Rule of Law, Economic Structure and Development

Roberto M. Samaniego[*] Juliana Y. Sun[†]

March 19, 2024

Abstract

If rule of law encourages relationship-specific investments, then industries that use intermediates which require relationship-specific investments should have a disproportionately low share of output or labor in countries where rule of law is weak. We find robust support for this prediction using data on industry composition for 189 countries. Using a standard preference framework to construct model-implied income values from the estimated coefficients, we find that the interaction between relationship specificity and rule of law may be an economically significant determinant of aggregate outcomes.

Keywords: Rule of law, relationship specificity, economic structure, economic development.

[*]Roberto M. Samaniego, George Washington University, 2115 G St NW Suite 340, Washington, DC 20052. Tel: (202) 994-6150. Fax: (202) 994-6147. Email: roberto@gwu.edu.

[†]Juliana Y. Sun, Monetary Authority of Singapore, 10 Shenton Way. Singapore 07911. Email: juliana_sun@mas.gov.sg.

JEL Codes: O11, O43.

1 Introduction

Two key goals of macroeconomic research are accounting for cross country differences in aggregate income and accounting for the link between economic structure and development. Productivity has been found to be a key factor behind both differences in income and differences in economic structure – see Ngai and Pissarides (2007), Restuccia, Yang and Zhu (2008) and Duarte and Restuccia (2010) among others. Independently, it is well known that measures of *rule of law* (ROL) are key empirical correlates of economic development – see for example Shleifer and Vishny (1998) and Ranasinghe and Restuccia (2018).

Suppose that weak ROL makes it easier for agents to break agreements, exacerbating the severity of hold-up or other contracting problems in the economy. This would lead to low productivity in the production of goods that require relationship-specific investments, suggesting one channel through which weak ROL might influence levels of economic development: low economic efficiency due to the discouragement of relationship-specific investments. Furthermore, the use of relationship-specific inputs varies across industries. This suggests that ROL would not only affect the level of economic development but also the *structure* of economies at different levels of development. As a result, we can detect and quantify the effect of ROL on economic development by exploiting *industry heterogeneity* in the use of relationship-specific inputs.

In this paper, we focus on the impact of ROL on the final goods' structure of the economy. First, we estimate a regression equation that includes an interaction of ROL and relationship specificity, using data on employment shares of 14 industries, covering the entire private economy in 189 countries. We use the Rauch (1999) measure of relationship specificity, whereby a good that is largely traded in an organized exchange is not "relationship specific", whereas a good that is not traded in an organized exchange is relationship specific – as it is less standardized, requires customization and thus is more vulnerable to flaws in the contracting relationship. We measure ROL using data from the World Bank Governance Indicators. We find that high ROL leads industries with more relationship specific intermediates to expand their share of economic activity. This finding is robust to a variety of controls and alternative specifications.

Then, we present a standard CES preference framework, where differences in productivity map into differences in economic structure as in, for example, Ngai and Pissarides (2007). If weak ROL makes it easier for buyers to renege on payment, this should particularly afflict the production of relationship specific intermediates. As a result, productivity in the production of relationship specific intermediates should be lower, particularly where ROL is weak. The model predicts that, when goods are substitutes, the output of goods that use relationship-specific inputs should be disproportionately low in countries with weak rule of law, as found earlier. Indeed, equilibrium behavior can be approximated as a linear regression equation where industry structure depends on an interaction of relationship-specificity and ROL identical to the empirical specification.

Finally, we use the preference framework to derive aggregate implications from our estimated interaction coefficients. We calibrate the model to exactly match the industry composition of the United States, and then use the model coefficients to produce counterfactual GDP per capita values for the 174 countries for which we have both Rule of Law data and GDP per capita data in the Penn World Tables 9.1. (PWT) – see Feenstra, Inklaar and Timmer (2015). We find that model-generated values of income are very highly correlated with those in the data. In addition, when we compare GDP per capita relative to the US in the data on the model-generated values, we find that the model is able to generate between a quarter and a half of the variation in levels of GDP per capita in the PWT 9.1.

A key input into this computation is the elasticity of substitution between the output of different sectors. We estimate this elasticity using price and expenditure data from the International Comparisons Program (ICP) of the World Bank. We find estimates of this elasticity to cluster around 1.8, a value intermediate to estimates found in disaggregated manufacturing data (e.g. Samaniego and Sun (2016)) and in 3-sector data comprising the private economy (Herrendorf et al (2013)). The estimates vary between 1.7 and 2.1 depending on the specification, but the extent to which the model accounts for variation in GDP per head around the world is larger for the estimates with more controls.

Our paper contributes to the literature linking economic development with institutions. For example, Ranasinghe (2017) and Ranasinghe and Restuccia (2018) relate rule of law to the risk of expropriation, so that low ROL leads to more resource misallocation, increased investments in protection and tighter financing constraints.

In Acemoglu and Johnson (2005), ROL is viewed as an institutional underpinning for the development of financial markets. In our case, we look at the interaction of ROL with relationship specificity, and show that ROL has an impact both on economic development and on economic structure.

Our paper also contributes to the literature on understanding the link between economic development and economic composition. This literature has tended to focus on composition defined in terms of agriculture, services and manufacturing – see Gollin, Parente and Rogerson (2007), Restuccia, Yang and Zhu (2008) and Duarte and Restuccia (2010). In contrast, we use more disaggregated data, which provides a sterner test. Nunn (2007) and Levchenko (2007) show that industries that use more relationship-specific intermediates make up a larger share of exports in countries with strong rule of law, identifying the interaction of relationship specificity and ROL as a key determinant of comparative advantage. We show the same pattern emerges in the distribution of economic activity overall and, notably, that this holds for non-tradeables as well as tradeables. Boehm (2022) shows that the input-output structure of economies is connected to contract enforcement costs, likely indicating sizable welfare gains from improvements in rule of law. What these papers do not address is to what extent such a mechanism might account for the final goods' composition of the economy – including non-tradeables – nor do they directly assess the aggregate impact of weak ROL on relationship specific activity. Indeed, it is not possible to assess the aggregate impact without including non-tradeables, which comprise most of economic activity. In encompassing all sectors of the economy, our work overcomes this limitation.

Our model framework assumes a closed economy context. This is consistent with the fact that our data on economic structure cover all sectors, most of which are non-tradeable. In addition, we find that our estimates with tradeables (manufacturing) are robust to conditioning on factors of trade. This is consistent with the survey of Goldberg and Pavcnik (2004), which finds that "empirical work has consistently documented a lack of major labor reallocation across sectors" in response to trade liberalization: thus, trade does not appear to be a key determinant of structural change. In addition, Święcki (2017) compares several determinants of structural change, finding that sector productivity growth rates are an important mechanism – whereas trade considerations are important only in selected economies, not across the board. This indicates that growth-theoretic considerations are likely to be more important determinants of structural transformation than trade, motivating our approach. Of course, it would be interesting to extend the model to an open economy context, but this is unlikely to overturn our main point: the impact of weak institutions on aggregates and on the economic structure of final goods and services is highly significant, including non-tradeables.

Section 2 discuses the data and presents empirical results on rule of law and economic structure. Section 3 describes the model framework, and derives the empirical specification from equilibrium behavior. Section 4 presents the calibration of the preference framework and studies the aggregate implications of our findings. Section 5 concludes.

2 Estimation

2.1 Empirical Specification

Suppose that RS_j is relationship-specificity measure for industry j, and suppose also that ROL_c is a measure of rule of law for country c. Let S_{jc} be the share of industry j in the output or employment of country c. We estimate the following specification:

$$\log S_{jc} = \delta_j + \delta_c + \hat{\beta}\,(RS_j \times ROL_c) + \varepsilon_{jc}. \tag{1}$$

The dummy variables δ_j and δ_c will soak up any industry- or country-specific factors respectively. The coefficient of interest is $\hat{\beta}$, which captures any interaction between relationship specificity and rule of law.[1]

In what follows, we will estimate specification (1) using economy-wide data, and also using disaggregated manufacturing data. In general, the literature on relative output shares often proxies for them using relative labor shares, as the latter tend to be more widely available (particularly outside of manufacturing). We will use labor shares where no alternative is available, and both where available for robustness.

We build on the extensive differences-in-differences literature that posits that the technology of production varies across industries in a systematic manner that is largely preserved across countries, at least in terms of rankings.[2] If this variation is more or less preserved across countries, then industry variation in, for example, the

[1] We use the natural logarithm of S_{jc} rather than S_{jc} because, as we will see later, equation (1) maps into the equilibrium of the preference framework we use later to derive the aggregate implications of $\hat{\beta}$. The results are robust to using S_{jc} as the dependent variable.

[2] See Rajan and Zingales (1998), Dell'Ariccia et al (2008) and Samaniego and Sun (2015) inter alia.

composition of the intermediate goods used would be preserved across countries.

In estimating (1), the dummy variables δ_j and δ_c are industry j and country c dummies. When estimated, they capture any industry- and country-specific factors that might affect industry shares that are not considered by the model. The error terms are estimated allowing for heteroskedasticity using the Huber-White method.

2.2 Data

2.2.1 Industry share data

For industry shares, S_{jc}, we use two distinct data sets. As a benchmark, we draw on a database that covers the entire economy for a large number of countries. This is the ILO employment dataset, which reports employment by industry for 14 industries in 189 countries over the years $1991 - 2018$. We use the average employment share during the sample period in our baseline regression (1), as output data are not available for broad sectors.

For robustness, we also draw on a more disaggregated database that focuses only on manufacturing industries. This is the INDSTAT 4, 2019 ISIC-revision 3 database, which ranges from $1973 - 2016$ and is collected and distributed by the United Nations Industrial Development Organization (UNIDO). We use the 28 manufacturing industries based on the ISIC-revision 2 classification, for which Nunn (2007) reports the values of the variables we use later to measure vulnerability RS_j. The advantage of these data is that it can be used to compute shares of employment, shares of output or shares of value added. We can thus check all three variables for robustness. We use average data for the years $1996 - 2016$, to overlap with the period for which we have data for measuring ROL_c. There are 84 countries reporting employment data, 77 reporting value-added data and 83 reporting output data.

2.2.2 Rule of law and Relationship specificity

We measure ROL_c using the Rule of Law index drawn from the Worldwide Governance Indicator dataset, maintained by the World Bank. The Rule of Law (ROL) measure captures the extent to which agents have confidence in and abide by the rules of society, especially the quality of contract enforcement, property rights, the police, and the courts, and the likelihood of crime and violence. The data set reports values ranging from $1996 - 2016$. We use the average ROL during the sample period in our regressions.[3]

To measure vulnerability to contract violations RS_j, we use the relationship-specificity indicator developed in Rauch (1999), as reported in Nunn (2007). It measures the extent to which inputs are dependent on relationship-specific investment between the supplier and the buyer. Nunn (2007) measures, for each good, the proportion of inputs that are not sold on an organized exchange nor reference-priced. If inputs are sold on an organized exchange, there must exist a large number of buyers and sellers, indicating this good is not dependent on relationship-specific investments. This means that these intermediate goods are relatively standardized. See Nunn (2007) for further discussion regarding how relationship-specific interme-

[3]The World Bank reports for each country the standard error of the ROL_c estimate. Because ROL_c enters our regression specification (1) as part of an interaction term, there is no easy way to adjust for these standard errors. However, in Appendix H we discuss the robustness of our estimates to these standard errors by drawing ROL_c values from the implied distribution and examining the distribution of the resulting estimated coefficients, a sort of simulated bootstrap.

diates depend more on the institutional environment for their delivery and quality.

Nunn (2007) only reports RS_j for manufacturing industries. As a result, in order to estimate equation (1) for non-manufacturing industries, we must extend his measure to more industries. Our procedure is as follows. First, we regress Nunn's RS_j measure in the manufacturing sector on the direct requirements variable from the US Input-Output (IO) 1997 tables. Thus, the contribution of each intermediate good to the measure of RS_j in all the manufacturing industries is given an estimated coefficient by the regression.[4] We use a fractional regression procedure to ensure all estimates have to be between zero and one – see Papke and Wooldridge (1996). Then, we use these coefficients to predict RS_j for other industries that are not in manufacturing, using *their* direct requirements for each of the intermediate goods.

The IO tables are much more disaggregated than the ILO data (which has only 14 industries, compared to 490). To aggregate up to the ILO level, we first generate a predicted RS_j measure at the level of disaggregation in the IO tables. Then, suppose we wish to compute RS_j for an ILO non-manufacturing industry i, which contains a subset of disaggregated industries I_i that feature in the IO tables. We use the 1997 make tables to compute the relative contribution of each of the elements of I_i to the output of i. Finally, we use the predicted RS_j at the disaggregated level in the first step to compute the weighted average RS_j at the ILO level of aggregation.

Nunn (2007) reports two measures of RS_j: one narrow, which is the proportion of intermediate inputs used in industry j that are neither sold on an organized exchange

[4] There are more potential inputs (over 458) than there are manufacturing industries in Nunn's data (381), this would be an overidentified regression. We cull the inputs by removing any that do not comprise at least 5 percent of the inputs used by some industry. This reduces the size of independent variables to 229. The R-squared values are extremely high, around 0.99.

Table 1: Estimates of the interaction between rule of law at the country level and relationship specificity at the industry level.

	ILO sectors (log emp shares)			
	Correl	Baseline	Income. eff.	No agric.
Coefficient	0.622***	1.057***	0.705***	0.660***
s.d.	(0.0408)	(0.0796)	(0.0655)	(0.0642)
Control on j, c dummies	no	yes	yes	yes

nor reference priced, and one broad, which is the proportion of components that are not sold on an organized exchange. In the paper we focus on the narrow measure. In the Appendix we report results for the broad measure, which are generally similar.

2.3 Baseline Results

The result with the ILO employment share data indicates the coefficient $\hat{\beta}$ in equation (1) is positive and significant at 1% level.[5] This is true without industry and country dummy variables; however, introducing the dummy variables increases the magnitude of the coefficient. See column (1) and (2) in Table 1.

The literature on structural transformation frequently raises the possibility of income effects affecting the share of agriculture. We consider whether our results are sensitive to the treatment of agriculture by including an interaction variable of an agriculture dummy variable with the log of GDP per head in each country average over the sample period, as measured in the Penn World Tables (PWT) 9.1,[6] and by dropping agriculture altogether. Introducing the interaction of an Agriculture dummy and log GDP per capita lowers the coefficient but it remains highly signifi-

[5]Throughout the paper, one, two and three asterisks represent statistical significance at the 10, 5 and 1 percent levels respectively.

[6]On the PWT 9.1, see Feenstra, Inklaar and Timmer (2015).

