. Network approaches have been very popular in recent decades, following Granovetter's (1985) emphasis that transactions are embedded in a broader social structure, the historical-theoretical analysis of Greif (1989), and case studies on the importance of trading networks in varied settings (Bernstein 1992; Landa 1981). However, as we point out in Section II, the term 'network' is used in diverse ways in the literature, some of which would imply that a network is present even when the reference is only to a large grouping of similar agents who are not necessarily invoked in any type of enforcement process. In fact, our results suggest that, in aggregate, non-paid private third parties are arguably the least important of the six agents in enforcement activities (Tables 1, 2a, and 2b).

V. Variations in the Importance of Governance Structures

In this section, we examine how governance structures vary with the characteristics of firms or their environments. We simply explore patterns in the data and do not attempt to isolate ceteris paribus, causal effects of single variables. To illustrate the type of thought experiment explored in this section, consider the prevalence of governance structures across the different countries where firms are located. We show the change in prevalence that would result if firms migrated from one place to another, simultaneously going through all the other changes associated with the differences between firms in the two places. As well as generating interesting descriptive results, the exercises in this section provide evidence on the validity of the data we have produced: if there were no significant association between chosen governance structures and firm characteristics, the meaningfulness of our estimates could be questioned. Section VI takes an additional step, producing evidence suggesting a causal effect of one determinant of governance structures, further validating the data and exemplifying its use.

We examine the correlates of the governance structures by implementing a maximum likelihood procedure developed by Vermunt (2010), Bakk et al. (2013), and Bakk et al. (2014). This procedure provides consistent estimates of the parameters defining the $\pi_c(.)$. The core intuition is simple: regress the estimated posterior probabilities of each governance structure on the Z_i of interest. However, using OLS leads to systematic underestimation of the strength of association (Bolck et al. 2004). The maximum likelihood procedure corrects this deficiency.[62]

[62] Mike and Kiss (2019) also conduct an exploratory analysis of the determinants of class membership. However, they employ a different statistical methodology than we do. Their methodology is described in Appendix B.2, where the relation between their and our own methodology is explained. In that appendix, we also provide a justification for our choice of methodology.

V.1 Cross-Country Variations in Governance Structures

We find a notable and statistically significant variation in the prevalence of governance structures across countries. Table 4 reports the p-values of Wald tests of the null hypothesis that there is no association between governance structure and country. For both upstream and downstream relations, the Wald tests indicate rejection at the 1% level. Figures 1a and 1b illustrate the cross-country variation. In these figures, darker shading indicates governance structures that are more complex, that is, those in which more third-party agents have a higher measure of effectiveness (reflecting the $\hat{\theta}_{kr|c}$ in Tables 2a and 2b).[63]

To illustrate the usefulness of our methodology in generating new findings, consider Bolivia in these figures. Governance structures including more than just bilateralism are relatively more prevalent in Bolivia than in all other countries. This is surprising. Bolivia is the least developed of the six countries and scores lowest on standard measures of the quality of the legal system (e.g., the World Governance Indicator's rule of law index).[64] Given the low levels of personal trust in Bolivia (Latin American Public Opinion Project, 2016/7), it is tempting to think that this might reflect comparative, rather than absolute, advantage in the legal realm. However, this cannot be a complete explanation, since LCA's estimation of a firm's governance class does not use data on a firm's country. Therefore, the greater effectiveness indicated for legal institutions in Bolivia than in, say, Uruguay, is inconsistent with the latter country's much higher ratings on indexes of legal institutions.[65] This is a puzzle, and one that is most clearly raised in an exercise like the current one.

[63] Thus the ordering of classes is different from that in Tables 2a and 2b, which followed LCA's somewhat arbitrary ordering.
[64] See Appendix Table A.1 for data on the general characteristics of the six countries.
[65] Bolivia has the highest estimated shadow-economy presence among the six countries studied in this paper (Medina and Schneider, 2019). This may have important implications for the share and the composition of the element of the economy that the WBES samples represent. However, Peru follows Bolivia closely in the estimated rate of the shadow economy, but has a markedly different pattern of estimated prevalences of governance structures. Thus the size of the shadow economy cannot be a complete explanation for the puzzling pattern found in Bolivia.

V.2 Cross-Regional Variations in Governance Structures

We next look at within-country, regional variation in governance structures, studying each country separately, using regional dummy variables as covariates. On the customer side, Wald tests of the null hypothesis that there is no variation across a country's regions in the patterns of governance structures are rejected only for Argentina and Ecuador (Table 4). On the supplier side, all countries except Bolivia exhibit significant cross-regional variation. This is another observation suggesting that Bolivia is anomalous, in this case by indicating that national institutions produce more regional homogeneity than in other countries.

Figures 2a and 2b show inter-regional variation in governance structures for those countries where we find statistically significant variation. The importance of bilateralism varies enormously. For example, an average firm in Rosario (Argentina) is 22 percentage points more likely to rate pure bilateralism as effective in its relations with suppliers than an average firm in its neighboring region of Cordoba. The difference on the customer side is even starker—44 percentage points. Regions even appear very different from their own countries: Piura (in Peru) has the lowest level of pure bilateralism amongst any of the 17 regions in Figure 2a, even though Peru has the highest level of pure bilateralism of the six countries in Figure 1a.

The conclusion is inescapable: inter-regional variation is even more important than cross-country variation. For example, the within-country, cross-regional standard deviation of the percentage of bilateralism in supplier relations is greater for each of Argentina, Ecuador, Peru, and Uruguay than the corresponding cross-country standard deviation. Even though legal systems are

country-level institutions in the six nations we study, regions, rather than countries, might be the best unit of analysis for understanding patterns of governance structures.[66]

V.3 Other Associations with Governance Structures

We find statistically significant associations between governance structures and firm size, sector, foreign ownership, and management practices. (See Table 4 for the corresponding Wald tests.) The nature of these associations is consistent with intuitively expected patterns. For example, the larger is a firm the more likely it is that it employs a comprehensive governance structure and one that makes use of the legal system. In particular, 67% of small firms (employing between 5 and 19 workers) use pure bilateralism when dealing with their suppliers, compared with 39% of very large firms (400 or more workers), who rely heavily on bilateralism with legal support (34% vs 15% of small firms). The pattern is the same on the customer side.

Sectoral variations, which are substantial, suggest that firms' governance structures are shaped by the nature of the good being exchanged. The patterns found here are intuitive. For example, the use of bilateralism varies from 74% when food processors interact with their suppliers to 24% when construction companies interact with their customers. This is consistent with Williamson's emphasis on the value of more complex governance when exchange is less frequent and involves idiosyncratic interactions (Williamson 1985).[67] Governance structures also vary substantially with the extent of firms' foreign ownership: firms with at least 10% foreign ownership are 17 percentage points less likely to use pure bilateralism than are other firms, mostly supplementing bilateralism

[66] In contrast, inter-regional variations in both GDP and GDP per capita are less than cross-country variation for these six countries. (We use the regional data in Gennaioli et al. (2013) and the corresponding country-level data from the World Development Indicators.)

[67] Lack of sufficient sample sizes at the appropriate level of sectoral granularity prevents us from exploring a more direct link with sector in ways similar to that in, e.g., Nunn (2007), Hidalgo et al. (2007), Boehm (2022).

with legal support. This pattern holds for both upstream and downstream relations, though the Wald test of the latter fails to reject the null hypothesis of no variation.

The association of governance structures with firms' management practices is sizeable, as illustrated in Figures 3a and 3b. As management practices improve, the prevalence of pure bilateralism falls: from 76% to 54% on the suppliers' side and from 72% to 41% on the customers' side.[68] This is intuitive since the improvement in firms' internal management practices is expected to be accompanied by the use of more complex methods of governance of external relations.

The exercises laid out in this section can be undertaken for any firm characteristics for which data are available: Appendix D reports many more such results, showing a significant association between governance structures and attitudes toward courts, export status, and membership in business associations.[69] The message of Appendix D is even stronger than the sum of its individual results: the patterns in the data suggest that our estimates of governance structures are substantively valid given the significant and intuitive associations between the use of governance structures and characteristics of firms or their environments.

VI. Determinants of Governance Structures: Uncertainty in Supplier Relations

This section presents an illustrative example of how our new data can be used to explore a specific hypothesis on the determinants of governance structure. In this empirical example, we use as our dependent variable the estimated posterior probability that a firm uses pure bilateralism for its supplier transactional governance. Given the nature of the estimated classes, which all embody bilateralism to some degree, a lower probability of membership in the pure bilateralism class also indicates a higher likelihood of membership in a class that supports bilateralism with some degree

[68] Firms' management practices are measured using a set of questions based on Bloom and Van Reenen (2007, 2010) and elaborated in, for example, Bloom et al. (2012).
[69] Appendix D contains additional details on the associations reported in this subsection.

of third-party involvement. The explanatory variable of chief interest is the degree of uncertainty in the conditions surrounding the exchange between the two parties. Given the empirical exercise in the paper—providing a specific example of the use of our data—we do not aim for definitive conclusions, but rather offer some suggestive insights.

We start with a brief summary of the existing literature that has explored the link between the uncertainty that firms face in their relations with suppliers and the choice of transactional governance structures. Williamson (1985) saw uncertainty as a prime factor that was conducive to opportunism in business exchanges. We focus on environmental uncertainty, which is exogenous to firms and their relationships with their suppliers. Noordewier et al. (1990) define such uncertainty as "unanticipated changes in circumstances surrounding an exchange". Environmental uncertainty is generally regarded as costly—often leading to the necessity of renegotiating the terms of an agreement.[70]

In investigating the effect of such unforeseen shocks on the choices of governance structures, researchers often follow Macneil's (1980) general perspective: 'discrete' (or contractual) exchanges involve clearly delineated and formalized arrangements, while 'relational' governance structures consist of more informal arrangements that can afford flexibility, through trust and intentions to cooperatively solve problems that might arise.[71] Notably, this conceptualization is often framed around substitutability between contractual and relational governance structures, as opposed to the complementarity that we found in Section IV. Using this dichotomy in a significant early contribution, Noordewier et al. (1990) found that relational governance structures mitigate the costs incurred under greater environmental uncertainty. While Noordewier et al. (1990) viewed

[70] We occasionally also refer to 'environmental uncertainty' as 'exogenous uncertainty' to reflect that it arises from unanticipated events. This is distinguished from 'behavioral' or endogenous uncertainty, which is due to the actions of a partner within an exchange.

[71] The prototypical, purely discrete transaction occurs on the spot market while the prototypical relational exchange occurs between family members.

contractual and relational as lying on opposite ends of a spectrum, with their use varying inversely, another body of work has considered whether greater uncertainty can increase the use of both contractual and relational governance. Zhou et al. (2008), for example, find that managers increase the use of both relational norms and more-detailed contractual terms under greater environmental uncertainty.

An important implication of this literature is that when firms face an uncertain environment, they increasingly favor adaptable governance structures (Poppo and Zenger 2002). When uncertainty implies that the complete decision-tree for the transaction cannot be mapped out, a third-party facilitates increased adaptability (Williamson 1975,1985). More complex governance structures, by invoking third parties, can mitigate the costs of unforeseen events for the continuation of the relationship by, for example, providing dispute resolution through a neutral referee, who could play a gap-filling role. Such neutral interpretations from third-party referees would be especially useful in situations where the occurrence and source of a specific exogenous shock might be unclear to one of the parties, and thus create room for opportunism. This refereeing can also safeguard against the failure of bilateral norms by instilling a credible belief that opportunism will be punished when it is undertaken under the guise of a necessary reaction to a shock. That is, engagement with third parties can add a dimension of adaptability to bilateralism alone. Nothing in this theory rules out the possibility that first- and second- party governance can be employed alongside the use of third parties. Thus, the empirical implication is simply that the role of third parties increases as uncertainty increases: the incidence of pure bilateralism will decrease.

We explore this prediction for supplier relations. Much of the existing literature on how firms adjust to downstream exogenous uncertainty focuses on the make-or-buy decision. (See

LaFontaine and Slade (2007) for a discussion.) Generally, studies show that vertical integration is more likely as uncertainty on the suppliers' end increases (e.g., Walker and Weber, 1984, Lieberman, 1991).[72] Even when using a strategy of vertical integration, uncertainty and the concomitant opportunism exacerbates principal-agent problems. While the standard agency models, which focus on risk-aversion, predict that decentralized incentives decrease when risk rises, many empirical findings suggest a positive relationship, potentially attributed to the strategy of firms relying on private, often local, knowledge by using decentralized decision-making in more uncertain environments (e.g., Prendergast 2002, Foss and Laursen, 2005).[73] A separate literature on firms' relations with external suppliers in cross-border trade finds that firms may shift toward lower-risk suppliers (Gervais, 2018), or withdraw from using foreign inputs altogether (Novy and Taylor, 2019) when uncertainty arises. An important takeaway for our purposes from these distinct branches of literature is that firms facing increased uncertainty, employ more adaptable governance structures. Using the framework of our agent-based classes, we take such adaptability to imply complementing bilateralism with third-party enforcement agents in efforts to reduce opportunism.

VI.1 The Empirical Framework

The generic form of the equation to be estimated regresses the probability that firms use pure bilateralism on environmental uncertainty:

$$B_{isr} = \alpha + \beta U_{isr} + \beta_X X_{isr} + \delta_s + \eta_r + \varepsilon_{isr} + \gamma\theta_{isr} \quad (5)$$

where B_{isr} is the posterior probability of membership in the pure bilateralism class for firm i in sector s and region r, U_{isr} is a measure of uncertainty, X_{isr} is a vector of observable characteristics,

[72] As noted by LaFontaine and Slade (2007), the source of this uncertainty need not be on the supplier side (that is, upstream). Others have found that even downstream uncertainty results in greater vertical integration (e.g., Hanson, 1995). A similar literature focusing on more endogenous forms of uncertainty, for instance the technological complexity of a firm's industry. has also found that such uncertain conditions make it more likely that firms vertically integrate (Acemoglu et al., 2007).
[73] Such mechanisms may explain why decentralized firms may fare better under unforeseen shocks (Aghion et al., 2021).

δ_s is a set of sector dummies, η_r is a set of region dummies, ε_{isr} is an i.i.d. disturbance term, and θ_{isr} is a vector of unobservable factors.

The variables used in estimation are defined in detail in Appendix E. Here, we provide a general description. For the dependent variable, we use the posterior probability that the firm employs pure bilateralism with its suppliers as estimated by LCA using the process described in Section III with the results of that process captured in class 1 in Table 2a. The main explanatory variable is a measure of the environmental uncertainty that firms face in their relations with suppliers. Uncertainty is measured using the survey responses to the question: "[p]lease indicate to what degree this establishment's suppliers are prevented from fulfilling agreements because of circumstances beyond their control." Note that the question's phrasing encourages respondents to focus on the elements of uncertainty that are exogenous to their relationship. The response categories were not at all, slightly, moderately, very much, and extremely. We create a dummy variable that equals 1 if firms choose anything other than "not at all" and 0 otherwise.

Our measure of uncertainty aims to capture Williamson's (1985) 'primary' uncertainty, which arises from exogenous sources, such as natural events, changes in regulations or prices, or other aspects of the business environment beyond the firms' control. This is a non-strategic form of uncertainty that is distinct from behavioral uncertainty related to the suppliers' opportunism. The more recent literature views environmental uncertainty as encompassing technological uncertainty (e.g., Handley and Angst, 2015), or external uncertainty (e.g., Katsikeas et al. 2009), or unanticipated shifts in demand (Gaur et al., 2011). This kind of uncertainty is often measured through survey questions relating to the firms' sector of operation in general (e.g., Abdi and Aulakh, 2017, Crocker and Masten, 1991) or by direct measures of the volatility of prices or other

disruptive shocks in the business environment (e.g., Aghion et al., 2021). Our measure aims to capture this kind of uncertainty as it relates to the firms' relations with their own suppliers.

The vector of control variables, X_{isr} in equation (5), contains elements that are natural candidates to reduce omitted variable bias, reflecting firm size, age, foreign ownership, whether the firm is a subsidiary, and whether the firm exports its product. Each of these elements can conceivably influence firms' exposure to uncertainty as well as their governance structures, necessitating their inclusion in the regression. Importantly, we use sector and region fixed effects, which will provide controls for many variables that are not idiosyncratic to the firm. For sector fixed effects, δ_s, we use two-digit ISIC codes, and for region fixed effects, η_r, we use the ES stratification regions. More detailed information about these variables and summary statistics are provided in Appendix E.

VI.2 Estimation and Results

Table 5 reports the results. In all specifications, sampling weights are used, and standard errors are calculated assuming clustering at the level of strata, which reflects sector, region, and firm size. Column (3) contains the core OLS regression, with columns (1) and (2) providing important information for robustness tests.

The estimate of the coefficient on uncertainty in column (3) is negative and highly significant, lying over four standard errors below zero. The coefficient's negative sign indicates that when exogenous uncertainty is high, governance structures are more likely to involve third-party agents. It is also economically significant—a one-standard deviation increase in uncertainty leading to a decline in the probability of bilateralism of 0.097. If all the firms in our sample not experiencing uncertainty were suddenly afflicted with uncertainty, then the rate of non-bilateralism in our sample would be 21% higher. These results are a significant contribution to the stock of evidence

on the effect of uncertainty on governance structures, simply because there are such few results in the literature that clearly isolate such effects.

Since θ_{isr} is unobservable, the composite error term in equation (5) is $\varepsilon_{isr} + \gamma\theta_{isr}$ and OLS estimates of β will be biased and inconsistent if γ is nonzero and θ_{isr} is correlated with U_{isr}. This type of endogeneity can arise, for example, if some firms are more forgiving than others in ascribing problems to sources outside their suppliers' control, while the same trait contributes to the development of bilateral relations with suppliers.

Such problems of endogeneity are rife in the literature on the determinants of governance transactions, and it would be fair to say that few papers have solved these problems satisfactorily. Often, endogeneity is simply ignored. We address the problem of endogeneity in two ways. First, we use the Altonji et al. (2005)-Oster (2019) method, which employs comparisons between OLS regressions with different covariates to cast some perspective on the likely sensitivity of OLS results to unmeasured, and therefore omitted, covariates. Second, we employ a standard IV approach.

The coefficients in columns (1), (2), and (3) of Table 5 are quite similar, with variations between them less than 5% of their absolute values. This is especially important because while (3) includes regional and sector fixed effects, (1) does not, indicating that a large source of variation that is common in such cross-sectional data has no effect on our coefficient of interest. The Altonji et al. (2005)-Oster (2019) method formalizes such intuition. A formal application following Oster (2019) provides estimates of an interval in which an estimated coefficient would almost certainly lie if all omitted-variable problems were solved. Comparing the effect of the addition of covariates when moving from (1) to (3), the estimated interval is [-0.24746 ,-0.1939], while the same comparison for (2) versus (3) provides an estimated interval of [-0.1939, -0.18781]. Hence, insofar

as comparisons between OLS results can provide evidence, there is no reason to think that the addition of all omitted covariates would change the conclusions reached when using the estimates in column (3).

To develop an instrumental-variable approach, we view U_{isr} as comprising two elements (Fisman, and Svensson 2007):

$$U_{isr} = u_{isr} + u_{sr} \tag{6}$$

where u_{sr} is the uncertainty common to firms in sector s and region r, while u_{isr} is the uncertainty that is idiosyncratic to the firm.[74] We use u_{sr} as an instrumental variable. Its validity as an instrument rests on the assumption that the sector-region specific part of the uncertainty (u_{sr}) is determined by underlying factors at the level of sector and region (in combination), and exogenous to a firm-specific source of endogeneity: u_{sr} and θ_{isr} are uncorrelated.

We create u_{sr} as follows. We first group observations at the level of two-digit ISIC sector interacted with region. We then calculate the weighted average of the uncertainty variable for each such sector-region group. We drop observations for firms where the weighted average reflects the responses of fewer than 5 firms located in the same sector-region cell.

The exclusion restriction on u_{sr} will be violated if there are unobserved factors at the sector-region level that are determinants of bilateralism. Importantly, neither sector-specific nor region-specific elements pertaining to bilateralism or uncertainty are present in θ_{isr}, because δ_s and η_r are included in (5). Thus, the processes that would violate the exclusion restriction would need to work systematically at the sector-region level.

Note however that our instruments do not have to fully satisfy the exclusion restriction if we do not focus here on pinning down the size of the uncertainty effect, but rather simply focus on

[74] In constructing these variables, we use a more granular classification of sectors than Fisman and Svensson (2007), and exclude sector-region averages that are based on fewer than 5 observations.

what can be learned about its sign and the differences from the OLS estimates. The thrust of our argument is that there are fewer endogeneity problems when using the sector-region averages (u_{sr}) as instruments than when using the variable itself (U_{isr}) as its own instrument.[75]

This logic is crucial when comparing the OLS results of column (3) of Table 5 to the IV results in column (4). The estimate of the coefficient on uncertainty increases in absolute value and is still significant, now two and one-half standard errors below zero. As we address problems of endogeneity, albeit imperfectly, by moving from (3) to (4) in Table 5, and the critical estimated coefficient increases in absolute value, we reach firmer evidence supporting the hypothesis that increasing uncertainty reduces pure bilateralism.[76] With this logic, an estimate produced by a method that solved all endogeneity problems would be expected to be negative.[77]

It would strain credibility if we claimed that the results in Table 5 pin down the exact size of any coefficients. Nevertheless, one can have more confidence that these results provide evidence for a negative sign of the uncertainty coefficient. As we have stated before, problems of endogeneity in estimating the determinants of inter-firm governance structures are legion and, in the case of the effect of uncertainty, have not received any definitive treatment in the literature. Therefore, our attempts to address the problem of endogeneity should be regarded as an approach to learning from imperfect estimates rather than solving the problems of endogeneity. Cast in that light, Table 5 provides evidence that uncertainty increases the propensity to use third parties beyond the two parties directly involved in the agreement.

This conclusion should also be cast in terms of the objective of this section in the context of the paper as a whole. We did not begin this section's empirical exercise with the intention of

[75] Recall that OLS is IV using the variable itself as its own instrument.
[76] We also used an alternative set of instrumental variables, reaching the same conclusions. These results are available upon request.
[77] The rough intuition is clear. These ideas can be backed up with more formal reasoning, which would require more space than is justified given the role of that reasoning in this paper. A more formal statement is available from the authors on request.

making strong statements about the precise effect of uncertainty. Rather the objective was to show that the use of our new data can lead to new insights into governance structures and their determinants. We believe that the current exercise fully satisfies this objective, despite the limitations discussed above.

VII. Lessons Learned and Avenues for Future Research

This paper uses an exploratory statistical approach to uncover governance structures that firms employ to support the implementation of their agreements. Those estimated governance structures capture the patterns in economy-wide, cross-country data that reflect responses to a survey question that considers a full range of enforcement agents. To our knowledge, it is the first paper whose results reflect such a comprehensive context. This past lacuna in the literature has partly been due to difficulties of data collection, a problem that we solve by designing and applying a new survey question to elicit consistent information on the conduct of transactions from firms of all types, functioning in very different environments. This survey question is based upon a conceptual framework that focuses on enforcement in its broadest sense and uses an agent-based approach, with the enforcement agents viewed as arrayed on a comprehensive spectrum.

Yet, obtaining such data provides only part of the solution to understanding how firms govern their transactions. There is also a need to summarize the patterns implicit in the data in a way that produces evocative measures. For this an exploratory method is necessary, given that there is no encompassing theory on how firms combine enforcement agents to form governance structures. LCA eminently suits this task: it is unsupervised in discovering patterns in the data, but relies on an underlying probabilistic, generative model. Thus, it combines the advantages of both classical statistical methodology and machine learning. The use of a generative model permits reliance on standard statistical techniques for model selection and evaluation of estimates. The unsupervised

learning offers the possibility of finding governance structures that are not even contemplated before the analysis. Nevertheless, the governance structures we uncover through this exploratory technique resonate comfortably with concepts that are standard in the existing analysis of transactions, lending credence to our analysis.

We find that pure bilateralism is the most prevalent governance structure that we observe, and that all governance structures involve bilateralism (necessitating the use of the term 'pure' to describe the governance structure that relies solely on the parties directly involved in the agreement). Our finding that no governance structure comprises purely arm's length transactions, where firms rely only on impersonal enforcement and formal institutions, is important because arm's length transactions have often been viewed as something of an ideal, the aspirational endpoint in the process of economic development.[78] This view implicitly casts bilateralism and formal institutions as substitutes, a relationship that does not appear in our estimates, which instead suggest that for many firms, bilateralism and formal institutions are complements.

Some of our findings provide a marked contrast to very common emphases in the literature. In particular, while the literature devotes much attention to various unpaid, third-party, mechanisms of supporting agreements, such as networks, social clubs, and culturally defined groups, our data suggests that their role is very limited.[79] We find that government officials, enforcement agents that have been almost completely ignored in the literature, do have a role in supporting agreements.

