Edinburgh School of Economics
Discussion Paper Series
Number 271

*Managerial Delegation, Law Enforcement,
and Aggregate Productivity*

Jan Grobovsek
(University of Edinburgh)

Date
August 2014

Published by
School of Economics
University of Edinburgh
30 -31 Buccleuch Place
Edinburgh EH8 9JT
+44 (0)131 650 8361
<http://edin.ac/16ja6A6>

THE UNIVERSITY *of* EDINBURGH

MANAGERIAL DELEGATION, LAW ENFORCEMENT, AND AGGREGATE PRODUCTIVITY

Jan Grobovšek

University of Edinburgh

August 2014

ABSTRACT. This paper proposes a novel conceptual framework to quantify how law enforcement shapes the internal organization of firms and thereby aggregate equilibrium outcomes. First, we present empirical evidence on a significant positive cross-country correlation between the aggregate share of managerial workers and the degree of law enforcement. Second, we construct a tractable model that captures benefits to managerial delegation in large organizations. The model features an agency problem between the owner of the firm and its middle managers. Ineffective law enforcement, allowing middle managers to divert revenue from the firm, constrains firm size by limiting the efficient delegation of managerial authority as well as managerial employment. Third, a calibrated version of the model measures the effect of deteriorating legal protection. Decreasing law enforcement from the U.S. benchmark to a level associated with countries at ten percent of U.S. GDP per capita reduces aggregate productivity by 18 percent. Auxiliary statistics on the mean employer business size, self-employment, productivity dispersion, skill premium and human capital all paint a picture characteristic of low-income countries.

I would particularly like to thank Juan Carlos Conesa, Tim Kehoe and Nezih Guner for their advice. This paper has benefited from the suggestions of Albert Marcet, Vasco Carvalho, Fabrizio Perri, Sevi Rodríguez Mora, Pau Pujolàs, David Wiczer, Bob Lucas, Tasso Adamopoulos, Román Fossati, Martí Mestieri, Chris Woodruff, Michael Peters, Philipp Kircher, and seminar participants at UAB, the Minneapolis Fed, Edinburgh, Toulouse, CREI, Notre Dame, Louvain, Cologne, Constance, Amsterdam, NES, the Mannheim Conference in Quantitative Macro, the EEA meeting, and the SED.

JEL codes: O10, O40, O43, O47.

Key words: Growth and Development, TFP, Misallocation, Management, Delegation, Law Enforcement.

1. INTRODUCTION

Compared to their peers in the rich world, firms in developing countries are on average badly managed. This follows from a series of papers building on the pioneering empirical work on managerial practices by Bloom and Van Reenen (2007).¹ They find that firms could boost productivity via minor and seemingly cheap changes to daily management. In particular, one major source of managerial inefficiency in less developed countries is insufficient delegation of decision-making. Many productivity-enhancing measures are left on the table as the workers who are best informed about specific problems are not endowed with sufficient authority to solve them. Why? Empirically, Bloom, Sadun and Van Reenen (2012) argue that social capital in the form of trust as well as the rule of law are important drivers of firm decentralization across countries, together with the attendant benefits of re-allocation and productivity gains. This squares neatly with the empirical investigation on Indian firms in Bloom et al. (2013) where the lack of trust in non-family members is found to be the main obstacle to the delegation of managerial tasks.

The novelty of the present paper is the joint formalization of these concepts and their quantification through the lens of a simple structural equilibrium model. The ineffective law enforcement characteristic of poor countries is assumed to allow middle managers - i.e. workers with a sufficient degree of autonomy - to divert resources from firms. This hinders delegation, especially in efficient and therefore larger organizations. The resulting misallocation of resources within the firm may be interpreted as poor management. Our main interest is to measure the corresponding aggregate productivity loss. In addition, we also quantify the extent to which the mechanism replicates other stylized facts about poor countries such as the small average firm size with a large concentration of own-account workers, a strong positive relationship between revenue productivity and firm size, large skill premia and lower levels of human capital.

We proceed in three steps. First, we uncover a robust positive cross-country relationship between the aggregate share of managerial workers and law enforcement. It hints at a causal link running from a proxy measure of the risk of firm expropriation by employees on a proxy measure of managerial quality and delegation. The aggregate share of managerial workers is admittedly a coarse indicator of managerial quality and delegation, but it has the benefit of being a tangible measure that is available across widely different countries. The relationship provides an empirical foothold in the form of the expected drop in the aggregate managerial share corresponding to law enforcement levels typical in less developed economies. For instance, compared to the U.S., that drop is estimated to be 5 percentage points for countries at one tenth of U.S. GDP per capita.

Second, we build a general equilibrium model of heterogeneous firms that bridges the notions of expropriation, delegation and the aggregate share of managerial workers. It features two key elements. One is the firm's production function that captures in the simplest possible manner the endogenous choice of delegation. The degree of delegation is the number of managerial layers overseeing production workers. The addition of a managerial layer allows entrepreneurs to increase their span of control by creating knowledge or supervision, which is traded off against the extra overhead cost per layer. Efficient firms choose to delegate more in order to direct a larger workforce, thereby increasing their share of managerial employees. The resulting organizational structure can be viewed as the generalized version of the Lucas (1978) span of control model since the span parameter

¹Bloom and Van Reenen (2010) extend the initial dataset to a larger number of developing countries. Bloom, Eifert, Mahajan, McKenzie and Roberts (2013) is a field experiment in India measuring the sensitivity of firms' management practices to external professional advice. See also Bloom, Mahajan, McKenzie and Roberts (2010) for a review of the findings and how they may relate to cross-country productivity differences.

in effect becomes a choice between zero and constant returns. Moreover, we assume that skill-intensity rises in the managerial hierarchy. While this is not a necessary ingredient for the functioning of the core mechanism, it does enrich the model's predictions on skill accumulation as well as improves its empirical application.

It is the institutional environment that is the other key element. The central hypothesis is that in each layer managerial employees have the option to divert the firm's revenue, up to a fraction governed by law enforcement. Firms offer their middle managers long-term wage contracts that are juicy enough to prevent expropriation, thus creating ex post heterogeneity between various types of managers and production workers. Imperfect contract enforcement generates managerial efficiency wages *à la* Shapiro and Stiglitz (1984), with production work replacing the role of unemployment as a market clearing mechanism. Formally, the degree of contracting friction turns out to be a simple linear tax rate on revenue that is increasing in the hierarchy length. The most efficient firms - those choosing a longer hierarchy - face a disproportional output wedge. The mechanism therefore delivers an endogenous foundation for the positive correlation between efficiency and generic output taxes stipulated in Restuccia and Rogerson (2008) and backed out empirically by Hsieh and Klenow (2009) and Bartelsman, Haltiwanger and Scarpetta (2013). In addition, the contracting friction affects the marginal cost of workers by altering equilibrium wages and compensation bonuses. There is misallocation along the *extensive* margin of hierarchy length and along the *intensive* margin in the choice between managerial versus production workers. This in turn produces misallocation on occupational *selection* between entrepreneurship and labor market participation. Finally, there is a dynamic *investment* distortion as the accumulation of worker skills is endogenous. In particular, it can lead to skill mismatch whereby skilled workers perform production tasks that require no skills. In that scenario, skill accumulation is wasteful as its only equilibrium purpose is to dampen the compensation of skilled managerial workers by confronting them with a queue of would-be managers willing to replace them.

Third, we calibrate the model to the U.S. and consider various counterfactuals. That exercise is most notably disciplined by matching the firm size-wage premium, both unconditionally and conditionally on worker skills, as larger firms in the model pay higher wages.² The quantitative version reveals several appealing properties. A smooth thick-tailed firm size distribution is generated from a thin-tailed distribution of projects, revenue productivity increases in firm size, while the progression of managerial layers up to a maximum level of six appears plausible. We then examine variations in law enforcement. In particular we focus on law enforcement levels corresponding to low-income countries, inferred from the accompanying drop in the aggregate managerial share. Consider the law enforcement level associated with countries at ten percent of the U.S. income level. As efficient firms curb delegation, the average size of employer businesses shrinks significantly from about 20 to 8 workers. The aggregate share of self-employment doubles, with the share of own-account workers (i.e. non-employers) climbing from 7 to 12 percent. That outcome is especially noteworthy as the model mechanism may well have generated fewer own-account workers by concentrating employment in small firms facing no contracting

²See for instance Idson and Oi (1999) and Troske (1999) for empirical investigations of the size-wage premium puzzle.

frictions.³ As skill-intensive managerial work becomes more expensive, the wage skill premium increases by about a third, and the supply of skilled workers drops.⁴ Both the elasticities of wages and of revenue productivity with respect to firm size almost double - compared to small firms, large and therefore constrained businesses now pay significantly higher wages and are a lot more productive. To sum up, the mechanism produces an economy that resembles representative developing countries in many respects. Above all, the corresponding GDP per capita loss is 18 percent. That certainly falls short of claiming a silver bullet to the TFP puzzle, but it is nonetheless substantial. The mechanism provides a *complementary* theory of TFP, and the exogenous variation derives from a fairly specific policy institution.⁵ Also, the productivity loss boils almost entirely down to pure misallocation rather than auxiliary leveraging channels since the dynamic investment effect via skill accumulation turns out to play second fiddle.⁶

This paper relates to a number of theoretical contributions on firm organization. The nexus of delegation and span of control is similar in spirit to Garicano (2000) as well as its various applied applications.⁷ The technological trade-off whereby delegation saves knowledge acquisition expenditure by increasing communication or supervision plays out similarly here, albeit with a production function that is somewhat less structural and more reminiscent of Lucas (1978). Our assumption of rising skill-intensity through layers leads to implications akin to those from the production process in Kremer (1993). There, countries with expensive human capital reduce the complexity of products, while in the present paper firms shorten their optimal hierarchy. As for the agency problem of expropriation arising from delegation, it is modeled along the lines of optimal dynamic debt contracts with one-sided limited commitment.⁸ The resulting prediction that efficiency wages are increasing in the hierarchical managerial position is closely related to that in Calvo and Wellisz (1978) and Calvo and Wellisz (1979). The main difference is the interpretation. There, higher layers are disproportionately incentivized to provide effort so as to maximize effort supervision of subordinate workers. Here, higher managerial layers are simply assumed to handle more revenue per manager and effort is not a concern. Its abstraction, on the other hand, allows the production function as well as the dynamic game to be highly tractable. Finally, the trade-off underlying delegation differs from other classic issues such as incentives for initiatives as in Aghion and Tirole (1997), risk of spin-offs as in Rajan and Zingales (1998), and noisy communication as in Dessein (2002).⁹

Methodologically, the paper follows in the footsteps of contributions measuring equilibrium misallocation of individual production units arising from institutional frictions.