Table 2: Estimates of the interaction between rule of law at the country level and relationship specificity at the industry level for employment share.

	ILO sectors (emp shares)			
	Correl	Baseline	Income. eff.	No agric.
Coefficient	0.0115***	0.0802***	0.0215***	0.0197***
s.d.	(0.00260)	(0.00810)	(0.00380)	(0.00265)
Control on j, c dummies	no	yes	yes	yes

cant, as it does when we remove Agriculture from the specification. See columns (3) and (4) in Table 1. These results indicate that industries with more dependence on relationship specific inputs tend to have higher shares if they are located in countries with sound rule of law.

We check whether the baseline results hold for employment share as the dependent variable, instead of log of employment share. Naturally the coefficients are smaller in magnitude since the variation in S_{jc} will be larger than that of its log. Nonetheless, the coefficients are still significant at 1% level, consistent with the baseline results. See Table 2.

In Appendix D, we show that our interaction of interest is robust to measuring rule of law using contracting institutions instead, and to allowing for endogeneity of rule of law using instrumental variables. Perhaps most importantly, we show that our findings are robust to controlling for potential income effects outside of agriculture in a variety of ways. We also show that the results hold for disaggregated manufacturing industries. Finally, noting that our ROL measure is an estimate, and that the data provider reports standard deviations, we run many thousands of versions of our baseline regression using values of ROL drawn from these distributions. The results are robust to all these tests.

To sum up, we find a robust link between economic structure and the interaction of ROL with relationship specificity. We find that the raw correlations tend to be smaller than the coefficients when we add control variables such as fixed effects, which tells us that industry shares are influenced by country- and industry-specific factors, as is widely assumed in the related literature, and that it is important to condition on such factors. We tend to find somewhat smaller coefficients when we allow for income effects in agriculture, drop agriculture, or account for income effects in general: this tells us that controlling for income effects is important but they do not overwhelm our results. The coefficients are smaller when we use shares rather than log shares, this is not surprising as the shares are one or two orders of magnitude larger than the log shares so these coefficients are not comparable. Finally, the coefficients are larger when we use instrumental variables, this is not unusual in the differences-in-differences literature (e.g. Rajan and Zingales (1998)).

3 Model Environment

We now present a model economy that articulates how the vulnerability to low-quality intermediates can raise the cost of final goods, particularly where ROL is weak. We use the model to provide a foundation for the simple and intuitive regression specification we estimate in the previous section, and to provide a quantitative framework for assessing whether the interaction of interest can have a significant macroeconomic impact. The interaction coefficient estimated in Section 2 will play a key role in calibrating the model economy. We will abstract from income effects in order to maintain tractability, because we found that income effects had little impact on our estimates of the interaction coefficient, and because income effect coefficients were uncorrelated with the interaction of interest.

3.1 Households

There are $J \in \mathbb{Z}$ final goods and C countries. Households maximize discounted utility defined over consumption $u(\mathbf{c})$, where u is increasing and quasiconcave and where the consumption aggregate $\mathbf{c}$ combines the J final goods using a CES function:

$$\mathbf{c} = \left(\sum_{j=1}^{J} \omega_j c_{cj}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}, \quad \sum_{j=1}^{J} \omega_j = 1. \tag{2}$$

where c_{cj} is consumption of good j in country c.

Households earn income from working, supplying up to one unit of labor per period to a competitive labor market in exchange for a wage w:

$$\sum_{j=1}^{J} p_{cj} c_{cj} \leq w \tag{3}$$

where p_{cj} is the price of good j in country c.

In equilibrium w will equal total value produced (nominal GDP). Let $w = 1$ so that labor is the numeraire.

We assume $u(\mathbf{c}) = \mathbf{c}$ in order to abstract from risk generated by failed contracts. It would be interesting to explore this *risk channel* of rule of law in future work.

We also assume a closed economy environment, so that c_{cj} for any good j equals production of good j. This is not a limiting assumption as most of the sectors we look at are non-tradeables, although we will look at tradeables also.

3.2 Final goods

Each good j is assembled from a combination of relationship-specific intermediates (RS) and non-relationship specific intermediates (NRS). Goods differ in their input requirements. Suppose that producing one unit of good j requires $a_j \in [0, 1]$ units of the RS input and $1 - a_j$ units of the NRS input. Final goods are assembled by a *principal*, which is a profit-maximizing technology owned and operated by households (i.e. a firm). Making one unit of an intermediate of either type requires one unit of labor supplied by an *agent*.

The price of the RS intermediate is p_r and the price of the NRS intermediate is p_n.

Production of the RS intermediate is subject to a hold-up problem. The agent has an incentive to provide low effort if they do not expect to be paid, or expect to be paid less than otherwise, leading to lower productivity or higher costs. Higher costs translate into higher prices for the RS intermediate.[7] Alternatively, p_r will be lower where contract enforcement is more effective.

The expectation of being paid depends on an institutional parameter γ_c. We think of γ_c as a contract enforcement parameter or an index of ROL, which may vary by country c. As a result, p_r is decreasing in γ_c: $p_r = p_r(\gamma_c)$, and $p_r' < 0$. On

[7] The assumption that goods tend to be more expensive if the cost of making them is higher (or, equivalently, productivity is lower) is a common feature of models of economic structure such as Ngai and Pissarides (2007), Samaniego and Sun (2020) or Duarte and Restuccia (2020).

the other hand, the value of p_n is not affected by γ_c. Note that this is conservative to assume p_n is unaffected by institutions, in the sense that it lowers the potential aggregate impact of γ_c.

The price of final good j is $p_{cj} = p_{cj}(a_j, \gamma_c)$. If final goods' markets are competitive, then $p_{cj}(a_j, \gamma_c) = a_j p_r(\gamma_c) + (1 - a_j) p_n$. This implies that:

Condition 1 *Under the above assumptions,*

$$\frac{d^2 p_{cj}(a_j, \gamma_c)}{da_j d\gamma_c} = p_r'(\gamma_c) < 0. \tag{4}$$

The idea that weaker ROL should disproportionately lower productivity (and thus raise prices) in industries that use RS intermediates is intuitive. Nonetheless, Appendix A presents a specific example of a production environment that satisfies this condition due to a hold-up problem. In the example, final good j is assembled from a_j units of the RS intermediate and $1 - a_j$ units of the NRS intermediate, produced by a principal and an agent. Whether the output of the RS intermediate is usable depends on uncontractible effort q by an agent who works with the principal – and whether a good is usable can only be verified with probability γ_c. If so, the principal and the agent bargain over the surplus. In this environment, Appendix A shows that equation (4) holds.

3.3 Consumer's solution

The household's optimal spending implies that:

$$\omega_j c_{cj}^{\frac{-1}{\sigma}} \left(\sum_{j=1}^{J} \omega_j c_{cj}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}-1} = p_{cj} \tag{5}$$

Consequently, for any two goods j and k,

$$\frac{c_{cj}}{c_{ck}} = \left(\frac{\omega_j}{\omega_k} \right)^{\sigma} \left(\frac{p_{ck}}{p_{cj}} \right)^{\sigma} \tag{6}$$

Define S_{jc} as the share of output produced by industry j in country c so $S_{jc} = c_{cj} p_{cj}$ divided by output. Equation (6) then implies that:

$$\frac{S_{cj}}{S_{ck}} = \left(\frac{\omega_j}{\omega_k} \right)^{\sigma} \left(\frac{p_{ck}}{p_{cj}} \right)^{\sigma-1} \tag{7}$$

Taking logs, we find that

$$\log S_{cj} - \log S_{ck} = \sigma \log \omega_j - \sigma \log \omega_k - (\sigma - 1) \log p_{cj} + (\sigma - 1) \log p_{ck}. \tag{8}$$

Choosing a benchmark industry k, setting $\delta_c \equiv (\sigma - 1) \log p_{ck} + \log S_{ck} - \sigma \log \omega_k$ and $\delta_j \equiv \sigma \log \omega_j$, we get that

$$\log S_{cj} = \delta_c + \delta_j - (\sigma - 1) \log p_{cj}. \tag{9}$$

This equation is essentially a differences-in-differences regression specification that we can use to generate and test the model predictions, as well as to quantify the aggregate impact of the effects highlighted by the model.

Proposition 1 *Consider an interior value of Rule of Law* γ^* *and relationship specificity* a_j^*. *Given equation* (4), *industry structure in the model economy is given by*

$$\log S_{jc} = \delta_c + \delta_j + \beta \times a_j \times \gamma_c + \varepsilon_{jc}. \tag{10}$$

where β *is a constant with the same sign as* $(\sigma - 1)$, δ_c *is a constant that varies by country,* δ_j *is a constant that varies by industry and* $\varepsilon_{jc} \to 0$ *as* $[\gamma_c, a_j] \to [\gamma^*, a^*]$.

Proof. See Appendix B for details. Equation (10) is a second-order Taylor approximation of (9), with all country- and industry-specific terms that depend on γ_c or a_j absorbed into the constants δ_c and δ_j respectively. All that remains is the cross-derivative of $\log p_{cj}$, which we show shares the sign of the cross-derivative of p_{cj}. ■

Notice that, if a_j is an index of relationship specificity RS_j and γ_c is a measure of rule of law, equation (10) is exactly the same as our estimation equation (1). Thus, our estimates of the interaction between relationship specificity and ROL are consistent with the model as long as $\sigma > 1$.

As an econometric matter, if we were to estimate equation (10), the Proposition does not prove that, even though small, the remainder ε_{jc} might not be correlated with the regressors, the empirical proxies for a_j and γ_c. This is why in our empirical analysis we allow for heteroskedasticity as well as using instrumental variables, among other robustness checks.

The assumption that $\sigma > 1$ is well-established among disaggregated manufacturing industries: for example, Ilyina and Samaniego (2012) find values of σ ranging from about 1.5 up to about 4. On the other hand, the value of σ when considering industries outside of manufacturing is less clear. One reason is that income effects are thought to be a strong influence on the share of agriculture – see for example Gollin et al (2002) and Restuccia et al (2008). Another is that the literature on structural transformation tends to decompose the economy into only 3 sectors – agriculture, manufacturing and services – so that estimates based on that classification may not be useful for a more detailed breakdown such as the one we use in this paper.[8] An exception is Duarte and Restuccia (2020), who find that service industries where the relative price declines with development appear to expand as a share of output – suggesting that $\sigma > 1$, given the strong relation between ROL and GDP per capita. Thus, we expect that $\hat{\beta} > 0$, whether we look at all industries, or whether we focus exclusively on disaggregated manufacturing industries.

4 Quantitative Experiments

4.1 Calibration

In this section we use the preference framework and the estimated interaction coefficients to quantify the impact of the interaction between rule of law and relationship specificity on aggregate output. We do so by using the model to back out what the impact on relative shares of the interaction between vulnerability and weak enforcement tells us about relative prices, and hence the efficiency of production in different

[8]For example, in some specifications Herrendorf et al (2013) estimate that $\sigma < 1$.

industries.

Recall that final output in the model is defined as a CES aggregate $\mathbf{c}$ of all the final goods industries c_j, with elasticity of substitution σ and weights ω_j. Given income I, households solve the problem

$$\max_{\{c_{j,n}\}} \left(\sum_{j=1}^{J} \omega c_j^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}},$$
$$s.t. \sum_{j=1}^{J} p_j c_j = I,$$

where p_j is the price of good j, and I the real income of the household in units of the numeraire (labor). Let λ equal the Lagrange multiplier on income in this problem, and define $P \equiv 1/\lambda$ as the price of aggregate consumption in terms of income (the inverse of the shadow value of income) – or, since income is measured in units of labor, it means $\lambda = 1/P$ is a measure of productivity. In Appendix C, we derive the result that:

Proposition 2 *The price index of aggregate consumption in terms of labor in model economy j is:*

$$P_j = \left(\sum_j \omega_j^{\sigma} p_{cj}^{1-\sigma} \right)^{\frac{1}{1-\sigma}}. \tag{11}$$

Thus, the way to measure the impact of rule of law is simply by comparing P across countries, as predicted by the model and the estimated coefficients as discussed below. If P is high (so λ is low) it means a given amount of labor produces correspondingly fewer goods. If P is low (λ is high), it means goods can be produced cheaply i.e. efficiently with a given amount of labor. Since the quantity of labor is constant and equal to the population, this provides us with an index of aggregate output.

Computing P requires values of p_{cj} for each country and industry. However, it turns out that this is relatively straightforward to obtain from the estimates, given one assumption. The assumption is that, in the industry with the lowest value of RS_j, productivity is insensitive to improvements in enforcement and, thus, is constant across countries. Since the numeraire is labor, this is equivalent to the assumption that productivity in industry k is insensitive to improvements in ROL. Let this be industry k.[9] Note that this is a conservative assumption: if the lowest-RS_j industry were sensitive to ROL, then the aggregate impact or variation in ROL would be *greater*. Then, we set the remaining parameters so as to exactly match industry structure in the United States, as described below.

Recall that, for all industries j, equation (7) links relative industry shares to relative prices. Equation (7) can be rearranged so that

$$p_{cj} = \left(\frac{S_{jc}}{S_{kc}} \frac{\omega_k^\sigma}{\omega_j^\sigma} \right)^{\frac{1}{1-\sigma}} p_{ck} \tag{12}$$

Thus, given the assumption that p_{ck} is constant across countries, we can create values of p_{cj} for all countries c and industries j in the data, and thus values of P for all countries, provided we have information on the shares of all industries predicted by the model.

[9] For the narrow measure of $SPEC$ that we use, this is Education. The assumption here is not that productivity in education is constant across countries, see for example Jedwab et al (2023). The assumption ensures that our findings are solely due to the interaction of ROL and RS through economic structure, net of any other effects of ROL that might affect overall productivity. In that sense this is a conservative assumption.

We proceed as follows. First, we will calibrate the model using data from the United States (US). The United States serves as a good benchmark as it has high ROL, is a large economy, and has a relative low share of trade in GDP. For robustness we will also use Canada as a benchmark later instead.

Taking the US shares as given, we can generate p_{cj} and P for the US using equation (12), given values of ω_j and σ. Then for other countries we need to compute the shares predicted by the model. We can generate them using this equation, which tells us what the industry shares would be solely based on the approximate interaction of RS_j and ROL_c in the model:

$$\log S_{c,j} = \delta_j + \delta_c + \hat{\beta}\left(RS_j \times ROL_c\right) \tag{13}$$

which rearranges to equal

$$S_{c,j} = e^{\delta_j} e^{\delta_c} e^{\hat{\beta}(RS_j \times ROL_c)} \tag{14}$$

For each country, the term $\hat{\beta}\left(RS_j \times ROL_c\right)$ is known: $\hat{\beta}$ was estimated earlier, and the values of RS_j and ROL_c for each industry and country respectively are known from the data. We would just need to have values of δ_j.