[78] To be sure, this is an aspiration not shared by all, or perhaps even a majority, of scholars contributing to the economics literature. It is even less popular among diverse groups of scholars studying the detailed workings of the legal system, for example, both the law and economics and the law and society schools.

[79] Our results are broadly consistent with those of Mike and Kiss (2019). Given that these authors use different survey questions, study a different context (Hungary), and implement LCA in a different way, such consistencies point to robust general conclusions about the landscapes of transactions. One difference is that Mike and Kiss (2019) find a latent class in which third-party reputational mechanisms are quite important. Whether this is a reflection of the different context, Hungary, or of different survey questions is an open issue, to be answered only by implementing a consistent cross-country methodology.

In Sections V and VI, we provide examples of further analyses that can be conducted once our estimates of governance structures are obtained. The observations are at the firm level and the variables of interest are the probabilities that the firm has chosen each of the four governance structures. Notably, in Section V, we find that regional variation in the effectiveness of different governance structures is more important than cross-country variation. This is somewhat puzzling given that institutional rules relevant to transactions are set at the national level in all the countries we analyze. It suggests that institutional implementation is at least as important as the quality of formal rules. In Section VI, we provide evidence that the exogenous uncertainty that firms face influences the firms' choices concerning reliance on third-party enforcement. This finding shows that our data can provide a wholly new way of testing existing hypotheses that have been the subject of much theoretical discussion but without any consensus arising in the empirical literature, as is the case with the effect of uncertainty on governance.

While explorations of the variations in and determinants of governance structures, as in Sections V and VI, have not been this paper's prime objective, these exercises show the validity of the methodology we have developed and demonstrate the potential of the datasets that we have generated.[80] Specifically, our methodology allows readers to test other hypotheses, using firms' governance structures as dependent variables as we have done, or as independent variables helping to explain variations in firms' behavior and performance. Moreover, given the information we provide, readers could add different countries, or cities, or sectors to those we have studied here. If the question that we lay out in Section II were to be implemented in a survey of any size, even a single firm, then the tools that we have made publicly available can be used to characterize the

[80] The data is posted at www.enterprisesurveys.org/portal. Users need an account (available free of charge) to gain access to the data. After signing in, go to the "Combined Data" tab. There, download the zip file "Landscape of Transactions Data and Estimates".

governance structure of the firms in the survey. Through new surveys, readers could produce results that are comparable with ours without repeating the laborious steps involving LCA that we undertook in this paper, thereby facilitating diagnosis of a firm's, or a sector's, or a country's strengths and weaknesses.

References

Abdi, Majid, and Preet S. Aulakh. "Locus of uncertainty and the relationship between contractual and relational governance in cross-border interfirm relationships." *Journal of Management* 43, no. 3 (2017): 771-803.

Acemoglu, Daron, Philippe Aghion, Claire Lelarge, John Van Reenen, and Fabrizio Zilibotti. "Technology, Information, and the Decentralization of the Firm." *The Quarterly Journal of Economics* 122, no. 4 (2007): 1759-1799.

Aghion, Philippe, Nicholas Bloom, Brian Lucking, Raffaella Sadun, and John Van Reenen. "Turbulence, firm decentralization, and growth in bad times." *American Economic Journal: Applied Economics* 13, no. 1 (2021): 133-69.

Alwin, Duane F. and Brett A. Beattie. 2016. "The Kiss Principle In Survey Design: Question Length And Data Quality" *Sociological Methodology*, 46 (August 2016): 121-152.

Argyres, Nicholas, Janet Bercovitz, and Giorgio Zanarone. "The role of relationship scope in sustaining relational contracts in interfirm networks." *Strategic Management Journal* 41, no. 2 (2020): 222-245.

Baker, George, Robert Gibbons, and Kevin J. Murphy. "Relational Contracts and the Theory of the Firm." *The Quarterly Journal of Economics* 117, no. 1 (2002): 39-84.

Bakk, Zsuzsa, Fetene B. Tekle and Jeroen K. Vermunt. 2013. Estimating the Association between Latent Class Membership and External Variables Using Bias-adjusted Three-step Approaches. *Sociological Methodology* 2013 43: 272.

Bakk, Zsuzsa, Oberski, Daniel L., and Vermunt, Jeroen K. 2014. Relating Latent Class Assignments to External Variables: Standard Errors for Correct Inference. *Political Analysis*, 22, 520-540.

Bernstein, L., 1992. Opting out of the legal system: Extralegal contractual relations in the diamond industry. *The Journal of Legal Studies*. 21(1), pp.115-157.

Bernstein, Lisa. "Beyond relational contracts: Social capital and network governance in procurement contracts." *Journal of Legal Analysis* 7, no. 2 (2016): 561-621.

Bolck, A., Croon, M. and Hagenaars, J., 2004. Estimating latent structure models with categorical variables: One-step versus three-step estimators. *Political Analysis*. 12(1):.3-27.

Cannon, Joseph P., Ravi S. Achrol, and Gregory T. Gundlach. "Contracts, norms, and plural form governance." *Journal of the academy of marketing science* 28, no. 2 (2000): 180-194.

Cao, Z. and Lumineau, F. 2015. Revisiting the interplay between contractual and relational governance: A qualitative and meta-analytic investigation. *Journal of Operations Management* 33: 15–42.

Carson, Stephen J., Anoop Madhok, and Tao Wu. "Uncertainty, opportunism, and governance: The effects of volatility and ambiguity on formal and relational contracting." *Academy of Management journal* 49, no. 5 (2006): 1058-1077.

Chao, John C., and Norman R. Swanson. "Consistent estimation with a large number of weak instruments." *Econometrica* 73, no. 5 (2005): 1673-1692.

Collins, L. M., & Lanza, S. T. 2010. *Latent class and latent transition analysis: With applications in the social, behavioral, and health sciences*. New York: Wiley.

Crocker, Keith J., and Scott E. Masten. "Pretia ex machina? Prices and process in long-term contracts." *The Journal of Law and Economics* 34, no. 1 (1991): 69-99.

Dekker, Henri C. 2008. "Partner selection and governance design in interfirm relationships." *Accounting, Organizations and Society* 33, no. 7-8: 915-941.

Devarakonda, Shivaram V., Brian T. McCann, and Jeffrey J. Reuer. "Marshallian forces and governance externalities: Location effects on contractual safeguards in Research and Development alliances." *Organization Science* 29, no. 6 (2018): 1112-1129.

Dixit, Avinash. "On modes of economic governance." *Econometrica* 71, no. 2 (2003): 449-481.

Dixit, Avinash. "Governance institutions and economic activity." *American economic review* 99, no. 1 (2009): 5-24.

Ebers, Mark and Leon Oerlemans. 2016. The Variety of Governance Structures Beyond Market and Hierarchy. *Journal of Management* 42(6), September: 1491– 1529.

Ellickson, Robert C. Comment on "Contracting, Enforcement, and Efficiency: Economics beyond the Law." In *Annual world bank conference on development economics*, vol. 1996, pp. 239-265. Washington, DC: The World Bank, 1997.

Fafchamps, Marcel. 2004. *Market Institutions in Sub-Saharan Africa: Theory and Evidence*. Boston: MIT Press.

Fisman, Raymond, and Jakob Svensson. "Are corruption and taxation really harmful to growth? Firm level evidence." *Journal of development economics* 83, no. 1 (2007): 63-75.

Foss, Nicolai J., and Keld Laursen. "Performance pay, delegation and multitasking under uncertainty and innovativeness: An empirical investigation." *Journal of Economic Behavior & Organization* 58, no. 2 (2005): 246-276.

Fowler and Cosenza 2008

Furnham, A. and Boo, H.C., 2011. A literature review of the anchoring effect. *Journal of Behavioral and Experimental Economic.*, 40(1), pp.35-42.

Gaur, Ajai S., Debmalya Mukherjee, Sanjaya S. Gaur, and Florian Schmid. "Environmental and firm level influences on inter-organizational trust and SME performance." *Journal of Management studies* 48, no. 8 (2011): 1752-1781.

Gennaioli, Nicola, Rafael La Porta, Florencio Lopez-de-Silanes, and Andrei Shleifer. "Human capital and regional development." *The Quarterly journal of economics* 128, no. 1 (2013): 105-164.

Gervais, Antoine. "Uncertainty, risk aversion and international trade." *Journal of International Economics* 115 (2018): 145-158.

Gil, Ricard, Myongjin Kim, and Giorgio Zanarone. "Relational contracting in developed economies: Lessons from slot exchanges in the US airline industry." *The Japanese Economic Review* 70, no. 3 (2019): 411-421.

Granovetter, Mark. 1985. Economic Action and Social Structure: The Problem of Embeddedness. *American Journal of Sociology* 91 (November): 481-510

Greif, Avner. 1989. Reputation and coalitions in medieval trade: evidence on the Maghribi traders. *The Journal of Economic History.* 49(4), pp.857-882.

Greif, Avner. "Cultural beliefs and the organization of society: A historical and theoretical reflection on collectivist and individualist societies." *Journal of political economy* 102, no. 5 (1994): 912-950.

Greif, Avner. "Contracting, enforcement and efficiency: Economics beyond the law." In *Annual world bank conference on development economics*, vol. 1996, pp. 239-265. Washington, DC: The World Bank, 1997.

Greif, Avner. "The fundamental problem of exchange: a research agenda in historical institutional analysis." *European Review of Economic History* 4, no. 3 (2000): 251-284.

Greif, Avner. 2001. The fundamental problem of exchange: A research agenda in Historical Institutional Analysis. *European Review of Economic History*, 40(3), February: 251-284.

Greif, Avner. "Commitment, coercion, and markets: The nature and dynamics of institutions supporting exchange." In *Handbook of new institutional economics*, pp. 727-786. Springer, Boston, MA, 2005.

Grewal, Rajdeep, Gary L. Lilien, and Girish Mallapragada. "Location, location, location: How network embeddedness affects project success in open source systems." *Management Science* 52, no. 7 (2006): 1043-1056.

Gulati, Ranjay. "Does familiarity breed trust? The implications of repeated ties for contractual choice in alliances." *Academy of management journal* 38, no. 1 (1995): 85-112.

Hadfield, Gillian K. and Bozovic, Iva. 2016. Scaffolding: Using formal contracts to support informal relations in support of innovation. *Wisconsin Law Review*, 981–1032.

Handley, Sean M., and Corey M. Angst. "The impact of culture on the relationship between governance and opportunism in outsourcing relationships." *Strategic Management Journal* 36, no. 9 (2015): 1412-1434.

Hanson, Gordon H. "Incomplete contracts, risk, and ownership." *International Economic Review* (1995): 341-363.

Hart, Oliver, and John Moore. "Incomplete contracts and renegotiation." *Econometrica: Journal of the Econometric Society* (1988): 755-785.

Hendley, K., Murrell, P. and Ryterman, R., 2000. Law, relationships and private enforcement: Transactional strategies of Russian enterprises. *Europe-Asia Studies*. 52(4), pp.627-656.

Hendley, Kathryn, and Murrell, Peter. 2003. Which mechanisms support the fulfillment of sales agreements? Asking decision-makers in firms. *Economics Letters*, 78(1), pp 49-54

Holbrook, A., M. Green, and J. Krosnick. 2003. Telephone versus Face-to-Face Interviewing of National Probability Samples with Long Questionnaires: Comparisons of Respondent Satisficing and Social Desirability Response Bias. *The Public Opinion Quarterly*. 67(1), pp. 79-125.

Holbrook, Allyson, Young Ik Cho, and Timothy Johnson. 2006. "The Impact Of Question And Respondent Characteristics On Comprehension And Mapping Difficulties." *Public Opinion Quarterly*, Vol. 70, No. 4: 565–595.

Homburg, Christian, Joseph P. Cannon, Harley Krohmer, and Ingo Kiedaisch. "Governance of international business relationships: A cross-cultural study on alternative governance modes." *Journal of International Marketing* 17, no. 3 (2009): 1-20.

Jedidi, Kamel, Harsharanjeet S. Jagpal, and Wayne S. DeSarbo. "Finite-mixture structural equation models for response-based segmentation and unobserved heterogeneity." *Marketing Science* 16, no. 1 (1997): 39-59.

Johnson, Simon, John McMillan, and Christopher Woodruff. "Courts and relational contracts." *Journal of Law, Economics, and organization* 18, no. 1 (2002): 221-277.

Kamakura, Wagner A., and Gary J. Russell. "A probabilistic choice model for market segmentation and elasticity structure." *Journal of Marketing Research* 26, no. 4 (1989): 379-390.

Katsikeas, Constantine S., Dionysis Skarmeas, and Daniel C. Bello. "Developing successful trust-based international exchange relationships." *Journal of international business studies* 40, no. 1 (2009): 132-155.

Krosnick, J. and D. Alwin. 1988. A Test of the Form-Resistant Correlation Hypothesis: Ratings, Rankings, and the Measurement of Values. *The Public Opinion Quarterly*, 52(4), pp. 526-538.

Krosnick, Jon A. 1991. "Response strategies for coping with the cognitive demands of attitude measures in surveys." *Applied Cognitive Psychology* 5, 213-236.

Krosnick, Jon and Stanley Presser. 2010. "Questionnaire design ." In J. D. Wright & P. V. Marsden (Eds.), *Handbook of Survey Research* (Second Edition). West Yorkshire, England: Emerald Group.

Lafontaine, Francine, and Margaret Slade. "Vertical integration and firm boundaries: The evidence." *Journal of Economic literature* 45, no. 3 (2007): 629-685.

Landa, J.T., 1981. A theory of the ethnically homogeneous middleman group: An institutional alternative to contract law. *The Journal of Legal Studies*, 10(2), pp.349-362.

Latin American Public Opinion Project. 2016/7. The Americas Barometer. Latin American Public Opinion Project (LAPOP), Vanderbilt University. http://www.LapopSurveys.org (accessed June 12, 2018).

Lieberman, Marvin B. "Determinants of vertical integration: An empirical test." In *Academy of Management Proceedings*, vol. 1991, no. 1, pp. 31-35. Briarcliff Manor, NY 10510: Academy of Management, 1991.

Macaulay, Stewart. 1963. Non-Contractual Relations In Business: A Preliminary Study. *American Sociological Review*, Vol. 28, No. 1, February.

Macchiavello, Rocco, and Ameet Morjaria. "The value of relationships: evidence from a supply shock to Kenyan rose exports." *American Economic Review* 105, no. 9 (2015): 2911-45.

MacLeod, W. Bentley. "Reputations, relationships, and contract enforcement." *Journal of economic literature* 45, no. 3 (2007): 595-628.

Macneil, I. R. 1980. *The new social contract: An inquiry into modern contractual relations*. New Haven, CT: Yale University Press.

Mahapatra, Santosh K., Ram Narasimhan, and Paolo Barbieri. 2010. "Strategic interdependence, governance effectiveness and supplier performance: A dyadic case study investigation and theory development." *Journal of Operations Management* 28, no. 6: 537-552.

Masyn, Katherine. 2013. Latent class analysis and finite mixture modeling. In T. D. Little (Ed.) *The Oxford handbook of quantitative methods in psychology. Volume 2*. pp. 551-611. New York, NY: Oxford University Press.

McMillan, J. and Woodruff, C., 1999. Interfirm relationships and informal credit in Vietnam. *The Quarterly Journal of Economics*, 114(4), pp.1285-1320.

Medina, L. and Schneider, F., 2019. Shedding light on the shadow economy: A global database and the interaction with the official one.

Mike, Károly, and Gábor Kiss. "Combining formal and informal contract enforcement in a developed legal system: a latent class approach." *Journal of Institutional Economics* 15, no. 3 (2019): 521-537.

Moran, Peter, and Sumantra Ghoshal. "Markets, firms, and the process of economic development." *Academy of Management Review* 24, no. 3 (1999): 390-412.

Noordewier, T. G., G. John and J. R. Nevin. 1990. “Performance Outcomes of Purchasing Arrangement in Industrial Buy-Vendor Relationships.” *Journal of Marketing* 54: 80-93.

North, Douglass C. 1990. *Institutions, Institutional Change, and Economic Development*, Cambridge: Cambridge University Press.

Novy, Dennis, and Alan M. Taylor. "Trade and uncertainty." *Review of Economics and Statistics* 102, no. 4 (2020): 749-765.

Nylund-Gibson, Karen, Katherine E. Masyn 2016 Covariates and Mixture Modeling: Results of a Simulation Study Exploring the Impact of Mis-specified Effects on Class Enumeration, *Structural Equation Modeling: A Multidisciplinary Journal*, 23:6, 782-797.

Oberski, Daniel. 2016. "Questionnaire Science." In *The Oxford Handbook of Polling and Survey Methods*, edited by Lonna Rae Atkeson and R. Michael Alvarez. Oxford: Oxford University Press.

Parmigiani, Anne and Will Mitchell. 2010. "The hollow corporation revisited: Can governance mechanisms substitute for technical expertise in managing buyer-supplier relationships?" *European Management Review* 7: 46-70.

Peng, Mike W. 2003. Institutional Transitions and Strategic Choices. *Academy of Management Review*, Vol, 28, No, 2, 275-296,

Poppo, L. and T. Zenger. 2002. "Do Formal Contracts and Relational Governance Function as Substitutes or Complements?" *Strategic Management Journal*, 23(8): 707- 725.

Poppo, Laura and Zheng Cheng. 2017. "Trust and Contracts: Complements versus substitutes in business-to-business exchanges." in Rosalind Searle, Ann-Marie Nienaber, and Sim Sitkin eds. *The Routledge Companion to Trust*. New York : Routledge.

Poppo, Laura, Kevin Zheng Zhou, and Julie J. Li. 2016. "When Can You Trust 'Trust'? Calculative Trust, Relational Trust, and Supplier Performance." *Strategic Management Journal*, v. 37 n. 4, p. 724-741.

Prendergast, Canice. "The tenuous trade-off between risk and incentives." *Journal of political Economy* 110, no. 5 (2002): 1071-1102.

Samuelsson, Mikael and Per Davidsson. 2009. Does Venture Opportunity Variation Matter? Investigating Systematic Process Differences between Innovative and Imitative New Ventures. *Small Business Economics*, Vol. 33, No. 2: 229-255.

Shaeffer, Eric M., Jon A. Krosnick, Gary E. Langer, and Daniel M. Merkle. 2005. "Comparing The Quality Of Data Obtained By Minimally Balanced And Fully Balanced Attitude Questions." *Public Opinion Quarterly*, Vol. 69. No. 3: 417-428.

Susarla, Anjana, Martin Holzhacker, and Ranjani Krishnan. "Calculative trust and interfirm contracts." *Management Science* 66, no. 11 (2020): 5465-5484.

Vermunt, Jeroen K. and Jay Magidson. 2002. Latent Class Cluster Analysis. In Applied Latent Class Analysis. Edited by Jacques A. Hagenaars and Allan L. McCutcheon. Cambridge University press, Cambridge, 2002: 89-106.

Vermunt, Jeroen K. and Jay Magidson. *Technical Guide for Latent GOLD 5.1: Basic, Advanced, and Syntax*. January 9, 2016. Statistical Innovations Inc. http://www.statisticalinnovations.com

Vermunt, Jeroen K., 2010. Latent class modeling with covariates: Two improved three-step approaches. *Political Analysis*, 18(4), pp.450-469.

Walker, Gordon, and David Weber. "A transaction cost approach to make-or-buy decisions." *Administrative science quarterly* (1984): 373-391.

Williamson, Oliver E. "Markets and hierarchies: analysis and antitrust implications: a study in the economics of internal organization." *University of Illinois at Urbana-Champaign's Academy for Entrepreneurial Leadership Historical Research Reference in Entrepreneurship* (1975).

Williamson, Oliver E. 1979. Transaction-Cost Economics: The Governance of Contractual Relations *The Journal of Law & Economics*, 22(2), Oct.: 233-261.

Williamson, Oliver E. 1985. *The Economic Institutions of Capitalism*. New York: Free Press.

Williamson, Oliver E. 1991. Comparative economic organization: The analysis of discrete structural alternatives. *Administrative science quarterly*, pp.269-296.

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

World Bank. 2017. *World Development Report 2017: Governance and the Law*. Washington, DC: World Bank.

World Bank. 2020. *World Development Report 2020: Trading for Development in the Age of Global Value Chains.* Washington, DC: World Bank.

Worldwide Governance Indicators (www.govindicators.org)

Yan, Ting and Roger Tourangeau. 2008. "Fast Times and Easy Questions: The Effects of Age, Experience and Question Complexity on Web Survey Response Times." *Applied Cognitive Psychology*. 22: 51–68.

Yang, Chenlung, John G. Wacker and Chwen Sheu. 2012. "What makes outsourcing effective? A transaction-cost economics analysis." *International Journal of Production Research*. 50(16): 4462–4476.

Zhou, Kevin Z., Laura Poppo, and Zhilin Yang. 2008. "Relational ties or customized contracts? An examination of alternative governance choices in China." *Journal of International Business Studies* 39, 526–534.

Zhou, Kevin Zheng, and Laura Poppo. "Exchange hazards, relational reliability, and contracts in China: The contingent role of legal enforceability." *Journal of International Business Studies* 41, no. 5 (2010): 861-881.

Zhou, Kevin Zheng, and Dean Xu. "How foreign firms curtail local supplier opportunism in China: Detailed contracts, centralized control, and relational governance." *Journal of International Business Studies* 43, no. 7 (2012): 677-692.

Tables and Figures

Table 1: Summary of Responses to the Questions

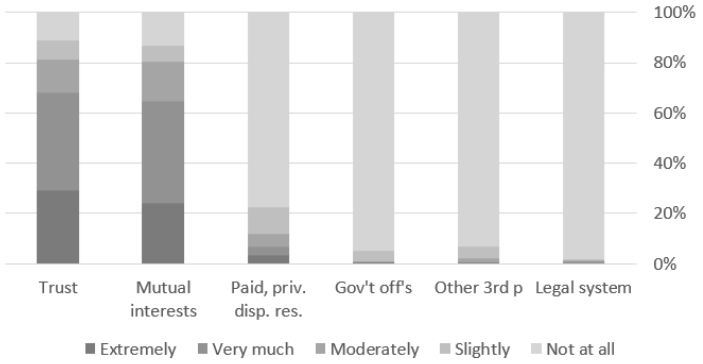
Relations with suppliers (n=3,350)						
	Not at all	Slightly	Moderately	Very much	Extremely	Total
Trust	8.3	8.0	14.4	41.8	27.6	100.0
Mutual interests	10.9	6.7	14.4	42.4	25.5	100.0
Paid private third-parties	61.9	15.4	13.9	5.6	3.1	100.0
Gov't officials	86.2	8.0	2.5	1.9	1.4	100.0
Other third-parties	81.5	12.5	3.8	1.7	0.4	100.0
Legal system	70.1	17.6	8.2	2.7	1.4	100.0
Relations with customers (n=3,339)						
	Not at all	Slightly	Moderately	Very much	Extremely	Total
Trust	6.5	7.8	12.2	39.3	34.2	100.0
Mutual interests	9.0	7.1	11.5	42.3	30.0	100.0
Paid private third-parties	62.0	19.9	10.9	5.2	2.1	100.0
Gov't officials	89.5	6.2	2.3	1.2	0.8	100.0
Other third-parties	82.0	12.3	3.7	1.6	0.3	100.0
Legal system	72.5	16.1	7.6	2.4	1.5	100.0

Table 2a: Predicted Response Probabilities, Relations with Suppliers

Class 1 – Pure Bilateralism

Method	Not at all	Slightly	Moderately	Very much	Extremely
Trust	0.11	0.08	0.13	0.39	0.29
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)
Mutual interests	0.13	0.06	0.16	0.41	0.24
	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)
Paid, priv. disp. res.	0.77	0.11	0.05	0.03	0.03
	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
Gov't off's	0.95	0.05	0.00	0.00	0.00
	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)
Other 3rd p	0.93	0.05	0.02	0.00	0.00
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)
Legal system	0.98	0.00	0.00	0.01	0.00
	(0.02)	(0.01)	(0.01)	(0.01)	(0.00)

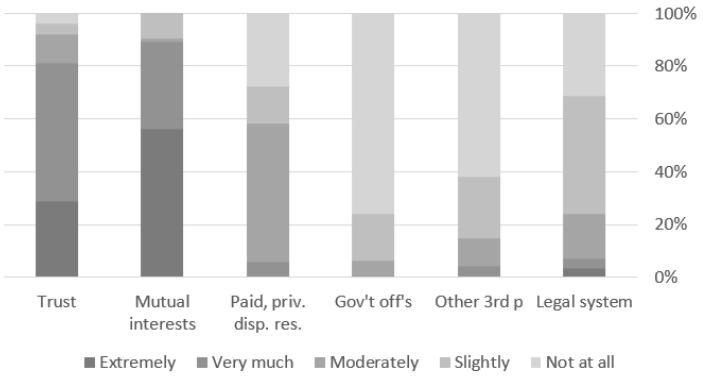
Standard errors in parenthesis.