³Low-income countries typically have small employer businesses, as reviewed in Tybout (2000) and recently confirmed by Poschke (2014). They are also characterized by a high share of self-employment, in particular the share of single own-account workers, as documented by Gollin (2008).

⁴The empirical findings on the joint existence of large skill premia and low human capital in poor countries are well summarized by Caselli (2005).

⁵See for instance Restuccia and Rogerson (2013) for a review on the difficulty of generating large misallocation from very concrete policy variations.

⁶For example, the literature on institution-induced credit constraints quoted below finds productivity losses that are by and large comparable in magnitude. Larger aggregate productivity losses are found in recent contributions focusing on the agricultural sector as in Adamopoulos and Restuccia (2014), or on human capital accumulation as in Gennaioli, La Porta, Lopez-de Silanes and Shleifer (2013) and Manuelli and Seshadri (2014), but these do not aim at uncovering particular institutional failures.

⁷See in particular Antràs, Garicano and Rossi-Hansberg (2006), Garicano and Rossi-Hansberg (2006), Garicano and Rossi-Hansberg (2012), Caliendo and Rossi-Hansberg (2012), and Caliendo, Monte and Rossi-Hansberg (forthcoming).

⁸See for instance Kehoe and Levine (1993) and Alvarez and Jermann (2000).

⁹See also Aghion, Bloom and Van Reenen (2014) for a synthesis of the theoretical and empirical literature on decentralization in firms.

Much of the literature has focused on credit frictions. Prime examples include Erosa and Hildago Cabrillana (2008), Greenwood, Sanchez and Wang (2010), Amaral and Quintin (2010), Buera, Kaboski and Shin (2011), Caselli and Gennaioli (2013), Moll (forthcoming), and Midrigan and Xu (forthcoming). These papers have in common a game between the capital provider and the entrepreneur/manager, with poor contract enforcement draining the flow of credit.¹⁰ The present paper, in contrast, sets the game inside the firm between the entrepreneur and her middle managers. As such it relates to a nascent literature on firm organization, institutions and aggregate outcomes. The closest contribution is Chen (2014). His exercise is equally concerned with the hierarchical decentralization of the firm as a function of institutions and the monitoring technology more generally, with the empirical application centering on international trade. Apart from modeling differences (including the treatment of labor markets and human capital), the present paper distinguishes itself primarily by an alternative quantification of income losses due to poor law enforcement. Another close contribution is Akcigit, Alp and Peters (2014) who examine the dynamics of innovation and firm growth in an environment of non-contractible managerial tasks. They find that compared to the U.S., ineffectual law enforcement in India explains much of the observed slow firm growth and substantial employment share of small firms. The present paper differs methodologically in its distinct treatment of various managerial layers within a hierarchy, and empirically in its concern with static rather than dynamic misallocation, focusing on the arising aggregate income loss. Powell (2013) also models a dynamic commitment game between firms and their managers. In contrast to the present paper, the commitment problem is two-sided so that more efficient firms can pledge more reputational collateral, leading them to over, rather than under-produce. The two papers generate distinct predictions as stricter law enforcement here disproportionately benefits larger firms while driving the least efficient producers out of the market. Finally, Boehm (2013) examines the effect of contract enforcement on the vertical integration of firms. His theory is highly complementary to ours as a lack of contract enforcement hinders the delegation of activities to intermediate input providers.¹¹

The next section describes the empirical motivation. Section 3 presents the model environment and Section 4 its theoretical implications. Section 5 covers the calibration as well as its quantitative implications, followed by simulation results in Section 6. Section 7 concludes.

2. CROSS-COUNTRY EVIDENCE

The first panel of Figure (1) presents a striking positive cross-country correlation between the share of managerial workers and GDP per capita, as previously noted by Eeckhout and Jovanovic (2011). The data on managerial workers are from the ILO and derive mainly from labor force surveys.¹² To control for the different sample periods these are

¹⁰Other frictions studied in the literature include size-dependent policies as in Guner, Ventura and Xu (2008) and García-Santana and Pijoan-Mas (2014), informational frictions as in Cole, Greenwood and Sanchez (2013), matching frictions as in Alder (2012), and market power as in Peters (2013).

¹¹Roy and Seshadri (2014) consider an environment of assortative matching between entrepreneurs and workers where the delegation of tasks is inversely related to human capital accumulation. Firm decentralization there is not driven by institutional variations but rather by exogenous TFP differences that affect human capital accumulation.

¹²For each country I sum the number of employees categorized as managers, professionals and administrative workers (categories 0, 1 and 2, as well as 3 in the ILO classification) and divide by the total working population, excluding agricultural and non-classifiable workers (categories 6 and X). Non-managerial workers - labeled production workers henceforth - are clerks, service and sales workers, craft and related trade workers, plant and machine operators and workers in elementary occupations (categories 4, 5, 7, 8 and 9).

FIGURE 1

averages over the years 1999 to 2008. GDP per capita is taken from the Penn World Tables and represents averages over the same time period, normalized to the U.S. Our focus here is rather on the second panel of Figure (1) that highlights the positive correlation between the managerial worker share and an indicator of . For that I use data from the World Bank Doing Business database which offers a series of variables capturing the quality of contract enforcement. The preferred indicator used here is the pecuniary cost of a lawsuit as a percentage of the value of the claim as it comes closest to the notion of contract enforcement employed in the theory. The measure of enforcement is constructed as 1 minus the World Bank indicator and hence stands for the percent of the claim that is expected to be recovered in a lawsuit.¹³ Finally, I use averages over the years of interest to wash out any one-off effects.

We resume by assuming the share of managerial workers to be a dependent variable. Table (1) presents OLS regressions of the managerial share on several potential explanatory variables. The variables are in logs and the sample of countries is the maximum number of available observations. The first three columns imply that while contract enforcement impacts the managerial share on its own, its importance vanishes when included side by side with GDP per capita. As such, contract enforcement on its own may not matter much, but this ignores many other characteristics that correlate strongly with income and that may help improve identifiability. The remaining columns of Table (1) contain some of the more obvious such variables and contrast the results of excluding and including contract enforcement.

First, consider the sectoral composition of the workforce as well as the share of government spending, which arguably acts as a proxy for the share of public employees.¹⁴ These seem to matter in the sense that the content of managerial work may well differ across economic activities and sectors. Notice that GDP per capita now loses significance, especially when contract enforcement is also accounted for. The latter acquires economic and statistical significance.

¹³For some of the countries the measure turns negative. These countries are dropped from the analysis, both to omit outliers as well as to be able to run regressions in logs. In any case, including them in a linear regression would strengthen the effect of enforcement on the share of managerial workers.

¹⁴The data are from the World Bank and the Penn World Tables, respectively, and represent averages over the period 1999 through 2008.

Regressors	(1)	(2a)	(2b)	(3a)	(3b)	(4a)	(4b)	(5a)	(5b)
Contract enforce. (%)	0.885*** (0.261)		0.350 (0.235)		0.495** (0.239)		0.645*** (0.235)		0.567** (0.248)
GDP/capita (% of U.S.)		0.230*** (0.028)	0.210*** (0.031)	0.164*** (0.059)	0.106* (0.062)	0.161** (0.072)	0.079 (0.074)	0.118 (0.083)	0.066 (0.085)
Empl. agriculture (%)				0.035 (0.047)	0.009 (0.052)	0.008 (0.050)	−0.026 (0.054)	0.017 (0.054)	−0.031 (0.060)
Empl. services (%)				0.348** (0.147)	0.384** (0.156)	0.089 (0.173)	0.166 (0.178)	0.170 (0.194)	0.175 (0.201)
Govern. spending (%)				0.101 (0.063)	0.082 (0.063)	0.171** (0.079)	0.155** (0.076)	0.183** (0.087)	0.164** (0.084)
Tertiary education (%)						0.160*** (0.044)	0.172*** (0.045)	0.162*** (0.051)	0.171*** (0.056)
Credit/GDP (%)						−0.053 (0.053)	−0.069 (0.051)	−0.061 (0.057)	−0.080 (0.056)
Self-employment (%)								−0.084 (0.068)	−0.047 (0.067)
Number obs.	118	125	118	122	115	105	99	98	93
R ²	0.090	0.349	0.354	0.393	0.403	0.459	0.496	0.483	0.503

TABLE 1. Regression on the managerial share

Next we consider two more candidates for the determination of the managerial share. The first is the population share with completed or attempted tertiary education.¹⁵ It is intuitive that countries with higher educational attainments have more managerial workers assuming that these are characterized by higher skills. The second variable is private credit over GDP, an indicator often employed to measure the quality of financial markets.¹⁶ A lack of credit could hinder the accumulation of managerial skills as well as keep firms from expanding, both of which may impact the share of managerial workers. As it is, private credit does not appear to matter in these regressions. Education, on the other hand, seems to pick up most of the significance of service employment as services are on average more skill-intensive. Above all, contract enforcement is again significant, to the detriment of the country's income level *per se*. The final two columns include one additional variable, namely the share of employers and self-employed in the population.¹⁷ This could be an important control statistic since our quantitative model creates cross-country differences in the share of self-employment.