Finally, the factor δ_c can be chosen so that the shares add up to one in each country, i.e. so that:

$$\delta_c = -\log\left(\sum_j e^{\delta_j} e^{\hat{\beta}(RS_j \times ROL_c)}\right).$$

To obtain values of δ_j, we again rearrange equation (13) to match US shares exactly. In other words, we set δ_j to satisfy

$$\delta_j = \log S_{US,j} - \hat{\beta}\,(RS_j \times ROL_{US}) \tag{15}$$

where $S_{US,j}$ is the share of industry j in the US data and ROL_{US} is the rule of law value for the US. Notice $\delta_c = 0$ in this case, because the US shares already add up to one and we are matching them exactly.

So, to recap, our procedure is to use US data to get δ_j, and then use equation (13) to get the shares predicted by the estimation equation. Then, we use equation (12) to get prices for all final goods in all countries, on the assumption that p_k is the same in all countries, where k is the lowest-SPEC industry. Finally, we use equation (11) to measure the relative price of consumption, and measure output in each country using $\lambda = 1/P$. More precisely, we can measure output in model units for any country for which we have data on ROL_c, and thus compare output in the model economy relative to the US for each country to output relative to the US in the data.

Having calibrated the model, we then measure GDP and compute economic structure in all countries in the world, using values of RS_j and ROL_c for each industry and country. The only factor leading countries and industries to differ in terms of GDP and economic structure is then variation in ROL_c, measured as the negative of the ROL index as before.

Note that this simulation procedure uses the mapping between economic structure and prices implied by the preference structure, and the estimated coefficients. Thus, it captures the influence on aggregates of any mechanism through which rule of law might interact with relationship specificity to change economic structure by lowering the cost of goods. It does not, however, depend on the specifics of the production side of the model.

The first step is to choose the set of industries J. We use as a benchmark the 14 industries in the ILO data, since they have broad country coverage and they cover the entire private economies of these countries. Of course this data source reports employment shares, not value added shares, but the two map into each other under common assumptions in the literature on economic structure and development.[10]

An alternative is to use the disaggregated manufacturing industries in the UNIDO database instead. The problem with that is that it will not map directly into GDP in the data, which includes sectors other than manufacturing. Nonetheless we will repeat the exercise with the UNIDO database for robustness. The advantage of this source is that we can use value added shares instead of labor shares.

Given our choice of J, this exercise requires values for the following parameters: $\hat{\beta}$, $\{\omega_j\}_j$ and σ.

We measure $\hat{\beta}$ using the coefficient from the ILO regressions. We use the median of the estimated coefficients, which is 0.67. Later we will examine the implications of alternative estimates.

We set $\omega_j = 1/J$ for all industries. The reason is that a choice of ω_j is equivalent to a choice of units for good j. However, all that matters for our procedure are the

[10]For example, Ngai and Pissarides (2007) show this is true if there is little variation across sectors in labor shares. Much of the literature on economic structure and development assumes labor is the only productive input, for example Duarte and Restuccia (2010).

shares. It can be readily verified that any arbitrary set of positive values of ω_j that add to unity will generate the same results, because ω_j enters equation (12) in the denominator, but enters equation (11) multiplicatively, so they cancel out when we compute the value of (11).[11]

To estimate σ, we observe that estimates from the literature are unlikely to be suitable. On the one hand, manufacturing data suggest that σ is likely somewhere between 3 and 6 – see Samaniego and Sun (2016). On the other hand, Herrendorf et al (2013) estimate that, using value-added data, $\sigma = 0$. The former estimates use highly disaggregated data for only a subset of the economy, and the latter estimates consider only three sectors – agriculture, manufacturing and services. Our data cover the entire private economy but are more disaggregated than theirs. This is important, as Duarte and Restuccia (2020) find evidence of structural change within the service sector that is consistent with $\sigma > 1$.

Thus, we generate our own estimates. Notice that, provided suitable data on prices and expenditures, we could estimate equation (9) directly – i.e. we could estimate

$$\log S_{cj} = \delta_c + \delta_j + \hat{\eta} \log p_{cj} + \epsilon_{cj} \tag{16}$$

The coefficient $\hat{\eta}$ is then equal to an estimate of $1 - \sigma$.

The International Comparisons Program of the World Bank (ICP) provides such data. While the data do not correspond exactly to our classifications, they are at

[11] To see this, note that combining (11) with (12) we have that $P = \left(\sum_j \omega_j^\sigma \left(\frac{S_{jc}}{S_{kc}} \frac{\omega_k^\sigma}{\omega_j^\sigma} \right) p_{ck}^{1-\sigma} \right)^{\frac{1}{1-\sigma}}$. Then replacing S_{jc} using (14) we get that $P_j = \omega_k^{\frac{\sigma}{1-\sigma}} \left(\sum_j \left(\frac{S_{jc}}{S_{kc}} \right) p_{ck}^{1-\sigma} \right)^{\frac{1}{1-\sigma}}$. In other words, the values of ω_j do not matter when $j \neq k$. However, raising or shrinking ω_k raises or shrinks P_j in all countries proportionately, so it is just a universal scaling factor that does not affect relative GDP.

Table 3: Estimates of epsilon obtained by regressing log industry shares against log price indices from the ICP. All standard errors are clustured by industry.

ICP round	Coefficient	s.d.	Obs	R2	Implied σ	Industry dummies	Income effects
2011	-0.908^{**}	0.380	2,804	0.1353	1.908	no	no
2017	-1.090^{**}	0.393	2,835	0.1491	2.090	no	no
Both	-0.954^{**}	0.355	5,639	0.1333	1.954	no	no
2011	-0.763^{***}	0.142	2,804	0.6633	1.763	yes	no
2017	-0.836^{***}	0.191	2,835	0.6683	1.836	yes	no
Both	-0.748^{***}	0.151	5,639	0.6577	1.748	yes	no
2011	-0.837^{***}	0.102	2,699	0.7368	1.837	yes	yes
2017	-0.861^{***}	0.093	2,669	0.7465	1.861	yes	yes
Both	-0.777^{***}	0.086	5,368	0.7333	1.777	yes	yes

roughly the same level of aggregation, and are thus suitable for our purposes. We pool data for the 2011 and 2017 rounds to obtain a panel, and also condition on income effects by including an interaction between industry dummy variables and log GDP per capita, to allow for separate income effects for each industry. Thus, in practice we estimate:

$$\log S_{cjt} = \delta_c + \delta_j + \delta_t + \hat{\eta} \log p_{cj} + \hat{\gamma}_j \times \log GDP_{ct} \times \delta_j + \epsilon_{cjt} \tag{17}$$

and cluster errors by industry. See Appendix G for further details regarding the ICP data.

Table 3 displays our findings. We look at each ICP round separately and also pool them. We find that, when industry dummies and income effects are left out, we obtain an estimate of about $\sigma = 2$. Industry dummies yield estimates closer to $\sigma = 1.8$, as does allowing for income effects. These methods also yield increasing precision in the form of significantly smaller standard errors. Our preferred estimate is in the ninth row, where both rounds of the ICP data are merged and all control variables are included. Thus, moving forward, we set $\sigma = 1.777$. We note that typical two-standard deviation confidence boundaries correspond to coefficients of $\sigma = 1.605$ and $\sigma = 1.949$: we will use these for robustness.

4.2 Quantitative results

We compute the variable "relative GDP" in the model economy, which is the value of λ relative to the corresponding value in the US. We also compute this variable in the data, which is the of average GDP per capita over the period as reported in the PWT 9.1 relative to the US over the period. Then, we compare log relative GDP in the data to log relative GDP in the model. The correlation is 0.76 and significant at the 1 percent level. This in itself might not be surprising as it is well known that ROL is correlated with GDP levels. The only mechanism through which GDP in the calibration may be affected is through the estimated interaction of ROL and SPEC. Given the sign of the regression coefficient, low ROL will only lower GDP. See Figure 1 for a visual sense of the goodness of fit. Even though we know ROL and GDP are correlated in the data, the match overall between model and data is striking.

What is much more telling is the regression coefficient. This gives a sense of how much the magnitude of variation in the model compares to the magnitude of variation in the data. The regression coefficient is 2.808^{***}, with a standard deviation of 0.186. Moreover, the estimated intercept of -0.191 has a standard deviation of 0.119: it is not statistically significantly different from zero. This implies that, for reasonable parameters, the model is capable of accounting for the magnitude of $1/2.808 = 36$

percent of the empirical variation in relative levels of GDP per capita around the world. This is a substantial amount.

Since this is a key result, one might ask whether it is sensitive to the choice of the baseline country, the United States. To look at this we repeated the exercise using Canada as the baseline country. Canada has a higher ROL value than the US, but is also a more open economy (e.g. exports are about a third of GDP, as opposed to around a tenth for the US). The regression coefficient was 2.711^{***} and the estimated intercept was again not significantly different from zero. This implies that the model is capable of accounting for the magnitude of $1/2.711 = 37$ percent of the empirical variation in relative levels of GDP per capita around the world, when calibrated using Canada as a benchmark.

We also repeat these result using the UNIDO data for manufacturing only. This exercise has downsides, since manufacturing output is only a part of GDP. At the same time, the UNIDO data allow us to use value added shares instead of employment shares. We find a regression coefficient of 3.039^{***}, which implies that the model is capable of accounting for the magnitude of $1/3.039 = 33$ percent of the empirical variation in relative levels of GDP per capita around the world. Again, the results are in the same ballpark.

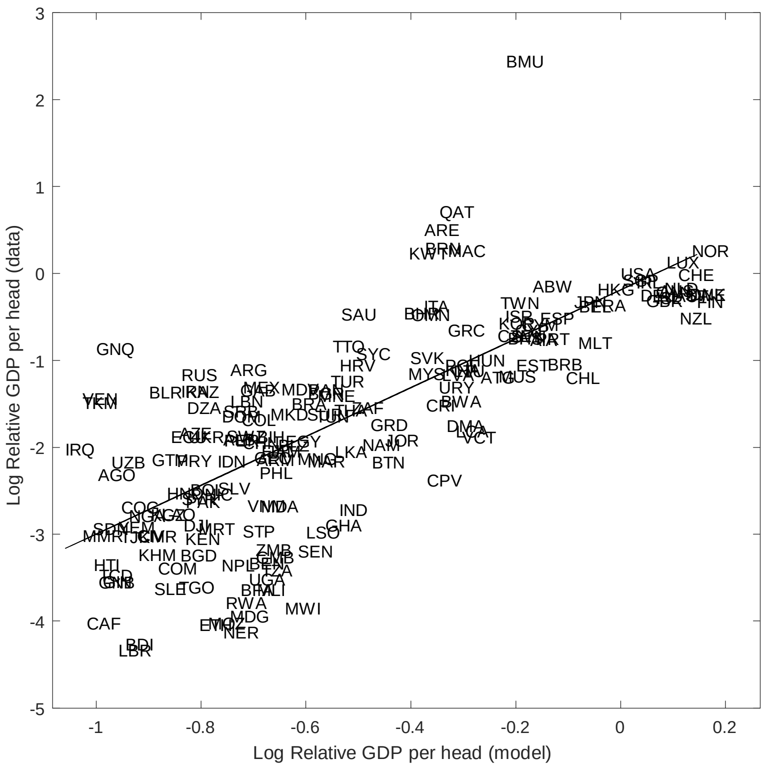

Figure 1 – Comparison of model GDP per capita to GDP per capita in the PWT 9.1. Values are the natural logs of the GDP per capita values relative to the US – so the US value is zero on both axes.

Naturally the extent of variation in model-generated GDP per capita is sensitive to the calibrated value of σ. We repeat this exercise using other values of σ in our range. When we set $\sigma = 1.605$ – the lowest value within the significance boundaries for our preferred estimate – we find a point estimate of 2.1863, so the model accounts for 46 percent of the empirical variation in GDP per head. On the other hand, if we set $\sigma = 1.949$ (the highest value within standard significance bounds), we obtain a point estimate of 3.4294 so the model accounts for 29 percent of the empirical variation in GDP per head. We conclude that the calibrated model supports an interaction of RS and ROL large enough to be of macroeconomic significance.

As well as σ, these findings are sensitive to the value of $\hat{\beta}$, the estimate of the interaction coefficient between rule of law and relationship specificity. To examine this sensitivity, we generate results with the somewhat lower value of $\hat{\beta} = 0.59$, which is the 20th percentile value among our estimated coefficients. In this case, our point estimate is 3.1509, so the model accounts for 32 percent of variation in log GDP per capita. Even using the smallest estimate of $\hat{\beta} = 0.44$, which we obtain when looking at the raw correlation between the interaction variable and industry structure in column (1) of Table 6, we find a coefficient of 4.1356, which implies that the model accounts for 24 percent of variation in GDP per head. This is smaller but still highly economically significant.

Another possible exercise is to use the model prices weighted by the actual shares in the data to generate alternative synthetic values of P_j. When we do this we find a regression coefficient of the model GDP on the data GDP of 1.259***, which implies that the model is capable of accounting for the magnitude of $1/1.259 = 79$ percent of the empirical variation in relative levels of GDP per capita around the world. At the same time, this exercise results in an estimate of the constant that is significantly different from zero, which contradicts this interpretation. We interpret this finding as telling us that there are determinants of economic structure other than the interaction studied in this paper that are important, and that ROL accounts for a significant proportion of the residual. This is not surprising since the literature on economic structure and development has suggested many different mechanisms relating economic structure to development e.g. financial development (Rajan and Zingales (1998), Ilyina and Samaniego (2011, 2012)), productivity growth (Ngai and Pissarides (2007), Samaniego and Sun (2020)), etc. The bottom line is that there is an economically significant link between the interaction of relationship specificity and macroeconomic outcomes through changes in economic structure that deserves further study.

Since ROL impacts real GDP via both changes in productivity and changes in structure, we regress the industry shares in the data on the model-generated shares. See Table 4. We find that the shares are highly correlated in all specifications, provided that we exclude agriculture. This is again consistent with the share of agriculture being related to income effects rather than price effects of the kind in this study.

Since our mechanism revolves around the impact of ROL on prices, it makes sense to ask whether there is an empirical link between prices in our model and prices in the data. Before doing so, it is worth noting that any comparison of model prices with prices in the data has to take into account that in general it is well known since Harrod (1933), Balassa (1964) and Samuelson (1964) that rich countries tend to have higher prices than poor countries. Since ROL is highly correlated with real GDP per capita, this suggests that in the background there are other factors that lead prices to be higher in rich, high-ROL countries (see Alessandria and Kaboski (2011) for a

Table 4: Comparison of shares between model and the ILO data.