Class 2 – Bilateralism with private support

Method	Not at all	Slightly	Moderately	Very much	Extremely
Trust	0.04	0.04	0.11	0.52	0.29
	(0.01)	(0.02)	(0.04)	(0.06)	(0.05)
Mutual interests	0.00	0.10	0.01	0.33	0.56
	(0.00)	(0.03)	(0.01)	(0.07)	(0.07)
Paid, priv. disp. res.	0.28	0.14	0.52	0.06	0.00
	(0.06)	(0.04)	(0.07)	(0.02)	(0.00)
Gov't off's	0.76	0.18	0.05	0.01	0.00
	(0.04)	(0.04)	(0.02)	(0.00)	(0.00)
Other 3rd p	0.62	0.23	0.10	0.04	0.00
	(0.05)	(0.05)	(0.03)	(0.02)	(0.00)
Legal system	0.31	0.45	0.17	0.04	0.03
	(0.06)	(0.06)	(0.04)	(0.02)	(0.02)

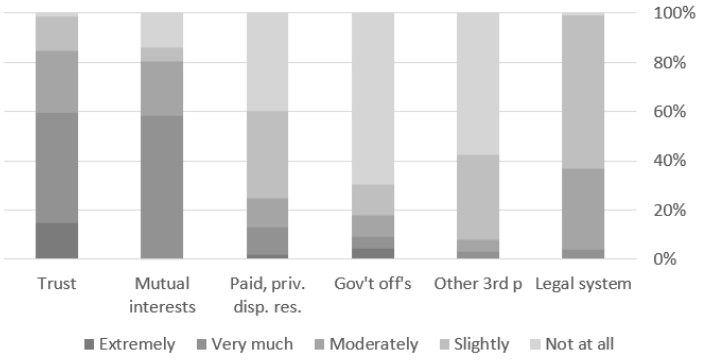
Standard errors in parenthesis.



Class 3 – Bilateralism with legal support

Method	Not at all	Slightly	Moderately	Very much	Extremely
Trust	0.02	0.14	0.25	0.45	0.15
	(0.01)	(0.04)	(0.04)	(0.05)	(0.03)
Mutual interests	0.14	0.06	0.22	0.59	0.00
	(0.03)	(0.03)	(0.04)	(0.05)	(0.00)
Paid, priv. disp. res.	0.40	0.35	0.12	0.11	0.02
	(0.06)	(0.05)	(0.05)	(0.03)	(0.02)
Gov't off's	0.69	0.13	0.08	0.05	0.05
	(0.05)	(0.03)	(0.03)	(0.02)	(0.02)
Other 3rd p	0.58	0.34	0.05	0.03	0.00
	(0.05)	(0.05)	(0.01)	(0.02)	(0.00)
Legal system	0.01	0.62	0.33	0.04	0.00
	(0.02)	(0.05)	(0.05)	(0.02)	(0.00)

Standard errors in parenthesis.



Class 4 – Strong comprehensive governance

Method	Not at all	Slightly	Moderately	Very much	Extremely
Trust	0.03	0.00	0.12	0.08	0.77
	(0.03)	(0.00)	(0.07)	(0.05)	(0.10)
Mutual interests	0.04	0.00	0.19	0.52	0.25
	(0.04)	(0.00)	(0.10)	(0.13)	(0.12)
Paid, priv. disp. res.	0.01	0.25	0.00	0.38	0.36
	(0.03)	(0.11)	(0.00)	(0.15)	(0.14)
Gov't off's	0.13	0.00	0.00	0.57	0.30
	(0.10)	(0.00)	(0.00)	(0.13)	(0.14)
Other 3rd p	0.50	0.00	0.16	0.31	0.03
	(0.15)	(0.00)	(0.10)	(0.13)	(0.03)
Legal system	0.03	0.00	0.00	0.47	0.50
	(0.02)	(0.00)	(0.00)	(0.15)	(0.15)

Standard errors in parenthesis.

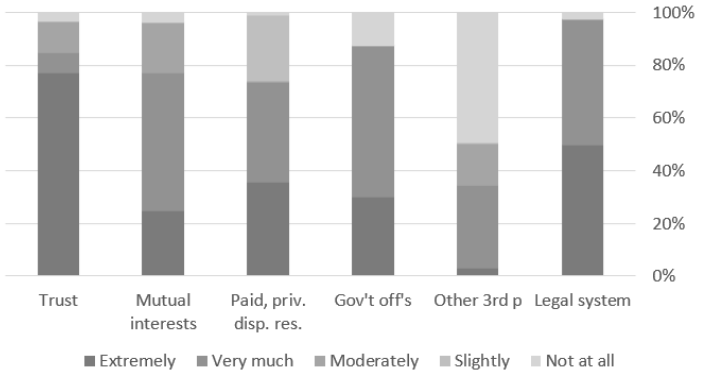
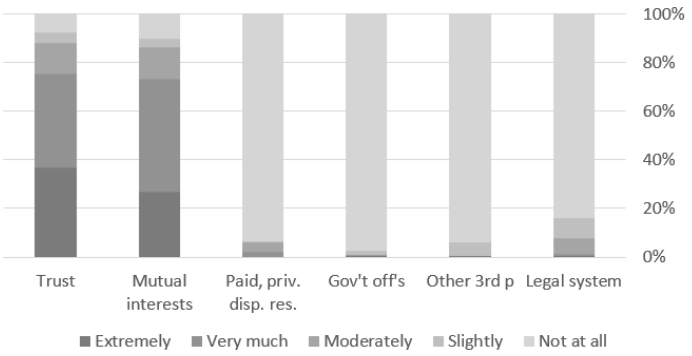


Table 2b: Predicted Response Probabilities, Relations with Customers

Class 1 – Pure Bilateralism

Method	Not at all	Slightly	Moderately	Very much	Extremely
Trust	0.08	0.04	0.13	0.39	0.37
	(0.02)	(0.01)	(0.02)	(0.03)	(0.03)
Mutual interests	0.10	0.03	0.13	0.46	0.27
	(0.02)	(0.01)	(0.02)	(0.03)	(0.03)
Paid, priv. disp. res.	0.94	0.00	0.04	0.02	0.00
	(0.04)	(0.00)	(0.03)	(0.02)	(0.00)
Gov't off's	0.98	0.02	0.00	0.00	0.00
	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)
Other 3rd p	0.94	0.05	0.00	0.00	0.00
	(0.02)	(0.02)	(0.00)	(0.00)	(0.00)
Legal system	0.84	0.08	0.06	0.01	0.00
	(0.02)	(0.02)	(0.01)	(0.01)	(0.00)

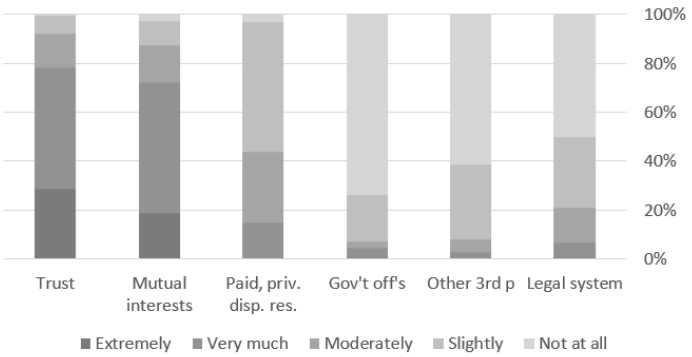
Standard errors in parenthesis.



Class 2 – Bilateralism with private support

Method	Not at all	Slightly	Moderately	Very much	Extremely
Trust	0.00	0.07	0.14	0.50	0.29
	(0.01)	(0.02)	(0.02)	(0.04)	(0.04)
Mutual interests	0.03	0.10	0.15	0.54	0.19
	(0.01)	(0.03)	(0.03)	(0.04)	(0.04)
Paid, priv. disp. res.	0.03	0.53	0.29	0.15	0.00
	(0.07)	(0.07)	(0.06)	(0.05)	(0.00)
Gov't off's	0.74	0.19	0.02	0.04	0.01
	(0.04)	(0.04)	(0.02)	(0.01)	(0.01)
Other 3rd p	0.61	0.31	0.05	0.03	0.00
	(0.05)	(0.05)	(0.01)	(0.01)	(0.00)
Legal system	0.50	0.29	0.15	0.06	0.01
	(0.05)	(0.05)	(0.03)	(0.02)	(0.01)

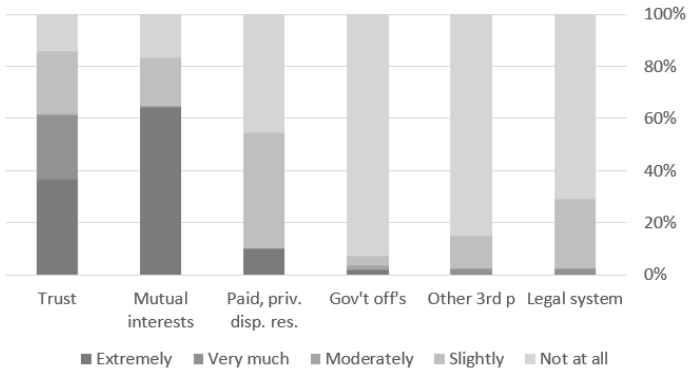
Standard errors in parenthesis.



Class 3 – Bilateralism with weak support

Method	Not at all	Slightly	Moderately	Very much	Extremely
Trust	0.14	0.24	0.00	0.25	0.37
	(0.05)	(0.06)	(0.00)	(0.06)	(0.07)
Mutual interests	0.17	0.18	0.00	0.00	0.64
	(0.05)	(0.05)	(0.00)	(0.01)	(0.07)
Paid, priv. disp. res.	0.45	0.45	0.00	0.00	0.10
	(0.09)	(0.08)	(0.00)	(0.00)	(0.04)
Gov't off's	0.93	0.04	0.02	0.00	0.02
	(0.03)	(0.02)	(0.01)	(0.00)	(0.01)
Other 3rd p	0.85	0.13	0.00	0.02	0.01
	(0.05)	(0.05)	(0.00)	(0.01)	(0.00)
Legal system	0.71	0.27	0.00	0.02	0.01
	(0.06)	(0.06)	(0.00)	(0.01)	(0.00)

Standard errors in parenthesis.



Class 4 – Weak comprehensive governance

Method	Not at all	Slightly	Moderately	Very much	Extremely
Trust	0.01	0.03	0.36	0.37	0.22
	(0.01)	(0.02)	(0.09)	(0.09)	(0.07)
Mutual interests	0.02	0.07	0.07	0.66	0.19
	(0.01)	(0.03)	(0.03)	(0.08)	(0.07)
Paid, priv. disp. res.	0.36	0.13	0.32	0.07	0.12
	(0.09)	(0.05)	(0.09)	(0.03)	(0.07)
Gov't off's	0.64	0.00	0.30	0.02	0.04
	(0.07)	(0.00)	(0.06)	(0.01)	(0.03)
Other 3rd p	0.37	0.02	0.49	0.12	0.01
	(0.10)	(0.01)	(0.10)	(0.05)	(0.01)
Legal system	0.54	0.12	0.10	0.01	0.23
	(0.09)	(0.05)	(0.04)	(0.01)	(0.09)

Standard errors in parenthesis.

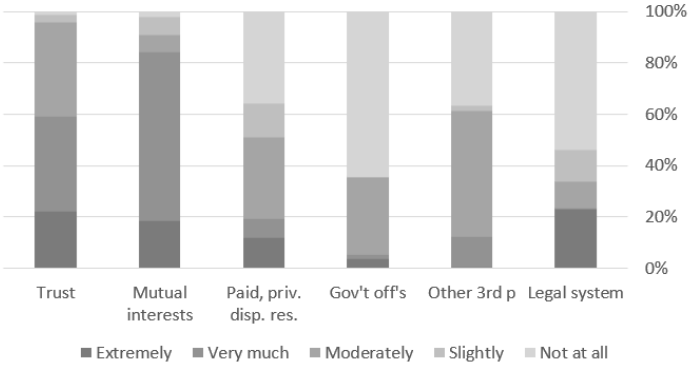


Table 3: Estimated Class Membership Probabilities

	$\hat{\pi}_c$	s.e.
Relations with Suppliers		
Pure bilateralism	0.657	0.024
Bilateralism with private support	0.166	0.022
Bilateralism with legal support	0.160	0.017
Strong comprehensive governance	0.017	0.004
Relations with Customers		
Pure bilateralism	0.565	0.032
Bilateralism with private support	0.242	0.000
Bilateralism with weak support	0.145	0.025
Weak comprehensive governance	0.050	0.008

Table 4: Tests of the Association of Country, Region, and Firm Characteristics with the Use of Governance Structures

	Tests of association with supplier governance structures	Tests of association with customer governance structures
	p-values for Wald tests	
Country	0.001***	0.000***
Sub-regions of:		
Argentina	0.000***	0.040**
Bolivia	0.390	0.140
Ecuador	0.000***	0.000***
Paraguay	0.000***	0.780
Peru	0.000***	0.270
Uruguay	0.019**	0.520
Firm size	0.082*	0.160
Sector of operation	0.000***	0.034**
Management practices	0.078*	0.031**
Proportion domestic private ownership	0.008***	0.200
At least 10% foreign owned	0.073*	0.019**

Significance levels: .01 - ***; .05 - **; .1 - *.

Table 5: Effect of Uncertainty on the Use of Pure Bilateralism

	(1)	(2)	(3)	(4)
Uncertainty	-0.1893***	-0.1943***	-0.1939***	-0.2078**
	(0.048)	(0.049)	(0.047)	(0.103)
Log of size			-0.030*	-0.030*
			(0.016)	(0.016)
Firm is part of a larger firm			-0.003	-0.003
			(0.058)	(0.057)
Firm age (years)			-0.000	-0.000
			(0.001)	(0.001)
Exporting directly at least 10% of sales			-0.032	-0.029
			(0.058)	(0.059)
At least 10% foreign-owned			-0.096	-0.096
			(0.068)	(0.067)
2-digit sector fixed effects	NO	NO	YES	YES
Region fixed effects	NO	YES	YES	YES
Number of observations	2,505	2,505	2,505	2,505
R^2	0.047	0.064	0.122	
Cragg-Donald Wald F statistic				194.19
Kleibergen-Paap Wald F statistic				155.22 {0.00}

Notes: OLS (columns 1-3) and IV regression (column 4) estimates are shown. Sector-region averages of uncertainty are used for the instrument in (4). The coefficient from the first-stage regression, which includes all exogenous variables included in the second-stage, is 0.9578*** (with standard error 0.077). Standard errors are in parentheses, calculated assuming clustering at the level of strata. All regressions use sampling weights. Braces are used for p-values. Significance levels: .01 - ***; .05 - **; .1 - *.

Figure 1a: Cross-country variation of the mix of governance structures used in relations with suppliers

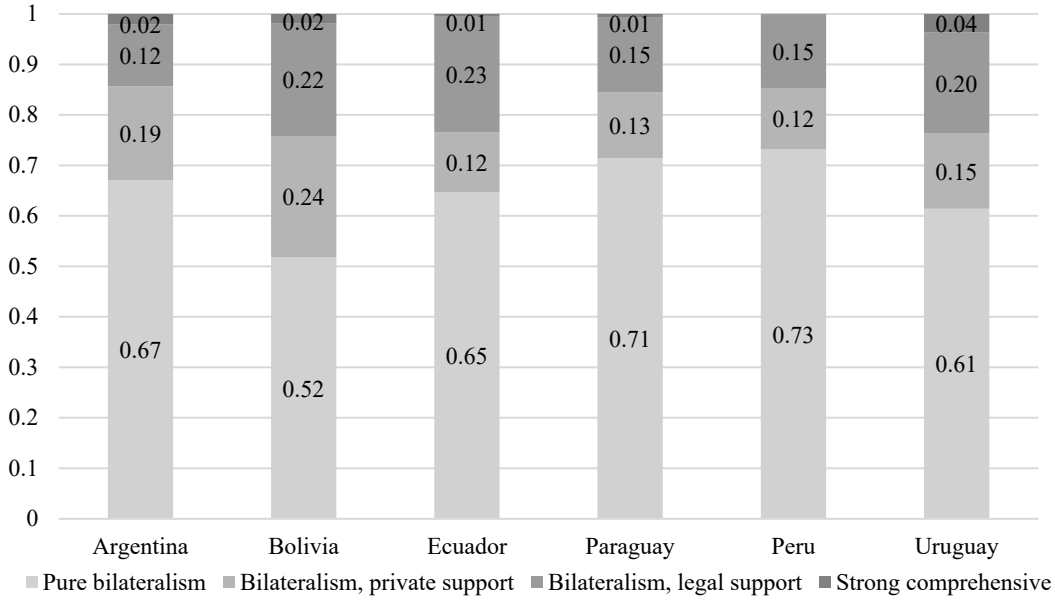


Figure 1b: Cross-country variation of the mix of governance structures used in relations with customers

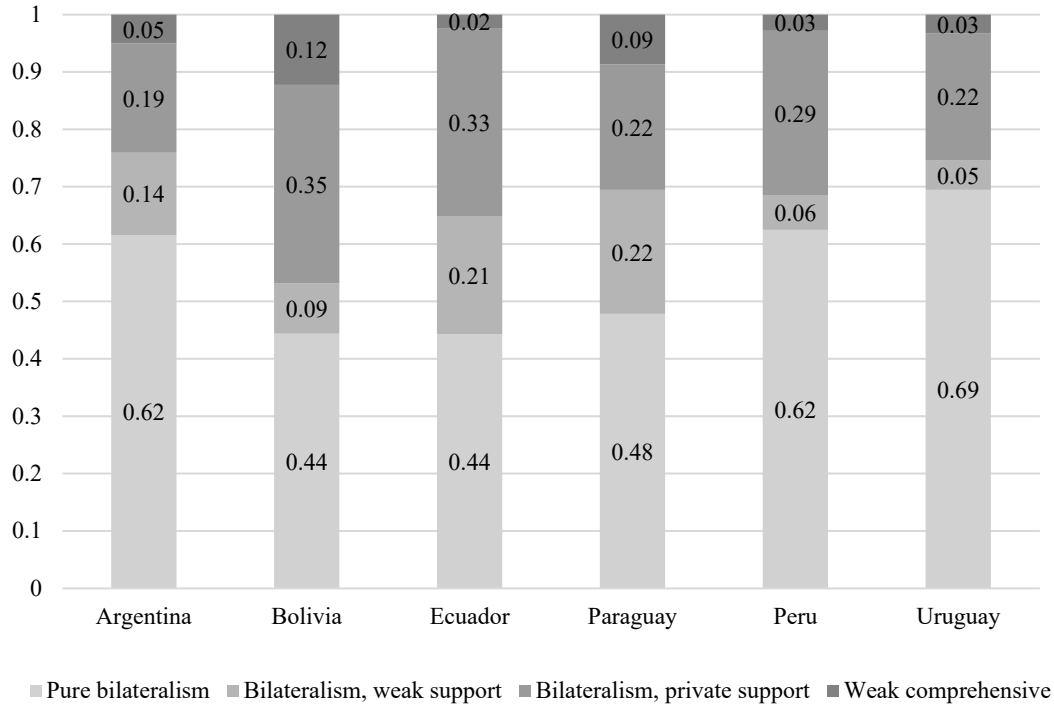


Figure 2a: Within-country variation of governance structures for supplier-relations

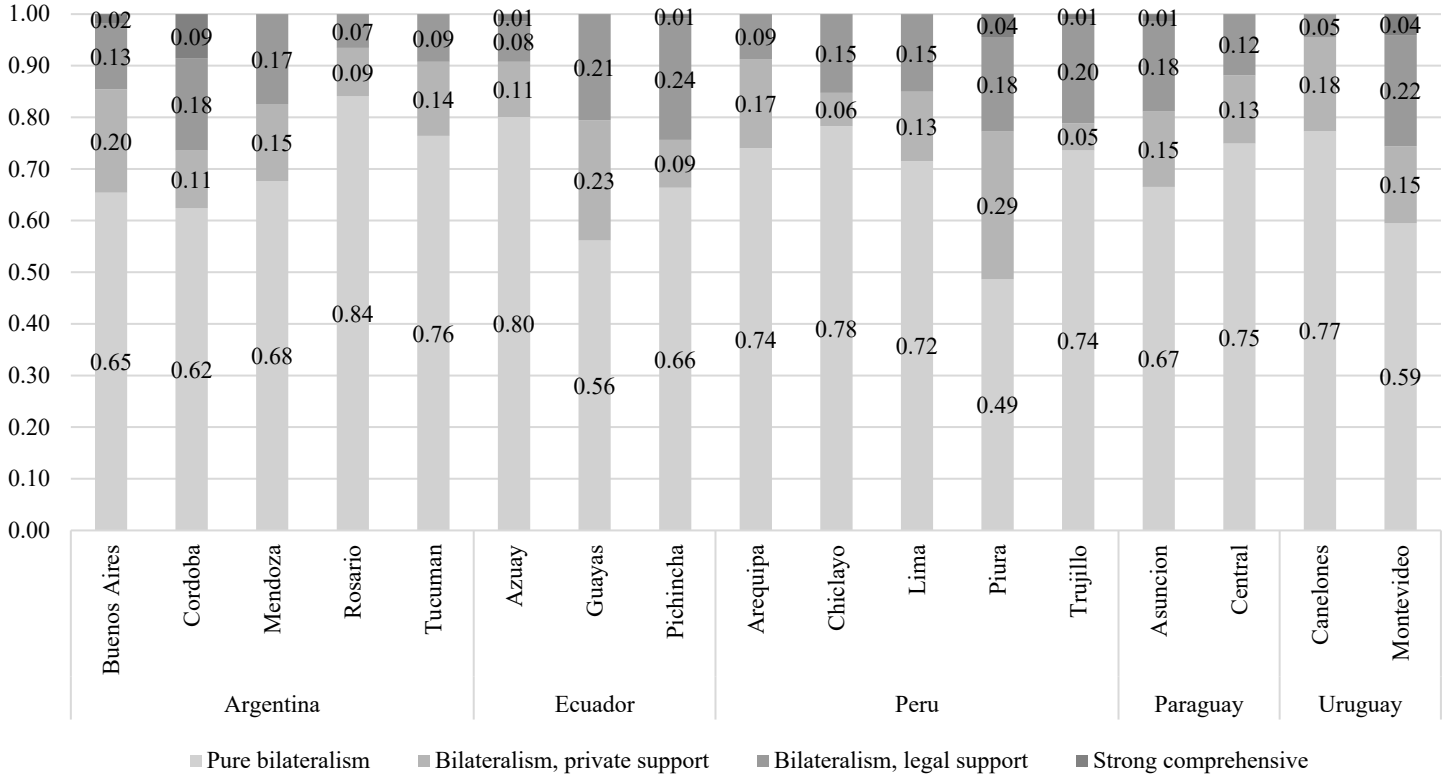


Figure 2b: Within-country variation of governance structures for customer-relations