These findings offer the following interpretation. The first is *qualitative*, namely evidence that contract enforcement positively impacts the managerial share over and beyond GDP per capita. Given the usual pitfalls of cross-country regressions, we may well stop here. If, however, we were to lay more trust in the resulting numbers, we can also make use of the *quantitative* strength of the relationship to extract information. One interpretation that we will return to below is the following. The employed indicator of contract enforcement is only a proxy measure and as such difficult to include explicitly in a quantitative model. For the quantitative analysis will we therefore choose an indirect approach. We will alter the model-based measure of law enforcement so as to match the corresponding managerial share of interest. What share is that? Regressing the log of contract enforcement on log GDP per capita yields that countries with income levels of 0.25, 0.1, and 0.05 of the U.S., respectively, have expected contract enforcement levels

¹⁵The variable is from the dataset of Barro and Lee (2010) and represents the 2005 figures for individuals aged 25 and above.

¹⁶The data are from the World Bank and averaged over 1999-2008.

¹⁷Computed from the ILO data and averaged over 1999-2008.

of 0.734, 0.663 and 0.614, as opposed to the U.S. empirical value of 0.856.¹⁸ Let 0.567 be the elasticity between enforcement and the managerial share. Applying that elasticity we can compute the expected drop in the managerial share originating exclusively from variations in enforcement. For representative income levels of one-fourth, one-tenth and one-twentieth of the U.S., the predicted managerial shares are 0.321, 0.303 and 0.290, respectively, compared to the U.S. empirical value of 0.35.¹⁹

3. MODEL

3.1. Basic environment

The economy is populated by a unit mass of individuals. These have period utility $u = (\int c(j)^\phi dj)^{1/\phi}$, $\phi \in (0, 1]$, over the consumption of an infinite amount of goods varieties j . Time is discrete and discounted at the rate $\beta(1 - \delta_a)$, where $\beta \in (0, 1)$ is a discount factor and $\delta_a \in (0, 1)$ represents the hazard rate of exiting the economy. As will become clear, the absence of overall curvature in the utility function is a major assumption that ensures tractability in our environment. We can gently sweep it under the rug by supposing the existence of complete credit markets in the background.

Each individual is born into the economy with a permanent project idea z , drawn from a cumulative distribution function $G(z)$ with continuous support $[0, \bar{z}]$. Armed with that idea, the individual starts the beginning of each period with an occupational choice. He can become an entrepreneur in charge of his own firm with associated value $V_f(z)$. Alternatively, he can enter the labor market to become an employee, which generally yields the expected value $V_w(z)$. Employees make no intrinsic use of their project idea, but its level could in principle influence their continuation value. In the stationary equilibrium, however, where aggregate variables are constant, the optimal occupational choice upon birth is permanent. Since the stationary equilibrium is the only focus of our analysis we assume from now on that $V_h(z) = V_h$, independently of z . It follows that the occupational choice is $V(z) = \max\{V_f(z), V_h\}$, where the cutoff value $\underline{z}$ is such that $V_f(\underline{z}) = V_h$.

3.1.1. Firms

Each firm is attached to one entrepreneur and characterized by the project quality z . The entrepreneur in principle solves $V_f(z) = \max\{\pi(z) + \beta(1 - \delta_a)V'(z)\}$. Again, in the stationary equilibrium the occupational choice is permanent. Moreover, the problem includes no additional state variables so we can restrict the analysis to $V_f(z) = 1/[\beta(1 - \delta_a)] \max \pi(z)$. The entrepreneur provides $\alpha > 0$ units of efficiency labour as opposed to regular employees whose efficiency is normalized to one. A value $\alpha > 1$ indicates that the entrepreneur's work is relatively more valuable, for instance because his incentives are more aligned. In contrast, $\alpha < 1$ can be interpreted as capturing a fixed time cost associated with operation. The market structure is monopolistic competition as each project is a distinct good variety. The price charged by the firm is therefore $p(z) = [R/y(z)]^{1-\phi}$. R is the economy's total revenue, $y(z)$ the quantity of variety z , and the implicit ideal price index P is normalized to unity. Moreover, firm operation requires a fixed per-period expenditure cost $\kappa \in \mathbb{R}$.²⁰ The gist of the model lies in two aspects, which are the organization of the firm and the compensation of employees.

¹⁸The regression over 173 observations gives a coefficient of 0.111 (standard error: 0.0167). Predicted contract enforcement is therefore a factor $x^{0.111}$ of the U.S. level where x is the country's income per capita with respect to the U.S.

¹⁹From $0.35 \times x^{0.111 \times 0.567}$ for income level x .

²⁰Both monopolistic competition and the fixed cost only feature for the empirical application. They are of no theoretical interest and may be thought away by setting $\phi = 1$ and $\kappa = 0$.

3.1.1.1. Organizational structure. There are three distinct organizational structures. The first type is the own-account firm, or zero-layer firm ($L = 0$). Here, the entrepreneur is the only worker - more precisely a *production* worker ($n = 1$) - and output is $y_0(z) = z\alpha$. There is no further choice so period profits are simply

$$\begin{aligned}\pi_0(z) &= r_0(z) - \kappa \\ &= R^{1-\phi} z^\phi \alpha^\phi - \kappa\end{aligned}\tag{1}$$

where $r(z) = p(z)y(z)$ stands for revenue.

The second type of firm is the single-layer firm ($L = 1$). The entrepreneur now becomes a *managerial* worker ($m_1 = 1$) and delegates production to n hired *production employees*. These produce zn units of output, but their efficiency is subject to a delegation adjustment. We assume that the employees' efficiency is now $(\alpha \frac{m_1}{n})^\theta = (\frac{\alpha}{n})^\theta$, $\theta \in [0, 1)$. The entrepreneur expends α units of efficiency labour to a task that can generally be interpreted as either monitoring, knowledge creation, or both. His effectiveness depends on time devoted to each production worker. As a result, effective output is $y_1(z) = zn \times (\frac{\alpha}{n})^\theta = z\alpha^\theta n^{1-\theta}$, and $1 - \theta$ becomes the standard span of control parameter. Production employees are paid a wage w and profits are

$$\begin{aligned}\pi_1(z) &= r_1(z) - wn - \kappa \\ &= R^{1-\phi} z^\phi \alpha^{\phi\theta} n^{\phi(1-\theta)} - wn - \kappa.\end{aligned}\tag{2}$$

The third organizational type is the multi-layered firm where $L \in \{2, 3, \dots\}$ denotes the total number of layers. Consider the two-layer firm, $L = 2$. The entrepreneur occupies the highest managerial position ($m_2 = 1$) and production employees again produce zn units of output. Their efficiency is now given by $\left[\frac{m_1}{n} \times \left(\alpha \frac{m_2}{m_1}\right)^\theta\right]^\theta = \left[\frac{m_1}{n} \times \left(\frac{\alpha}{m_1}\right)^\theta\right]^\theta$. Knowledge and supervision are processed by a mass m_1 of hired *managerial employees* in layer 1. Their effectiveness depends on the time devoted to each production employees, $\frac{m_1}{n}$, as well as on their own efficiency, $\left(\frac{\alpha}{m_1}\right)^\theta$. Analogously, production employees in a three-layer firm ($L = 3$) have efficiency $\left(\frac{m_1}{n} \times \left[\frac{m_2}{m_1} \times \left(\alpha \frac{m_3}{m_2}\right)^\theta\right]^\theta\right)^\theta$ with $m_3 = 1$, and so forth for $L \geq 4$. Generally, output of a multi-layer firm $L \geq 2$ is therefore

$$y_L(z) = z\alpha^{\theta^L} n^{1-\theta} \prod_{\ell=1}^{L-1} m_\ell^{(1-\theta)\theta^\ell}.$$

Managerial employees are paid firm-specific wages $\tilde{w}_l(z)$, depending on their position in the hierarchy $l \in \{1, 2, \dots, L-1\}$. Profits are

$$\begin{aligned}\pi_L(z) &= p(z)y_L(z) - wn - \sum_{l=1}^{L-1} \tilde{w}_l m_l - \kappa \\ &= R^{1-\phi} z^\phi \alpha^{\phi\theta^L} n^{\phi(1-\theta)} \prod_{\ell=1}^{L-1} m_\ell^{\phi(1-\theta)\theta^\ell} - wn - \sum_{l=1}^{L-1} \tilde{w}_l(z) m_l - \kappa.\end{aligned}\tag{3}$$

What is the trade-off associated with delegation? The gain consists of dampening the rate at which returns to scale decrease. As in the standard Lucas setting returns in output are decreasing because the firm is constrained by the entrepreneur's fixed input and the fact that the information flow across layers is imperfect as $\theta < 1$. Adding managerial layers permits the firm to increase its span of control by spreading supervision (or knowledge)

along the managerial hierarchy. In the limit where $L \rightarrow \infty$ we end up in a setting of constant returns to scale. The technological cost of delegation, on the other hand, stems from the necessity to employ more workers. Consider passing from $L = 0$ to $L = 1$ and let $n_e \equiv n/\alpha$ denote effective production employees. By writing $\pi_1(z) - \pi_0(z) = r_0(z)[n_e^{1-\theta} - 1] - \alpha w n_e$ two things become clear. First, adding an additional layer can only be profitable if, at the very least, $n_e > 1 \Leftrightarrow n > \alpha$. There is an implicit fixed labor cost of setting up an additional layer. Second, the implicit average (and marginal) cost of employing production workers is increasing in α . This mechanism holds for any increase in layers from $L-1 \geq 1$ to L as $\pi_L(z) - \pi_{L-1}(z) = r_{L-1}(z)[m_{e,L-1}^{(1-\theta)\theta^{L-1}} - 1] - \alpha w_{L-1}(z)m_{e,L-1}$ where $m_{e,L-1} \equiv m_{L-1}/\alpha$. The role of the entrepreneur's efficiency units of labor acquires economic meaning. The higher it is, the more costly is delegation. We label the overall mechanism of the organizational structure the generalized Lucas span of control model.²¹