Specification	(1)	(2)	(3)	(4)	(5)	(6)
Coefficient	0.085*	0.576***	.085*	0.574***	.0851*	.574***
S.e.	0.053	0.023	0.051	0.022	0.053	0.022
Coefficient on ROL	–	–	0.000	0.013***		–
S.e.	–	–	0.002	0.001		–
Country dummies?	No	No	No	No	Yes	Yes
Excluding Agriculture?	No	Yes	No	Yes	No	Yes
Obs	2310	2145	2310	2145	2310	2145
R^2	0.01	0.23	0.01	0.29	0.01	0.74

Notes: In specification (1) there are country dummies. In specifications (2) and (4) we condition instead on ROL. In specification (3) there are also industry dummies.

survey). As a result, when we compare prices in the model with prices in the data we need to condition either on ROL or on country fixed effects.[12]

First of all, we find that the coefficient of variation in the ICP data, once it is aggregated to match the ILO sectors, is 0.55. In contrast, the coefficient of variation for the same sectors in the model-generated data is 0.47, quite close. In this sense, the variation generated by the model is reasonable.

We find that, when we regress the price indices from the data on the model-generated indices with country dummies, there is a positive and significant coefficient with a P-value below 1 percent. We do this separately for 2011 and 2017 ICP data. A similar result obtains when we condition on ROL instead of country dummies. As expected the coefficient on ROL in this experiment is positive, consistent with the prior literature. In any case, these results indicate that the model accounts for

[12] In addition, since the ICP industries do not correspond exactly to the ILO sectors, we need to omit some sectors and aggreage some industries using value added shares.

Table 5: Comparison of prices between model and the ICP data. The dependent variable is prices in the data in 2011 and 2017 separately.

ICP Year	2011			2017		
Specification	(1)	(2)	(3)	(1)	(2)	(3)
Coefficient	0.05***	0.08***	0.28***	0.04***	0.06***	0.21***
S.e.	0.01	0.01	0.03	0.01	0.01	0.02
Coefficient on ROL	–	39.1***	53.8***	–	37.3***	48.0***
S.e.	–	1.41	2.15	–	1.30	2.00
Country dummies?	Yes	No	No	Yes	No	No
Industry dummies?	No	No	Yes	No	No	Yes
Obs	1183	1183	1183	1162	1162	1162
R^2	0.72	0.45	0.52	0.74	0.48	0.54

Notes: In specification (1) there are country dummies. In specifications (2) and (4) we condition instead on ROL. In specification (3) there are also industry dummies.

variation in relative prices aside from the effects identified in Balassa (1964) and related references. See Table 5.

How should we interpret the quantitative exercise? The exercise builds solely on the preference structure of the model and the estimated interaction coefficients. Thus, the key assumption is that differences in economic structure predicted by the estimated coefficients are due to price effects stemming from the relative efficiency with which relationship specific intermediates are produced. The precise details of the production side of the economy are not important for this finding. The estimates condition on income effects in a variety of ways, as well as conditioning on other determinants of economic structure. As a result, we believe our quantitative findings are robust.

The literature on structural transformation in general equilibrium links economic structure to either price effects (as here, generally linked to productivity) or income effects. Thus, the alternative explanation for our findings would be for ROL to affect economic structure (and thus P) through income effects, in which case our mapping between the interaction coefficients and prices would not be accurate. However, it is not clear why income effects would disproportionately affect highly relationship-specific industries – and our estimates condition in multiple ways on income effects. Indeed, relationship specificity is arguably due to technological characteristics of the goods in question (see Nunn (2007)), whereas income effects have to do with preferences. From a theoretical perspective, assuming a coincidence between the parameters of technology and preferences is problematic (see Ngai and Pissarides (2007)) but, ultimately, the question is whether *empirically* there is a link between income effects and relationship specificity – e.g. Caron et al (2020) find a link between the demand for skill-intensive goods and income. In our empirical section we accounted for income effects in various ways, including by estimating income elasticities using both microeconomic and macroeconomic data, finding that the elasticities were not correlated with relationship specificity and that the significance and magnitude of our coefficients were not sensitive to whether or not we accounted for income elasticity in one way or another. As a result, we are confident that our quantitative results indicate that the interaction of relationship specificity with the institutional capacity to support contract enforcement has highly significant macroeconomic impact.

5 Concluding Remarks

We argue that, if weak rule of law particularly impacts productivity in the production of relationship-specific intermediates, then this can be detected through its impact on economic structure by estimating a simple differences in differences regression specification. We find significant empirical support for an interaction of relationship specificity and rule of low via changes in economic structure that is robust to several controls, including controlling for income effects in a variety of ways. Finally, we use a standard preference framework to assess the aggregate impact of the effects we identify, finding the aggregate impact to be economically significant.

Ranasinghe and Restuccia (2018) find that the property rights motive for wanting rule of law interacts with financial frictions to roughly double the direct impact of rule of law. In this context, the efficiency mechanism of ROL could also interact with financial frictions. If intermediate good producers are financially constrained, they might find themselves unable to make relationship specific investments for financial reasons, possibly exacerbating the mechanisms in the paper.

An environment with weak ROL, one possible outcome that we do not study is that idiosyncratic risk might be greater, stemming from the potential of one's partners to renege. We leave the study of this possible "risk channel" of rule of law for future work.

Finally, the preference framework abstracts from income effects. This is a natural assumption given that income effects are not the subject of the paper: our focus is on understanding the influence of institutions on economic structure through price effects, and that the empirical findings regarding the interaction of ROL with relationship specificity are robust to accounting for income effects. Of course, given that our interaction of interest has an impact on aggregate income, there could be secondary income effects, and it would be interesting to extend the framework to allow for income effects. Quantitative frameworks with income effects are considerably more complex,[13] and augmenting them with industry-institution intereactions would be a significant but worthwhile challenge.

6 References

Acemoglu, Daron, Pol Antràs and Elhanan Helpman. 2007. Contracts and Technology Adoption. American Economic Review 97(3), 916-943.

Acemoglu, Daron and Simon Johnson. 2005. Unbundling Institutions. Journal of Political Economy 113(5), 949-995.

Alessandria, George and Joseph P. Kaboski. 2011. Pricing-to-Market and the Failure of Absolute PPP. American Economic Journal – Macroeconomics 3(1), 91-127.

Balassa, Bela. 1964. The Purchasing-Power Parity Doctrine: A Reappraisal. Journal of Political Economy, 72(6): 584–96.

Boehm, Johannes. 2022. The Impact of Contract Enforcement Costs on Value Chains and Aggregate Productivity. Review of Economics and Statistics 104 (1), 34–50.

Boppart, Timo. 2014. Structural Change and the Kaldor Facts in a Growth

[13] See Boppart (2014) and Comin et al (2021), the focus of which is to explore the influence of income effects on economic structure in frameworks with no institutional interactions.

Model With Relative Price Effects and Non-Gorman Preferences. Econometrica 82, 2167-2196.

Ciccone, Antonio and Elias Papaioannou. 2016. Estimating Cross-Industry Cross-Country Interaction Models Using Benchmark Industry Characteristics. NBER Working Paper 22368.

Comin, Diego, Danial Lashkari and Martí Mestieri. 2021. Structural Change with Long-run Income and Price Effects. Econometrica 89(1), 311-374.

Caron, Justin, Thibault Fally & James Markusen. 2020. Per Capita Income and the Demand for Skills. Journal of International Economics

Volume 123, 103306

Dell'Ariccia, Giovanni, Enrica Detragiache and Raghuram Rajan. 2008. The Real Effect of Banking Crises. Journal of Financial Intermediation 17, 89-112.

Duarte, Margarida and Diego Restuccia. 2010. The Role of the Structural Transformation in Aggregate Productivity. Quarterly Journal of Economics 125(1), 129-173.

Duarte, Margarida and Diego Restuccia. 2020. Relative Prices and Sectoral Productivity. Journal of the European Economic Association, forthcoming.

Feenstra, Robert C., Robert Inklaar and Marcel P. Timmer. 2015. The Next Generation of the Penn World Table. American Economic Review, 105(10), 3150-3182, available for download at www.ggdc.net/pwt

Geary, Roy C. 1950. A Note on 'A Constant-Utility Index of the Cost of Living'. Review of Economic Studies 18 (2), 65–66.

Goldberg, Pinelopi K., Pavcnik, Nina, 2004. Trade, inequality, and poverty:

what do we know? Evidence from recent trade liberalization episodes in developing countries. NBER Working Papers 10593. National Bureau of Economic Research, Inc.

Gollin, Douglas, Stephen Parente and Richard Rogerson. 2002. The Role of Agriculture in Development. American Economic Review 92(2), 160-164.

Harrod, Roy F. 1933. International Economics. London: James Nisbet & Co.

Herrendorf, Berthold, Richard Rogerson and Ákos Valentinyi. 2013. Two Perspectives on Preferences and Structural Transformation. American Economic Review 103(7), 2752-89.

Ilyina, Anna O. and Roberto M. Samaniego. 2011. Technology and Financial Development. Journal of Money, Credit and Banking 43(5), 899-921.

Ilyina, Anna O. and Roberto M. Samaniego. 2012. Structural Change and Financing Constraints. Journal of Monetary Economics 59(2), 166-179.q

Kongsamut, Piyabha, Sergio Rebelo and Danyang Xie. 2001. Beyond Balanced Growth. Review of Economic Studies 68(4), 869-882.

Levchenko, Andrei A. 2007. Institutional Quality and International Trade. Review of Economic Studies 74(3), 791–819.

Mulligan, Casey B., and Xavier Sala-i-Martin. 1997. A Labor Income-Based Measure of the Value of Human Capital: An Application to the States of the United States. Japan and the World Economy, 9, 159–91.

Ngai, L. Rachel and Christopher A. Pissarides. 2007. Structural Change in a Multisector Model of Growth. American Economic Review 97(1), 429-443.

Nunn, Nathan. 2005. Domestic institutions, international trade and economic development. Dissertation: University of Toronto.

Nunn, Nathan. 2007. Relationship-Specificity, Incomplete Contracts and the Pattern of Trade. Quarterly Journal of Economics 122 (2) : 569-600.

Papke, Leslie E. and Jeffrey M. Wooldridge. 1996. Econometric methods for fractional response variables with an application to 401(k) plan participation rates. Journal of Applied Econometrics 11(6), 619-632.

Rajan, Raghuram G. and Luigi Zingales. 1998. Financial Dependence and Growth. American Economic Review 88(3), 559-586.

Ranasinghe, Ashantha and Diego Restuccia. 2018. Financial frictions and the rule of law. Journal of Development Economics 134(C), 248-271.

Ranasinghe, Ashantha. 2017. Property Rights, Extortion and the Misallocation of Talent. European Economic Review 98, 86-110.

Rauch, James E. 1999. Networks versus markets in international trade." Journal of International Economics, 48(1), 7-35.

Restuccia, Diego, Dennis Tao Yang and Xiaodong Zhu. 2008. Agriculture and aggregate productivity: A quantitative cross-country analysis. Journal of Monetary Economics 55 (2), 234-250.

Samaniego, Roberto M. 2013. Knowledge Spillovers and Intellectual Property Rights. International Journal of Industrial Organization 31(1), 50-63.

Samaniego, Roberto M. and Juliana Sun. 2015. Technology and Contractions: Evidence from Manufacturing. European Economic Review 79, 172-195.

Samaniego, Roberto M. and Juliana Sun. 2020. The Relative Price of Capital and Economic Structure. Review of Economic Dynamics 37, 127-155.

Samuelson, Paul A. 1964. Theoretical Notes on Trade Problems. Review of Economics and Statistics, 46(2): 145–54.

Shleifer, Andrei. and Robert W. Vishny 1998. The Grabbing Hand: Government Pathologies and Their Cures. Harvard University Press.

Stone, Richard. 1954. Linear Expenditure Systems and Demand Analysis: An Application to the Pattern of British Demand. Economic Journal 64 (255), 511–527.

Święcki, Tomasz, 2017. Determinants of structural change. Review of Economic Dynamics 24, 95–131.

Appendix to "Rule of Law, Economic Structure and Development"
by Roberto Samaniego and Juliana Sun

A Production Structure: An Example

In what follows we display a production structure that maps into prices such that $\beta > 0$ in the main text. We build on the model of Nunn (2005). An alternative would be to build on the framework of Acemoglu, Antràs and Helpman (2007), who develop a rich contracting environment and show that weak enforcement lowers the productivity of sectors to the extent that they rely on inputs where contracting is difficult.[14] The key result in both of these cases is that ROL particularly raises productivity in the production of final goods that disproportionately use relationship-specific inputs.

Making an intermediate requires one unit of labor supplied by an *agent*, who is contracted based on a payment b_r, made by the principal after the intermediate is sold.

Production of the RS intermediate is subject to a hold-up problem. The agent decides which proportion of his labor q will be devoted to standard tasks or to customized tasks. RS intermediates may not be usable, with a probability $f(q)$ that depends on q. More customization increases the probability that the intermediate will be usable. However, the supply of customized labor can only be verified imperfectly,

[14] This would require having two intermediates, one where all tasks are contractible, and another where only some share μ of tasks are contractible.

as described below. The price of the RS intermediate is p_r. The principal and the agent bargain over a payment b_r to be made after the good is produced and sold. Similarly, the price of the NRS intermediate is p_n, and the principal and the agent bargain over a payment b_n.

A.1 RS intermediate

There is a court that provides costless verification and enforcement services. The court may see if a usable good was delivered, but cannot see if it was not,[15] thus payments can only be enforced if the good was usable. This rules out payments if the good was not delivered.

If the good is usable, the court can distinguish this with probability γ and fully enforce the contract. Parameter γ is an indicator of Rule of Law. We assume that disputes are costless, so that there is no direct dissipation of resources in the economy depending on the Rule of Law. Alternatively, with probability $1-\gamma$ the court fails to fully enforce the contract, in which case the agent can only recover a share $1-g(q)$ of the payment. We assume $g' > 0$: higher customization makes identification and/or enforcement more difficult.

Furthermore, suppose that $g(0) = 0$ and $g(1) = 1$. Then, without loss of generality we assume that $g(q) = q$.[16]

[15] Alternatively the principal has property rights over the output. More generally the presumption is that property rights are determined by possession in this environment of weak ROL.

[16] This is without loss of generality because for more general $g(q)$ we can define $\tilde{q} = g(q)$ and define $\tilde{f}(\tilde{q})$ as:

$$\tilde{f}(\tilde{q}) = f\left(g^{-1}(\tilde{q})\right).$$

We assume that the likelihood of a successful fit $f(q) \in [0,1]$ has these properties:

Condition 3 $f' > 0$, $f'' \leq 0$ *and* $\lim_{q \to 1} f'(q) = 0$.