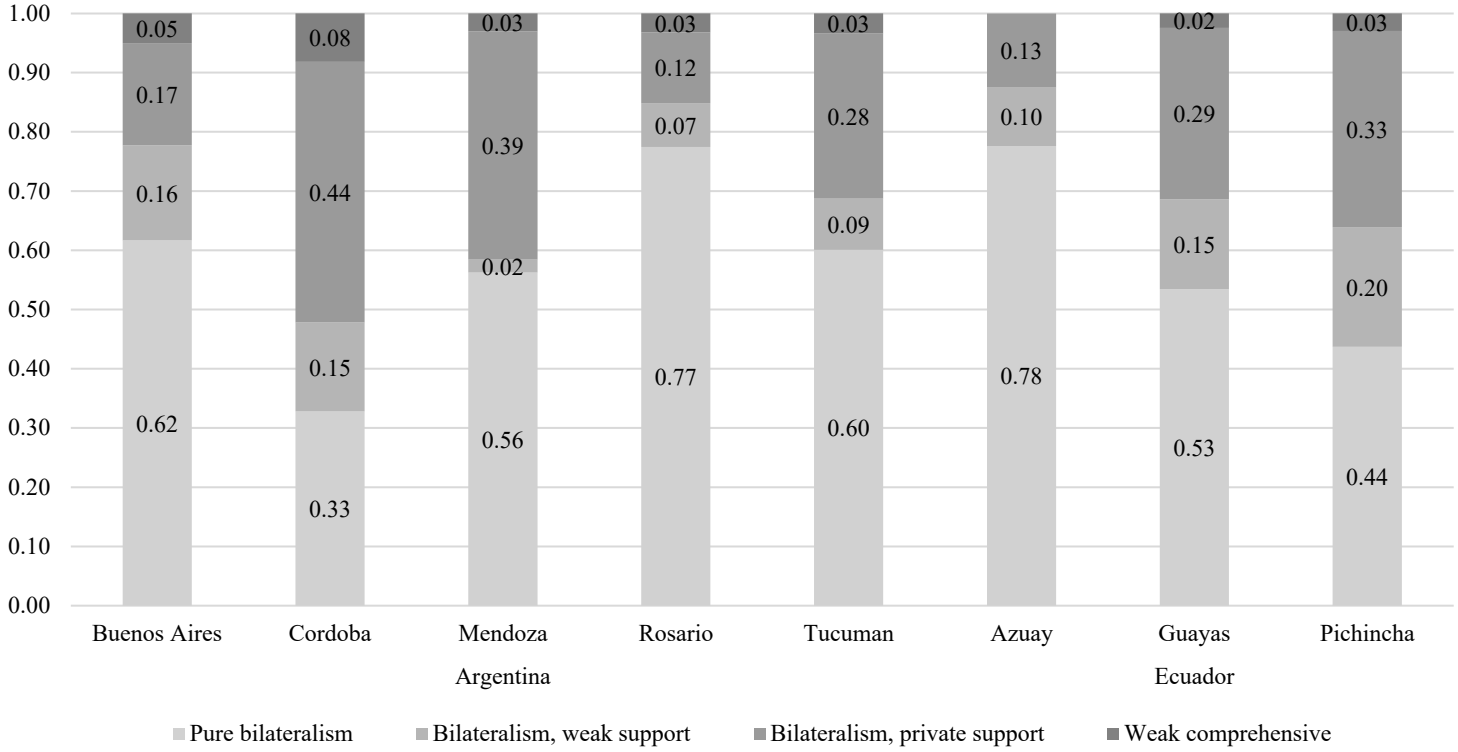


Figure 3a: Management practices and governance structures towards suppliers

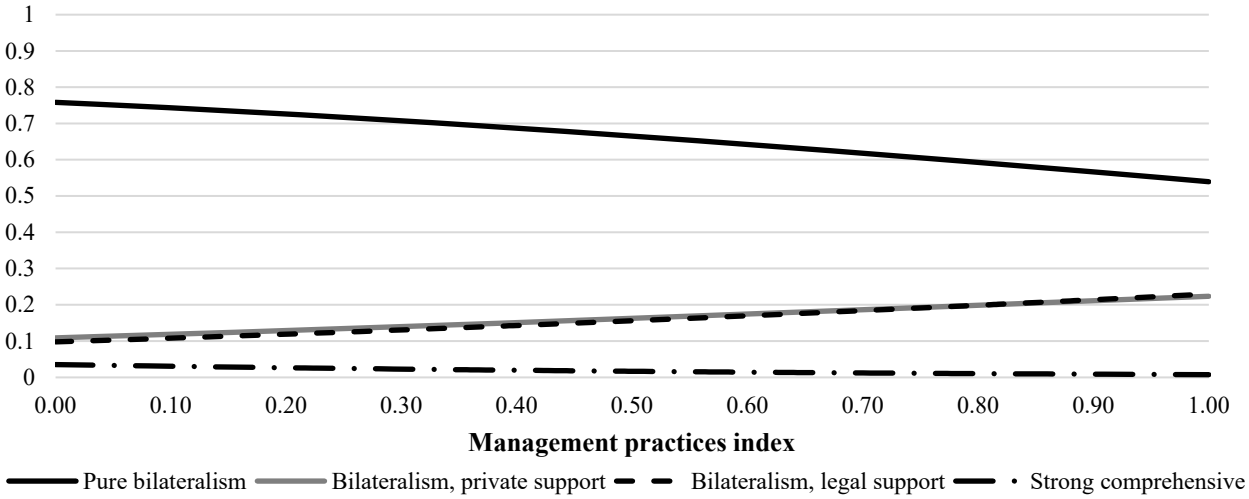
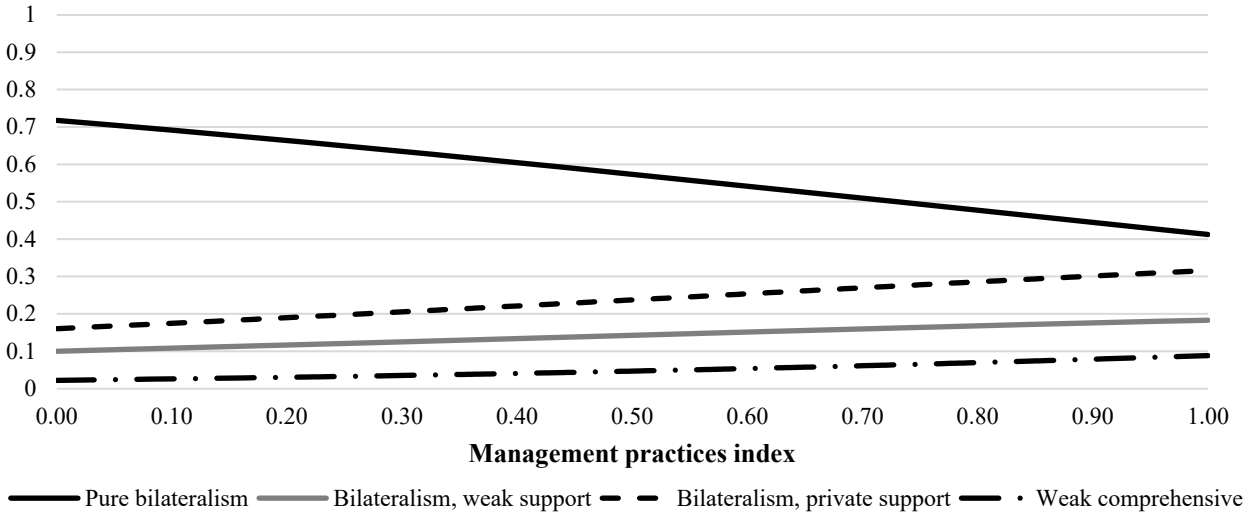


Figure 3b: Management practices and governance structures towards customers



Supplementary Appendixes

Transactional Governance Structures: New Cross-Country Data and an Application to the Effect of Uncertainty

Peter Murrell Nona Karalashvili David C. Francis

October 18, 2022

Appendix A: Details of Question Design, Survey Implementation, and Item Responses Rates

A.1 Wording of Questions in English and Spanish, and Show Card

ASCD.9	When making agreements with **suppliers**, please indicate to what degree each of the following is **effective in resolving or preventing** problems. **[READ OUT OPTIONS] SHOW CARD 8**

	Not at all	Slightly	Moderately	Very much	Extremely	**DON'T KNOW (SPONTANEOUS)**
Personal relationship and trust **ASCd9a**	1	2	3	4	5	**-9**
Mutual interest in maintaining business relationship, without involving others **ASCd9b**	1	2	3	4	5	**-9**
Paid, private dispute resolution **ASCd9d**	1	2	3	4	5	**-9**
Assistance of government officials **ASCd9e**	1	2	3	4	5	**-9**
Intervention of other third-parties (excluding paid, private dispute resolution and government officials) **ASCd9c**	1	2	3	4	5	**-9**
Legal system **ASCd9f**	1	2	3	4	5	**-9**

ASCD.9	Por favor indique en qué medida las siguientes circunstancias son efectivas para resolver o evitar problemas en los acuerdos con proveedores. **LEER OPCIONES -MOSTRAR TARJETA 8**

	Para nada	Ligera-mente	Modera-damente	Bastante	Extreme-damente	**NO SABE (ESPONTÁNEO)**
Relaciones personales y confianza **ASCd9a**	1	2	3	4	5	**-9**
El interés mutuo de mantener una relación de negocios sin tener que involucrar a terceros **ASCd9b**	1	2	3	4	5	**-9**
Mecanismos de resolución privados ofrecidos por terceros y que son pagados **ASCd9d**	1	2	3	4	5	**-9**
Ayuda de funcionarios del gobierno **ASCd9e**	1	2	3	4	5	**-9**
Intervención de otros terceros (excluyendo entes privados y pagados y personas del gobierno **ASCd9c**	1	2	3	4	5	**-9**
Recurso al sistema legal **ASCd9f**	1	2	3	4	5	**-9**

ASCD.18	When making agreements with **customers**, please indicate to what degree each of the following is **effective in resolving or preventing** problems. **[READ OUT OPTIONS]** **SHOW CARD 8**

	Not at all	Slightly	Moderately	Very much	Extremely	DON'T KNOW (SPONTANEOUS)
Personal relationship and trust **ASCd18a**	1	2	3	4	5	**-9**
Mutual interest in maintaining business relationship, without involving others **ASCd18b**	1	2	3	4	5	**-9**
Paid, private dispute resolution **ASCd18d**	1	2	3	4	5	**-9**
Assistance of government officials **ASCd18e**	1	2	3	4	5	**-9**
Intervention of other third-parties (excluding private dispute resolution and government officials) **ASCd18c**	1	2	3	4	5	**-9**
Legal system **ASCd18f**	1	2	3	4	5	**-9**

ASCD.18	Por favor indique en qué medida las siguientes circunstancias son efectivas para resolver o evitar problemas en los acuerdos con clientes. **[LEER OPCIONES] MOSTRAR TARJETA 8**

	Para nada	Ligera-mente	Moderada-mente	Bastante	Extreme-damente	NO SABE (ESPONTANEO)
Relaciones personales y confianza **ASCd18a**	1	2	3	4	5	**-9**
El interés mutuo de mantener una relación de negocios sin tener que involucrar a terceros **ASCd18b**	1	2	3	4	5	**-9**
Mecanismos de resolución privados ofrecidos por terceros y que son pagados **ASCd18d**	1	2	3	4	5	**-9**
Ayuda de funcionarios del gobierno **ASCd18e**	1	2	3	4	5	**-9**
Intervención de otros terceros (excluyendo entes privados y pagados y personas del gobierno **ASCd18c**	1	2	3	4	5	**-9**
Recurso al sistema legal **ASCd18f**	1	2	3	4	5	**-9**

The Show Card Used When Administering the Questions

Relaciones personales y confianza	Para nada	Ligeramente	Moderadamente	Bastante	Extremadamente
El interés mutuo de mantener una relación de negocios sin tener que involucrar a terceros	Para nada	Ligeramente	Moderadamente	Bastante	Extremadamente
Mecanismos de resolución privados ofrecidos por terceros y que son pagados	Para nada	Ligeramente	Moderadamente	Bastante	Extremadamente
Ayuda de funcionarios del gobierno	Para nada	Ligeramente	Moderadamente	Bastante	Extremadamente
Intervención de otros terceros (excluyendo entes privados y pagados y personas del gobierno)	Para nada	Ligeramente	Moderadamente	Bastante	Extremadamente
Recurso al sistema legal	Para nada	Ligeramente	Moderadamente	Bastante	Extremadamente

A.2 Translation of Questions

The questions were designed in English and then translated into Spanish. In translation to Spanish, the use of the phrase "which of the following" necessitated a noun, with "circunstancias" used, most directly translated as "circumstances" but also possibly understood as "situations". This phrasing also led to the translation of sub-question (or enforcement agent) 6 as "recurso al sistema legal" or "recourse to the legal system". Both adjustments merit some comment as they may affect our analysis.

The use of "circumstances" as well as the word "recourse" may result in respondents' understanding questions in terms of the realized circumstances in which they found themselves or the actions they had actively undertaken (for instance, through the legal system). On the other hand, the expression "resolving or preventing problems" does appear in the question, suggesting that respondents should have borne in mind pre-emptive acts that did not go as far as, for example, filing a legal action. Piloting of these questions indicated that some respondents understood these questions as referring to their own actions, rather than indicating problems being prevented by the threat, but not use, of an action. To the extent that these questions are understood as referring to actions or realized experience, enforcement agents that involve only the threat of action rather than the action itself, such as the shadow of the law, will be under-reported or rated lower.

During translations, surveyors reviewed the questionnaire, commenting on any problematic terms, which would be changed before implementation. Standard back-translation procedures were applied. No conceptual issues were identified in this process.

A.3 Survey Implementation

After a draft questionnaire and its translation was completed, WBES implementers (including one of the paper's authors) met with the local surveyors for in-country training on the survey methodology and a question-by-question review of the questionnaire. The survey was piloted with firms that were not in the sample, allowing for the possibility of changes in wording even at this stage. Following these pilot interviews, a report of potential issues (such as question wording) was provided: no important problems were reported on these questions, in any of the separate reports provided for each of the six countries.

One product of this preparation was an interviewer manual, an important part of which is instructions about the interviewer should react to respondent inquiries about how to interpret specific questions. Help in interpreting questions has been shown to increase measurement reliability (Conrad and Schober 2000). WBES procedure confines interviewer comments to those that appear in the manual. The relevant section of the manual is included in the next subsection of this appendix.

A 'cognitive audit' of the questions was also undertaken, using open-ended interviews with managers of a heterogenous collection of firms. One inquiry made in these audits was on how managers understood the use of the word "effective". Interviewees did interpret that term as referring to a combination of both the extent of reliance and productiveness of the enforcement agent: the lowest ranking on the scale would be chosen if either the enforcement agent was not relied upon or if it had been relied upon and it did not work for the firm. The qualitative interviews also revealed that the question was not interpreted in a conjectural way but rather was taken to refer to the firm's actual experience. Thus, if a respondent had no knowledge of how effective an enforcement agent could be because the respondent's firm had not used the agent, it was given the lowest rating on the scale.

Feedback during and after survey implementation could have potentially provided indications of whether problems had arisen with any question. Interviewers and their supervisors could ask the researchers for clarification of issues that surfaced during fieldwork. No important issues arose that were related to the questions used in this paper. After fieldwork concluded, the local survey administrators provided a report with comments on implementation. There were no specific comments on the questions used in this paper.

A.4 The Relevant Sections of the Interviewer Manual

Interviewers must first read instructions and questions as written. Then, on respondent inquiries, interviewers use their manual to explain the meaning of some types of questions to the respondents. The following contains the (translated) sections of the manual relevant to the questions used in this paper.

"Resolving or preventing problems in agreements incudes negotiations, how agreements are finalized, and how other available means are implemented to resolve problems when they appear. Each one of the following six enforcement agents includes not only their use but also the promise or threat of their use when a problem emerges.

Personal relationship and trust: self-explanatory

Mutual interest in maintaining business relationships, without involving others: When the establishment and its suppliers both know that each will complete their part of an agreement or that disagreements will be resolved because the two have an interest in continuing doing business in the future.

Paid, private dispute resolution: includes only dispute resolution mechanisms that are private and paid, including the threat of using those mechanisms. It includes arbitration, the use of legally constituted private security, and/or the involvement of criminal groups. This option does not include third parties that are not paid, for example the use of business associations that sometimes help resolve problems without charge.

Assistance of government officials: includes the formal or informal involvement of representatives of local or central governments as well as the threat of involving officials in a position to resolve or avoid problems.

Intervention of other third-parties (excluding paid, private dispute resolution and government officials): includes all interventions (or associated threats) of third parties excluding paid third parties or the government. This includes the participation of business organizations or other customers or suppliers, as well as other respected community members like business leaders or community organizations.

Legal system: includes the direct use of the Legal System, including the courts or civil action, as well as the threat of legal action in the case of problems. This also includes the design of agreements in such a way that it facilitates legal recourse in the case that it is necessary. "

A.5 Some Characteristics of the Six South American Countries, Together with Global Comparisons

Table A.1: Some Characteristics of the Six South American Countries

Country	GDP per capita, PPP (current int. $, 2016)	WGI: Rule of Law, 2016	WJP: Civil Justice Index, 2017	Doing Business DTF: Enforcing Agreements 2017 (higher is better)	Interpersonal trust (% very or somewhat trustworthy)	Fractionalization		
						Ethnic	Language	Religion
Argentina	19,939	-0.35	0.58	55.6	71.6%	25.5%	6.2%	22.4%
Bolivia	7,234	-1.20	0.34	54.6	46.0%	74.0%	22.4%	20.8%
Ecuador	11,242	-0.69	0.46	56.0	55.2%	65.5%	13.1%	14.2%
Paraguay	9,567	-0.67		59.7	69.0%	16.9%	59.8%	21.2%
Peru	13,018	-0.49	0.44	60.7	46.7%	65.7%	33.6%	19.9%
Uruguay	21,619	0.63	0.74	54.4	75.9%	25.0%	8.2%	35.5%
Latin America & Caribbean	15,210	-0.06	0.53	54.0	65.7%	40.5%	18.9%	44.4%
Europe & Central Asia	31,361	0.53	0.63	65.7		32.7%	31.2%	40.0%
East Asia & Pacific	17,021	0.21	0.58	56.4		29.6%	39.3%	52.0%
World	16,214	0.00	0.56	55.3	67.4%	43.9%	38.6%	43.7%

Notes:
WGI = World Governance Indicators; WJP = World Justice Project; Doing Business DTF = the distance to the frontier measure of Doing Business.
Countries included in regional averages vary by the respective data availability (e.g. the interpersonal trust world average includes only Canada and the United States of America in addition to the Latin American and Caribbean countries).
The WGI rule of law index captures perceptions of the extent to which agents have confidence in and abide by the society's rules.
The WJP Civil Justice Index measures whether disputes can be resolved peacefully and effectively through the civil justice system.
The DB distance to frontier score is measured on a scale of 0 to 100, where 100 is best practice on enforcing agreements and 0 represents the lowest performance.
Interpersonal trust is based on Americas Barometer and shows the percentage answering 'Very trustworthy' or 'Somewhat trustworthy' to the following question: "And speaking of the people from around here, would you say that people in this community are very trustworthy, somewhat trustworthy, not very trustworthy or untrustworthy?"
The data on fractionalization is from Alesina et al. (2003) and is available for different countries in different years, ranging from 1981 to 1998. The data measures 100 times the probability that a random member of the population is not from the same group.

A.6 Data Collection

All interviews were conducted face-to-face with business owners and top managers using tablet devices. Table A.2 provides information on the dates of fieldwork and the total number of interviews conducted in each country. Fieldwork started in each country following a three- or four-day training and piloting phase.

Table A.2: Dates of Fieldwork and Sample Sizes

Country	Dates of Fieldwork	Number of Firms Surveyed
Argentina	March 2017 through March 2018	991
Bolivia	January 2017 through June 2017	364
Ecuador	March 2017 through October 2018	361
Paraguay	February 2017 through August 2017	364
Peru	March 2017 through March 2018	1,003
Uruguay	March 2017 through December 2017	347
Total		3,430

A.7 Item Non-response

Table A.3 shows item non-response rates. These rates consider (spontaneously given) "Don't Know" responses as non-response. "Don't Know" is not displayed as a possible option in the 'show card' that lists possible responses. More than 97% of respondents answered all the six sub-questions about the methods of governing relations with both suppliers and customers, i.e., not once saying "Don't Know" to any of the 12 sub-questions. The question with the most frequent occurrence of "Don't Know" for relations with suppliers is on paid private dispute resolution (1.4% of sample). For relations with customers, the question about personal trust had the highest item non-response (1.2% of sample). Given the low item non-response rates, in our application of LCA we drop observations that have at least one "Don't Know" in the relevant series of questions. This leaves 3,350 observations on relations with suppliers (97.7% of the sample), and 3,339 observations on relations with customers (97.3% of the sample).

Table A.3: Item Non-response Due to "Don't Know" Responses

Country	Total	Share (%) of respondents with different numbers of "Don't Know" responses (of 6 questions)					
		Relations with suppliers			Relations with customers		
		0	1	2+	0	1	2+
Argentina	991	**96.5**	2.1	1.4	**94.9**	2.7	2.4
Bolivia	364	**95.6**	3.3	1.1	**96.7**	1.6	1.6
Ecuador	361	**99.7**	0.3	0.0	**99.4**	0.6	0.0
Paraguay	364	**98.1**	1.1	0.8	**98.6**	1.1	0.3
Peru	1,003	**99.1**	0.6	0.3	**98.8**	0.6	0.6
Uruguay	347	**96.5**	1.4	2.0	**97.4**	0.9	1.7
Overall N	3,430	**97.7**	1.4	0.9	**97.3**	1.4	1.3

A.8 Most Common Response Patterns

Table A.4: Most common response patterns on questions about relations with suppliers

	Effectiveness of:						Weighted Share	Cumul.	N
	Trust	Mutual interests	Paid third parties	Gov't officials	Other third parties	Legal system			
1	Vm	Vm	No	No	No	No	12%	12%	428
2	Ex	Ex	No	No	No	No	10%	22%	253
3	No	No	No	No	No	No	5%	27%	146
4	Ex	Vm	No	No	No	No	4%	30%	96
5	Mo	Mo	No	No	No	No	3%	33%	118
6	Vm	Mo	No	No	No	No	2%	36%	65
7	Vm	Ex	No	No	No	No	2%	38%	45
8	Vm	Vm	Sl	No	No	No	2%	40%	64
9	Vm	Vm	Vm	No	No	No	2%	41%	52
10	Vm	Vm	Mo	No	No	No	1%	43%	51
11	Sl	Sl	No	No	No	No	1%	44%	63
12	Vm	Vm	Mo	No	No	Sl	1%	45%	21
13	Mo	Vm	No	No	No	No	1%	46%	52
14	Sl	No	No	No	No	No	1%	47%	49
15	Ex	Ex	Sl	No	No	No	1%	48%	23
16	Ex	Ex	Ex	No	No	No	1%	49%	13
17	Vm	No	No	No	No	No	1%	50%	29

Table A.5: Most common response patterns on questions about relations with customers

	Effectiveness of:						Weighted Share	Cumul.	N
	Trust	Mutual interests	Paid third parties	Gov't officials	Other third parties	Legal system			
1	Vm	Vm	No	No	No	No	14%	14%	497
2	Ex	Ex	No	No	No	No	13%	27%	336
3	Ex	Vm	No	No	No	No	4%	31%	107
4	No	No	No	No	No	No	3%	34%	134
5	Mo	Mo	No	No	No	No	2%	36%	101
6	Ex	Ex	Sl	No	No	No	2%	37%	33
7	Vm	Vm	Sl	No	No	Sl	2%	39%	48
8	Vm	Vm	Mo	No	No	No	2%	41%	52
9	Vm	Vm	Vm	No	No	No	1%	42%	41
10	Ex	Ex	No	No	No	Sl	1%	43%	36
11	Vm	Vm	Sl	No	No	No	1%	45%	69
12	Sl	Sl	Sl	No	No	No	1%	46%	42
13	Mo	Vm	No	No	No	No	1%	47%	46
14	Vm	Ex	No	No	No	No	1%	48%	39
15	Vm	Vm	No	No	No	Sl	1%	50%	48

Ex	Vm	Mo	Sl	No
Extremely	Very much	Moderately	Slightly	Not at all

References to Appendix A

Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S. and Wacziarg, R., 2003. Fractionalization. *Journal of Economic Growth*, *8*(2), pp.155-194.

Doing Business, The World Bank (http://www.doingbusiness.org).

Latin American Public Opinion Project. 2016/7. The Americas Barometer. Latin American Public Opinion Project (LAPOP), Vanderbilt University. http://www.LapopSurveys.org (accessed June 12, 2018).

World Development Indicators, The World Bank

World Justice Project. 2018. Rule of Law Index (http://worldjusticeproject.org)

Worldwide Governance Indicators (www.govindicators.org)

Appendix B: Notes on LCA

B.1 An Intuitive Introduction to LCA

LCA is analogous to factor analysis (or principal components). In the simplest case of factor analysis, a continuous, cardinal, latent variable is estimated using a set of observed measures that reflect the variable with error. LCA is used when estimating a discrete, nominal, latent variable from a set of measures that reflect the variable with error. The key difference, then, lies in the measurement characteristics of the estimated latent variable. For both factor analysis and LCA, all types of observed variables (categorical, continuous, etc.) can be used (perhaps with slight modifications in the details of the statistical procedures). In our application, the measures are the survey responses to the sub-questions on the use of the six different enforcement agents. Each latent class is a transactional governance structure, reflecting a combination of a set of enforcement agents.

For our simplified example, suppose that a researcher visits the country of Erewhon and asks the following questions to representatives of 500 firms:

1. When making agreements with suppliers, please indicate to what degree personal trust is effective in resolving or preventing problems: 'not at all', 'moderately', or 'extremely'.
2. The same question with 'legal system' substituted for 'personal trust'.

The fictitious responses appear in Table B.1, a 3x3 contingency table.

Systematic patterns in this table are not obvious. A standard approach in first parsing the data would be to assume that the probability of choosing one of the three answers for personal trust is independent of the probability of choosing any one of the answers for legal system; this is the independence assumption. But that assumption is obviously incorrect: a standard chi-squared test rejects it at the 0.001 level.