3.1.1.2. Employee compensation. The other feature of interest is the incentive compatibility of middle managers. First, note that in this basic environment managerial workers are not *per se* different from their production peers. The assumption is that the nature of their work is different in that managerial employees can abuse their autonomy (e.g. in knowledge creation or in supervision) to divert resources from the firm, which production employees cannot. Concretely, we postulate that the firm's revenue $r(z)$ flows up the managerial hierarchy and that in every layer l each managerial employee handles a proportional part of it, namely $r(z)/m_l(z)$. Any such middle manager can then divert up to a fraction $1 - \lambda$ of that value without facing charges. The parameter $\lambda \in [0, 1]$ measures the quality of law enforcement or property protection in the economy. For instance, a low λ would well capture costly legal procedures that push up the threshold at which the firm is willing to take its employees to court. As for the revenue flow, take a supermarket chain to visualize that process. The cashiers collect revenue, but cannot divert any of it because their task is fully automated which allows for a perfect monitoring of their activity. Within a given supermarket, the cashiers pass revenue on to the store's group of sales managers whose autonomy allows them to divert a part of the receipts, for example by fudging their reports. The total revenue of a number of supermarkets is subsequently collected by the regional headquarters whose employees (including accountants, lawyers, sales directors) may again grab a fraction, and so forth up the residual claimant, the owner/entrepreneur.

We further assume that firms must take action to prevent stealing. The timing is as follows. At the beginning of the period the firm z offers workers an employment contract. Workers in production positions (n) are offered a wage w . Workers in middle management positions (m_l), meanwhile, are offered a contract that stipulates three elements. First, a period wage, $\tilde{w}_l(z)$, to which the firm can commit. Second, the threat of firing the manager if diversion is detected. Third, the promise to retain the manager when diversion is not detected. That promise can only be kept imperfectly as matches (conditional on the manager's survival) may break up exogenously with probability $\delta_m \in (\delta_a, 1]$.²² Commitment to the wage is a crucial assumption while the threat and the promise (up to δ_m) are credible and need not be assumed, as will become clear shortly. The manager may accept or decline the contract. Upon acceptance, the manager decides whether to

²¹Superficially, this structure is also reminiscent of Williamson (1967) for the case $\alpha > 1$, i.e. where the compliance of the workers in the layer adjacent to the entrepreneur is imperfect. However, in that paper compliance is imperfect between each layer of managers and subordinates, so that an infinite hierarchy leads to zero compliance of production workers and hence zero output.

²²The lower limit δ_a refers to the entrepreneur's (or firm's) probability of exiting the economy, rather than the manager's.

divert resources within the period. At the end of the period the firm pays out wages and perfectly detects any potential diversion of resources.

For the middle manager in position l at firm z the value of stealing is given by

$$V_l^{out}(z) = \tilde{w}_l(z) + \frac{(1-\lambda)r(z)}{m_l(z)} + \beta(1-\delta_a)V_h. \quad (4)$$

He amasses the period wage as well as diverted revenue, following which he is fired and enters the labor market once again. The value of not stealing is $V_l(z) = \tilde{w}_l(z) + \beta(1-\delta_a)[\delta_m V_h + (1-\delta_m) \max\{V_l(z), V_h\}]$. Barring exogenous separation, the manager can expect to keep the same position unless he prefers to re-join the labor market in expectation of better opportunities. In equilibrium - as demonstrated below - we have $V_l(z) \geq V_h$, so that

$$V_l(z) = \frac{\tilde{w}_l(z) + \beta(1-\delta_a)\delta_m V_h}{1-\beta(1-\Delta)} \quad (5)$$

where $1-\Delta \equiv (1-\delta_a)(1-\delta_m)$.

The firm offers middle managers just enough to prevent stealing, setting $V_l(z) = V_l^{out}(z)$. Combining (4) and (5) yields

$$\begin{aligned} \tilde{w}_l(z) &= \frac{1-\beta(1-\Delta)}{\beta(1-\Delta)} \frac{(1-\lambda)r(z)}{m_l(z)} + [1-\beta(1-\delta_a)]V_h \\ &= \frac{1-\beta(1-\Delta)}{\beta(1-\Delta)} \frac{(1-\lambda)r(z)}{m_l(z)} + (1+b)w. \end{aligned} \quad (6)$$

For convenience, we define $(1+b)w \equiv [1-\beta(1-\delta_a)]V_h$ where b is the expected managerial premium over the production wage. The wage offered to middle managers includes two parts. First, the period return flow of potentially diverted revenue where the factor includes the worker's time discount adjusted for his own probability of exiting the economy as well the separation rate from the firm. It is increasing in the amount that can be diverted in location (z, l) , i.e. the responsibility of a particular manager. Notice that it is implicitly also increasing in the quality of detection which is normalized to one time period. The second part is the opportunity cost of not participating in the labour market, which is the same for all managers as they have equal access in that market.

3.1.1.3. Summary. We are now ready to combine the organizational structure of the firm with employee compensation into one single profit function for $L \geq 1$. Let $B \equiv 1/[1-\beta(1-\Delta)] - 1 \geq 0$. Using (2), (3) and (9) obtains:

$$\begin{aligned} \pi_{L \geq 1}^*(z) &= \max_{L, n, m_l} \left\{ [1 - B(1-\lambda)(L-1)] R^{1-\phi} y_{L \geq 1}^\phi(z) - wn - w(1+b) \sum_{l=1}^{L-1} m_l \right\} - \kappa, \\ y_{L \geq 1}(z) &= z\alpha^{\theta^L} n^{1-\theta} \prod_{\ell=1}^{L-1} m_\ell^{(1-\theta)\theta^\ell}. \end{aligned} \quad (7)$$

This formulation clarifies the role of the contractual friction. On the external margin of optimal organizational choice it acts directly as a revenue tax on the firm. For $\lambda < 1$ that tax burden increases with each added layer beyond 1. Moreover, we will show that it is imperfect credit enforcement that can lead to $b > 0$ so that it also affects the firm's choice on the internal margin by driving a wedge between the relative cost production and managerial employees.

3.1.2. Labor market

Workers entering the labor market meet jobs randomly according to a set of endogenous probabilities. Managerial employment at firm z in layer $l \in \{1, 2, \dots, L(z) - 1\}$ yields value $V_l(z) = (1 - \lambda)/[\beta(1 - \Delta)]r(z)/m_l(z) + V_h$, obtained by plugging (9) back into (5). Note that the condition $V_l(z) \geq V_h$ - assumed previously - is indeed satisfied. Let $q_l(z)$ denote the probability of finding such a job. The other type of employment consists of production jobs resulting in value V_0 . Production employees are offered a period wage w and are free to search for other opportunities in the labor market in the following period, so $V_0 = w + \beta(1 - \delta_a) \max\{V_0, V_h\}$. This is independent of z . Moreover, since the value of a managerial job is weakly preferred to the expected value of entering the labor market, it must be that the value associated with production work is lower than the expected outcome in the labor market. It follows that $V_0 = w + \beta(1 - \delta_a)V_h$ and all managerial workers prefer their jobs to those of production employees. Let q_0 be the probability of ending up in production job. As search is frictionless, $q_0 + \int \sum_{l=1}^{L(z)-1} q_l(z) dz = 1$ and the expected value in the labor market is simply $V_h = q_0 V_0 + \int \sum_{l=1}^{L(z)-1} q_l(z) V_l(z) dz$.

The assumption that firms must prevent stealing is crucial. It leads to an ex post partitioning of workers who are ex ante identical. Consider the set of contracts $\mathcal{V}$ including V_0 as well as $V_l(z)$ for all equilibrium combinations $(z, l \geq 1)$, and order them in a sequence such that $V^1 \geq V^2 \geq V^3 \geq \dots \geq V_0$. As in any competitive market we assume that workers instantaneously meet firms at the beginning of each period. The lucky worker who is offered V^1 readily accepts that managerial contract as she knows that there is no better offer around. Meanwhile, the assumption that stealing must be prevented implies that the firm is willing to pay V^1 as any lower compensation would lead to diversion. Similarly, the individual that receives the second most lucrative contract V^2 also accepts the offer as the only more interesting contract is already taken. That logic continues until all managerial positions are filled instantaneously. All the remaining workers accept production jobs that offer the competitive wage w . This also elucidates why firms can credibly threaten to fire workers as well as retain them. They are indifferent between the exact identity of the worker they are dealing with. Yet, the combination of stick and carrot does lower the managerial wage bill which - from (9) - decreases in the extent to which firms can commit to keeping managers, Δ . If, on the other hand, firms could allow stealing to happen in equilibrium, we are back in the first-best world of the Coase conjecture. Firms would then treat all workers equally by negatively compensating diversion through wages and potentially even make some managerial workers pay for the right to work at the firm. One can argue that such an outcome is not possible, for instance by appealing to the fact that detection may be costly or because stealing in equilibrium would further depress profits through the loss of oversight. As it is, the equilibrium result here resembles in no small part to the efficiency wage argument of Shapiro and Stiglitz (1984) with the difference that the lower floor of unemployment is replaced by production employment.