With probability $f(q)\gamma$ the match is good and the contract is enforced, in which case the agent receives a payment b_c. With probability $f(q)(1-\gamma)$ the match is good but the court finds in favor of the principal and the agent recovers $(1-q)b_r$. With probability $1 - f(q)$ the match is bad and no payment is made.

The expected payoff of the agent is:

$$\begin{aligned}\pi_A &= f(q)(\gamma b_r + (1-\gamma) b_r (1-q)) \\ &= f(q)(1 - q + \gamma q) b_r\end{aligned}$$

The expected payoff of the principal is:

$$\begin{aligned}\pi_P &= f(q)(p_r - \gamma b_r - (1-\gamma) b_r (1-q)) \\ &= f(q)(p_r - b_r + b_r q - \gamma b_r q)\end{aligned}$$

The agent chooses q so as to maximize π_A, so the optimal choice of q is defined implicitly by:

$$f'(q)(\gamma + (1-\gamma)(1-q)) = f(q)(1-\gamma).$$

Let $q(\gamma)$ be the solution. Under Condition 3, it it is straightforward to show that:

$$\frac{\partial q(\gamma)}{\partial \gamma} = \frac{f(q) + f'(q) q}{2f'(q)(1-\gamma) - f''(q)(\gamma + (1-\gamma)(1-q))} > 0$$

Assume Nash bargaining over b_r. Define the surplus function as follows, where μ is the bargaining power of the principal:

$$S(b_r) = [\pi_P(b_r)]^{\mu} [\pi_A(b_r)]^{1-\mu}$$

This is maximized by

$$[\pi'_P(b_r)] \mu [\pi_P(b_r)]^{\mu-1} [\pi_A(b_r)]^{1-\mu} + (1-\mu) [\pi_P(b_r)]^{\mu} [\pi_A(b_r)]^{-\mu} [\pi'_A(b_r)] = 0$$

which delivers that:

$$b_r = p_r \frac{1-\mu}{1-(1-\gamma) q(\gamma)}.$$

A.2 NRS intermediate

The NRS intermediate production is the same except that customization is not important, so that $f(q) = 1$ and $g(q) = 0$ for all q. The agent gets b_n and the principal gets $p_n - b_n$. Nash bargaining now implies:

$$\begin{aligned} S(b_n) &= [p_n - b_n]^{\mu} [b_n]^{1-\mu} \\ &\Rightarrow b_n = p_n (1-\mu). \end{aligned}$$

A.3 Occupational choice

Now since agents can enter both markets it must be that

$$\pi_A = f\left(q\left(\gamma\right)\right)\left(1-\left(1-\gamma\right)q\left(\gamma\right)\right)p_r\frac{1-\mu}{1-\left(1-\gamma\right)q\left(\gamma\right)} = p_n\left(1-\mu\right)$$

which implies that

$$\frac{p_r}{p_n} = f\left(q\left(\gamma\right)\right)^{-1}. \tag{18}$$

Since the production technology for final goods is linear, the price of one unit of final good j will equal its cost, so that

$$p_{cj} = \left(1-a_j\right)p_n + p_r a_j.$$

Define the wage w_t as the expected return to labor – which will be convenient for normalization. This implies that $w_t = \pi_A = p_n\left(1-\mu\right)$. Setting $w_t = 1$ implies that:

$$p_{cj} = \left(1-\mu\right)^{-1}\left[\left(1-a_j\right) + f\left(q\left(\gamma\right)\right)^{-1}a_j\right]. \tag{19}$$

Using (18), equation (7) can be written in terms of the determinants of equilibrium prices:

$$\frac{S_j}{S_k} = \left(\frac{\omega_j}{\omega_k}\right)^{\sigma}\left(\left[1-a_j+f\left(q\left(\gamma\right)\right)^{-1}a_j\right] \div \left[1-a_k+f\left(q\left(\gamma\right)\right)^{-1}a_k\right]\right)^{\sigma-1}. \tag{20}$$

Finally, taking logs,

$$\begin{aligned}\log S_j - \log S_k &= \sigma \log \omega_j - \sigma \log \omega_k \\ &\quad + (\sigma - 1) \log \left[1 - a_j + f(q(\gamma))^{-1} a_j\right] \\ &\quad - (\sigma - 1) \log \left[1 - a_k + f(q(\gamma))^{-1} a_k\right].\end{aligned} \tag{21}$$

Suppose there are many countries which vary in terms of rule of law. Let γ_c be the rule of law in country c. Let S_{jc} be the value of S_j in country c.

Proposition 2 *Consider an interior value of Rule of Law γ^* and relationship specificity a_j^*. Industry structure in the model economy is given by*

$$\log S_{jc} = \delta_c + \delta_j + \beta \times a_j \times \gamma_c + \varepsilon_{jc}.$$

where β is a constant with the same sign as $(\sigma - 1)$, δ_c is a constant that varies by country, δ_j is a constant that varies by industry and $\varepsilon_{jc} \to 0$ as $[a_j, \gamma_c] \to [a^*, \gamma^*]$.

Proof. We show that $\frac{d^2 p_{cj}(a_j,\gamma)}{da_j d\gamma} < 0$ in this model, so Proposition 1 applies. To see this, consider the case were $\Gamma(\gamma, a) = -\log\left[1 - a_j + f(q(\gamma))^{-1} a_j\right]$. We simply need to compute the sign of $\Gamma_{\gamma a}(\gamma^*, a^*)$: ■

$$\Gamma_a(a, \gamma) = -\frac{f(q(\gamma))^{-1} - 1}{\left[1 - a_j + f(q(\gamma))^{-1} a_j\right]}$$

Using the quotient rule,

$$
\begin{aligned}
\Gamma_{a\gamma}(a,\gamma) &= -\frac{q_\gamma f'(q(\gamma))^{-1}\left[1-a_j+f(q(\gamma))^{-1}a_j\right]-\left[f(q(\gamma))^{-1}-1\right]q_\gamma\left[f'(q(\gamma))^{-1}a_j\right]}{\left[1-a_j+f(q(\gamma))^{-1}a_j\right]^2} \\
&= \frac{-q_\gamma f'(q(\gamma))^{-1}}{\left[1-a_j+f(q(\gamma))^{-1}a_j\right]^2}<0.
\end{aligned}
$$

B Proof of Proposition 1

Let $\Gamma(\gamma_c,a_j)\equiv\log p_{cj}(\gamma_c,a_j)$. A second order Taylor approximation around interior points (γ^*,a^*) implies that

$$
\begin{aligned}
\Gamma(\gamma,a) &= \Gamma(\gamma^*,a^*)+\Gamma_a(\gamma^*,a^*)(a-a^*)+\Gamma_\gamma(\gamma^*,a^*)(\gamma-\gamma^*) \\
&\quad+\frac{1}{2}\Gamma_{aa}(\gamma^*,a^*)(a-a^*)^2+\frac{1}{2}\Gamma_{\gamma\gamma}(\gamma^*,a^*)(\gamma-\gamma^*)^2 \\
&\quad+\Gamma_{\gamma a}(\gamma^*,a^*)(a-a^*)(\gamma-\gamma^*)+\tilde{o}(\|[\gamma,a]-[\gamma^*,a^*]\|)
\end{aligned}
$$

where Γ_x equals the derivative of Γ with respect to variable x and Γ_{xy} equals the derivative of Γ with respect to variables x and y. Since a_j depends on the industry and γ_c depends on the country, all the terms in this equation are country or industry-specific constants, except for the cross-derivative term and the small-o term. In addition, expanding $\Gamma_{\gamma a}(\gamma^*,a^*)(a-a^*)(\gamma-\gamma^*)$, the terms $\Gamma_{\gamma a}(\gamma^*,a^*)a^*\gamma_c$, $\Gamma_{\gamma a}(\gamma^*,a^*)a_j\gamma^*$ and $\Gamma_{\gamma a}(\gamma^*,a^*)a^*\gamma^*$ are also industry or country-specific constants (or just constants). As a result we have that

$$
\log S_{jc}=\delta_c+\delta_j-(\sigma-1)\Gamma_{\gamma a}(\gamma^*,a^*)a_j\gamma_c+(\sigma-1)\tilde{o}(\|[a,\gamma]-[a^*,\gamma^*]\|) \tag{22}
$$

where δ_c and δ_j absorb all the country- and industry-specific constants respectively.

Defining $\beta = -(\sigma-1)\,\Gamma_{a\gamma}(a,\gamma)$ and $\varepsilon_{jc} = o(\|[\gamma_c, a_j] - [\gamma^*, a^*]\|) = (\sigma-1)\,\tilde{o}(\|[\gamma_c, a_j] - [\gamma^*, a^*]\|)$, the result of Proposition 1 follows since

$$
\begin{aligned}
\Gamma_a(\gamma_c, a_j) &= \frac{p_r(\gamma_c)}{a_j p_r(\gamma) + (1-a_j)\,p_n} \\
\Gamma_{a\gamma}(\gamma_c, a_j) &= \frac{p_r'(\gamma_c)\left[a_j p_r(\gamma) + (1-a_j)\,p_n\right] - a_j p_r'(\gamma)\,p_r(\gamma_c)}{\left[a_j p_r(\gamma) + (1-a_j)\,p_n\right]^2} \\
&= p_r'(\gamma_c)\,\frac{(1-a_j)\,p_n}{\left[a_j p_r(\gamma) + (1-a_j)\,p_n\right]^2} < 0.
\end{aligned}
$$

C Derivation of P

The household's composition problem has the following first order condition:

$$
\omega_j c_{j,n}^{\frac{\sigma-1}{\sigma}-1}\left(\sum_{j=1}^{J}\omega c_{j,n}^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}-1} = \lambda p_{cj}
$$

or

$$
\omega_j c_{j,n}^{\frac{\sigma-1}{\sigma}-1}\mathbf{c}^{\frac{1}{\sigma}} = \lambda p_{cj}
$$

After a few more derivations, we get

$$
\begin{aligned}
\omega_j^{1-\sigma} c_{j,n}^{\frac{\sigma-1}{\sigma}}\mathbf{c}^{\frac{1-\sigma}{\sigma}} &= \lambda^{1-\sigma} p_{cj}^{1-\sigma} \\
\omega_j c_{j,n}^{\frac{\sigma-1}{\sigma}}\mathbf{c}^{\frac{1-\sigma}{\sigma}} &= \omega_j^{\sigma}\lambda^{1-\sigma} p_{cj}^{1-\sigma} \\
\mathbf{c}^{\frac{1-\sigma}{\sigma}}\sum_{j=1}^{J}\omega_j c_{j,n}^{\frac{\sigma-1}{\sigma}} &= \lambda^{1-\sigma}\sum_{j=1}^{J}\omega_j^{\sigma} p_{cj}^{1-\sigma}
\end{aligned}
$$

Now

$$\begin{aligned}
\sum_{j=1}^{J} \omega_j c_{j,n}^{\frac{\sigma-1}{\sigma}} dj &= \mathbf{c}^{\frac{\sigma-1}{\sigma}}, \\
\mathbf{c}^{\frac{1-\sigma}{\sigma}} \mathbf{c}^{\frac{\sigma-1}{\sigma}} &= \lambda^{1-\sigma} \sum_{j=1}^{J} \omega_j^{\sigma} p_{cj}^{1-\sigma} dj \\
&\Rightarrow 1/\lambda^{1-\sigma} = \sum_{j=1}^{J} \omega_j^{\sigma} p_{cj}^{1-\sigma} dj \\
P &\equiv 1/\lambda = \left(\sum_{j=1}^{J} \omega_j^{\sigma} p_{cj}^{1-\sigma} dj \right)^{\frac{1}{1-\sigma}}
\end{aligned}$$

so

$$\begin{aligned}
\mathbf{c}^{\frac{1-\sigma}{\sigma}} \mathbf{c}^{\frac{\sigma-1}{\sigma}} &= \lambda^{1-\sigma} \sum_{j=1}^{J} \omega_j^{\sigma} p_{cj}^{1-\sigma} dj \\
&\Rightarrow 1/\lambda^{1-\sigma} = \sum_{j=1}^{J} \omega_j^{\sigma} p_{cj}^{1-\sigma} dj \\
P &\equiv 1/\lambda = \left(\sum_{j=1}^{J} \omega_j^{\sigma} p_{cj}^{1-\sigma} dj \right)^{\frac{1}{1-\sigma}}.
\end{aligned}$$

D Robustness of the empirical results

D.0.1 Robustness: manufacturing industries

We repeat the estimation using disaggregated manufacturing data from UNIDO. This is important for several reasons. First, the UNIDO data contain several measures of economic structure – employment, value added and gross output. Thus, they will

Table 6: Estimates of the interaction between rule of law at the country level and relationship specificity at the industry level.

	UNIDO Manuf. industries		
	log emp share	log VA share	log output share
Coefficient	0.439***	0.731***	0.596***
s.d.	(0.143)	(0.161)	(0.158)
Control on j, c dummies	yes	yes	yes

serve to show that the results do not depend on how we define structure. Second, they will demonstrate that the results continue to hold at a greater level of disaggregation. Third, the UNIDO data contain information on wages. We can use this information, along with information about schooling in different countries, to condition on human capital differences across countries. Fourth, we can control for capital intensity which, together with controlling for human capital differences, controls for key determinants of international trade, as in Nunn (2007), and possible determinants of productivity differences also. Controlling for trade determinants, as well as the fact that our baseline results using ILO data are mainly for non-tradeables, gives us confidence that our findings are not simply the result of changes in trade structure, and that we are conditioning on other potential technological determinants of productivity differences.

First, we find that estimates without these additional controls are positive. This is true whether we use shares of employment, shares of industry value added or output shares: the coefficient β is positive and significant at 1% level too. The median coefficient is about 0.60, not very different from that obtained from the ILO regression. See Table 6.

The structure of the manufacturing sector may be more affected by factors of international trade than the overall economy which is largely comprised of non-tradeables. As a result, we include some additional controls based on the link between endowments and trade emphasized in the trade literature. This will assess whether the results are driven by trade factors rather than a direct interaction between ROL and relationship specificity.