LCA is a method of uncovering a simple structure in such data. It begins by postulating that there are distinct classes of firms. In so doing, it suggests that the failure of the independence assumption in the aggregate data arises from the fact that the responses reflect a mixture of different classes of firms. Firms within a class are viewed as all having the same data-generating process for the survey responses. In the simplest application of LCA, the independence assumption is applied within classes, and hence is usually referred to as local independence.

Let us suppose that there are two classes of firms, each class having a different approach to the governance of agreements. Very roughly speaking, LCA uses correlations in the answers to the two different questions to estimate the row and column probabilities for each class and the proportion of firms falling into each class. This leads to two separate contingency tables (B.2 and B.3) the first reflecting the responses of 200 firms and the second for 300 firms. (The numbers of firms in each category are a product of the estimation and are not imposed a priori. The number of classes is an a priori assumption.) The local independence assumption is satisfied exactly within each table: the number in each cell is a product of its row probability, column probability, and the

number of firms in the class. Thus, within each table, the standard chi-squared test-statistic is zero. Table B.1 is simply a cell-by-cell summation of Tables B.2 and B.3, which shows the essence of LCA—the aggregate data are assumed to arise from a mixture of simple distributions.

Now, the patterns in the data stand out starkly and are easy to describe. For the firms in Table B.2, both enforcement agents are highly effective. For Table B.2b, the legal system is ineffective and personal trust is effective. We could label the strategies of those in Table B.2a as 'comprehensive governance', while the strategies of those in Table B.2b are 'purely bilateral'. The litmus test of an insightful LCA application is an evocative description of the behavior typical within each class, and the identification of stark differences between the behaviors of each class.

Note that in this example LCA estimates nine parameters, requiring more estimates than the eight that would directly reflect Table B.1. Nevertheless, it adds a rich understanding of the data generating process by identifying two quite separate and meaningful patterns of behavior. The full benefit of LCA arrives only when the complexity of the problem increases. The number of parameters to be estimated by LCA increases linearly in the number of questions asked. In contrast, the number of cells in the contingency matrix analogous to Table B.1 increases exponentially. In the data analyzed in this paper, the number of cells is 5^6 (15,625) and nearly this many parameters would have to be estimated without the imposition of a simple structure. In contrast, a 2-class LCA model applied to the same data would require estimation of 49 parameters.

Table B.1: An example of the use of LCA: whole sample

		Effectiveness of personal trust			Row probability
		Not at all	Moderately	Extremely	
Effectiveness of legal system	Not at all	26	30	204	**0.52**
	Moderately	7	15	48	**0.14**
	Extremely	17	45	108	**0.34**
	Column probability	**0.10**	**0.18**	**0.72**	

Table B.2: An example of the use of LCA: class 1

		Effectiveness of personal trust			Row probability
		Not at all	Moderately	Extremely	
Effectiveness of legal system	Not at all	2	6	12	**0.10**
	Moderately	4	12	24	**0.20**
	Extremely	14	42	84	**0.70**
	Column probability	**0.10**	**0.30**	**0.60**	

Table B.3: An example of the use of LCA: class 2

		Effectiveness of personal trust			Row probability
		Not at all	Moderately	Extremely	
Effectiveness of legal system	Not at all	24	24	192	**0.80**
	Moderately	3	3	24	**0.10**
	Extremely	3	3	24	**0.10**
	Column probability	**0.10**	**0.10**	**0.80**	

B.2 Estimating the Correlates of Class Membership

Sections V and VI conduct some exploratory exercises in relating class membership to characteristics of firms or their environments. The following paragraphs outline how the estimation methodology of those sections integrates with the estimation methodology for the latent classes that is laid out in Section III.

Equation (2) of Section III is easily modified to take into account the dependence of class membership on firm characteristics. Let $\pi_c(Z_i)$ be the probability of membership in latent class c given that the firm has characteristics Z_i. These Z_i may include characteristics of the firm (e.g., size) as well as the environment or context in which it operates (e.g., culture). Then the probability of observing a specific response vector, Y_i, for firm i with characteristics Z_i is:

$$P(Y_i \,|Z_i) = \sum_{c=1}^{C} \pi_c(Z_i) \prod_{h=1}^{H} f(Y_{ih}|c) \qquad \text{(B. 1)}$$

With this model, one estimates the functions $f(\cdot \mid \cdot)$ and $\pi_c(.)$.

If (B.1) is the preferred model, there are two routes to estimation. One obvious choice is to form a likelihood from (B.1) and estimate the $f(\cdot \mid \cdot)$ and the $\pi_c(.)$ directly. Alternatively, one could proceed in three steps. First, maximize the likelihood (3) of Section III, estimating $f(\cdot \mid \cdot)$ and the π_c. Then, use Bayes theorem to estimate firm-specific class membership probabilities for each firm, $\hat{\pi}_{ci}$. Finally, use regression techniques to estimate the functions $\pi_c(.), c = 1, \ldots C$, using as data the $\hat{\pi}_{ci}$ and Z_i.

There is a large literature, both theoretical and applied, reflecting on the choice between the two routes to estimation. From theory, there are procedures to obtain consistent estimates of $f(\cdot \mid \cdot)$ and $\pi_c(.)$ using the 3-step process (Vermunt 2010, Bakk et al. 2013, Bakk et al. 2014). The applied literature suggests that using the 3-step process is advisable unless one has great confidence in the specification of (B.1), especially understanding which Z_i to include and exclude (Nylund-Gibson and Masyn 2016).[1] For that reason, we use the 3-step process because many of the relevant Z_i's remain unmeasured or unknown when using cross-country data. This is the case because of the

[1] Jedidi et al. (1997) make a similar comment in a more general modeling context than LCA.

inherent difficulty of collecting cross-country data on transaction-related activities and because the discriminating-alignment research program—identifying the Z_i's—is still a work in progress.

Our decision to use the 3-step procedure also rests on our central goal of understanding the nature of the classes. A prime contribution of this paper is the characterization of the most common governance structures (step 1, implemented in Section IV). By estimating the classes in general—independently of the correlates of class membership—we can focus on this goal and provide readers with results unencumbered by any more ambitious objectives, together with the strong assumptions that would be necessary to attain these objectives. Then by producing the Bayesian posterior probabilities of step 2 of the 3-step procedure (the $\hat{\pi}_{ci}$), we provide examples in Sections V and VI of an approach that could be easily followed by others. They could use our data on posterior probabilities and select their own Z_i's from the copious data available from the WBES or other sources.

References to Appendix B

Bakk, Zsuzsa, Fetene B. Tekle and Jeroen K. Vermunt. 2013. Estimating the Association between Latent Class Membership and External Variables Using Bias-adjusted Three-step Approaches. *Sociological Methodology* 2013 43: 272.

Bakk, Zsuzsa, Oberski, Daniel L., and Vermunt, Jeroen K. 2014. Relating Latent Class Assignments to External Variables: Standard Errors for Correct Inference. *Political Analysis*, 22, 520-540.

Jedidi, Kamel, Harsharanjeet S. Jagpal, and Wayne S. DeSarbo. "Finite-mixture structural equation models for response-based segmentation and unobserved heterogeneity." *Marketing Science* 16, no. 1 (1997): 39-59.

Nylund-Gibson, Karen, Katherine E. Masyn 2016 Covariates and Mixture Modeling: Results of a Simulation Study Exploring the Impact of Mis-specified Effects on Class Enumeration, *Structural Equation Modeling: A Multidisciplinary Journal*, 23:6, 782-797.

Vermunt, Jeroen K., 2010. Latent class modeling with covariates: Two improved three-step approaches. *Political Analysis*, 18(4), pp.450-469.

Appendix C: Methods for Choosing the Preferred Model

This Appendix details the process of choosing one LCA model for customer relations and one for supplier relations when implementing step-1 of the 3-step procedure described in Appendix B.2. Implementation of step-1 involves estimation of $f(\cdot \mid \cdot)$ and the π_c by maximizing the likelihood at equation (3) of Section III. However, implementation of that step requires choosing one version of the DGP, by specifying the number of classes and the particular way in which the assumption of local independence is relaxed. In terms of the notation in the text, this entails choosing *C*, dividing the *K* elements of the response vector ($Y_i = (y_{i1}, \ldots, y_{iK})$) into *H* subsets, and placing structure on the form of the $f(Y_{ih}|c)$ to be used for the estimation.

C.1 The Subset of Models to be Considered

It is not possible to conduct a search over all possible models in order to find a single optimum. For example, when considering specific forms of the relaxation of local independence, our 6 separate responses can form a total of 15 unique pairs of responses, with 32,766 combinations of these pairs each generating a different model as a possibility for consideration.[2]

It is natural to include for consideration models that use the assumption of local independence in all possible situations, but there are also reasons to consider a relaxation of that assumption, as described in Section III.2. As noted previously, respondents might not clearly distinguish between trust (sub-question 1) and mutual interest (sub-question 2). In the cognitive interviews conducted prior to the survey, we learned that individuals sometimes did not clearly distinguish assistance of government officials (agent 4) from intervention of other third-parties (agent 5). As a result, it seems natural to consider model specifications that allow dependence of these two responses as well. This gives us four types of models to consider. These are: (i) the basic specification with local independence for all pairs of indicators, (ii) allowing a correlation between the answers on agents 1 and 2 (in brief a 1-2 correlation), (iii) allowing a 4-5 correlation, and (iv) allowing both 1-2 and 4-5 correlations. Invoking parsimony, we focused on 3-, 4-, 5-, and 6-class specifications for each of these four types of model structures. When beginning to explore model selection, this gives 16 models to estimate for each side, customers and suppliers.

With this starting point, we conducted an empirical exploration of whether there was a need to further relax the local independence assumption. To do this, we estimated the 16 models and examined the size of bivariate residual correlations, a measure of the marginal increase in the log-likelihood function that could be obtained by any specific relaxation of the local independence assumption (Vermunt and Magidson, 2016: 83-5). We then observed which particular combinations of enforcement agents had bivariate residual correlations that were prominent in this set of models. Table C.1 reports the patterns that we found.

[2] $32{,}766 = \sum_{k=1}^{14} \binom{15}{k}$. LCA needs at least one local independence assumption for identification, hence the 14 in the summation.

Based on the correlation patterns reported in Table C.1, a model with the correlation structure 1-2, 2-3, 3-4, 3-5, 4-5 was added to the original four model structures for transactions with suppliers. And for transactions with customers, a model with the correlation structure 1-2, 4-5, 4-6, 5-6 was added to the same four original structures. In sum, for each side of business relations (with suppliers and with customers), we chose to consider a total of 20 models, that is five correlation structures each with 3-, 4-, 5-, and 6-class specifications.[3]

C.2: Criteria for Model Selection

Model-selection criteria employ a number of standard statistical measures. All measures begin with the log likelihood (LL), which reflects goodness-of-fit without any adjustment for the number of estimated parameters. The measures, other than LL itself, then add extra terms to the LL, where those terms reward parsimony and penalize classification uncertainty. The likelihood-ratio χ^2 goodness-of-fit statistic (referred to as L^2 in Vermunt and Magidson (2016)) is used to test the null hypothesis that the estimated model fits the data. In the tables that follow, we present only the p-values for L^2, since its distribution varies across models, precluding comparisons of absolute values.[4] The Bayesian information criterion (BIC), the consistent Akaike information criterion (CAIC), and the approximate weight of evidence criterion (AWE) are varieties of information criteria, all reflecting the log likelihood, and thus goodness-of-fit, plus a penalty term that is a function of the number of estimated parameters and the number of observations.[5] As a consequence of the specification of the penalty terms, AWE favors more parsimonious models than does CAIC, followed, in terms of favoring parsimony, by BIC and then log likelihood. Lower values of the information criteria indicate preferred models.

Entropy R^2 is a measure of classification certainty. It has not been traditionally used as a model selection criterion but rather as an ex-post check on the model's results (Masyn 2013). An entropy R^2 that is close to zero indicates that the estimated latent classes are not well-distinguished. Two additional information criteria add a term based on the entropy R^2, thus penalizing classification uncertainty (in addition to rewarding goodness-of-fit and parsimony). These are the 'classification AWE' and the 'integrated classification likelihood' (ICL-BIC)'. Again, lower values indicate preferred models.[6]

Statistics on homogeneity and separability provide a final check on acceptability of a model. In terms of the notation of subsection III.1, homogeneity is characterized by estimated $\theta_{kr|c}$ that

[3] As noted immediately above, the specifics of the 20 models differs between supplier- and customer-relations.
[4] For background and formulae see Collins and Lanza (2010: 83) or Vermunt and Magidson (2016: 68).
[5] We use the BIC and CAIC based on the log likelihood, not the alternatives that are based on L^2. The formulae are standard (Vermunt and Magidson 2016: 70). See Banfield and Raftery (1993) for the statistic we label AWE in Section IV, which is the standard one employing this label (Masyn 2013: 568). This is not directly reported by Latent GOLD, but is easily derived from the LL, the number of estimated parameters, and the number of observations.
[6] See Biernacki et al. 2000 for a discussion of the ICL-BIC. The version of the approximate weight of evidence criterion reported by Latent GOLD (Vermunt and Magidson 2016: 72) is different from the more standard one in the literature (Banfield and Raftery 1993; Masyn 2013: 568). Thus, when we report Latent GOLD's statistic we refer to it (idiosyncratically) as the "classification AWE" to distinguish the two different concepts. The "classification AWE" modifies the standard AWE taking into account entropy in a manner exactly analogous to the Biernacki et al. (2000) modification of the BIC to obtain the ICL-BIC.

are not too close to 1/R. For binary ($R = 2$) response variables, one standard implementation of this criterion is that the $\theta_{kr|c}$ should not be in the interval [0.3, 0.7] (Masyn 2013). When we evaluate homogeneity, we aggregate responses into binary categories and apply this criterion.

The statistical measures related to separability are less ad hoc. Roughly speaking, in terms of the notation of subsection III.1, the measures assess whether the estimated $\hat{\pi}_{ci}$ are close to 0 or 1, that is classification certainty. These measures use modal class assignments—setting respondent i's class assignment to the j that maximizes $\hat{\pi}_{ji}$. Average posterior class probability for class c ($AvePP_c$) is the mean value of $\hat{\pi}_{ci}$ for all i classified in c using modal class assignment. Satisfactory values are close to 1. Odds of correct classification (OCC_c) is a ratio of two odds ratios. The numerator reflects $AvePP_c$ and the denominator uses $\hat{\pi}_c$, the estimated class membership probability for c derived at step-1 of the 3-step procedure. OCC_c equals 1 if class membership assignment is no better than random. A rule-of-thumb is that OCC_c should be at least 5.0, for all c. The modal class assignment proportion ($mcaP_c$) is the proportion of respondents in class c when respondent i's class assignment is set using modal class assignment. If respondents are assigned to classes with certainty, then $mcaP_c = \hat{\pi}_c$. Since step-1 of the LCA estimation gives standard errors for $\hat{\pi}_c$, a natural diagnostic is to examine whether $mcaP_c$ lies in a small confidence interval of $\hat{\pi}_c$.

C.3 Choosing the Preferred Model

In selecting one model from the 20 estimated, we use the statistical measures of model-fit and parsimony. At this stage, the implementation of the lattermost criterion meant a preference, but not a constraint, for describing the two sides of business relations with the same number of latent classes. Tables C.2a and C.2b present the measures of model fit for the two sets of 20 estimated models. In addition, the column listing the number of parameters is included to reflect parsimony.[7]

In both tables, the numbers in bold highlight the three best-performing models according to the statistic noted in the relevant column. A glance at Tables C.2a and C.2b already suggests that the models with complex correlation structures generally perform better for a variety of statistics. This is hardly surprising given the steps leading up to the consideration of this specific correlation structure (i.e., relaxation of local independence based on bivariate residual correlations).

For relations with suppliers, Table C.2a indicates that the model with 4 classes and correlation structure 1-2, 2-3, 3-4, 3-5, 4-5 performs well across most statistics. It is included in the best three models across all statistics except AWE; it is the best-performer on BIC, CAIC, and ICL-BIC; it is the second-best on Entropy-R^2 and third-best on LL and classification AWE. Note that among the Bayesian statistics both AWE statistics penalize an increase in the number of parameters most strongly and therefore, not surprisingly, the first- and second-best models on the classification AWE are far more parsimonious than those classified as best by other statistical criteria. However,

[7] Vermunt and Magidson (2016: 68) state that asymptotic p-values of L^2 cannot be trusted with sparse tables, which is why we report the p-values obtained from the $Bootstrap\ Chi^2$ option of Latent GOLD software. Our data does have sparse tables since we only observe 711 from the total of 15,625 possible distinct response combinations in questions about the relations with suppliers, and for customers we observe even less—631 (see Section II for response patterns).

given the strong performance on most statistics for the 4-class, 1-2, 2-3, 3-4, 3-5, 4-5 model for relations with suppliers, it is difficult to argue for a more parsimonious model. The 5-class model with the same correlation structure is the next best model.

Model selection for the relations with customers is less clear-cut. Since the 4-class model is preferred for suppliers, it is worth focusing first on 4-class models for the customer-side as well. Among these, the best performers are the one with no correlations and the one with the most complex correlation structure. While the model with no correlations performs better on some statistics (Entropy-R^2 and classification AWE), it underperforms the correlation structure 1-2, 4-5, 4-6, 5-6 on all other Bayesian statistics. Importantly, both BIC statistics are lower for the more complex model. Consequently, among the 4-class models, the correlation structure 1-2, 4-5, 4-6, 5-6 is preferred. Comparing the performance of this model with other models more broadly, the 5- and 6-class models with the same correlation structure are the closest in performance. However, the 4-class model is the best-performer on CAIC and ICL-BIC, and is only slightly inferior on the BIC and other measures. Combining this statistical evidence and an a priori preference to select models with the same number of latent classes across the two types of relations, we select the 4-class, 1-2, 4-5, 4-6, 5-6 correlation model to describe relations with customers. Here too, the 5-class model with the same correlation structure is the next best alternative.

Note that in all these steps leading up to selecting one model for each type of the two types of business relation, we did not examine the behavioral patterns reported by each of the 40 estimated models. This was entirely intentional as we followed the standard model-selection steps separating the process of selection from the analysis and interpretation of its findings.

C.4: Robustness: Comparison of the Chosen Models with the Next-Best Alternatives

We now provide further checks on the validity of our choices of LCA models. These checks use terminology and graphical formats that are laid out in Section IV of the paper, and we therefore recommend reading this part of the Supplementary Appendixes after completing Section IV.

We examine whether the behavioral patterns suggested by our chosen models differ from the behavioral patterns suggested by the next best alternatives—the 5-class models with the same correlation structure across questions as our chosen 4-class models (see Tables C.2a and C.2b). Tables C.3a and C.3b illustrate the governance structures of the 5-class models in the same format as Tables 2a and 2b from Section IV. Even a quick glance at these tables and figures is enough to recognize the same governance structures we already saw in Section IV, with no new behavioral pattern meriting a distinct name.

Table C.4a presents the firm-by-firm correspondence between governance structures assigned (modally) by the 4-class models with those of 5-class models, for supplier-relations. Four of the five classes in the 5-class model have a near perfect mapping with the original four classes. The additional 5th class can be safely described as using pure bilateralism, albeit with a tinge of legal support (see Table C.3a). It comprises mostly the firms that were assigned to the group using pure bilateralism in the 4-class classification. A close mapping between the class assignments is also

reflected in the estimates of class membership probabilities, with the prevalence of governance structures similar whether we apply the 4- or 5-class LCA.

While the governance structures for customer-relations suggested by the 5-class model (Table C.3b) do not contain a structure that is qualitatively different from the structures of the 4-class model (see Table 2b), the firm-by-firm correspondence exhibited in Table C.4b is less straightforward than it was for supplier-relations. Three governance structures in the 5-class group are clearly mapped into single classes in the 4-class group. The rest of the mapping is straightforwardly derived from the figures that illustrate the underlying behavior of classes 4 and 5. Namely, class 4 comprises firms that were assigned to pure bilateralism, or bilateralism with private support, or weak comprehensive. However, examining Table C.3b, class 4 is substantively indistinguishable from bilateralism with private support. Similarly, class 5 comprises firms that were assigned across all possible classes, but in terms of the behavioral pattern given in Table C.2b it is a close version of weak comprehensive governance.

To summarize, the 4-class and the 5-class models produce very similar overall estimates of governance structures. For supplier relations, nearly all firms are assigned to the same governance structures across the two models. For customer relations, the firm-by-firm assignments are clear-cut only for some governance structures. In cases with a more noisy mapping of the firm-by-firm assignments, the governance structures of the 5-class models have a structure that closely corresponds to ones already suggested by the 4-class model. Such a close correspondence between the governance structures across our chosen and the next-best models indicates that our findings are robust to small changes in model selection.

C.5: Class Homogeneity and Separability

As a final check on our chosen models, we examined measures of class homogeneity and separability, as laid out in Subsection C.2 above. As noted by Masyn (2013), a class has a high degree of homogeneity if there are both high and low probabilities predicted response probabilities that class (that is, high and low $\hat{\theta}_{kr|c}$ within each c).[8] A standard rule-of-thumb is to consider a category homogeneous if these probabilities are either below 0.3 or above 0.7, but this rule-of-thumb is applicable only to binary responses. Therefore, for this exercise alone, we converted the probability data given in Tables 2a and 2b into two binary categories—'Not at all', 'Slightly' and 'Moderately' versus 'Very much' and 'Extremely'. Table C.5 reports counts of the estimated probabilities of responses in our model that qualify as homogeneous by this criterion. (Note that we now have 12 categories = 6 questions × binary responses). All four classes in both types of relations appear highly homogeneous.

Because all classes could be highly homogenous but very similar, it is also important to check whether one can reliably distinguish between the classes. This is the notion of separability, several

[8] As already noted, the $\hat{\theta}_{kr|c}$ as referenced here should be interpreted as the marginal probability that a firm in latent class c chooses answer r on question k.

measures of which are introduced in appendix subsection C.2. Tables C.6a and C.6b report the estimates of these measures for our classes.

$AvePP_c$ (Average Posterior Class Probability) measures average class membership probability across all respondents classified into c by modal class assignment (i.e., using the maximum posterior class probability). If the class memberships are assigned with certainty, then this measure equals 1. As Tables C.6a and C.6b show, $AvePP_c$ is very close to 1, comfortably exceeding the minimum rule-of-thumb rule.

OCC_c (Odds of Correct Classification Ratio) is a ratio of odds ratio, with the denominator reflecting the $\hat{\pi}_c$ and the numerator reflecting $AvePP_c$. It equals 1 if average posterior probabilities are no better than a random application of the estimated class membership probabilities (that is, if Bayes theorem using firm-specific responses for class assignment does no better than class assignment ignoring the firm-specific data). Again, the tables show that our model exhibits a high degree of class separation, well above the rule-of-thumb minimum.

$mcaP_c$ (Modal Class Assignment Proportion) is the proportion of respondents in each class when firms are assigned to classes modally. If respondents were assigned with certainty, then $mcaP_c$ would exactly equal the directly estimated class membership probabilities ($\hat{\pi}_c$). To assess any discrepancy, one rule of thumb is whether $mcaP_c$ lies within a 95% confidence interval (CI) of the corresponding class membership probability estimates. Tables C.6a and C.6b demonstrate clearly separate classes as our $mcaP_c$'s are close to the estimated class membership probabilities falling within the 95% CIs. Indeed, all $mcaP_c$ lie in a 33% CIs of the corresponding $\hat{\pi}_c$.

All classes in both upstream and downstream relations are homogeneous and well separated.

References to Appendix C

Banfield, Jeffrey D. and Adrian E. Raftery. 1993. Model-Based Gaussian and Non-Gaussian Clustering. *Biometrics*. Vol. 49, No. 3 (Sep.,), pp. 803-821

Biernacki, C.; G. Celeux; G. Govaert. 2000. Assessing a mixture model for clustering with the integrated completed likelihood. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. Volume: 22, Issue: 7. Pages: 719 - 725

Collins, L. M., & Lanza, S. T. 2010. *Latent class and latent transition analysis: With applications in the social, behavioral, and health sciences*. New York: Wiley.

Masyn, Katherine. 2013. Latent class analysis and finite mixture modeling. In T. D. Little (Ed.) *The Oxford handbook of quantitative methods in psychology. Volume 2*. pp. 551-611. New York, NY: Oxford University Press.