3.2. Extension to human capital

Managerial workers are clearly not identical to production workers, nor are workers in distinct managerial positions the same. Jack Welch's talent may well be misallocated when running a burger joint. We will now introduce human capital to address that issue. We do so both to improve the empirical application, but also to show how the accumulation of human capital may itself react to contract enforcement.

Agents again enter the economy with an idea z . They are born unskilled and may decide, at the start of life, to become skilled at a cost $\psi \geq 0$. Skills are acquired instantaneously and do not depreciate through time. They open the door to jobs that require skills,

but skilled workers may just as well perform tasks for which no skills are necessary. Following the schooling choice individuals again decide between running their own firm procuring $V_f(z)$ and entering the labor market. The value associated with labor market entry is now distinct for unskilled vis-à-vis individuals: $V_{u,h}$ as opposed to $V_{s,h}$. We make the simplifying assumption that entrepreneurship does not require skills. Since skill accumulation is costly, entrepreneurship is compared with V_u and individuals with good ideas $z \geq \underline{z}$ become entrepreneurs. When there is positive demand for strictly skilled workers arbitrage requires that $V_{s,h} - \psi = V_{u,h}$ and newly born agents are indifferent between the two options.

3.2.1. Firms

As mentioned earlier, the entrepreneur's ability is independent of his skills. What is different now is that the effectiveness of employees depends on their skill content. The contribution of production employees to output - as opposed to n - is a bundle $n_u^\mu n_s^{1-\mu}$ where n_u (n_s) refers to the mass of employees performing unskilled (skilled) tasks, and $\mu \in [0, 1]$ expresses the relative importance of unskilled tasks. For firms that employ managerial workers in layer 1 we will similarly assume that rather than m_1 their contribution to output is a combination $m_{u,1}^\gamma m_{s,1}^{1-\gamma}$. The parameter $\gamma \in [0, 1]$ parsimoniously captures the idea that managing tasks are more skill-intensive than production tasks. Analogously, the contribution of managerial employees in the second layer is $m_{u,2}^{\gamma^2\mu} m_{s,2}^{1-\gamma^2\mu}$, and so forth for higher layers. The further management is situated from the production process, the higher is the necessary skill content. Notice that the interpretation of the organizational structure must be modified a bit. Rather than supervising individuals, managers now supervise bundles of individuals. Production equals:

$$y(z) = \begin{cases} y_0(z) = z\alpha & \text{if } L = 0, \\ y_{L \geq 1}(z) = z\alpha^{\theta^L} (n_u^\mu n_s^{1-\mu})^{1-\theta} \prod_{l=1}^{L-1} \left(m_{u,l}^{\gamma^l\mu} m_{s,l}^{1-\gamma^l\mu} \right)^{(1-\theta)\theta^l} & \text{if } L = 1, 2, \dots \end{cases} \quad (8)$$

The limitation to two skill levels obviously eases the exposition and the numerical computations. But it does capture in a nutshell the progression of skills through layers and hence reproduces the central result of Kremer (1993), namely that economies where human capital is relatively expensive choose organizational forms that are less complex.

As before we assume that managerial employees may divert resources while production employees may not. Equally, each managerial employee can steal a fraction of revenue that is proportional to the mass of managers in that layer, i.e. $(1-\lambda)r(z)/[m_u(z) + m_s(z)]$. The possibility to steal is only linked to managerial autonomy rather than to the skills of a particular manager. What distinguishes unskilled from skilled employees is their outside option ($V_{u,h}$, $V_{s,h}$) and so the compensation levels $\tilde{w}_{u,l}(z)$ and $\tilde{w}_{s,l}(z)$ are distinct. Let w_u and w_s denote the production wages of the respective skill level, and define by $(1+b_i)w_i \equiv [1 - \beta(1 - \delta_a)]V_{i,h}$, $\forall \{i = u, s\}$ each of the discounted expected period wages in the labor market. Repeating the game from the previous section yields the wages associated to each equilibrium pair $(z, l \geq 1)$:

$$\tilde{w}_{i,l}(z) = \frac{1 - \beta(1 - \Delta)}{\beta(1 - \Delta)} \frac{(1 - \lambda)r(z)}{m_{u,l}(z) + m_{s,l}(z)} + (1 + b_i)w_i, \forall i = s, u. \quad (9)$$

Finally, the profit function for employer firms ($L \geq 1$) is

$$\pi_{L \geq 1}^*(z) = \max_{L, n_u, n_s, m_{u,l}, m_{s,l}} \left\{ [1 - B(1 - \lambda)(L - 1)] R^{1-\phi} y_{L \geq 1}^\phi(z) - w_u n_u - w_s n_s - w_u(1 + b_u) \sum_{l=1}^{L-1} m_{u,l} - w_s(1 + b_s) \sum_{l=1}^{L-1} m_{s,l} \right\} - \kappa \quad (10)$$

where $y_{L \geq 1}(z)$ is obtained from (8). Optimal profits are $\pi^*(z) = \max\{\pi_0(z), \pi_{L \geq 1}^*(z)\}$.

3.2.2. Labor market

Analogous to the previous section, the value of entering the labor market is $V_{i,h} = q_{i,0}V_{i,0} + \int \sum_{l=1}^{L(z)-1} q_{i,l}(z)V_{i,l}(z)dz$, $\forall i = \{u, s\}$. The probabilities associated to obtaining different types of jobs are now conditional on the individual's skill level. Managerial positions demanding skills are filled up from the most to the least lucrative ones by skilled individuals as described in the previous section. The same is true for managerial positions that do not require skills. Remember that skilled individuals are supposed to be able to perform unskilled tasks as well. In equilibrium, however, firms will not fill unskilled managerial jobs with skilled workers because these are more expensive. This can be seen from (9) where the outside option for skilled workers is higher, $(1 + b_s)w_s \geq (1 + b_u)w_u$, given that $V_{s,h} - \psi = V_{u,h}$. Finally, the value of production employees is $V_{i,0} = w_i + \beta(1 - \delta_a)V_{i,h}$. The lower bound for the wage of skilled workers is given by $w_s \geq w_u$. If there is little demand for skilled production employees, markets clear by having skilled individuals perform production tasks that require no skills and $w_s = w_u$. Summarizing the different types of jobs yield, $\forall i \in \{u, s\}$ and $\forall z$:

$$V_{i,l}(z) = \begin{cases} V_{i,0} = w_i + \beta(1 - \delta_a)V_{i,h} & \text{if } l = 0, \\ V_{i,l}(z) = \frac{1-\lambda}{\beta(1-\Delta)} \frac{r(z)}{m_{u,l}(z) + m_{s,l}(z)} + V_{i,h} & \text{if } l = 1, 2, \dots, L(z) - 1. \end{cases} \quad (11)$$

4. CHARACTERIZATION

We can now consider more in detail several equilibrium outcomes. The analysis pertains to the extended environment with human capital, but applies equally to the basic environment for the case where skills are in zero demand ($\mu = \gamma = 1$).

4.1. Aggregate variables

There are five aggregate equilibrium objects: R , w_u , w_s , b_u and b_s . Aggregate revenue is simply $R = \int r(z)dF(z)$. The wage w_u ensures that the overall labor market clears so that all workers are employed or self-employed, namely $\int \mathbb{1}_{L \geq 1}(z)[n_u(z) + n_s(z)] + \sum_{l=1}^{L(z)-1} [m_{u,l}(z) + m_{s,l}(z)]dF(z) = F(z)$. The remaining three variables are determined by arbitrage on the acquisition of human capital and the consistency of the labor market premia b_u and b_s . From the arbitrage condition $V_{s,h} - \psi = V_{u,h}$ we have $(1 + b_s)w_s = (1 + b_u)w_u + [1 - \beta(1 - \delta_a)]\psi$. Moreover, there exists a closed form expression for the premia. Let $N_i^{e,S}$ denote the aggregate supply of production employees of skill level $i \in \{u, s\}$ and let $N_i^{e,D} = \int \mathbb{1}_{L \geq 1}(z)n_i(z)dF(z)$ denote the aggregate demand for production employees performing tasks of type $i \in \{u, s\}$.

Proposition 1. *In the stationary equilibrium the expected labor market premium over production work for worker skill type $i \in \{u, s\}$ is*

$$b_i = \frac{1 - \lambda}{w_i} \frac{\Delta}{\beta(1 - \Delta)} \frac{1}{N_i^{e,S}} \int \left(\sum_{l=1}^{L(z)-1} \frac{m_{i,l}(z)}{m_{u,l}(z) + m_{s,l}(z)} \right) r(z)dF(z). \quad (12)$$

Moreover, $N_u^{e,S} = (1 - a)N_u^{e,D}$ and $N_s^{e,S} = N_s^{e,D} + aN_u^{e,D}$ where the variable $a \in [0, 1)$ is such that $a = 0$ if $w_s > w_u$.