We now include two additional interaction variables. One is the human capital intensity of industries times the country stock of human capital. We measure industry human capital intensity using the average wage bill of each industry divided by the number of workers as reported in the UNIDO data, following Mulligan and Sala-i-Martin (1997), and as measured in Ilyina and Samaniego (2011). The country capital stock is the variable *schooling* in the PWT 9.1. The other control is the interaction of the country capital-labor ratio and the capital intensity of industry. We measure industry capital intensity using one minus the labor share, computed in Ilyina and Samaniego (2011) as the total wages and salaries divided by value added in as reported in UNIDO in current USD.[17] We measure country capital-labor ratio using the PWT 9.1. The results are robust to the inclusion of these additional control variables – see Table 7. In this case the median coefficient is about 0.59. This also means our results are robust to accounting for factors the trade literature identifies as being key determinants of comparative advantage. With these additional controls in manufacturing data that are more likely to be affected by trade considerations than most of the ILO industries, we nonetheless find coefficients of similar magnitude and

[17] Note that the correlation between $RS_j$and human capital intensity is -0.117, and the correlation between RS_j and capital intensity is -0.338, neither of which is statistically significant at conventional levels. This indicates that an interaction of institutions with either of these variables is not responsible for our findings.

Table 7: Estimates of the interaction between rule of law at the country level and relationship specificity at the industry level. Results condition on interactions of industry human capital intensity with the country stock of human capital, and on interactions of industry capital intensity with the country capital-labor ratio.

	UNIDO Manuf. industries		
	log emp share	log VA share	log output share
Coefficient	0.491***	0.680***	0.586***
s.d.	(0.144)	(0.162)	(0.163)
Control on j, c dummies	yes	yes	yes
Control on K/L, SCHOOL interactions	yes	yes	yes

Table 8: Estimates of the interaction between rule of law at the country level and relationship specificity at the industry level. Results condition on interactions of industry human capital intensity with the country income per capita, and on interactions of industry human capital intensity with the country rule of law.

	UNIDO Manuf. industries		
	log emp share	log VA share	log output share
Coefficient	0.693***	0.920***	0.782***
s.d.	(0.138)	(0.163)	(0.158)
Control on j, c dummies	yes	yes	yes
Control on HC interactions	yes	yes	yes

significance.

One might notice that industries that use relationship-specific industries are also fairly human-capital intensive, suggesting it might be good to interact human capital intensity with rule of law or with GDP per capita to see whether it is human capital intensity that is really behind our results. In Table 8, we condition on interactions of human capital intensity with rule of law and average GDP per capita over the period. We find that the results remain robust once more.

D.0.2 Robustness: contracting institutions

Returning to the ILO data for broad sectors, so far we have measured contracting institutions using measures of ROL. This is consistent with Acemoglu and Johnson (2005), who argue that rule of law is more fundamental for contract enforcement than more refined indicators of contracting institutions, since in the absence of ROL contracting institutions cannot function. Still, this literature suggests checking whether our results are robust to conditioning on direct measures of contracting institutions. To distinguish the effects of these two institutions, we condition on an interaction of more narrow measures of contract institutions and our industry specificity indicator. We measure the quality of contracting institution with two measures: the cost of enforcing contract as the share of the claim (Enf), which is averaged over $2004-2019$ from World Bank Doing Business (2019), and legal formalism ($Legf$) which measures formality in legal procedures for collecting on a bounced check, drawn from Djankov et al. (2003).[18] Results are reported in Tables 9 and 10. Two findings stand out. First, the coefficient on the interaction between ROL and RS_j remains significant at the 1 percent level in both cases. Second, the interaction of narrow contracting measures with RS_j is significant only in the absence of industry and country dummies, indicating it does not appear to have a significant independent effect, which is consistent with the results of Acemoglu and Johnson (2005).

[18] Note that both of these are measures of institutional weakness, so we expect them to carry negative coefficients.

Table 9: Estimates of the interaction between rule of law at the country level and relationship specificity at the industry level control on contract enforcement

	ILO sectors (log emp shares)			
	Correl	Baseline	Income. eff.	No agric.
$ROL_c \times RS_j$	0.548***	1.172***	0.859***	0.858***
s.d.	(0.0599)	(0.122)	(0.0937)	(0.0934)
$Enf_c \times RS_j$	−0.00345	−0.00362	0.00347	0.00341
s.d.	(0.00436)	(0.00973)	(0.00580)	(0.00583)
Control on j, c dummies	no	yes	yes	yes

Table 10: Estimates of the interaction between rule of law at the country level and relationship specificity at the industry level control on legal formalism

	ILO sectors (log emp shares)			
	Correl	Baseline	Income. eff.	No agric.
$ROL_c \times RS_j$	0.508***	0.999***	0.635***	0.643***
s.d.	(0.0503)	(0.104)	(0.0870)	(0.0893)
$Legf_c \times RS_j$	0.167***	0.0868	0.0245	0.0530
s.d.	(0.0359)	(0.0910)	(0.0768)	(0.0757)
Control on j, c dummies	no	yes	yes	yes

Table 11: Estimates of the interaction between rule of law at the country level and relationship specificity at the industry level with instruments.

	ILO sectors (log emp shares)			
	Correl	Baseline	Income. eff.	No agric.
Coefficient	0.634***	0.937***	0.556***	0.533***
s.d.	(0.0509)	(0.0873)	(0.0800)	(0.0763)
Control on j, c dummies	no	yes	yes	yes
Instrument $RS_j \times \hat{1}_{1c}$ and $RS_j \times \hat{1}_{2c}$	yes	yes	yes	yes

D.0.3 Robustness: ILO industries with instruments

Ciccone and Papaioannou (2019) argue that the coefficient $\hat{\beta}$ estimated using cross-country cross-industry differences-in-differences approach may be biased if the technological indicator (in our case RS_j) varies across countries, and this variation is correlated with the institutional variable (in our case ROL_c) – although the sign of any bias is undetermined. They suggest a 2SLS estimator by instrumenting the $RS_j \times ROL_c$ interaction with $RS_j \times \hat{1}_{1c}$ and $RS_j \times \hat{1}_{2c}$, where $\hat{1}_{1c}$ is a dummy variable for countries with high rule of law, and $\hat{1}_{2c}$ is a dummy variable based on whether or not a country-specific estimate of $\hat{\beta}$ is high or low. We re-estimated the baseline equation (1) with these instruments. Our baseline results still hold: the coefficients actually increase in value, and remain significant at 1 the percent level. See Table 11.

D.0.4 Robustness: Income effects outside agriculture

So far we have interpreted our coefficient $\hat{\beta}$ as indicating that weak ROL harms productivity and thus raises costs particularly in high-RS_j industries. However, the literature on structural transformation also contemplates the possible importance of income effects as a factor of economic structure.[19] There are several reasons why income effects are unlikely to influence our results. First, income effects are mainly thought to be important for agriculture, and we account for the possibility of income effects related to agriculture in several ways as described above. Second, unlike the literature on structural transformation with income effects, which tends to distinguish between agriculture, manufacturing and services, our study uses much more disaggregated data, so most of our data points are for industries where income effects are not thought of as being very important for structural transformation. Third, even if income effects were important for economic structure at our level of disaggregation, we would need income elasticities to be correlated with relationship specificity for it to bias our estimates. We find this unlikely given the centrality for the literature on economic growth of the assumption that preferences and technology are independent: see Ngai and Pissarides (2007). This is not to suggest that income effects are not important for economic structure: our point is that there is no evidence in the literature that income effects might be related to an interaction between rule of law and relationship specificity, which is the topic of this paper.

That said, ultimately whether we are capturing a result that is due to income effects rather than our preferred interpretation is an empirical question. In this subsection, we pursue several further approaches to conditioning on income effects in our estimation.

One additional way to deal with the possibility of income effects is to consider the literature on legal origin – see for example La Porta et al (1997, 1998). In this litera-

[19] E.g. Kongsamut, Rebelo and Xie (2001), Gollin, Parente and Rogerson (2002), Restuccia, Yang and Zhu (2008) and Comin, Lashkari and Mestieri (2021).

Table 12: Estimates of the interaction between rule of law at the country level and relationship specificity at the industry level with legal origin as instrument

	ILO sectors (log emp shares)			
	Correl	Baseline	Income. eff.	No agric.
Coefficient	0.643***	1.794***	1.299***	1.376***
s.d.	(0.0893)	(0.458)	(0.486)	(0.490)
Control on j, c dummies	no	yes	yes	yes
Instrument $LegO_c \times RS_j$	yes	yes	yes	yes

ture, legal origin is viewed as a generally exogenous determinant of institutions that itself should be unrelated to the determinants of GDP *except through* institutions. We thus repeat our estimation of equation (1) using interactions of legal origin dummies $LegO_c$ and RS_j as instruments for the interaction between ROL_c and RS_j.[20] Our baseline results still hold, and remain significant at least 1 percent level. The coefficients in Table 12 are even larger than in our baseline results in Table 1.

Another approach to accounting for income effects is to include interactions of the log of income per capita $\log GDP_c$ times industry dummies in the estimation of (1) for all industries, not just agriculture. The specification becomes:

$$\log S_{jc} = \delta_j + \delta_c + \hat{\beta}\,(RS_j \times ROL_c) + \hat{\gamma}_j \log GDP_c \times \delta_j + \varepsilon_{jc}. \tag{23}$$

We find that several of the interactions of log income and the industry dummies in (23) are statistically significant, consistent with the existence of income effects. At the same time, the coefficient of the interaction of ROL_c and RS_j is 0.622, still significant at the 1% level. In addition, the correlation between RS and the coefficients γ_j on $\log GDP_c \times \delta_j$ is only 26.5%, far from significance even at the 10 percent level.

[20]See Rajan and Zingales (1998), Ilyina and Samaniego (2011) and Samaniego (2013).

Table 13: Estimates of the interaction between rule of law at the country level and relationship specificity at the industry level, controlling for the interaction of logGDP per capita and industry dummies.

	ILO sectors (log emp shares)	
Interaction	Coefficient	s.d.
$ROL_c \times RS_j$	0.622***	(0.0977)
$\hat{\gamma}_1$	−0.848***	(0.0810)
$\hat{\gamma}_2$	0.163	(0.112)
$\hat{\gamma}_3$	0.228***	(0.0783)
$\hat{\gamma}_4$	0.530***	(0.0728)
$\hat{\gamma}_5$	0.404***	(0.0777)
$\hat{\gamma}_6$	0.0457	(0.0772)
$\hat{\gamma}_7$	0.351***	(0.0783)
$\hat{\gamma}_8$	0.475***	(0.0792)
$\hat{\gamma}_9$	0.559***	(0.0889)
$\hat{\gamma}_{10}$	0.785***	(0.0916)
$\hat{\gamma}_{11}$	0.553***	(0.0771)
$\hat{\gamma}_{12}$	0.455***	(0.0714)
$\hat{\gamma}_{13}$	0.525***	(0.0849)
$\hat{\gamma}_{14}$	0.212***	(0.0802)

This suggests again that, to the extent that income effects exist, they are unrelated to the effect we are focusing on. See Table 13.

Yet another approach to accounting for income effects is to use *microeconomic* data to estimate the income elasticity of each sector's output. While it is not necessarily the case that microeconomic estimates are congruent to estimates using macroeconomic data, we think it is worth checking for additional robustness to income effects. Given income elasticity coefficients ϵ_j, we estimate the following specification, which includes an interaction of GDP with the microeconomic estimates of income effects:

$$\log S_{jc} = \delta_j + \delta_c + \hat{\beta}\left(RS_j \times ROL_c\right) + \hat{\gamma}\log GDP_c \times \epsilon_j + \varepsilon_{jc}. \tag{24}$$

We estimate the income elasticities ϵ_j for ILO sectors using personal and family surveys from the Panel Study of Income Dynamics (PSID) over the years 1999-2015. The income elasticities can be identified as the slope coefficient in a regressions of a logarithm of the expenditure share of goods on the logarithm of expenditure level of each household, with control on time fixed effects. The estimation regression is:

$$\log(Share_{ijt}) = \alpha_t + \epsilon_j \log(exp_{ijt}) + Controls_{it} + \varepsilon_{ijt} \tag{25}$$

where $Share_{ijt}$ is the expenditure share by household i on the output of ILO sector j at time t, and exp_{ijt} is household i's expenditure on the value added produced by sector j in year t. Control variables include a set of family and family head's characteristics (sex, race, education level, place of residence, and household size and age composition),[21] as well as time fixed effects. We instrument expenditure exp_{ijt} by household income, as is common in the related literature (see Blundell, Pashardes, and Weber (1993) and Boppart (2014)).

To implement this strategy, we proceed as follows. The PSID database reports certain expenditure categories which do not correspond to the sectors in the ILO database.[22] However, input-output tables provide a link between expenditure data

[21] To be more specific, the expenditure categories include: the age, the sex, marital status, education level (whehter attended college/high school, whether complete college/high school, year of last attend college, other training, years of scholing outside U.S.) and working experience, the race, the ethnic group, the religious preference of the family head, family location (state, size of largest city, urban or rural area), family size and children share of total family size.

[22] The categories are: Food at Home, Car Purchase Cash-down Payments, Car Loan Payments,

and industry value-added – see Herrendorf et al (2013). We use the U.S total requirements input-output tables to compute for each dollar spent on each of these PSID categories how much originates from the sectors in the ILO database. Then, given this mapping, we can compute each household's expenditure on the output of each of the ILO categories. Since we use US data, the operational assumption is that the ranking of income elasticities across sectors is roughly common across different countries.

The results of estimating specification (24) for each ILO sector are reported in Table (14). We make five observations. First, the estimates are significant in most cases suggesting that, at the household level at least, income elasticities may affect the demand for goods and services outside of the agricultural sector. Second, the correlation between the elasticities estimated in Table (14) and the coefficients estimated using aggregate data in Table (13) is 31.4%, suggesting some weak overlap between the microeconomic and macroeconomic approaches to measuring income elasticity (even though this relationship is not statistically significant at the 10 percent level). Third, the correlation between the sectoral income elasticities and RS_i is only 33.7 percent and far from standard statistical significance. Again, this suggests again that, to the extent that income effects exist, they are unrelated to the interaction we are focusing on. Fourth, when we repeat our estimation of equation (1) including as

Car Lease Payments, Additional Car/Lease Payments, Gasoline,Household Furnishings, and Clothing, Property Tax Expenses, Owner Insurance, Mortgage Payments, Rent, Home Repairs, Utilities , Health Insurance, Hospital/Nursing Home Expenses, Doctor/Outpatient Surgery/Dental Expenses, Prescriptions/Other Medical Services, Donations, Child Care , Food Delivered to Home , Food Eaten Out ,Car Insurance , Car Repair , Parking , Bus/Train Fares, Cab Fares, Other Transportation Expenses, School Expenses, Other School Expenses, Trips/Vacations, Other Recreation, Support of Others, and Voluntary Pension Contributions of Head and Wife.