Vermunt, Jeroen K. and Jay Magidson. *Technical Guide for Latent GOLD 5.1: Basic, Advanced, and Syntax*. January 9, 2016. Statistical Innovations Inc. http://www.statisticalinnovations.com

Table C.1: The most prominent correlation structures found in initial model estimates

order in which question appears	1: Trust	2: Mutual interests	3: Paid third parties	4: Gov't officials	5: Other third parties	6: Legal system
1: Trust		S, C				
2: Mutual interests			S			
3: Paid third-parties				S	S	
4: Gov't officials					S, C	C
5: Other third-parties						C
6: Legal system						

Notes:
S indicates frequent occurrence of large bivariate residual correlations in the models estimated for relations with suppliers. C indicates the same phenomenon for customer-relations models.

Table C.2a: Statistics on goodness-of-fit for estimated models for relations with suppliers

Model	Npar	LL	p-value of L^2	BIC	CAIC	AWE	Entropy R^2	Classification AWE	ICL-BIC
3 classes: no correlations	74	-18580.0	1.00	37760.7	37834.7	38583.3	0.791	39998.9	39176.2
4 classes: no correlations	99	-18174.4	1.00	37152.4	37251.4	38252.9	0.811	39952.7	38852.2
5 classes: no correlations	124	-17964.1	1.00	36934.6	37058.6	38313.1	**0.814**	40151.0	38772.5
6 classes: no correlations	149	-17776.5	1.00	36762.4	36911.4	38418.8	0.812	40445.5	38789.2
3 classes: 1-2 corr	90	-17834.5	1.00	36399.5	36489.5	**37400.0**	0.728	**38653.4**	37652.9
4 classes: 1-2 corr	115	-17667.0	1.00	36267.3	36382.3	**37545.7**	0.720	39825.6	38547.2
5 classes: 1-2 corr	140	-17541.5	1.00	36219.3	36359.3	37775.6	0.735	39967.4	38411.1
6 classes: 1-2 corr	165	-17421.5	1.00	36182.2	36347.2	38016.4	0.724	40567.7	38733.4
3 classes: 4-5 corr	90	-18400.7	1.00	37531.8	37621.8	38532.3	0.776	40091.1	39090.5
4 classes: 4-5 corr	115	-18013.3	1.00	36960.0	37075.0	38238.4	0.798	40099.6	38821.2
5 classes: 4-5 corr	140	-17799.3	1.00	36735.0	36875.0	38291.4	0.812	40225.6	38669.3
6 classes: 4-5 corr	165	-17595.8	1.00	36530.9	36695.9	38365.1	0.814	40372.2	38537.9
3 classes: 1-2, 4-5 corr	106	-17699.4	1.00	36259.2	36365.2	**37437.6**	0.718	**38664.5**	37486.1
4 classes: 1-2, 4-5 corr	131	-17518.6	1.00	36100.5	36231.5	37556.8	0.709	39809.4	38353.1
5 classes: 1-2, 4-5 corr	156	-17410.9	1.00	36088.0	36244.0	37822.2	0.721	40120.5	38386.3
6 classes: 1-2, 4-5 corr	181	-17321.2	1.00	36111.6	36292.6	38123.7	0.740	40465.6	38453.5
3 classes: 1-2, 2-3, 3-4, 3-5, 4-5 corr	154	-17365.1	1.00	**35980.2**	**36134.2**	37692.1	0.713	38919.6	**37207.6**
4 classes: 1-2, 2-3, 3-4, 3-5, 4-5 corr	179	**-17222.0**	1.00	**35897.0**	**36076.0**	37886.9	**0.846**	**38850.9**	**36861.0**
5 classes: 1-2, 2-3, 3-4, 3-5, 4-5 corr	204	**-17140.4**	1.00	**35936.6**	**36140.6**	38204.4	**0.847**	39389.4	**37121.6**
6 classes: 1-2, 2-3, 3-4, 3-5, 4-5 corr	229	**-17063.8**	1.00	35986.3	36215.3	38532.0	0.776	40724.8	38179.1

Note: p-values of L^2 were obtained using the bootstrap chi^2 procedures of Latent GOLD (Vermunt and Magidson 2016, p. 52).

Table C.2b: Statistics on goodness-of-fit for estimated models for relations with customers

Model	Npar	LL	p-value of L^2	BIC	CAIC	AWE	Entropy R^2	Classification AWE	ICL-BIC
3 classes: no correlations	74	-17791.0	1.00	36182.5	36256.5	37004.9	0.795	38274.2	37451.8
4 classes: no correlations	99	-17232.9	1.00	35269.0	35368.0	36369.2	**0.881**	**37397.2**	**36297.0**
5 classes: no correlations	124	-17023.1	1.00	35052.3	35176.3	36430.3	**0.847**	37934.9	36556.8
6 classes: no correlations	149	-16863.9	1.00	34936.8	35085.8	36592.7	0.837	38431.0	36775.1
3 classes: 1-2 corr	90	-17537.6	1.00	35805.4	35895.4	**36805.6**	0.826	38019.6	37019.4
4 classes: 1-2 corr	115	-17115.7	1.00	35164.4	35279.4	**36442.4**	0.847	**37801.5**	36523.4
5 classes: 1-2 corr	140	-16923.7	1.00	34983.2	35123.2	36539.1	0.838	38151.5	36595.7
6 classes: 1-2 corr	165	-16765.3	1.00	34869.3	35034.3	36703.0	**0.850**	38375.4	36541.7
3 classes: 4-5 corr	90	-16969.3	0.946	34668.7	34758.7	35668.9	0.719	**36705.8**	**35705.6**
4 classes: 4-5 corr	115	-16801.2	1.00	34535.5	34650.5	35813.5	0.635	38458.7	37180.6
5 classes: 4-5 corr	140	-16678.4	1.00	34492.6	**34632.6**	36048.5	0.642	39227.5	37671.6
6 classes: 4-5 corr	165	-16568.8	1.00	34476.3	34641.3	36310.0	0.659	39426.6	37592.8
3 classes: 1-2, 4-5 corr	106	-16878.0	1.00	34616.1	34722.1	**35794.1**	0.596	38353.2	37175.2
4 classes: 1-2, 4-5 corr	131	-16723.8	1.00	34510.4	34641.4	35966.2	0.653	38335.1	36879.3
5 classes: 1-2, 4-5 corr	156	-16608.2	1.00	34482.0	34638.0	36215.7	0.631	38954.5	37220.8
6 classes: 1-2, 4-5 corr	181	**-16502.6**	1.00	34473.7	34654.7	36485.3	0.651	39712.8	37701.3
3 classes: 1-2, 4-5, 4-6, 5-6 corr	138	-16697.7	1.00	34515.1	34653.1	36048.8	0.632	38488.3	36954.6
4 classes: 1-2, 4-5, 4-6, 5-6 corr	163	-16561.5	1.00	**34445.6**	**34608.6**	36257.0	0.754	38054.9	**36243.4**
5 classes: 1-2, 4-5, 4-6, 5-6 corr	188	**-16450.4**	1.00	**34426.0**	**34614.0**	36515.4	0.673	39157.9	38891.2
6 classes: 1-2, 4-5, 4-6, 5-6 corr	213	**-16366.0**	1.00	**34460.1**	34673.1	36827.3	0.665	40385.1	38018.0

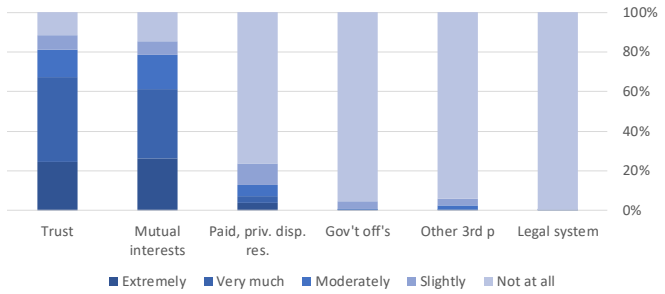
Note: p-values of L^2 were obtained using the bootstrap chi^2 procedures of Latent GOLD (Vermunt and Magidson 2016, p. 52)

Table C.3a: Predicted Response Probabilities, Suppliers Second-Best Model

Class 1

Method	Not at all	Slightly	Moderately	Very much	Extremely
Trust	0.12	0.07	0.14	0.43	0.24
	(0.02)	(0.01)	(0.02)	(0.02)	(0.03)
Mutual interests	0.14	0.07	0.17	0.35	0.26
	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)
Paid, priv. disp. res.	0.76	0.11	0.06	0.04	0.04
	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)
Gov't off's	0.95	0.04	0.00	0.00	0.00
	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)
Other 3rd p	0.94	0.04	0.02	0.00	0.01
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)
Legal system	1.00	0.00	0.00	0.00	0.00
	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)

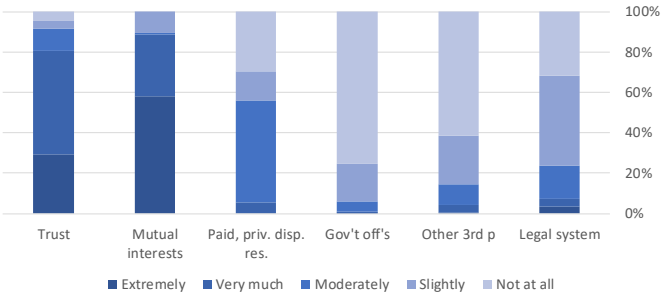
Standard errors in parenthesis.



Class 2

Method	Not at all	Slightly	Moderately	Very much	Extremely
Trust	0.04	0.04	0.11	0.51	0.29
	(0.01)	(0.02)	(0.04)	(0.06)	(0.05)
Mutual interests	0.00	0.10	0.01	0.31	0.58
	(0.00)	(0.03)	(0.01)	(0.07)	(0.08)
Paid, priv. disp. res.	0.30	0.14	0.50	0.06	0.00
	(0.06)	(0.04)	(0.07)	(0.02)	(0.00)
Gov't off's	0.75	0.19	0.05	0.01	0.00
	(0.04)	(0.04)	(0.02)	(0.00)	(0.00)
Other 3rd p	0.61	0.24	0.10	0.04	0.00
	(0.05)	(0.05)	(0.03)	(0.02)	(0.00)
Legal system	0.32	0.45	0.16	0.04	0.03
	(0.07)	(0.06)	(0.04)	(0.02)	(0.02)

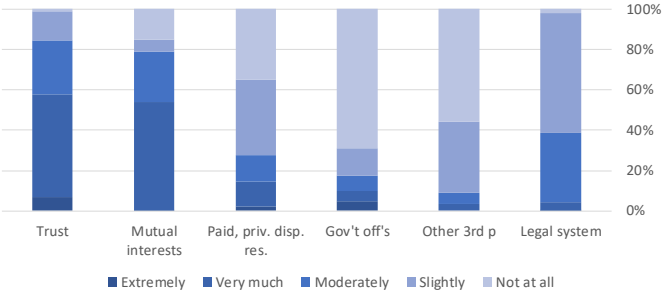
Standard errors in parenthesis.



Class 3

Method	Not at all	Slightly	Moderately	Very much	Extremely
Trust	0.01	0.14	0.27	0.51	0.07
	(0.01)	(0.04)	(0.05)	(0.06)	(0.03)
Mutual interests	0.15	0.06	0.25	0.54	0.00
	(0.04)	(0.03)	(0.04)	(0.06)	(0.00)
Paid, priv. disp. res.	0.35	0.37	0.13	0.12	0.02
	(0.05)	(0.06)	(0.06)	(0.04)	(0.02)
Gov't off's	0.69	0.13	0.08	0.05	0.05
	(0.05)	(0.03)	(0.03)	(0.02)	(0.03)
Other 3rd p	0.56	0.36	0.05	0.04	0.00
	(0.05)	(0.05)	(0.02)	(0.02)	(0.00)
Legal system	0.02	0.59	0.34	0.05	0.00
	(0.03)	(0.06)	(0.05)	(0.02)	(0.00)

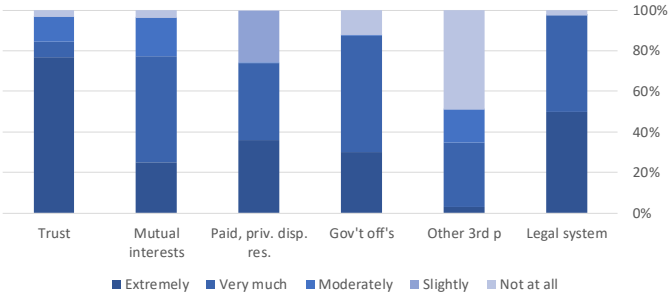
Standard errors in parenthesis.



Class 4

Method	Not at all	Slightly	Moderately	Very much	Extremely
Trust	0.03	0.00	0.12	0.08	0.77
	(0.03)	(0.00)	(0.06)	(0.05)	(0.10)
Mutual interests	0.04	0.00	0.19	0.52	0.25
	(0.04)	(0.00)	(0.10)	(0.13)	(0.12)
Paid, priv. disp. res.	0.01	0.25	0.00	0.38	0.36
	(0.01)	(0.11)	(0.01)	(0.15)	(0.14)
Gov't off's	0.12	0.00	0.00	0.57	0.30
	(0.10)	(0.00)	(0.00)	(0.13)	(0.13)
Other 3rd p	0.49	0.00	0.16	0.32	0.03
	(0.15)	(0.00)	(0.10)	(0.13)	(0.03)
Legal system	0.03	0.00	0.00	0.48	0.50
	(0.02)	(0.00)	(0.00)	(0.15)	(0.15)

Standard errors in parenthesis.



Class 5

Method	Not at all	Slightly	Moderately	Very much	Extremely
Trust	0.06	0.12	0.01	0.01	0.80
	(0.03)	(0.06)	(0.06)	(0.01)	(0.09)
Mutual interests	0.00	0.00	0.00	0.99	0.01
	(0.00)	(0.00)	(0.00)	(0.07)	(0.07)
Paid, priv. disp. res.	0.85	0.09	0.06	0.00	0.00
	(0.09)	(0.08)	(0.03)	(0.01)	(0.00)
Gov't off's	0.89	0.07	0.03	0.00	0.01
	(0.07)	(0.07)	(0.03)	(0.00)	(0.01)
Other 3rd p	0.86	0.13	0.00	0.00	0.00
	(0.06)	(0.06)	(0.00)	(0.00)	(0.00)
Legal system	0.72	0.17	0.04	0.07	0.00
	(0.08)	(0.07)	(0.03)	(0.04)	(0.00)

Standard errors in parenthesis.

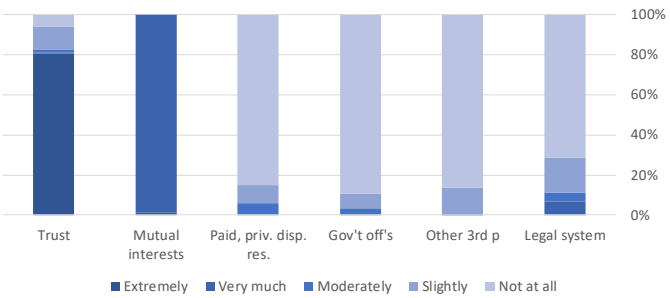
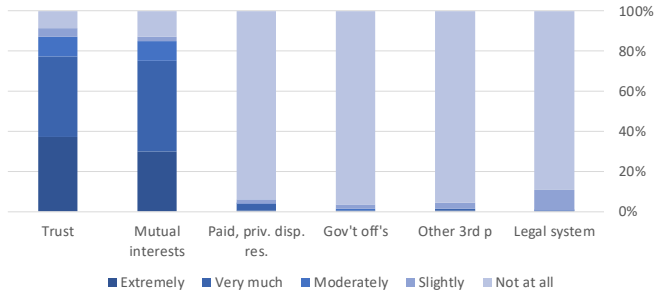


Table C.3b: Predicted Response Probabilities, Customers Second-Best Model

Class 1

Method	Not at all	Slightly	Moderately	Very much	Extremely
Trust	0.09	0.04	0.10	0.40	0.37
	(0.02)	(0.01)	(0.02)	(0.03)	(0.03)
Mutual interests	0.13	0.02	0.10	0.45	0.30
	(0.02)	(0.01)	(0.02)	(0.04)	(0.03)
Paid, priv. disp. res.	0.94	0.02	0.00	0.03	0.00
	(0.05)	(0.04)	(0.00)	(0.02)	(0.00)
Gov't off's	0.97	0.02	0.01	0.00	0.00
	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)
Other 3rd p	0.96	0.03	0.00	0.01	0.00
	(0.02)	(0.02)	(0.00)	(0.01)	(0.00)
Legal system	0.89	0.10	0.00	0.00	0.00
	(0.02)	(0.02)	(0.00)	(0.00)	(0.00)

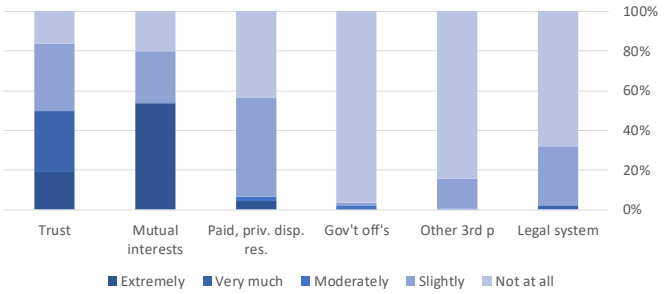
Standard errors in parenthesis.



Class 2

Method	Not at all	Slightly	Moderately	Very much	Extremely
Trust	0.16	0.34	0.00	0.31	0.19
	(0.06)	(0.11)	(0.02)	(0.09)	(0.13)
Mutual interests	0.20	0.26	0.00	0.00	0.54
	(0.07)	(0.11)	(0.00)	(0.01)	(0.11)
Paid, priv. disp. res.	0.44	0.50	0.02	0.00	0.04
	(0.09)	(0.09)	(0.04)	(0.00)	(0.03)
Gov't off's	0.96	0.01	0.02	0.00	0.00
	(0.02)	(0.01)	(0.02)	(0.00)	(0.00)
Other 3rd p	0.84	0.15	0.01	0.00	0.00
	(0.06)	(0.06)	(0.01)	(0.00)	(0.00)
Legal system	0.68	0.30	0.00	0.02	0.00
	(0.09)	(0.09)	(0.00)	(0.02)	(0.00)

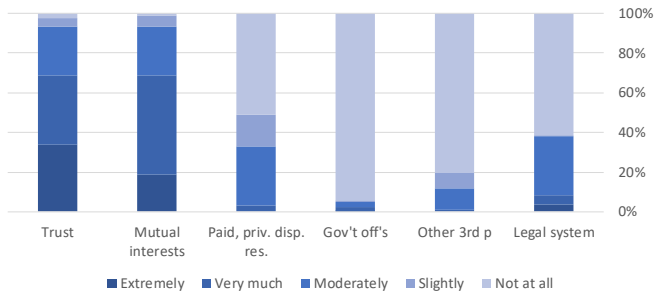
Standard errors in parenthesis.



Class 3

Method	Not at all	Slightly	Moderately	Very much	Extremely
Trust	0.02	0.05	0.25	0.35	0.34
	(0.01)	(0.02)	(0.04)	(0.05)	(0.05)
Mutual interests	0.01	0.06	0.24	0.50	0.19
	(0.01)	(0.02)	(0.04)	(0.05)	(0.04)
Paid, priv. disp. res.	0.51	0.16	0.30	0.03	0.00
	(0.06)	(0.04)	(0.05)	(0.02)	(0.00)
Gov't off's	0.94	0.01	0.02	0.02	0.00
	(0.02)	(0.01)	(0.01)	(0.01)	(0.00)
Other 3rd p	0.81	0.07	0.11	0.01	0.00
	(0.04)	(0.02)	(0.03)	(0.01)	(0.00)
Legal system	0.61	0.00	0.30	0.05	0.04
	(0.06)	(0.01)	(0.05)	(0.01)	(0.02)

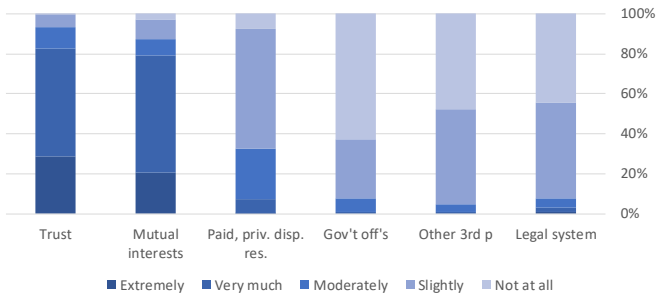
Standard errors in parenthesis.



Class 4

Method	Not at all	Slightly	Moderately	Very much	Extremely
Trust	0.00	0.06	0.10	0.54	0.29
	(0.01)	(0.04)	(0.03)	(0.06)	(0.06)
Mutual interests	0.03	0.10	0.08	0.59	0.21
	(0.02)	(0.04)	(0.02)	(0.06)	(0.05)
Paid, priv. disp. res.	0.07	0.60	0.25	0.07	0.00
	(0.06)	(0.07)	(0.06)	(0.04)	(0.00)
Gov't off's	0.63	0.29	0.07	0.01	0.00
	(0.07)	(0.06)	(0.03)	(0.01)	(0.00)
Other 3rd p	0.48	0.48	0.03	0.01	0.00
	(0.10)	(0.09)	(0.01)	(0.01)	(0.00)
Legal system	0.44	0.48	0.05	0.02	0.01
	(0.09)	(0.08)	(0.03)	(0.02)	(0.01)

Standard errors in parenthesis.



Class 5

Method	Not at all	Slightly	Moderately	Very much	Extremely
Trust	0.00	0.00	0.13	0.18	0.69
	(0.00)	(0.00)	(0.09)	(0.16)	(0.15)
Mutual interests	0.00	0.07	0.08	0.14	0.71
	(0.00)	(0.05)	(0.08)	(0.12)	(0.16)
Paid, priv. disp. res.	0.00	0.00	0.00	0.54	0.45
	(0.00)	(0.01)	(0.01)	(0.17)	(0.16)
Gov't off's	0.60	0.05	0.00	0.16	0.19
	(0.12)	(0.04)	(0.00)	(0.10)	(0.09)
Other 3rd p	0.51	0.00	0.21	0.25	0.02
	(0.14)	(0.00)	(0.11)	(0.11)	(0.02)
Legal system	0.55	0.03	0.10	0.22	0.12
	(0.12)	(0.03)	(0.07)	(0.09)	(0.06)

Standard errors in parenthesis.

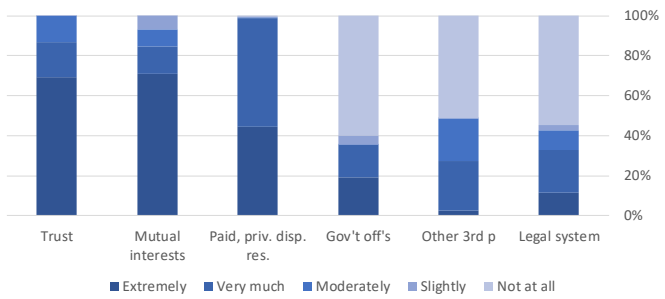


Table C.4a: Comparison of modal class assignments across the 4-class and the 5-class models for relations with suppliers

Read by rows: from 4-class (row) to 5-class (column)

	Class 1	Class 2	Class 3	Class 4	Class 5	Total
Pure bilateralism	**89.04**	0.84	0.29	0	9.83	100
Bilateralism with private support	9.23	**88.93**	0.03	0	1.81	100
Bilateralism with legal support	0	1.06	**88.53**	0.01	10.4	100
Strong comprehensive governance	0	0	0	**100**	0	100
Total	60.69	14.88	14.32	1.62	8.49	100

Read by columns: from 5-class (column) to 4-class (row)

	Class 1	Class 2	Class 3	Class 4	Class 5	Total
Pure bilateralism	**97.58**	3.77	1.34	0.07	77.06	66.52
Bilateralism with private support	2.42	**95.09**	0.03	0	3.4	15.91
Bilateralism with legal support	0	1.13	**98.63**	0.09	19.54	15.95
Strong comprehensive governance	0	0	0	**99.84**	0	1.62
Total	100	100	100	100	100	100

Estimated class membership probabilities and standard errors

Pure bilateralism	0.657	0.0243	Class 1	0.5932	0.0273
Bilateralism with private support	0.166	0.0218	Class 2	0.1655	0.0244
Bilateralism with legal support	0.160	0.0171	Class 3	0.1496	0.0171
Strong comprehensive governance	0.017	0.0043	Class 4	0.0165	0.0043
			Class 5	0.0751	0.0204

Table C.4b: Comparison of modal class assignments across the 4-class and the 5-class models for relations with customers

Read by rows: from 4-class (row) to 5-class (column)

	Class 1	Class 2	Class 3	Class 4	Class 5	Total
Pure bilateralism	**89.42**	0.65	0.1	8.87	0.97	100
Bilateralism with private support	8.7	**51.75**	0.48	33.86	5.21	100
Bilateralism with weak support	1.54	6.28	**84.1**	0.15	7.94	100
Weak comprehensive governance	15.73	10.13	1.1	56.73	16.32	100
Total	54.58	14.17	11.65	15.94	3.66	100

Read by columns: from 5-class (column) to 4-class (row)

	Class 1	Class 2	Class 3	Class 4	Class 5	Total
Pure bilateralism	**94.43**	2.65	0.47	32.06	15.28	57.64
Bilateralism with private support	3.84	**87.98**	0.99	51.16	34.33	24.09
Bilateralism with weak support	0.38	6.02	**98.1**	0.12	29.52	13.59
Weak comprehensive governance	1.35	3.35	0.44	16.65	20.88	4.68
Total	100	100	100	100	100	100

Estimated class membership probabilities and standard errors

Pure bilateralism	0.565	0.032	Class 1	0.4739	0.0422
Bilateralism with private support	0.145	0.025	Class 2	0.1152	0.0267
Bilateralism with weak support	0.242	0.000	Class 3	0.2205	0.0305
Weak comprehensive governance	0.050	0.008	Class 4	0.1579	0.0264
			Class 5	0.0325	0.0088

Table C.5: Degree of homogeneity of classes

	Count	Share
Relations with suppliers		
Pure bilateralism	8	67%
Bilateralism with private support	12	100%
Bilateralism with legal support	8	67%
Strong comprehensive governance	10	83%
Relations with customers		
Pure bilateralism	12	100%
Bilateralism with private support	12	100%
Bilateralism with weak support	8	67%
Weak comprehensive governance	10	83%

Table C.6a: Degree of separation of classes, relations with suppliers

	$AvePP_c$	OCC_c	$mcaP_c$	class membership probabilities ($\hat{\pi}_c$)	95% CI of the $\hat{\pi}_c$	
Pure bilateralism	0.963	13.730	0.665	0.657	0.609	0.705
Bilateralism with private support	0.883	37.990	0.159	0.166	0.123	0.209
Bilateralism with legal support	0.940	81.934	0.159	0.160	0.127	0.194
Strong comprehensive governance	0.977	10016.817	0.016	0.017	0.008	0.025
Rule-of-thumb minimum	0.7	5				

Note: See Appendix C.2 for definitions of the statistical measures.