Proof. Appendix. □

The first part of the proposition has a neat interpretation. When there are no frictions, $\lambda = 1$, the premium b_i is naught as there is nothing to steal. This also holds for $\Delta \rightarrow 0$. If firms were perfectly capable to commit to retaining their managerial employees and if these had no prospect of exiting the economy ($\delta_m, \delta_a \rightarrow 0$) then workers entering the labor market would not expect to find any employment type other than production jobs. For $\lambda < 1$ the premium is increasing in the number of layers offering managerial employment, i.e. it is higher when many of the active firms choose a large number of layers $L(z)$. We gain a sense that b_i is likely not to be monotonic in $\lambda = 1$. As λ drops, there is more to steal conditional on obtaining a managerial job, but such jobs then become increasingly few and far between. This is reflected not only in the fact that the number of layers $L(z)$ is likely to drop, but also in that there is an increased supply of workers in non-managerial activities, $N_i^{e,S}$, which lengthens the queue to managerial positions and keeps a lid on b_i . What distinguishes the premium of unskilled from that of skilled workers is the relative demand for their services in managerial positions, $m_{i,l}(z)/[m_{u,l}(z) + m_{s,l}(z)]$. To the extent that we can expect the fraction of skilled managerial employees to be relatively high compared to skilled production employment it (and therefore $N_u^{e,S}$ to be relatively large compared to $N_s^{e,S}$) it is likely that $b_s > b_u$. Finally, observe the second part of the proposition. The equilibrium is unique, but it can be of two distinct types. The first is such that $a = 0$ as long as the arbitrage condition for human capital acquisition implies that $w_s \geq w_u$. Alternatively, if that condition implied $w_s < w_u$, an equilibrium with $a = 0$ is not feasible. Rather, the equilibrium demands an adjustment such that skilled and unskilled production employees earn the same period wage ($w_s = w_u$) with a fraction $a > 0$ of unskilled tasks being performed by skilled workers. Notice that there always exists an $a \in [0, 1)$ that makes this possible as a high enough value of a can always lead to a sufficient convergence between b_u and b_s . Such an equilibrium produces an additional type of misallocation and cannot be the outcome in a frictionless environment ($\lambda = 1$).

4.2. Firm choice

Proposition 2. *The optimal choice of layers $L^*(z)$ is finite, $\forall z$, in either of the following cases: (a) contract enforcement is imperfect ($\lambda < 1$); (b) the distribution of ideas $F(z)$ has finite support ($\bar{z} < \infty$). Moreover, $L^*(z)$ is weakly increasing in z .*

Proof. Appendix. □

The second part of the proposition captures the notion that it is the most efficient firms, i.e. those that optimally choose many layers, that are likely to be most negatively affected by a rise in contractual frictions. We cannot establish this for sure, however, because λ not only impacts the layer choice directly in the profit function (10) but also feeds through the aggregate variables discussed above. As for the first part of the proposition, notice that the organizational structure of the firm yields a non-trivial choice of layers even in the absence of any contracting issues. In that sense the employed production function may prove useful in other applications as well.

Proposition 3. *In the stationary equilibrium, managerial employees of skill level $i \in \{u, s\}$ at firm z in layer $l \geq 1$ earn a wage*

$$\tilde{w}_{i,l}(z) = \frac{1}{\theta^l} \frac{1}{1 + \gamma^l \mu \left[\frac{(1+b_s)w_s}{(1+b_u)w_u} - 1 \right]} \frac{1}{1 - B(1-\lambda)[L(z)-1]} \frac{(1-\lambda)[1-\beta(1-\delta)]}{\beta(1-\delta)\phi(1-\theta)} (1+b_s)w_s + (1+b_i)w_i.$$

It is increasing in the manager's skill level i , hierarchical position l as well as the total number of layers $L(z)$.

Proof. Appendix. □

Skilled managers earn more than unskilled managers because their outside option is higher, $(1-b_s)w_s \geq (1+b_u)w_u$. For a given level of skills, the managerial wage is then determined by the first three terms of the above equation. The first one determines the managerial layer. Higher managerial layers optimally demand relatively fewer employees leaving more revenue per person to be diverted by each of them. The second term captures the fact that higher managerial layers demand relatively more skilled workers. Again, because these are relatively expensive, firms optimally employ even fewer managerial workers in higher echelons of the hierarchy. The third term relates to the fact that a higher number of total layers $L(z)$ constrains the firm's profitability per worker. Firms that optimally choose many layers $L(z)$ employ relatively few workers (including managerial workers) per unit of revenue, which again increases the amount that each managerial worker can grab. The last term is in some sense the most interesting one as it delivers higher managerial wages in large firms (those with more layers), conditional on any given managerial position. Store managers at Walmart must be compensated more than those at smaller supermarket chains.

Proposition 4. *Let $x(z)$ denote the fraction of managerial workers. In the stationary equilibrium employer firms, $L(z) \geq 1$, choose*

$$x(z) = \frac{1}{1 + \frac{\mu + (1-\mu)\frac{w_u}{w_s}}{\frac{\mu}{n_u} + \sum_{l=1}^{L(z)-1} \theta^l \left[\frac{\gamma^l \mu}{1+b_u} + \frac{(1-\gamma^l \mu)w_u}{(1+b_s)w_s} \right]}}.$$

Revenue productivity, i.e. revenue per worker, is then proportional to $\frac{1-x(z)}{1-B(1-\lambda)[L(z)-1]}$.

Proof. Appendix. □

This provides us with two more observables to be tested empirically. The firm's size pushes the fraction of managerial workers in two opposite directions. It is decreasing in n_u because of the vanishing overhead cost associated with the entrepreneur's time, and increasing in the choice of $L(z)$ as each further layer acts as an additional overhead cost. As firms grow larger and both n_u and $L(z)$ rise, the qualitative evolution of the managerial share is unclear and remains an empirical question. The model is ultimately allowed to fail as our goal consists of measuring the drop in the (aggregate) managerial share due to a decline in enforcement. Now, turn to the second part of the proposition. In the absence of contracting frictions ($\lambda = 1$), revenue productivity is decreasing in the fraction of managerial workers. If the managerial share rises as firms add layers we would expect revenue productivity to fall. The presence of contracting frictions ($\lambda < 1$) and the resulting implicit revenue tax in the number of layers can overturn that result. The well-known fact that larger firms have higher revenue productivity may be related to the presence of higher efficiency wages in large organizations with limit the optimal amount of revenue per worker.

5. CALIBRATION

5.1. *Parameter choice*

The model is calibrated to the U.S. around the year 2005, proceeding in two steps. First, I fix three parameters. The length of a period is set to a year, assuming that to be the time required by firms to detect expropriation. For instance, a year is a suitable time unit to evaluate managerial performance and adjust medium-term projects. We fix the average length of a worker's labor market participation to be roughly 40 years, which corresponds to an exit rate of $\delta_a = 0.025$. To match an implicit annual interest rate of 4 to 5 percent, the time discount factor β is then set to 0.975. The exogenous separation rate is set to $\delta_m = 0.152$. This yields an average job tenure (conditional on the worker's survival) of 6.6 years as computed for the U.S. over the nineties by Auer, Berg and Coulibaly (2005).²³

In a second step I jointly calibrate ten parameters to minimize the sum of squared (percent) errors between model-generated moments and targeted values, reported in Table (2). The distribution $F(z)$ is assumed to be log-normal such that $\log z \sim \mathcal{N}(0, \sigma^2)$. In the absence of the possibility to add managerial layers, the firm size distribution would directly inherit the statistical properties of z and therefore feature a thin tail. This is at odds with the evidence as the right tail of the U.S. firm size distribution closely follows a Pareto distribution. The possibility of adding layers, however, implies that firms can locally increase returns to scale, which thickens the right tail. According to Luttmer (2007) the proportion of firms with more than x workers approximately equals $x^{-1.06}$. To match this I regress the log of the inverse distribution of firms on the log of firm size to back out the slope. If the distribution were perfectly Pareto, the associated R^2 would equal 1; we content ourselves with matching the average slope of -1.06 . The choice of σ strongly impacts that outcome. To discipline the size of the largest firms I vary $\bar{z}$ to match the fact that firms with more than 10,000 employees account for about 25.4 percent of overall employment as computed from summary statistics of the U.S. Census for 2008.²⁴

The share of managerial workers is mainly shaped by the span of control parameter θ . I use the empirical definition of managerial workers from Section 2 that yields 35 percent for the U.S. based on ILO data. The model definition includes all managerial entrepreneurs and employees.²⁵ The resulting $\theta = 0.499$ implies rather strong overall decreasing returns to scale $(1 - \theta)$ and substantial gains from adding layers. Turning to human capital, the cost of ψ regulates the expected skill premium, i.e. $w_s(1 + b_s)/[w_u(1 + b_u)] - 1$. As in Section 2 I define skilled workers as those with attempted tertiary education and above. A reasonable target is therefore the college premium as our definition includes both worker with more and less than completed college. The targeted value is 0.311, implied from a typical expected premium associated with a year of schooling derived from

²³The calibration is sensitive to the choice of the time period and the associated total discount factor. It is not, however, strongly sensitive to the exact values of δ_a and δ_m because variations in these parameters are mostly picked up by changes in the parameters ψ and λ . To a first-order approximation these can be viewed as normalizations.

²⁴Available at: <https://www.census.gov/econ/smallbus.html>. According to the Business Census the employment share of employer firms with at least 10,000 workers is 27.3 percent. We adjust this number downward to account for the additional 7 percent of own-account workers not included there.

²⁵This implies that we treat own-account workers as non-managerial workers. Since in the data many own-account workers are actually managers according to our occupational definition, an alternative would be to target the fraction of managerial employees only. According to that definition, the model delivers a managerial share of 0.345, i.e. practically the same.