Table 14: Estimates of sectoral income elasticities using PSID personal and family survey data from 1997-2015 waves

	ILO 1-Digit Classification	Elasticity	s.d.
1	Agriculture; forestry and fishing	-0.302^{***}	(0.006)
2	Mining and quarrying	-0.156^{***}	(0.006)
3	Manufacturing	0.087^{***}	(0.006)
4	Utilities	-0.473^{***}	(0.013)
5	Construction	0.475^{***}	(0.003)
6	Wholesale and retail trade; repair of motor	-0.175^{***}	(0.005)
7	Transport; storage and communication	-0.234^{***}	(0.009)
8	Accommodation and food service activities	-0.056^{***}	(0.008)
9	Financial and insurance activities	-0.217^{***}	(0.007)
10	Real estate; business and administrative	0.165^{***}	(0.004)
11	Public administration and defence	-0.13^{***}	(0.003)
12	Education	-0.007	(0.010)
13	Human health and social work activities	0.31^{***}	(0.004)
14	Other services	0.013^{***}	(0.006)

an additional control variable an interaction of the estimated income elasticities and the log of GDP per capita, we find that the coefficient on the interaction of income elasticities and log GDP per capita is in fact positive and significant, suggesting the existence of income effects for broader sectors (not just agriculture) as before. Fifth, nonetheless, our estimates of the interaction coefficient of ROL and RS_j remains highly significant as before – see Table 15. In fact, the estimates are very similar to our baseline results in Table 1.

Taking stock, there is some variation across specifications in the magnitude of the coefficients, but we do not find this particularly troubling. For example, we find that the raw correlations tend to be smaller than the coefficients when we add control variables such as fixed effects, which tells us that industry shares are influenced by

Table 15: Estimates of the interaction between rule of law at the country level and relationship specificity at the industry level control on interaction of income elasticity and logGDP

	ILO sectors (log emp shares)			
	Correl	Baseline	Income. eff.	No agric.
$ROL_c \times RS_j$	0.616***	0.902***	0.673***	0.645***
s.d.	(0.0418)	(0.0753)	(0.0654)	(0.0649)
$\log GDP_c \times elasticity_j$	0.101***	0.439***	0.102**	0.0988**
s.d.	(0.0109)	(0.0601)	(0.0444)	(0.0444)
Control on j, c dummies	no	yes	yes	yes

country- and industry-specific factors, as is widely assumed in the relate literature, and that it is important to condition on such factors. We tend to find somewhat smaller coefficients when we allow for income effects in agriculture, drop agriculture, or account for income effects in general: this tells us that controlling for income effects is important but they do not overwhelm our results. The coefficients are smaller when we use shares rather than log shares, this is not surprising as the shares are one or two orders of magnitude larger than the log shares. Finally, the coefficients are larger when we use instrumental variables, this is not unusual in the differences-in-differences literature (e.g. Rajan and Zingales 1998).

D.1 Extension of the model with income effects

Finally, we explore further the robustness of our findings to the potential presence of income effects by deriving an empirical specification from the preference structure in Comin et al (2022), which has non-vanishing income effects. This specification turns out to resemble (23) with a few additional control variables. Once more, the findings are robust.

The key hypothesis of the paper is that weak rule of law hinders productivity, particularly so in the production of goods that use relationship-specific intermediates. This should be reflected in them having a higher relative price and, as long as there is a non-unitary price elasticity between them, this should be reflected in differences in the composition of output. This should be identified by a regression on the share of different industries that includes an interaction term between relationship-specificity and ROL, regardless of whether or not income effects are present.

To see this, consider the literature on structural transformation, where a key determinant of the composition of output is the form of preferences. Consider the utility function in Comin et al (2022), which is a generalization of a standard CES utility function that accommodates both substitution effects and non-vanishing income effects. Let $\mathbf{C} = \{C_j\}_{j=1}^{J}$ be a vector of J goods where C_j is the consumption of goods. Then, utility $U(\mathbf{C})$ is defined implicitly by the constraint

$$\sum_{j=1}^{J} \Upsilon_j^{\frac{1}{\sigma}} \left(\frac{C_j}{g(U)^{\varepsilon_j}} \right)^{\frac{\sigma-1}{\sigma}} = 1, \tag{26}$$

where $g(\cdot)$ is a positive-valued, continuously differentiable and monotonically increasing function, $\Upsilon_j > 0$, $\varepsilon_j > 0$ and $\sigma \in \mathbb{R}^+ \setminus \{1\}$. This class of utility functions has a parameter that governs substitutability across goods σ, as well as non-vanishing income effects governed by the parameters $\{\varepsilon_j\}_{j=1}^{J}$.[23]

[23]Note that if $\omega_j \equiv \Upsilon_j^{\frac{1}{\sigma}}$, $\sum \Upsilon_j^{\frac{1}{\sigma}} = 1$ and $\varepsilon_j = \varepsilon \forall j$, setting $u(x) \equiv g^{-1}(x^{\varepsilon})$ we have a typical CES utility function with no income effects where:

$$U(\mathbf{C}) = u\left[\left(\sum_{j=1}^{J} \omega_j C_j^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} \right].$$

Let S_{jc} be the share of consumption of industry j in country c. Let E_c equal expenditure, and let k be a benchmark industry, chosen arbitrarily. Comin et al (2022) show that a consumer with these preferences optimally chooses a bundle such that:[24]

$$\log S_{jc} = (1-\sigma)\log\left(p_{jc}/p_{kc}\right) + (1-\sigma)\left(\epsilon_j - 1\right)\log\left(E_c/p_{kc}\right) + \epsilon_j \log S_{kc} + \log \Omega_j \quad (27)$$

where

$$\epsilon_j \equiv \frac{\varepsilon_j}{\varepsilon_k},\ \Omega_j \equiv \frac{\Upsilon_j}{\Upsilon_k^{\varepsilon_j/\varepsilon_k}}.$$

We can rewrite equation (27) as:

$$\log S_{jc} = (1-\sigma)\log p_{jc} - (1-\sigma)\log p_{kc} + (1-\sigma)\left(\epsilon_j - 1\right)\log\left(E_c/p_{kc}\right) + \epsilon_j \log S_{kc} + \log \Omega_j \quad (28)$$

We can use this preference specification to motivate an empirical approach to identifying the interaction of RS and ROL. First, note that in any particular country $\log p_{kc}$ is a country-specific constant, as are E_c and S_{kc}, whereas in any particular industry Ω_j is a constant. Setting $\delta_c = (1-\sigma)\left(\epsilon_j - 1\right)\log\left(E_c/p_{ck}\right) + \epsilon_j \log S_{ck} - (1-\sigma)\log p_{ck}$ and $\delta_j = \log \Omega_j$, equation (28) becomes:

$$\log S_{jc} = \delta_c + \delta_j - (\sigma - 1)\log p_{jc} \quad (29)$$

Note that this is the same as equation (21), just that the country and industry

[24] This is equation (11) in their paper.

dummies have different interpretations.

If instead we set $\delta_c = -(1-\sigma)(\epsilon_j - 1)\log p_{ck} + \epsilon_j \log S_{ck} - (1-\sigma)\log p_{ck}$ and $\delta_j = \log \Omega_j$, equation (28) becomes:

$$\log S_{jc} = \delta_c + \delta_j - (\sigma - 1)\log p_{jc} + (1-\sigma)(\epsilon_j - 1)\log E_c \tag{30}$$

This specification has an expicit coefficient on log income.

As per Proposition 2, let $\Gamma(\gamma_c, a_j) \equiv \log p_{cj}(\gamma_c, a_j)$. A second order Taylor approximation around interior points (γ^*, a^*) implies that

$$\begin{aligned}\Gamma(\gamma, a) &= \Gamma(\gamma^*, a^*) + \Gamma_a(\gamma^*, a^*)(a - a^*) + \Gamma_\gamma(\gamma^*, a^*)(\gamma - \gamma^*) \\ &\quad + \frac{1}{2}\Gamma_{aa}(\gamma^*, a^*)(a - a^*)^2 + \frac{1}{2}\Gamma_{\gamma\gamma}(\gamma^*, a^*)(\gamma - \gamma^*)^2 \\ &\quad + \Gamma_{\gamma a}(\gamma^*, a^*)(a - a^*)(\gamma - \gamma^*) + \tilde{o}(\|[\gamma, a] - [\gamma^*, a^*]\|)\end{aligned}$$

where Γ_x equals the derivative of Γ with respect to variable x and Γ_{xy} equals the derivative of Γ with respect to variables x and y. Since a_j depends on the industry and γ_c depends on the country, all the terms in this equation are country or industry-specific constants, except for the cross-derivative term and the small-o term. In addition, expanding $\Gamma_{\gamma a}(\gamma^*, a^*)(a - a^*)(\gamma - \gamma^*)$, the terms $\Gamma_{\gamma a}(\gamma^*, a^*)\, a^*\gamma_c$, $\Gamma_{\gamma a}(\gamma^*, a^*)\, a_j\gamma^*$ and $\Gamma_{\gamma a}(\gamma^*, a^*)\, a^*\gamma^*$ are also industry or country-specific constants (or just constants). As a result we have that

$$\Gamma(\gamma, a) \simeq \delta_j + \delta_c + \Gamma_{\gamma a}(\gamma^*, a^*)\, a_j\gamma_c$$

We find that

$$\begin{aligned}\log S_{jc} \simeq\ & \delta_j + \delta_c + (1-\sigma)\,\Gamma_{\gamma a}(\gamma^*, a^*)\, a_j \gamma_c \\ & + (1-\sigma)(\epsilon_j - 1)\log E_c \\ & - (1-\sigma)(\epsilon_j - 1)\,\Gamma_{\gamma a}(\gamma^*, a^*)\, a_k \gamma_c + \epsilon_j \log S_{kc}\end{aligned} \tag{31}$$

The resulting expression, which can be used to motivate a regression specification, is that, for all $j \neq k$:

$$\begin{aligned}\log S_{jc} =\ & \delta_j + \delta_c + \beta_1 \times (RS_j \times ROL_c) \\ & + \beta_{2j} \times (\delta_j \times \log Exp_c) \\ & + \beta_{3j} \times (\delta_j \times RS_k \times ROL_c) + \beta_{4j} \times (\delta_j \times \log S_{kc}) + \varepsilon_{jc}\end{aligned} \tag{32}$$

On the left we have $\log S_{jc}$ as in our baseline specification. On the right, we have an industry fixed effect, a country fixed effect, and our interaction of interest, as in our main specification. There are some additional controls, however: the interaction of log expenditure with an industry dummy, an interaction of the RS of our base industry k with ROL, and finally an interaction of industry dummies with the log share of our base industry. We select the base industry to be the industry with the lowest value of RS, noting that the choice of the base industry is arbitrary based on the theory.

When we estimate equation (32), we find that $\beta_1 = 0.591$, significant at the 1% level as before. In other words, our key empirical finding is robust to conditioning on income effects in the manner suggested by a preference framework with non-vanishing income effects.

One might wonder whether it is possible to use the estimated coefficients in equation (32) to calibrate the preference framework in (26) while allowing for income effects, since (31) links the estimated coefficients with model parameters. In particular, we have the following relationships between coefficients and model parameters:

$$\begin{aligned}
\beta_1 &= (1-\sigma)\,\Gamma_{\gamma a}\left(\gamma^*, a^*\right) \\
\beta_{3j} &= -\left(1-\sigma\right)\left(\epsilon_j - 1\right)\Gamma_{\gamma a}\left(\gamma^*, a^*\right) \\
\beta_{4j} &= \epsilon_j
\end{aligned}$$

Unfortunately this is not possible, however. The reason is that equation (31) also implies certain *restrictions* on the coefficients in (32), and the data reject these restrictions in most cases. For example, notice that, according to (31),

$$\begin{aligned}
\frac{\beta_{3j}}{\beta_1} &= \frac{-\left(1-\sigma\right)\left(\epsilon_j - 1\right)\Gamma_{\gamma a}\left(\gamma^*, a^*\right)}{\left(1-\sigma\right)\Gamma_{\gamma a}\left(\gamma^*, a^*\right)} = 1 - \epsilon_j \ \forall j \\
&\Rightarrow \ \frac{\beta_{3j}}{\beta_1} = 1 - \beta_{4j} \ \forall j
\end{aligned}$$

Furthermore, notice that $\beta_{2j} = (1-\sigma)\left(\epsilon_j - 1\right)$, so that

$$\frac{\beta_{2j}}{\left(\beta_{4j} - 1\right)} = (1-\sigma) \ \ \forall j$$

In other words, the equation (31) has the following restrictions on coefficients:

$$\frac{\beta_{3j}}{\beta_1} = 1 - \beta_{4j} \forall j \tag{33}$$

$$\frac{\beta_{2j}}{(\beta_{4j} - 1)} = \frac{\beta_{2j'}}{(\beta_{4j'} - 1)} \forall j, j' \tag{34}$$

After estimating (32), we can test each of these constraints separately. We find that constraint (33) is rejected for industries $4, 5, 6, 7$ and 9. We also find that constraint (34) is rejected for 42 out of 78 industry pairs. See Table (16), where the symbol $\times$ represents rejection for constraint (34), and $\bigcirc$ represents that constraint (34) is not rejected. The fact that these restrictions are rejected does not necessarily falsify the preference framework in (26), but it does indicate that we cannot use these coefficients to determine the values of any parameters. There may be other forms of income effects that could be at play – such as income effects that interact with the income *distribution* – or it could be that there are vanishing income effects as in Geary (1950) and Stone (1954) on top of those in (26). Determining the empirically relevant form of income preferences is an interesting topic for future work, but it is beyond the scope of this paper.

E Relationship specificity estimates

Here we report the relationshihp specificity measures used in our main specification.

Table 16: Tests for restrictions. Industry 12 is omitted as it is the "base" industry.

Industry	2	3	4	5	6	7	8	9	10	11	13	14
1	×	×	×	×	×	×	×	×	×	×	×	×
2		○	○	○	○	○	○	○	○	○	○	○
3			×	×	○	○	○	×	×	×	×	○
4				○	×	○	×	○	○	○	○	×
5					×	×	×	○	○	○	○	×
6						×	×	×	×	×	×	×
7							○	×	×	×	○	○
8								×	×	×	○	○
9									○	○	○	×
10										○	○	×
11											○	×
13												×

Table 17: Estimates of relationship specificity for ILO industries. The estimates are obtained by regressing the specificity values from Nunn (2007) on the direct requirements from the 1997 US IO tables, and using the coefficients to predict specificity values for the ILO industries based on their direct requirements.