Table C.6b: Degree of separation of classes, relations with customers

	$AvePP_c$	OCC_c	$mcaP_c$	class membership probabilities ($\hat{\pi}_c$)	95% CI of the $\hat{\pi}_c$	
Pure bilateralism	0.915	8.310	0.576	0.565	0.503	0.626
Bilateralism with private support	0.874	21.833	0.241	0.242	0.184	0.299
Bilateralism with weak support	0.810	25.188	0.136	0.145	0.096	0.193
Weak comprehensive governance	0.965	3412.999	0.047	0.050	0.034	0.065
Rule-of-thumb minimum	0.7	5				

Note: See Appendix C.2 for definitions of the statistical measures.

Appendix D

This appendix explores associations between the governance structure of firms and their characteristics or environments. It repeats and complements to the exercises laid out in subsection V. As noted there, this is an exploratory venture, not an attempt to isolate ceteris paribus, causal effects of single variables. Most importantly, this is an examination of the validity of the data: if there were no significant association between the use of governance structures and firm characteristics then there could be justified doubts about the meaningfulness of the estimates.

To reiterate the type of thought experiments explored here, consider firm size. We compare the pattern of governance structures used by small firms with that used by large firms. We show the resultant change in the choice of governance structures as a firm becomes large for any reason and then simultaneously goes through all other changes associated with the differences between small and large firms. We do this analysis using the 3-step method outlined in Appendix B.2.

With the richness of the WBES data and the complex origins of the governance structures used by firms, it is challenging to select a manageable set of covariates that are particularly germane. While some covariates are obviously crucial to examine, e.g., sector, others are less so, e.g., a firm's experience of corruption. Recognizing the exploratory nature of the exercise, we selected a set of variables that piqued our curiosity, without requiring a precise theory. Our interest is mainly in checking the validity of the estimated transactional governance structures by examining whether there are significant associations between governance choices and potential covariates. Table D.1 lists the covariates, together with summary statistics. For ease of exposition, the variables are organized in seven broad categories, also listed in this table. Section V reports on a subset of the covariates of Table D.1.

We study the associations between the use of the four governance structures and each of the covariates, one covariate at a time. We report Wald p-statistics in the second columns of Tables D.2a and D.2b, subset of which is included in Table 4. Each of these statistics are constructed to test the compound hypotheses that the variable listed in column 1 (the Z_i of Appendix B.2) has no explanatory power for the estimated posterior class-membership probabilities, $\hat{\pi}_{ci}$, $c = 1,\ldots 4$.

Interpreting these statistics involves a multiple comparisons problem, which entails deciding on the methods to use when judging statistical significance. Choice of methods depends upon the insights that the reader hopes to gain. One natural question to ask is whether the estimated class probabilities are no better than random in terms of the effects on them of all the variables listed in Table D.1. Roughly speaking, this question asks whether the estimated class probabilities are simply random numbers. This is a comprehensive hypothesis on the whole set of effect sizes, and not a hypothesis on the individual effect sizes. The family-wise error rate (FWER) tests this comprehensive hypothesis by appropriately adjusting the numerical values used to judge the significance levels of each of the test statistics on individual effect sizes. When a probability level of α is chosen, the adjustments are made so that α becomes the probability of erroneously concluding that at least one relationship is significant. Therefore, α is the type I error for the

comprehensive hypothesis. We use the Holm-Bonferroni method (Holm 1979), reporting criteria for statistical significance in the rightmost three columns of Tables D.2a and D.2b. A significant value for even one p-statistic in these columns is evidence of better-than-random for the LCA procedure.

The overall hypothesis that our estimated class probabilities are no better than random is rejected decisively. This is the case for both relations with suppliers and with customers. This rejection provides overall support for the validity of the method developed in this paper, including the formulation of the survey questions and the interpretation of the data that follows from LCA estimates.

If the focus turns to tests on individual variable-effects, rather than the overall implications to be drawn from the complete set of results on all variables, one can use procedures that have more power than the FWER. A natural approach is to control the expected proportion of errors made when examining all of the effect sizes individually by testing whether each is significant. In the case of a multiplicity of comparisons, standard procedures do not accomplish this because applying these procedures to the smallest values of a set of p-statistics violates the conditionality assumptions of standard tests. Instead, we use the false discovery rate (FDR) to implement this approach. In that approach the numerical values used to judge the significance levels of each of the individual test statistics on effect sizes are again adjusted, but in a different way from the FWER adjustments. If the FDR is set at 5%, for example, the adjustments are made so that 95% of the statistically significant individual-variable effect sizes are correctly judged to be significant. We use the Benjamini–Hochberg (1995) version of FDR in columns 3 through 5 of Tables D.2a and D.2b. These columns are most relevant to readers who are interested in the results for many of the variables but have no prior theoretical hypotheses on specific variables.

The differences in insights and corresponding methods of judging statistical significance underlies our choice to report all statistics in Tables D.2a and D.2b (i.e., p-values, as well as FDR and FWER), while reporting only p-values for a specific subset of covariates in Table 4. The remainder of Appendix D details broad categories of covariates that we explored, including the ones that are briefly reported in Section V.

D.1 Attitudes Towards Courts

We examine two standard questions that appear in every WBES and have often been used as measures of court performance. The first ("fair-court") asks whether the respondent agrees or disagrees with the statement "the court system is fair, impartial and uncorrupted". The second asks whether the courts are an obstacle to the current operations of the firm ("court-as-obstacle").

Figures D.1a and D.1b show the patterns in the data. On the suppliers' side, consistent with our intuition, firms considering the court fair are more likely to employ governance structures with a stronger legal element. This relation is weaker on the customers' side, where there is little association between attitudes about the courts and the use of the law. The conclusion is that the

fair-court question is not a reliable indicator of a firm's commitment to a legally-oriented governance strategy.

For the court-as-obstacle question, the firms that do not consider the court as an obstacle are the least likely to rely on the legal system. In contrast, as the assessment of the court as an obstacle increases, there is more reliance on governance structures that involve legal systems. If one viewed this question as a measure of court quality, one would expect exactly the opposite association.[9] The most likely explanation of this apparent paradox is reverse causality: if firms do not choose to use the legal system, then the courts are not an obstacle. The firms that need the legal system are more likely to be hindered by its flaws. That is, the interpretation of answers to this court-as-obstacle question in the literature seems to be diametrically opposite to what it actually reflects. Our conclusion here is consistent with observations on data on Russia's early transition made by Hendley et al. (2000) and explored thoroughly in papers by Hendley (2016; 2017): because going to court is inherently an unpleasant experience, attitudes to the courts are not good predictors of the use of the law.

D.2 Interactions with Business Associations

Figures D.2a and D.2b illustrate the correlation between business membership and governance structures. In all cases, firms with stronger ties to business associations are more likely to rely on bilateralism with private support. But, as shown in Tables D.2a and D.2b, this observation is backed by only weak statistical support, and only on the customer side. Perhaps what the data is showing here is that business associations are important in somewhat niche activities within particular sectors (Bernstein 2001), but not important generally in those sectors. Our test is too low-power to reflect such niche relationships.

D.3 Sectors

As the statistical tests in Tables D.2a and D.2b show, governance classes do vary significantly between sectors.[10] Figures D.3a and D.3b illustrate this variation, which is substantial. For example, the use of bilateralism varies from 74% when food processors interact with their suppliers to 24% in the sales of construction companies. One conjecture on this difference immediately follows from Williamson's emphasis on frequency: the more frequent are exchanges, the easier it is to construct purely bilateral governance. In their sales, construction companies use governance structures that employ private, paid, dispute resolution and the legal system. This is consistent with

[9] Gutmann and Voigt (2017) use the courts-as-obstacles question as a dependent variable that is viewed as a proxy for the quality of the courts. See also the following from World Bank (2014) on survey results for the Kyrgyz Republic: "Courts are perceived as one of the least problematic areas for doing business…In 2013, only 13 percent of firms saw courts as a problem, and only 4 percentage points of respondents saw it as major or very severe problem…. This is a significant improvement compared to 2008 when 60 percent of firms saw courts as a problem and 29 percent saw them as a major/severe problem." Note that over the same period, there were declines in the percentages of firms believing that the court system is fair, impartial, and uncorrupted, quick, and able to enforce its decisions.

[10] The WBES contains four-digit ISIC Rev.3.1 information on the main product and activity of each establishment. We used two-digit codes and grouped sectors as follows: Food (codes 15,16), Textiles and Garments (17,18,19), Wood and Wood Products (20-22, 36), Chemicals, Plastics, Non-metallic Mineral Products (24-26), Metals (27, 28), Machinery and Equipment (29, 31-35), Construction (45), Retail, Wholesale, Tourism (50-52, 55, 63), Transport (60-62), Telecom and IT (64, 72).

Williamson's emphasis on more complex governance when exchange is infrequent and involves idiosyncratic interactions (Williamson 1985).

D.4 Management Practices

The effects of firms' management practices are an important avenue of investigation currently in economics (Bloom et al. 2012, Bloom et al. 2013). To quantify the role of management practices, Bloom and Van Reenen (2007, 2010), in coordination with the US Census Bureau, developed a set of survey questions, which the World Bank's Enterprise Analysis Unit modified and implemented as part of the standard WBES.[11] We examine the association between the responses to these questions and the governance structures chosen by firms. Figures 3a and 3b in the main text illustrate this association. As firms' management practices improve, the prevalence of pure bilateralism falls (from 76% to 54% on the suppliers' side and from 72% to 41% on the customers' side), indicating that the improvement in internal management practices is accompanied by the use of more complex methods of governance of external relations.

D.5 Miscellaneous Firm Characteristics

We follow the WBES indicators in calling firms "foreign owned" if they are at least 10% owned by foreign private entities. Similarly, we call firms "exporters" if at least 10% of their total sales are in foreign markets. As Tables D.2a and D.2b indicate, the associations between these measures and the choice of governance structures are weak. But the direction of association is intuitive, as illustrated in Figures D.4 and D.5. Foreign-owned firms and exporters use pure bilateralism less than firms that are domestically owned and oriented.

Lastly, we examine firm size, which only has a weak association with the choice of governance structures. As Figure D.6 shows, this association reflects the distinctive behavior of very large firms, which have a greater tendency to use governance structures that are comprehensive and make use of the legal system.

We have not commented so far on the variables that fail to reach statistical significance in Tables D.2a and D.2b. There are also insights there. For example, we find no association between measures of corruption and governance structure, indicating that the effect of the quality of the legal system on these two might be orthogonal. Similarly, there are few connections between the governance of transactions and the type of ownership of the firm (apart from that of foreigners). Finally, there seems to be no difference between the transactional governance structures of the firms who trade locally and those of the firms who trade nationally, a result not to be expected from the existing literature (McMillan and Woodruff 1999).

[11] The original survey instrument implemented by the Census is known as the Management and Organizational Practices Survey (MOPS). Subsequent surveys building on these instruments include the World Management Survey (WMS), both with origins in the Management, Organisation and Innovation (MOI), involving the European Bank for Reconstruction and Development (EBRD).

References to Appendix D

Benjamini, Y. and Hochberg, Y., 1995. *Controlling the false discovery rate: a practical and powerful approach to multiple testing.* Journal of the Royal Statistical Society. Series B (Methodological). *pp.289-300.*

Bernstein, L., 2001. Private commercial law in the cotton industry: Creating cooperation through rules, norms, and institutions. *Michigan Law Review*, 99(7), pp.1724-1790.

Bloom, N. and Van Reenen, J., 2007. Measuring and explaining management practices across firms and countries. *The Quarterly Journal of Economics*. 122(4), pp.1351-1408.

Bloom, N. and Van Reenen, J., 2010. Why do management practices differ across firms and countries?. *Journal of Economic Perspectives*. 24(1), pp.203-24.

Bloom, N., Eifert, B., Mahajan, A., McKenzie, D. and Roberts, J., 2013. Does management matter? Evidence from India. *The Quarterly Journal of Economics*. 128(1), pp.1-51.

Bloom, N., Sadun, R. and Van Reenen, J., 2012. Americans do IT better: US multinationals and the productivity miracle. *American Economic Review*, 102(1), pp.167-201.

Bloom, N., Schweiger, H. and Van Reenen, J., 2012. The land that lean manufacturing forgot? Management practices in transition countries. *Economics of Transition,* 20(4), pp.593-635.

Doing Business, The World Bank (http://www.doingbusiness.org).

Gutmann, J. & S. Voigt 2017. Is judicial independence good for business? Paper presented at the Annual Meeting of the Public Choice Society, New Orleans.

Hendley, K., 2016. Justice in Moscow?. *Post-Soviet Affairs*, 32(6), pp.491-511.

Hendley, K., 2017. Contempt for Court in Russia: The Impact of Litigation Experience. *Review of Central and East European Law*. 42(2-3), pp.134-168.

Hendley, K., Murrell, P. and Ryterman, R., 2000. Law, relationships and private enforcement: Transactional strategies of Russian enterprises. *Europe-Asia Studies*. 52(4), pp.627-656.

Holm, S., 1979. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, pp.65-70.

McMillan, J. and Woodruff, C., 1999. Interfirm relationships and informal credit in Vietnam. *The Quarterly Journal of Economics*, 114(4), pp.1285-1320.

Williamson, Oliver E. 1985. *The Economic Institutions of Capitalism*. New York: Free Press.

World Bank 2014. BEEPS At-A-Glance 2013. Kyrgyz Republic. Washington D.C.

Table D.1: Summary statistics of covariates of the governance structure

		Supplier Relations			Customer Relations		
Topic and variable	Type	N	mean	s.d.	N	mean	s.d.
Cross- and within-country variation							
Country	6 categ.	3350			3339		
Regions within country	2-5 categ.	3350			3339		
Attitudes towards courts							
The extent of agreement with the statement "the court system is fair, impartial and uncorrupted"	4 categ. [1: strongly agree - 4: strongly disagree]	3278	1.62	0.77	3266	1.62	0.77
The degree to which courts are an obstacle to the firm's current operations	5 categ. [0: no obstacle - 4: very severe obstacle]	3259	1.63	1.37	3249	1.62	1.38
Relations with business associations							
Currently belong to an industry organization or business association?	Yes/No	3298	0.38	0.49	3289	0.38	0.49
Does the senior management regularly interact with a main Business Association to which the firm belongs?	Yes/No	3281	0.31	0.46	3271	0.31	0.46
Sector of operation							
Manufacturing, retail, or other services	3 categ.	3350			3339		
Disaggregated sector	10 categ.	3347			3336		
Management practices							
Index for management practices (larger equals better practices)	Cont.	3350	0.53	0.17	3339	0.52	0.17
Top manager's years of experience working in this sector	Cont.	3318	24.22	12.52	3307	24.25	12.54

Topic and variable	Type	Supplier Relations			Customer Relations		
		N	mean	s.d.	N	mean	s.d.
Firm characteristics							
Size	4 categ. [1: small – 4: very large]	3350	1.49	0.72	3296	1.49	0.72
Age	Cont.	3329	23.83	17.57	3318	23.81	17.60
Proportion of domestic private ownership	Cont.	3347	0.94	0.22	3336	0.94	0.23
Dummy variable for at least 10% foreign ownership	Yes/No	3347	0.06	0.25	3336	0.06	0.25
Dummy variable for exporting directly at least 10% of sales	Yes/No	3347	0.06	0.24	3336	0.06	0.24
Proportion of female ownership	Cont.	3202	0.23	0.31	3194	0.23	0.31
Dummy variable for a female top manager	Yes/No	3339	0.14	0.35	3328	0.14	0.35
Main Market – local, national, international	3 categ. [1-3]	1602	1.59	0.56	1602	1.59	0.55
Part of a multi-establishment firm?	Yes/No	3350	0.12	0.32	3339	0.12	0.33
Legal form	4 categ.	3303			3292		
Dummy variable for the legal form "Sole Proprietorship"	Yes/No	3350	0.09	0.29	3339	0.09	0.29
Dummy variable for shareholding company	Yes/No	3346	0.50	0.50	3335	0.49	0.50
Proportion of transactions with suppliers that were fulfilled smoothly	Cont.	3232	0.85	0.25	3184	0.85	0.25
Proportion of transactions with customers that were fulfilled smoothly	Cont.	3242	0.88	0.21	3283	0.88	0.21

Topic and variable	Type	Supplier Relations			Customer Relations		
		N	mean	s.d.	N	mean	s.d.
Corruption and security							
Bribery depth (among 8 types of interactions with government officials, share where a gift or informal payment was requested)	Cont.	2537	0.08	0.23	2533	0.08	0.23
Bribery incidence (dummy variable if experienced at least one gift or informal payment request across 8 types of interactions with government officials)	Yes/No	2537	0.13	0.34	2533	0.13	0.33
Dummy variable for paying for security	Yes/No	3344	0.65	0.48	3333	0.65	0.48

Table D.2a: Tests of the Association Between Governance Structures and a Variety Of Variables, Relations With Suppliers

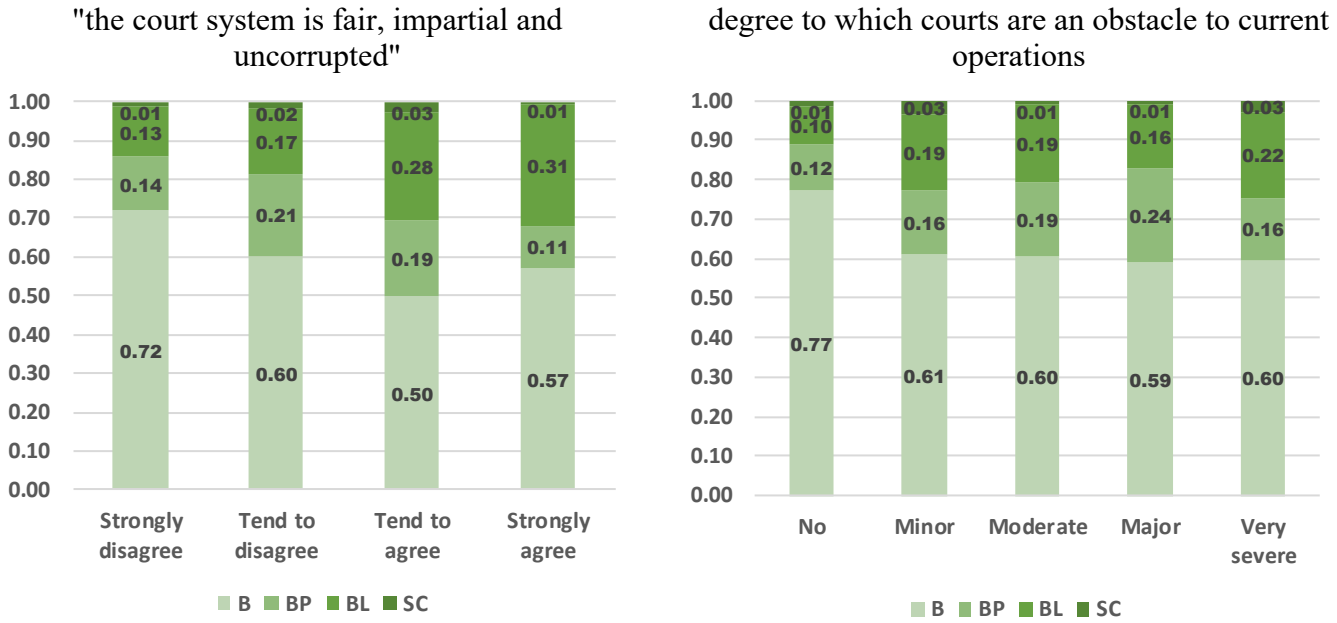
		FDR			FWER		
Description	p-value	1%	5%	10%	1%	5%	10%
Country	0.001***	0.0008	0.0040**	0.0080*	0.0004	0.0021**	0.0042*
Attitudes towards courts							
"the court system is fair…"	0.004***	0.0012	0.0060**	0.0120*	0.0004	0.0022	0.0044*
Courts as an obstacle	0.016**	0.0020	0.0100	0.0200*	0.0005	0.0024	0.0048
Relations with business associations							
Belong to a business association?	0.240	0.0056	0.0280	0.0560	0.0008	0.0042	0.0083
Regularly interact with a business association?	0.390	0.0064	0.0320	0.0640	0.0010	0.0050	0.0100
Sector of operation							
Manufacturing, retail, or other services	0.180	0.0052	0.0260	0.0520	0.0008	0.0039	0.0077
Disaggregated sector	0.000***	0.0004***	0.0020**	0.0040*	0.0004***	0.0020**	0.0040*
Management practices							
Management practices (higher means better)	0.078*	0.0036	0.0180	0.0360	0.0006	0.0029	0.0059
Top manager's experience working in sector	0.470	0.0068	0.0340	0.0680	0.0011	0.0056	0.0111
Firm characteristics							
Size	0.082*	0.0040	0.0200	0.0400	0.0006	0.0031	0.0063
Age	0.770	0.0092	0.0460	0.0920	0.0033	0.0167	0.0333
Proportion domestic private ownership	0.008***	0.0016	0.0080**	0.0160*	0.0005	0.0023	0.0046
At least 10% foreign owned	0.073*	0.0032	0.0160	0.0320	0.0006	0.0028	0.0056
Exporting directly at least 10% of sales	0.070*	0.0028	0.0140	0.0280	0.0005	0.0026	0.0053
Proportion owned by females	0.820	0.0096	0.0480	0.0960	0.0050	0.0250	0.0500
Female top manager	0.530	0.0076	0.0380	0.0760	0.0014	0.0071	0.0143
Main Market – local, national, international	0.350	0.0060	0.0300	0.0600	0.0009	0.0046	0.0091
Multi-establishment	0.740	0.0084	0.0420	0.0840	0.0020	0.0100	0.0200
Legal form	0.130	0.0048	0.0240	0.0480	0.0007	0.0036	0.0071
Sole Proprietorship	0.033**	0.0024	0.0120	0.0240	0.0005	0.0025	0.0050
Shareholding company	0.760	0.0088	0.0440	0.0880	0.0025	0.0125	0.0250
Share of transactions with suppliers that were fulfilled smoothly	0.130	0.0044	0.0220	0.0440	0.0007	0.0033	0.0067
Corruption and security							
Bribery depth	0.680	0.0080	0.0400	0.0800	0.0017	0.0083	0.0167
Bribery incidence	0.510	0.0072	0.0360	0.0720	0.0013	0.0063	0.0125
Paying for security	0.870	0.0100	0.0500	0.1000	0.0100	0.0500	0.1000

* significance at 10%, ** at 5%, *** at 1%.