Paramater	Value	Target	Data	Model
Discount factor (β)	0.975	Interest rate	$\approx 5\%$	-
Exit rate (δ_a)	0.025	Worklife	≈ 40 yrs	-
Separation rate (δ_m)	0.152	Average job tenure	6.6 yrs	-
Std deviation of $\log z$ (σ)	0.534	Pareto tail of firm CDF	-1.060	-1.074
Maximum efficiency ($\bar{z}$)	15.65	Empl. share large firms	0.254	0.254
Span of control (θ)	0.499	Managerial share	0.350	0.350
Skill acquisition cost (ψ)	2.136	Expected skill premium	0.311	0.310
Unskilled intensity (μ)	0.662	Share skilled employees	0.531	0.531
Skill progression (γ)	0.198	Uncond. vs. cond. premium	0.007	0.00696
Expropriation (λ)	0.759	Uncond. size-wage premium	0.026	0.0264
Goods substitutability (ϕ)	0.902	Profit share	0.20	0.191
Entrepreneur efficiency (α)	0.193	Mean employer firm size	20.4	20.53
Fixed cost (κ)	0.036	Share own-account entre.	0.07	0.071

TABLE 2. Benchmark calibration

Mincer regressions, and employed in Hall and Jones (1999).²⁶ According to the Barro-Lee dataset, in the U.S. in 2005 there were about 53.1 percent of workers that fit out empirical definition. In the model, of course, entrepreneurs are all unskilled, which would be clearly at odds with the data. To circumvent that problem I target the ratio of skilled to total employees. Moreover, I target the supply of such employees rather than the demand for skills. There is a differences between these two as in equilibrium some skilled workers end up working as unskilled production employees. The parameter μ strongly influences that outcome.

The most tricky bit of the calibration is the parameter choice of contract enforcement, λ , and the progression of skills through layers, γ . I use the following method. One of the most complete empirical studies on the firm size-wage premium is Troske (1999). There, the use of employer-employee matched data allows to establish the elasticity of wages with respect to firm size, both by controlling for and by abstracting from employee characteristics (0.026 versus 0.033). I similarly run two regressions on model-generated data. I compute average wages for skilled and unskilled employees for each employer firm ($L \geq 1$) and regress them on firm size, first by controlling for skills and then by ignoring them. All variables are in logs and the method is weighted least squares where the observation weights are the number of employees by skill type and firm. The resulting coefficients β_0 and β_1 are then targeted to yield $\beta_0 = 0.026$ and $\beta_1 - \beta_0 = 0.033 - 0.026 = 0.007$. The parameter λ (set to 0.759) influences the residual of the wage-firm size premium that is not accounted for by skills, β_0 , while the difference $\beta_1 - \beta_0$ stems mostly from γ (set to 0.198), which regulates the average increase in skill-intensity by firm size. For $\lambda = 1$ we would therefore end up with $\beta_0 = 0$. That method of course implies that λ encompasses reasons other than the risk of expropriation that may push towards compensating wage differentials in larger firms. This is not too much of a concern here because our ultimate interest consists of measuring the economy's response to a *change* in the parameter λ that is assumed to originate entirely from differences in contractual frictions, keeping other things equal. Note that the calibrated value of λ is not too far off from the U.S. value of the proxy measure of contract enforcement discussed in Section 2, which is 0.856. The intensity of unskilled labor, on the other hand, drops rather quickly through layers at the calibrated value of γ . This is likely to be a by-product of the very parsimonious

²⁶These authors uses a premium of 0.07 for an extra year of tertiary education based on findings in OECD countries, so $1.07^4 = 1.311$.

choice of only two skill levels. Finally, the wage premia associated with various levels of tertiary education found in Troske (1999) have a mean value of 0.33, which is just about the same as the skill premium targeted here. That gives us confidence that his dataset is representative for the economy overall.

There are three remaining parameters. I set $\phi = 0.902$, which has its natural empirical counterpart in the average profit share, $\int \pi(z) dF(z) / [\int y(z) dF(z)]$. I aim at a profit share of 0.2, which lies in between 0.10-0.15 found in Atkeson and Kehoe (2005) and higher mark-ups based on measures of substitutability between varieties such as the value of 0.33 used in Hsieh and Klenow (2009). The parameter α governs to a large extent the profitability of passing from own-account to employer firms. It is set to 0.193 to match the average employer firm size, 20.4, following U.S. Census data. As for κ , it is set to 0.036 to match the average fraction of own-account workers, which is 7 percent according to Hipple (2010).

5.2. Cross-sectional outcomes

The quantitative model delivers the following results. The left panel of Figure (2) presents the relationship between firm efficiency z and its size in terms of number of workers. Abstracting from overhead costs, it is strictly log-linear in standard span of control models. Here, it is also log-linear for any given amount of layers. With the endogenous layer choice, however, the slope of the relation changes and firm size becomes log-convex in z . This is due to the larger span of control as well as the change in the skill intensity across layers. Now consider the threshold levels at which firms choose to add layers, depicted in the right panel. The smallest employer firms have slightly more than two workers, i.e. they hire one additional production worker. Delegation to an intermediate management level occurs at firms with more than six workers. The next thresholds are then roughly 10, 40, 300, and 10,000. According to that quantification, additional delegation becomes profitable very quickly as firms grow larger and then slowly flattens off, with the largest firms employing five intermediate management levels between the entrepreneur and production workers. Whether such a progression of layers is plausible depends of course on the exact definition of a layer. One empirical counterpart is the classification that businesses themselves use, as for instance in Baker, Gibbs and Holmstrom (1994) who describe a large firm that is organized in eight layers. Another empirical counterpart is to follow statistical occupational codes as in Caliendo et al. (forthcoming) who identify five distinct layers in the largest firms from French census data. Our result of a maximum of seven distinct layers lies safely in between these two interpretations. The log of the inverse cumulative firm size distribution is depicted in the right panel of Figure (2), together with the approximate empirical Pareto distribution targeted in the calibration. The model is neatly capable of reproducing a thick-tailed distribution, up to the last layer where the properties of log-normality kick in. In the data the last bit of the tail also thins, but somewhat less than here.²⁷

One interesting aspect of this quantitative exercise is the wage premium associated with large firms. The left panel of Figure (3) presents the average wages of skilled and unskilled employees at each firm. The first thing to note is that production workers earn the same amount regardless of their skills, i.e. $w_u = w_s$. Our calibration hence suggests that the market for skilled workers does not clear and that a fraction of unskilled productions tasks are performed by workers who are overqualified (precisely, 30 percent).²⁸ The skill

²⁷The largest firm employs 360 thousand workers while empirically the largest private firm is Walmart with 2.2 million employees. Yet, even so, in the data there are only ten firms that employ more than 300,000 workers (including overseas workers).

²⁸That fraction is assumed to be the same across all firms.

FIGURE 2

premium arises entirely from managerial positions and the hold-up associated with them. The expected premium over and above the production wage is $b_u = 0.19$ for unskilled and $b_s = 0.56$ for skilled workers. The difference is visualized by the fact that expected wages of skilled employees increase substantially more in firm size than those of unskilled employees. This is because skilled workers dominate in managerial positions. The actual wages of skilled vis-à-vis unskilled managerial workers can be read from the right panel of Figure (3) which depicts the wages associated with each layer. From here we see that the highest managerial position pays skilled (unskilled) workers a wage that is 10.7 (10.2) times higher than the production wage. Finally, note how the wages associated with any given managerial layer rise in the total amount of layers chosen by the firm. That increase is non-negligible, but there are no reversal in the sense that none of the managers prefers to work in a lower layer at a larger firm.

FIGURE 3

The left panel of Figure (4) presents the share of managerial workers and skilled employees by firm. By definition, own-account firms have no managers. From there, the fraction of managerial workers is increasing, barring the overhead cost of the entrepreneur himself.

The rise in the fraction of managerial workers is not immense, however, with the largest firms employing a bit less than 40 percent of their workforce in managerial activities. As for the fraction of skilled employees (right panel), the increase is quite steep for smaller firms before plateauing at around 55 percent. Finally, the right panel shows the evolution of revenue productivity across firm size. Note that beyond own-account and single-layer firms, productivity increases in firm size. The largest firms beyond ten thousand employees are on average roughly 20 percent more productive than small firms of worker size ten. While in the data that productivity difference is larger - about 80 percent according to U.S. Census data - the fact that there is a perceptible increase in productivity is deemed a successful robustness feature of the calibration. It is not targeted, and the model's mechanism in absence of contracting frictions would lead to a negative correlation between productivity and firm size. With our choice of parameter combination the contracting friction dominates and pushes the correlation in the right (qualitative) direction.

FIGURE 4

6. SIMULATIONS

6.1. Variations in law enforcement

The following graphs depict the aggregate effect of varying enforcement λ away from the benchmark calibration. For informational purposes we consider economies on the whole support of $\lambda \in [0, 1]$ while the main focus is on the outcome from lowering λ within a plausible range. Figure (5) summarizes the main quantitative findings of the paper. Consider first the aggregate share of managerial workers in the left panel. The model neatly reproduces the qualitative positive correlation between managers and contract enforcement from Section 2. The lowest possible share of 22 percent means that even in the absence of any legal protection, managerial work is well bounded away from zero. This is because employer-entrepreneurs are counted as managers, but also because the relatively quick detection rate coupled with the firm's contracting game keeps managerial wages in check. What is a suitable range for lowering λ ? We use the findings from Section 2 providing us with the expected managerial share associated with the "typical" legal quality in low-income countries. The vertical lines in the Figure correspond to the values of countries at income levels of 25 percent (dotted), 10 percent (dotted-dashed), and 5 percent (dashed) of the U.S. benchmark (solid). The corresponding values of λ are

0.637, 0.556, and 0.495 as opposed to the benchmark case of 0.759. The drop in contract enforcement is therefore not predicted to be ridiculously high. All the remaining outcomes discussed below are highlighted at precisely these levels of λ . The main such outcome is GDP, depicted in the right panel of Figure (5). The output loss is not negligible. Compared to the U.S., the predicted GDP percentage point gap is about 11, 18, and 23, respectively, for representative middle, low and very low-income economies. It is a loss that results almost entirely from misallocation of resources. The model does feature endogenous accumulation of skills that provides leverage through a propagation channel, but that mechanism is weak as we will see shortly.