	ILO 1-Digit Classification	RS_i
1	Agriculture; forestry and fishing	0.3326
2	Mining and quarrying	0.3100
3	Manufacturing	0.5338
4	Utilities	0.2423
5	Construction	0.5749
6	Wholesale and retail trade; repair of motor	0.6941
7	Transport; storage and communication	0.6269
8	Accommodation and food service activities	0.4393
9	Financial and insurance activities	0.9009
10	Real estate; business and administrative	0.6774
11	Public administration and defence	0.4607
12	Education	0.1453
13	Human health and social work activities	0.7804
14	Other services	0.5980

Table 18: The specificity values from Nunn (2007): the proportion of intermediates that are neither sold on an organized exchange nor reference priced.

ISIC code	Manuf. Industry	RS_i
311	Food products	.3306358
313	Beverages	.7128567
314	Tobacco	.316615
321	Textiles	.3760778
322	Wearing apparel, except footwear	.7454111
323	Leather products	.5706084
324	Footwear, except rubber or plastic	.6504076
331	Wood products, except furniture	.5161881
332	Furniture, except metal	.5676599
341	Paper and products	.3481136
342	Printing and publishing	.7128221
351	Industrial chemicals	.2402836
352	Other chemicals	.4897071
353	Petroleum refineries	.0576543
354	Misc. petroleum and coal products	.3952399
355	Rubber products	.4073072
356	Plastic products	.4077337
361	Pottery, china, earthenware	.3287548
362	Glass and products	.5574157
369	Other non-metallic mineral products	.3765819
371	Iron and steel	.2422237
372	Non-ferrous metals	.1603983
381	Fabricated metal products	.4346565
382	Machinery, except electrical	.7635783
383	Machinery, electric	.7400185
384	Transport equipment	.8587404
385	Professional and scientific equipment	.7846673
390	Other manufactured products	.5467549

Table 19: Esitmates of relationship specificity for ILO industries. The estimates are obtained by regressing the specificity values from Nunn (2007) on the direct requirements from the 1997 US IO tables, and using the coefficients to predict specificity values for the ILO industries based on their direct requirements.

	ILO 1-Digit Classification	RS_i
1	Agriculture; forestry and fishing	0.6800
2	Mining and quarrying	0.8722
3	Manufacturing	0.8831
4	Utilities	0.8625
5	Construction	0.9869
6	Wholesale and retail trade; repair of motor	0.9541
7	Transport; storage and communication	0.6411
8	Accommodation and food service activities	0.6168
9	Financial and insurance activities	0.9882
10	Real estate; business and administrative	0.9259
11	Public administration and defence	0.9176
12	Education	0.7061
13	Human health and social work activities	0.9550
14	Other services	0.9443

F Robustness with alternative measure for relationship specificity

In the Table 21 and 22, we use another measure of relationship specificity, RS, developed by Nunn (Table 20) and calculated by authors for ILO sectors (Table 19), which counts the fraction of inputs not sold on exchange. The results show that coefficients are all positive, and most are significant at least 5%,except when we drop agriculture sector. The coefficients are mostly smaller than those in 1, corresponding to the fact that this measure of relationship specificity is highly correlated with the other but generally larger in magnitude.

Table 20: The specificity values from Nunn (2007): the proportion of intermediates that are no sold on an organized exchange.

ISIC code	Manuf. Industry	$SPEC_i$
311	Food products	.5572861
313	Beverages	.9485962
314	Tobacco	.4831527
321	Textiles	.8204015
322	Wearing apparel, except footwear	.9754047
323	Leather products	.8479057
324	Footwear, except rubber or plastic	.9339786
331	Wood products, except furniture	.6698167
332	Furniture, except metal	.9100205
341	Paper and products	.8850983
342	Printing and publishing	.9952528
351	Industrial chemicals	.8836988
352	Other chemicals	.9458457
353	Petroleum refineries	.7593222
354	Misc. petroleum and coal products	.8945789
355	Rubber products	.9229651
356	Plastic products	.9847885
361	Pottery, china, earthenware	.9458082
362	Glass and products	.9671793
369	Other non-metallic mineral products	.9634284
371	Iron and steel	.8162452
372	Non-ferrous metals	.4601733
381	Fabricated metal products	.9446488
382	Machinery, except electrical	.9747962
383	Machinery, electric	.9601691
384	Transport equipment	.9845868
385	Professional and scientific equipment	.9807606
390	Other manufactured products	.8634105

Table 21: Estimates of the interaction between rule of law at the country level and relationship specificity at the industry level.

	ILO sectors (log emp shares)			
	Correl	Baseline	Income. eff.	No agric.
Coefficient	0.366***	1.183***	0.354***	0.190*
s.d.	(0.0311)	(0.151)	(0.104)	(0.0981)
Control on i, c dummies	no	yes	yes	yes

Table 22: Estimates of the interaction between rule of law at the country level and relationship specificity at the industry level.

	UNIDO Manuf. industries		
	log emp share	log VA share	log output share
Coefficient	0.424**	0.698***	0.580***
s.d.	(0.177)	(0.200)	(0.193)
Control on i, c dummies	yes	yes	yes

G On the use of ICP data for estimating σ

The International Comparisons Program of the World Bank (ICP) provides data on prices and expenditure. While the data do not correspond exactly to our classifications, they are at roughly the same level of aggregation, and are thus suitable for our purposes. We draw on the 2011 and 2017 rounds, which are comparable, although a few countries are missing from the 2011 round. Price indices are normalized so that the value for the World is 100 in each year. As a result, we examine the rounds separately but also pool them.

The ICP data include data at various levels of aggregation. We use the most aggregated level, as these correspond roughly to the ILO classifications.[25] This is

[25] In addition, the food, manufacturing and transport sectors are reported at a disaggregated level, so estimates using a lot of data points from those sectors might not be representative of the economy as a whole.

defined in terms of the number of significant figures in the clasification code. The classification code has 7 digits, so we used only the 4 digit classifications. The resulting sectors are Food and non-alcoholic beverages, Alcoholic beverages and tobacco, Clothing and footwear, Furnishings and household equipment, Transport equipment and services, Communication, Restaurants and Hotels, Machinery and equipment, Construction, Other products, Housing and utilities, Health expenditure, Recreation and culture, Education and Other. This yields 14 sectors, like the ILO data, although the partition over disaggregated industries is not the same. In addition the ICP data do not include information on agriculture and mining as these sectors are not comparable across countries (e.g. there is a lot of Copper in Zambia but none in Turkmenistan, and there is a lot of rice in Thailand but not in Iceland).

Finally, the ICP data include information on total expenditure, which allows us to compute the share of spending S_{jct} for each sector j in each country c in each year t.

H On Measurement Error in Rule of Law

As well as reporting the Rule of Law variable ROL_c, the World Bank reports for each country the standard error of the ROL_c estimate – call it σ_c^{ROL}. As ROL_c enters specification (1) as part of an interaction term, there is no easy way to adjust for these standard errors. Here we discuss the robustness of our estimates to these standard errors by drawing ROL_c values from the implied distribution and examining the distribution of the resulting estimated coefficients, a sort of simulated bootstrap.

To do this, we assume that ROL_c is normally distributed with mean equal to the reported value and s.d. equal to σ_c^{ROL}. We then estimate (1) repeatedly with different values of ROL_c drawn from this distribution. We do this $100,000$ times, and compare the resulting distribution of coefficients on the interaction term of interest, $\hat{\beta}$, with the estimates in Table 1.

Note that this exercise requires an assumption of how mismeasurement in ROL_c is correlated across countries. On the one hand, the errors in measurement that are captured by σ_c^{ROL} could be completely independent across countries. On the other hand, if the errors are partly due to the common measurement *procedure*, then the mismeasurement could be correlated. Thus we examine three scenarios:

1. The errors are uncorrelated. The variance-covariance matrix is

$$\begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \times \overrightarrow{\sigma}_c^{ROL}$$

   where $\overrightarrow{\sigma}_c^{ROL}$ is the vector of values of σ_c^{ROL}

2. The errors are moderately correlated. The variance-covariance matrix is

$$\begin{bmatrix} 1 & 0.3 & \cdots & 0.3 & 0.3 \\ 0.3 & 1 & \cdots & 0.3 & 0.3 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0.3 & 0.3 & \cdots & 1 & 0.3 \\ 0.3 & 0.3 & \cdots & 0.3 & 1 \end{bmatrix} \times \overrightarrow{\sigma}_c^{ROL}.$$

3. The errors are strongly correlated. The variance-covariance matrix is

$$\begin{bmatrix} 1 & 0.5 & \cdots & 0.5 & 0.5 \\ 0.5 & 1 & \cdots & 0.5 & 0.5 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0.5 & 0.5 & \cdots & 1 & 0.5 \\ 0.5 & 0.5 & \cdots & 0.5 & 1 \end{bmatrix} \times \overrightarrow{\sigma}_c^{ROL}.$$

The resulting distribution of coefficients is displayed in Figures 2, 3 and 4 below. Figure 2 assumes no correlation between errors, while Figures 3 and 4 assume moderately and strongly correlated errors. Compared to the baseline results in Table 1, we observe the following. First, the means and modes of the distributions are not very different from the baseline coefficients – i.e. the range of estimates is fairly tight and roughly normal. Second, based on these simulated distributions, the baseline coefficients are within the 95% percentile ranges of the estimated coefficients in 9 of 12 cases – see Table 23. Third, only in one of the 12 cases is the baseline coeffi-

Table 23: Confidence bounds based on the distribution of simulated bootstrapped regression coefficients. Coefficients outside standard cofidence bounds are marked with an asterisk.

	Regression			
Uncorrelated errors	Corr	Base	Inc Eff	No Ag
Coefficient	0.6220	1.0570*	0.7050	0.6600
Upper confidence bound	0.6227	1.0515	0.7094	0.8420
Lower confidence bound	0.5689	0.9725	0.6448	0.6562

Moderately correlated	Corr	Base	Inc Eff	No Ag
Coefficient	0.6220	1.0570	0.7050	0.6600*
Upper confidence bound	0.6251	1.0574	0.7121	0.8316
Lower confidence bound	0.5784	0.9889	0.6568	0.6740

Highly correlated	Corr	Base	Inc Eff	No Ag
Coefficient	0.6220	1.0570	0.7050	0.6600*
Upper confidence bound	0.6263	1.0612	0.7134	0.8226
Lower confidence bound	0.5838	1.0003	0.6653	0.6877

cient *above* the upper confidence bound, and this is only when we assume that there is no correlation in the measurement error across countries. We conclude that our estimates of the interaction term of interest, $\hat{\beta}$, are not significantly biased due to potential measurement error in ROL_c.

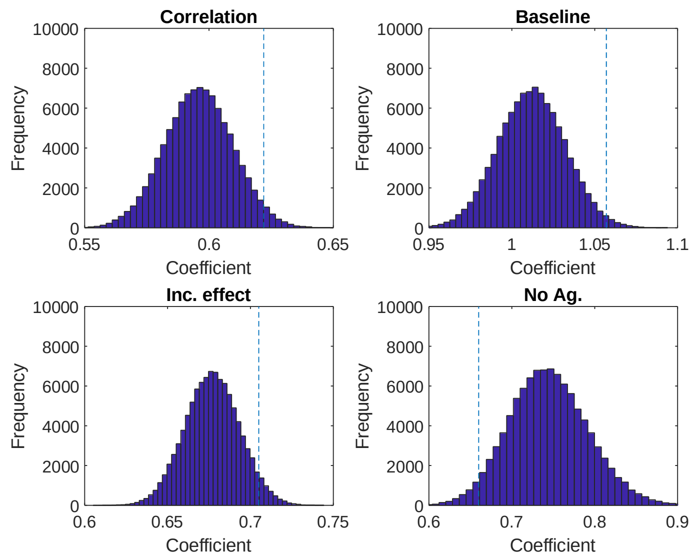

Figure 2 – Histogram of estimated coefficients from the bootstrap simulation. Assumes the correlation between errors is zero. The dashed lines represent the estimated coefficients in Table 1. Histogram has 50 bins.

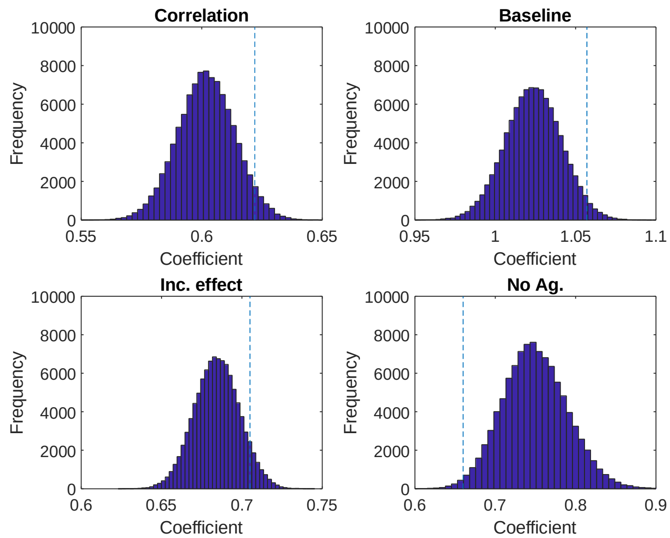

Figure 3 – Histogram of estimated coefficients from the bootstrap simulation. Assumes the correlation between errors is 0.3. The dashed lines represent the estimated coefficients in Table 1. Histogram has 50 bins.

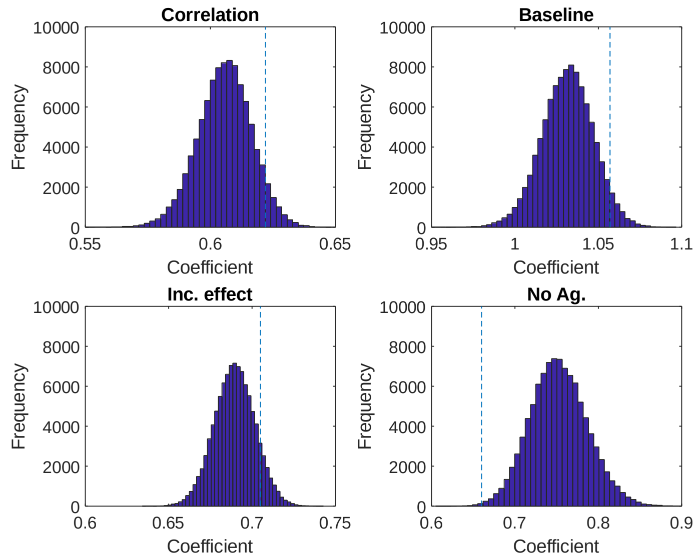

Figure 4 – Histogram of estimated coefficients from the bootstrap simulation. Assumes the correlation between errors is 0.5. The dashed lines represent the estimated coefficients in Table 1. Histogram has 50 bins.