Table D.2b: Tests of the Association Between Governance Structures and a Variety Of Variables, Relations With Customers

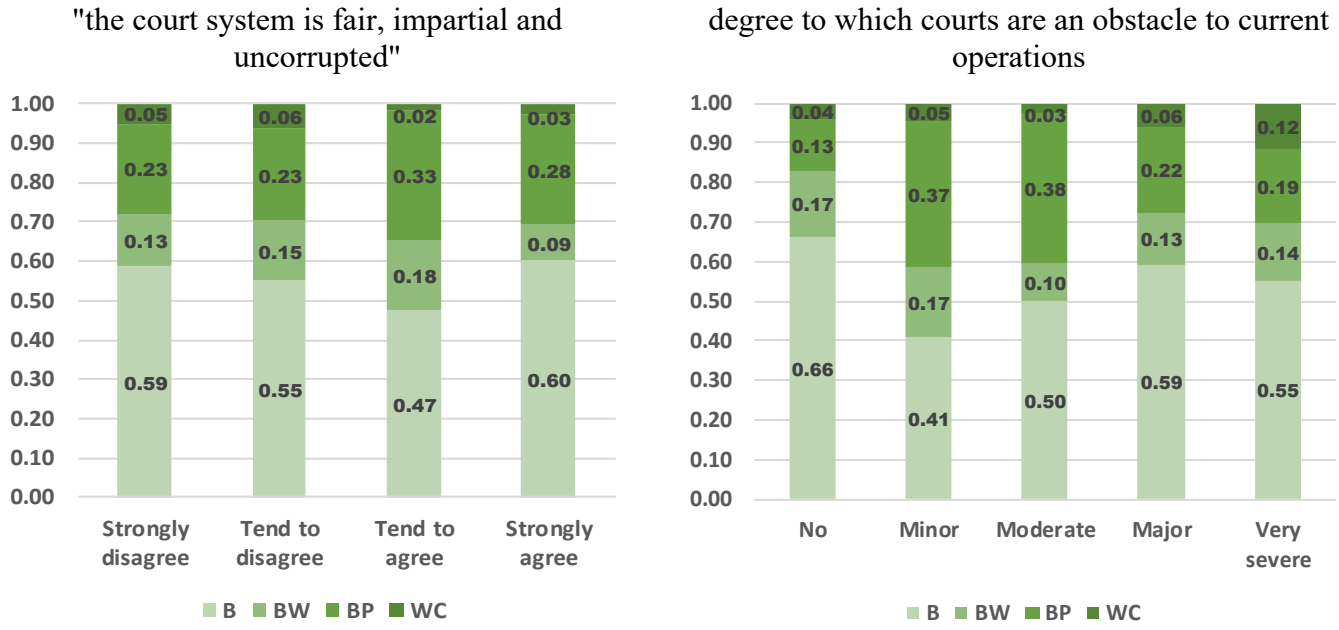
		FDR			FWER		
Description	p-value	1%	5%	10%	1%	5%	10%
Country	0.000***	0.0004***	0.0020**	0.0040*	0.0004***	0.0020**	0.0040*
Attitudes towards courts							
"the court system is fair…"	0.280	0.0076	0.0380	0.0760	0.0014	0.0071	0.0143
Courts as an obstacle	0.000***	0.0008***	0.0040**	0.0080*	0.0004***	0.0021**	0.0042*
Relations with business associations							
Belong to a business association?	0.041**	0.0032	0.0160	0.0320	0.0006	0.0028	0.0056
Regularly interact with a business association?	0.024**	0.0020	0.0100	0.0200	0.0005	0.0024	0.0048
Sector of operation							
Manufacturing, retail, or other services	0.000***	0.0012***	0.0060**	0.0120*	0.0004***	0.0022**	0.0044*
Disaggregated sector	0.034**	0.0028	0.0140	0.0280	0.0005	0.0026	0.0053
Management practices							
Management practices (higher means better)	0.031**	0.0024	0.0120	0.0240	0.0005	0.0025	0.0050
Top manager's experience working in sector	0.056*	0.0036	0.0180	0.0360	0.0006	0.0029	0.0059
Firm characteristics							
Size	0.160	0.0044	0.0220	0.0440	0.0007	0.0033	0.0067
Age	0.180	0.0056	0.0280	0.0560	0.0008	0.0042	0.0083
Proportion domestic private ownership	0.200	0.0064	0.0320	0.0640	0.0010	0.0050	0.0100
At least 10% foreign owned	0.019**	0.0016	0.0080	0.0160	0.0005	0.0023	0.0046
Exporting directly at least 10% of sales	0.180	0.0060	0.0300	0.0600	0.0009	0.0046	0.0091
Proportion owned by females	0.240	0.0068	0.0340	0.0680	0.0011	0.0056	0.0111
Female top manager	0.720	0.0092	0.0460	0.0920	0.0033	0.0167	0.0333
Main Market – local, national, international	0.620	0.0088	0.0440	0.0880	0.0025	0.0125	0.0250
Multi-establishment	0.910	0.0100	0.0500	0.1000	0.0100	0.0500	0.1000
Legal form	0.240	0.0072	0.0360	0.0720	0.0013	0.0063	0.0125
Sole Proprietorship	0.170	0.0048	0.0240	0.0480	0.0007	0.0036	0.0071
Shareholding company	0.450	0.0080	0.0400	0.0800	0.0017	0.0083	0.0167
Share of transactions with customers that were fulfilled smoothly	0.180	0.0052	0.0260	0.0520	0.0008	0.0039	0.0077
Corruption and security							
Bribery depth	0.900	0.0096	0.0480	0.0960	0.0050	0.0250	0.0500
Bribery incidence	0.530	0.0084	0.0420	0.0840	0.0020	0.0100	0.0200
Paying for security	0.110	0.0040	0.0200	0.0400	0.0006	0.0031	0.0063

Figure D.1a: Attitudes towards courts and governance structures in relations with suppliers



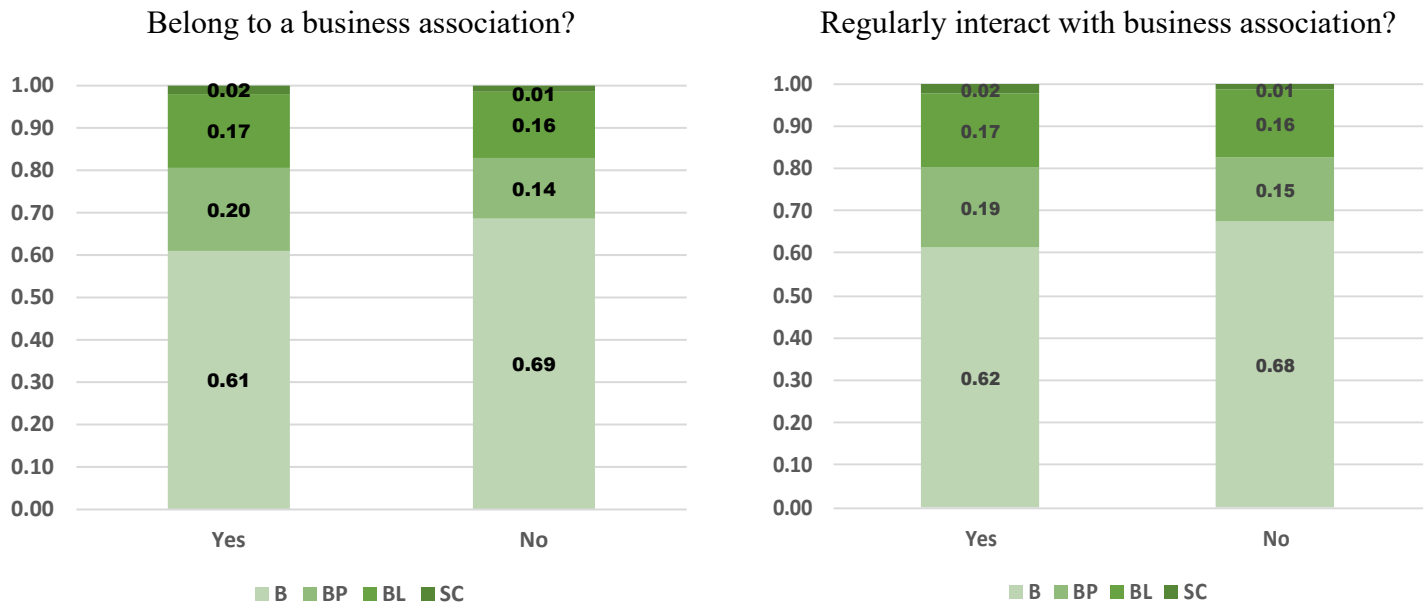
B = bilateralism, BP = bilateralism with private support, BL = bilateralism with legal support, SC = strong comprehensive

Figure D.1b: Attitudes towards courts and governance structures in relations with customers



B = bilateralism, BW = bilateralism with weak support, BP = bilateralism with private support, WC = weak comprehensive

Figure D.2a: Membership and interactions with business associations as they relate to the governance structures in relations with suppliers



B = bilateralism, BP = bilateralism with private support, BL = bilateralism with legal support, SC = strong comprehensive

Figure D.2b: Membership and interactions with business associations as they relate to the governance structures in relations with customers

B = bilateralism, BW = bilateralism with weak support, BP = bilateralism with private support, WC = weak comprehensive

Figure D.3a: Sectors and governance structures in relations with suppliers

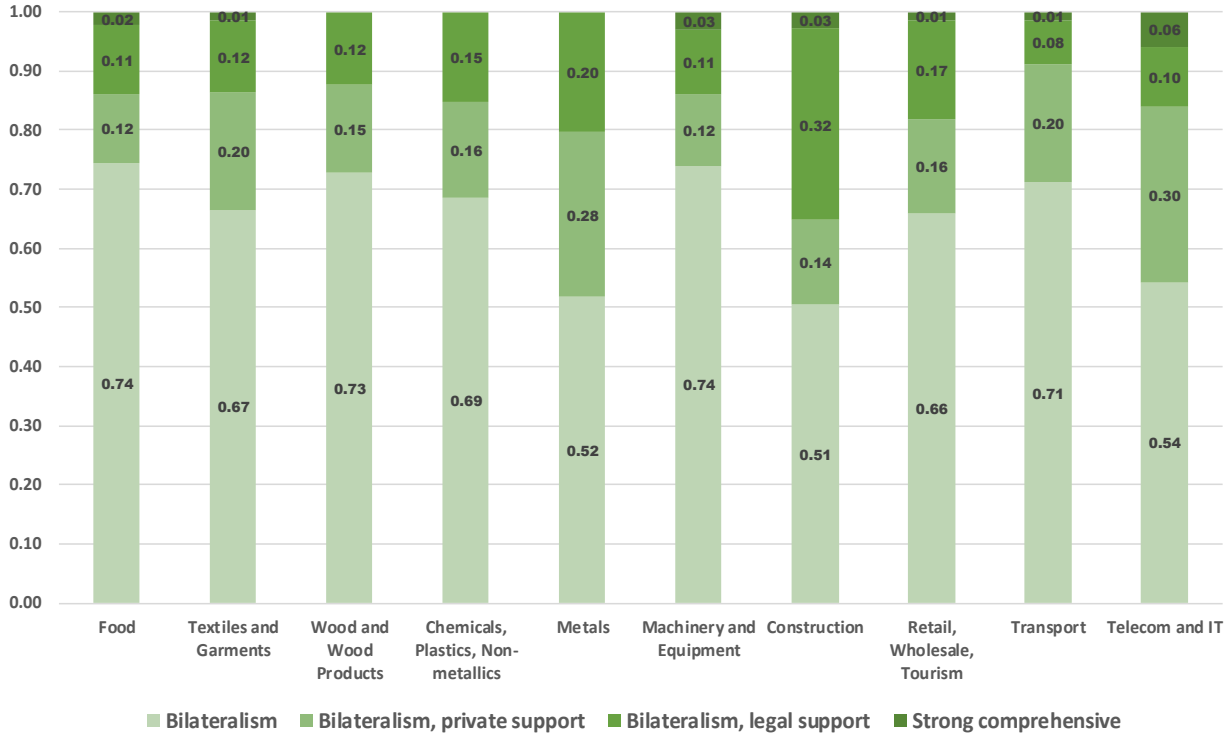


Figure D.3b: Sectors and governance structures in relations with customers

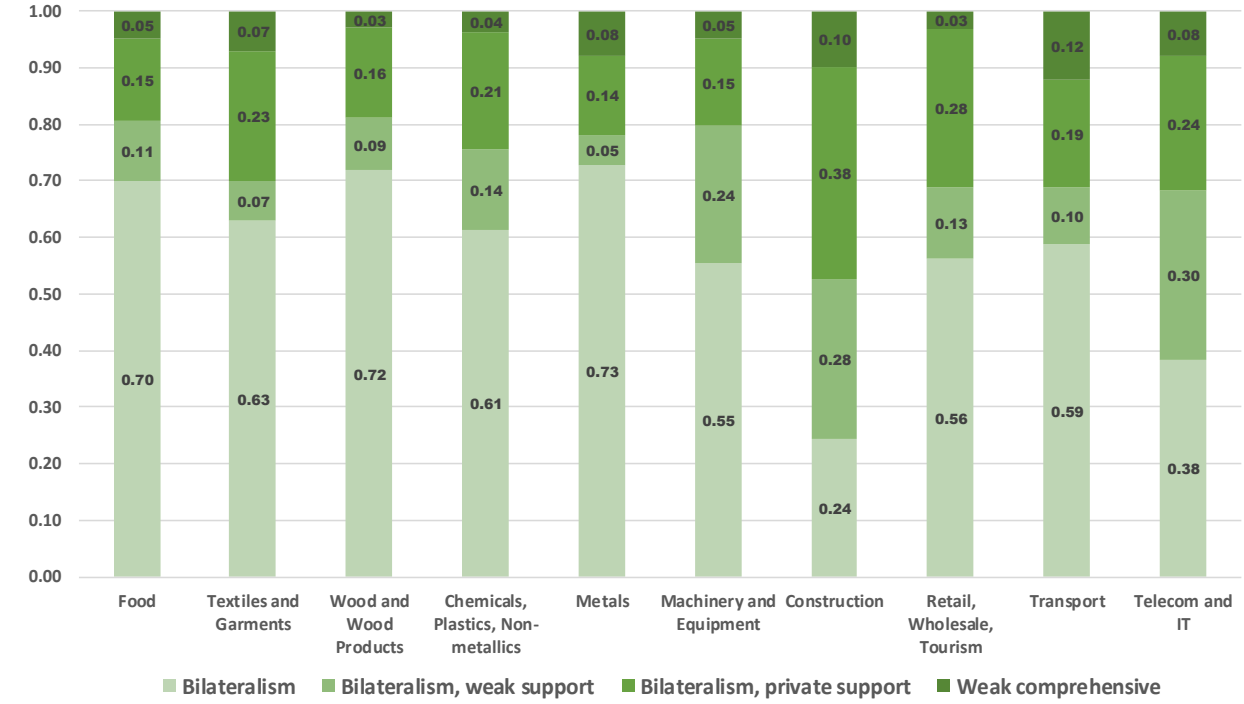


Figure D.4: Foreign ownership and governance structures

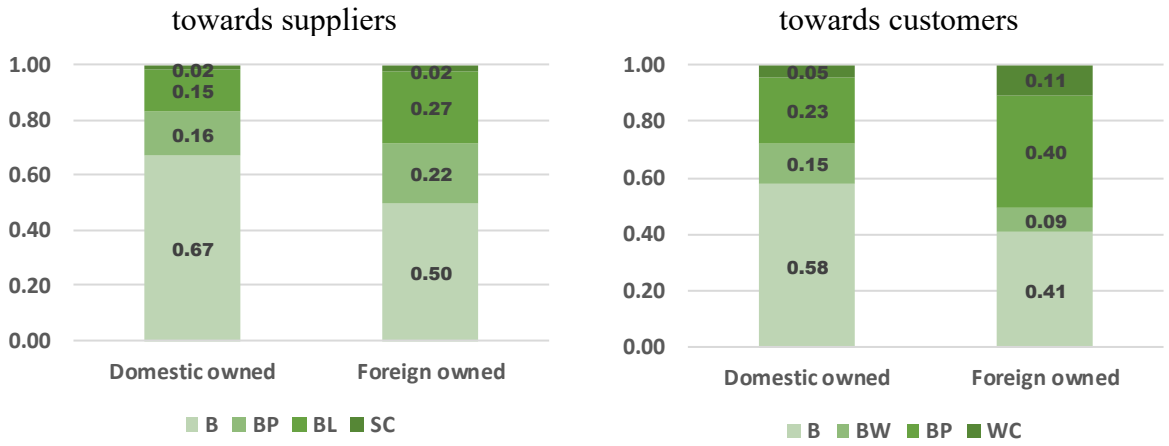


Figure D.5: Exporting status and governance structures

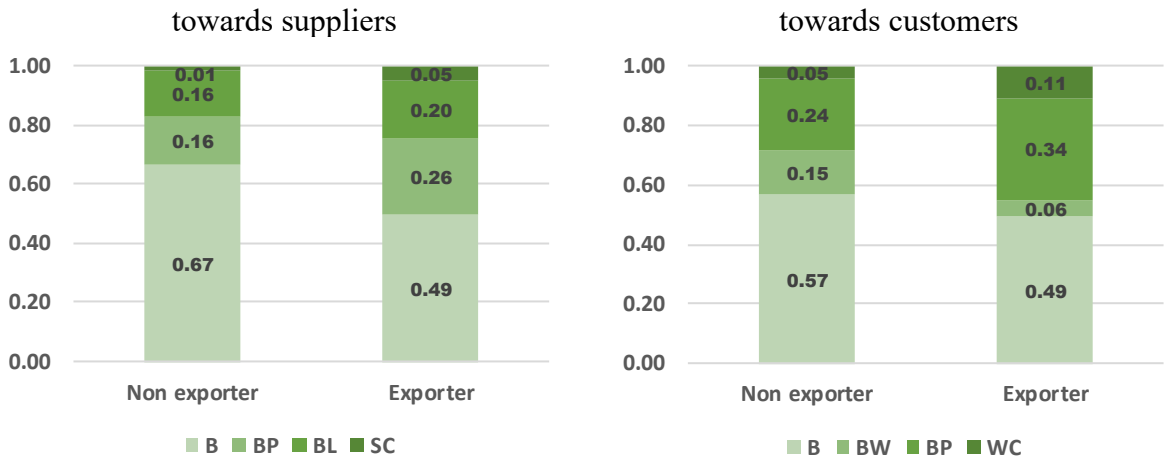
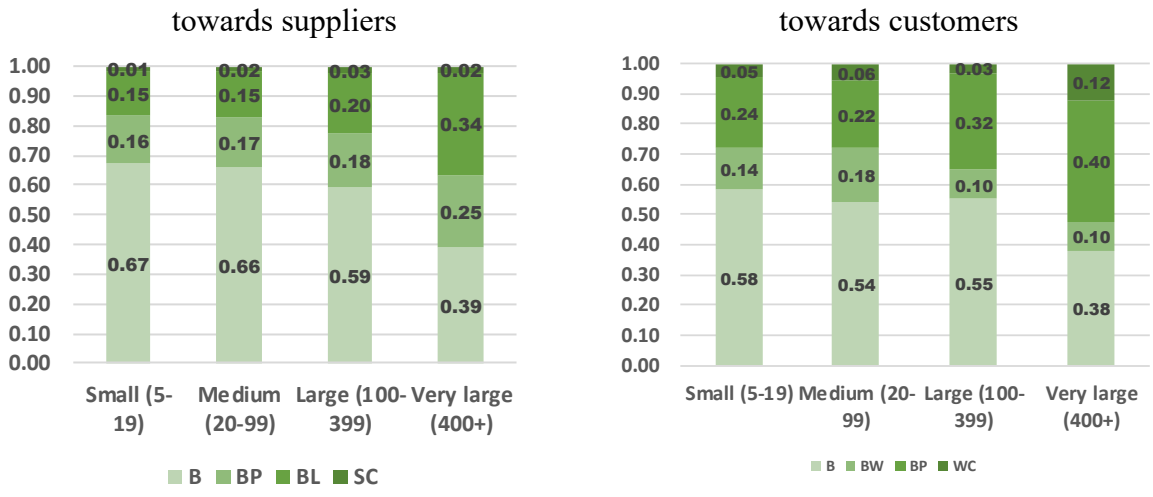


Figure D.6: Firm size and governance structures



Appendix E

This appendix provides more details on the variables used in Section VI to study the link between uncertainty and bilateralism. As the dependent variable, we use the posterior probability that the firm employs pure bilateralism for its transactional governance structure. This posterior probability is estimated as an integral part of the LCA technique, described in Section III. The posterior probability is a continuous variable varying over the interval (0,1).

Uncertainty is our main explanatory variable, and it is measured using the following question implemented as part of the ES survey: "[p]lease indicate to what degree this establishment's suppliers are prevented from fulfilling agreements because of circumstances beyond their control." The respondents were shown a card with the following scale: not at all, slightly, moderately, very much, and extremely. This question was classified as an opinion-based question, meaning that interviewers were not to provide any clarification. In case a respondent asked for clarification, the interviewers were instructed to read the question again without any additional information in order to avoid misunderstanding or confusion. Our main explanatory variable is a dummy equal to 1 if firms choose anything other than "not at all" from the above options and 0 otherwise.

The control variables included as part of the vector X_{isr} are as follows. For firm size, we use the log of a firm's total number of full-time equivalent permanent workers. It is plausible that larger firms have more capacity to employ lawyers and use more sophisticated governance structures, or that a larger scale of operations changes a firm's exposure to uncertainty. The ES is an establishment-level survey, and respondents are asked whether the establishment is part of a larger firm. The dummy variable included in X_{isr} equals 1 if the response is Yes, and 0 if No. A more complex firm structure may cushion against uncertainty, while changing the types of governance arrangement that is optimal for the firm. For firm age, we use responses to the question "[i]n what year did this establishment begin operations?". We apply an outlier removal procedure as implemented for the corresponding standard ES indicator.[12] In particular, we log-transform the variable and exclude observations that are three standard deviations above or below the unweighted mean for each country. This outlier removal procedure removes a total of six observations. We included firm age because older firms are more likely to have built relationships with peers, suppliers, or customers that affect their decisions on governance structures. At the same time, older firms may also be distinctly perceptive or knowledgeable about the uncertainties faced by their suppliers. Finally, exporting and foreign-ownership statuses are captured in two dummy variables. The exporter dummy variable equals 1 if the firm directly exports at least 10% of its sales and 0 otherwise. The foreign-owned dummy variable equals 1 if the firm is at least 10% foreign-owned and 0 otherwise. Firms that are foreign-owned or export across borders are likely to have a sophisticated internal structure that might change the types of governance structures they choose. A foreign orientation can also change exposure to uncertainty, perhaps through additional steps or processes required to comply with cross-border regulations. It has been widely noted that

[12] See the detailed description here: https://www.enterprisesurveys.org/content/dam/enterprisesurveys/documents/Indicator-Description.pdf

operating across distinct environments (with differing cultures, governments, and contexts) raises uncertainty (see, e.g., Handley and Angst 2015). For sector fixed effects, δ_s, we use the two-digit ISIC (Rev. 3.1). (The ES contains four-digit ISICs for each firm, which are based on the descriptions of the firms' main activity and product given by the owner or top manager. We use its two-digit version because at any higher level of granularity the number of ES observations in each cell drops dramatically.) For regional dummies, we use the ES stratification regions within each of the six countries.[13]

As noted in Section VI.2, to construct u_{sr}, we first group observations at the level of the two-digit ISIC sector and region. We then calculate the weighted average of the uncertainty variable for each such sector-region group. If fewer than 5 observations are available in a cell, then we do not use the corresponding observation in the regression. In total, there are 420 sector-region cells, 4 of which have no observations on uncertainty, and 237 have fewer than 5 observations. The mean (median) number of firms in the sector-regions groups used in the analysis is 16.0 (9).[14]

Summary statistics for all the variables used in Section VI are in Table E.1.

[13] For details of stratification regions, see Implementation Reports available as part of the documentation for each country on the ES data portal (login.enterprisesurveys.org). We used variable a3ax from the cross-country ES indicators database which corresponds to these regions.

[14] We also used an alternative set of instrumental variables, reaching the same conclusions. These results are available upon request.

References to Appendix E

Handley, Sean M., and Corey M. Angst. "The impact of culture on the relationship between governance and opportunism in outsourcing relationships." *Strategic Management Journal* 36, no. 9 (2015): 1412-1434.

Table E.1: Summary Statistics

	Mean	S.D.	Min	Max
Pure bilateralism	0.667	0.436	0	1
Uncertainty	0.529	0.499	0	1
Log of size	2.860	1.130	0	9.488
Firm is part of a larger firm	0.119	0.324	0	1
Firm age (years)	25.107	17.842	2	187
Exporting directly at least 10% of sales	0.070	0.256	0	1
At least 10% foreign-owned	0.050	0.219	0	1