FIGURE 5

There are several statistics that yield verifiable prediction. First, consider firm size. The left panel of Figure (6) draws the average size of employer firms ($L \geq 1$) and the share of own-account workers ($L = 0$).²⁹ The average number of workers in employer firms drops steeply as λ decreases. As contractual frictions rise, firms opt for fewer layers, which in turn lowers demand for workers and hence wages w . Observe from the right panel that the largest firms now choose 5 rather than 6 layers, and that the largest single firm decreases in size from 360,000 to a plausible 60,000 in the bottom scenario. With the reduction in wages the supply of workers shrinks as entrepreneurship becomes more lucrative at the margin. Yet this need *not* imply an increase in the share of own-account firms. After all, the jump from own-account to single-layer employer firms also becomes increasingly profitable - these firms face lower production wages and are not exposed to the risk of expropriation. Combined with the reduction in worker supply, the average employer firm size thus passes from 20 in the benchmark to 11, 8 and 7, respectively, in the range of interest. Meanwhile, the force pushing towards a higher share of own-account workers dominates, driving it from 7 to 10, 12 and 14 percent, respectively. In summary, the model delivers a powerful mechanism that helps explain the low average size of employer firms and the prevalence of own-account businesses in poorer countries. Finally, note that in absence of contract frictions, the maximum number of layers in the largest firms sharply increases to 20. This is not, however, accompanied by an unrealistic explosion in the size of the largest firm, which increases to just under a million workers.

²⁹For reference, the total share of self-employed entrepreneurs e equals $e = o + (1 - o)/x$ where o is the share of own-account workers and x the average employer firm size. The average firm size *including* own-account workers is simply $1/e = x/[1 + o(x - 1)]$.

FIGURE 6

Next we turn to managerial compensation. The left panel of Figure (7) depicts the evolution of the expected managerial wage premium for skilled (b_s) and unskilled (b_u) employees. Both are hump-shaped. Compensation is zero when enforcement is perfect, and it is relatively low when there is no enforcement so that managerial jobs are scarce. The action happens in the middle. For the range of interest, the model predicts a reduction in b_u and a slight increase in b_s . What is perhaps of more empirical interest is the firm size elasticities depicted in the right panel. The wage-firm size elasticity conditional on skills - i.e. the pure premium for working in a larger firm - is predicted to be monotonically decreasing in λ . To give an example, consider the implicit wage gain in switching from a firm with ten workers to one of ten thousand. It is 26 percent in the benchmark, and 46, 65 and 79 percent, respectively, in the other three scenarios. Such values are not outlandish, but at this point cannot be verified given the lack of detailed empirical studies of the firm size-wage premium over a wide range of countries. Another highly interesting statistic is the revenue productivity-firm size elasticity (across employer firms only). Surprisingly, it is almost exactly naught at $\lambda = 1$, which is an artefact of the parameter choice. Presumably, the prevalence of large firms with many layers washes away minor differences in revenue productivity across firms. As λ drops the elasticity generally increases, albeit not monotonically due to jumps in the maximum number of layers chosen by the largest firms. Within our range of interest we have that the productivity difference between firms with 10,000 rather than 10 workers ought to pass from 17 to about 34 percent. Again, data on this are hard to get by. But one indication that the span of revenue productivity between small and large firms is indeed larger in developing countries can be found in Hsieh and Klenow (2009) who find that their measure of TFPR increases more strongly in size in China and India than in the U.S.

Our final concern is the accumulation of skills, depicted Figure (8). The left panel shows that the share of skilled employees is an increasing function of enforcement, except at the very end. This happens while the expected skill premium $w_s(1 + b_u)/[w_u(1 + b_s)]$ moves steadily in the opposite direction. There are two forces at work. On the one hand, poorer contract enforcement makes managerial workers, and therefore disproportionately skilled ones, increasingly expensive. The demand for them drops. On the other hand, the reduction in wages w coupled with the assumed expenditure (or utility) cost of skill accumulation imply that the skill premium is required to rise further to make investing

FIGURE 7

in skills worthwhile. The model framework can therefore add to our understanding of two intertwined stylized facts in development accounting, namely the low level of human capital and the high skill premium in poorer countries. Low human capital accumulation, however, is a relatively weak propagator. This can be seen in the right panel, which presents the relative share of skilled employees engaged in unskilled tasks. Remember that in the benchmark calibration 30 percent of unskilled production tasks are undertaken by skilled workers as $w_u = w_s$. Altogether, this implies that about 37 percent of skilled workers are “underemployed” in unskilled tasks in the benchmark. Notice that over most of the support of λ the economy exhibits skill mismatch, except at the low and high ends. As for our range of interest, we have that the skill mismatch persists and then drops by about 8 percentage points at the lower limit. In summary, very poor economies are thus predicted to have fewer skilled workers (45 as opposed to 53 percent), but the fraction of these that are mismatched is lower as well (29 as opposed to 37 percent). Overall, the accumulation channel through skills is therefore not an important driver behind GDP losses.

FIGURE 8

6.2. Variations in law enforcement and other parameters

Our final interest is the interplay between variations in contractual frictions and variations in other characteristics typically associated with poorer economies. We consider two. The first is the cost of human capital accumulation (ψ). Apart from the benchmark we study three distinct levels of ψ that increase the expected skill premium $w_s(1 + b_s)/[w_u(1 + b_u)]$ to 1.33, 1.66 and 2 times the calibrated level, respectively, keeping other parameters constant. Such returns to skills are in the range of the findings by Psacharopoulos (1994) who uncovers significantly higher premia in low-income countries. Parameter ψ may capture many factors that influence the expenditure cost of schooling such as financial constraints on human capital accumulation, the quality of public schooling, school subsidies, and fertility. The other variation is TFP. Let the entrepreneur's efficiency be $\tilde{z} = A \times z$ where z follows the established statistical properties. The benchmark case is $A = 1$, and the other values are such that GDP matches 0.25, 0.1 and 0.05, respectively, of the calibrated level. These variations are coupled with changes in λ , more precisely to the values of interest found in the previous section.

6.2.1. Cost of skill acquisition

Statistic	$\psi =$ 2.136	$\psi =$ 2.697	$\psi =$ 3.218	$\psi =$ 3.730		$\psi =$ 2.136	$\psi =$ 2.697	$\psi =$ 3.218	$\psi =$ 3.730
Managerial workers (%)					GDP (benchmark=1)				
$\lambda = 0.759$	35.0	33.9	32.6	31.4	$\lambda = 0.759$	1.000	0.992	0.981	0.969
$\lambda = 0.637$	32.1	30.7	29.4	28.2	$\lambda = 0.637$	0.881	0.876	0.862	0.846
$\lambda = 0.556$	30.3	28.8	27.6	26.6	$\lambda = 0.556$	0.820	0.805	0.789	0.773
$\lambda = 0.495$	29.0	27.6	26.5	25.6	$\lambda = 0.495$	0.773	0.758	0.742	0.728
Avg. employer firm size					Own-account workers (%)				
$\lambda = 0.759$	20.5	19.6	18.3	16.8	$\lambda = 0.759$	7.1	9.2	10.8	12.1
$\lambda = 0.637$	10.9	10.3	9.7	9.0	$\lambda = 0.637$	9.6	12.2	14.0	15.3
$\lambda = 0.556$	8.3	7.8	7.3	6.8	$\lambda = 0.556$	12.1	14.7	16.3	17.3
$\lambda = 0.495$	7.0	6.6	6.2	5.9	$\lambda = 0.495$	14.3	16.5	17.7	18.7
Max. layers					Largest firm ('000)				
$\lambda = 0.759$	6	6	6	6	$\lambda = 0.759$	360	348	336	325
$\lambda = 0.637$	5	5	5	5	$\lambda = 0.637$	163	154	146	137
$\lambda = 0.556$	5	5	5	5	$\lambda = 0.556$	102	93	86	79
$\lambda = 0.495$	5	5	4	4	$\lambda = 0.495$	60	54	39	36
Premium b_u (%)					Premium b_s (%)				
$\lambda = 0.759$	19.4	13.5	10.0	7.9	$\lambda = 0.759$	56.5	60.7	66.9	75.1
$\lambda = 0.637$	19.5	13.7	10.4	8.5	$\lambda = 0.637$	62.9	69.0	77.4	87.5
$\lambda = 0.556$	16.2	11.4	8.9	7.5	$\lambda = 0.556$	63.1	71.2	81.4	93.0
$\lambda = 0.495$	13.1	9.4	7.6	6.7	$\lambda = 0.495$	62.1	72.1	83.7	95.0
Elast. wage (%)					Elast. productivity (%)				
$\lambda = 0.759$	3.34	3.43	3.56	3.75	$\lambda = 0.759$	2.31	2.35	2.41	2.48
$\lambda = 0.637$	5.53	5.72	5.98	6.28	$\lambda = 0.637$	3.26	3.34	3.45	3.56
$\lambda = 0.556$	7.25	7.55	7.91	8.29	$\lambda = 0.556$	4.41	4.51	4.62	4.74
$\lambda = 0.495$	8.41	8.81	9.24	9.67	$\lambda = 0.495$	4.17	4.17	4.19	4.38
Skilled empl. (%)					Skill premium (%)				
$\lambda = 0.759$	53.1	52.5	51.7	50.8	$\lambda = 0.759$	31.0	41.5	51.8	62.2
$\lambda = 0.637$	49.7	48.8	47.8	46.8	$\lambda = 0.637$	36.4	48.7	60.6	72.8
$\lambda = 0.556$	47.5	46.4	45.3	44.1	$\lambda = 0.556$	40.4	53.7	66.6	79.5
$\lambda = 0.495$	45.9	44.7	43.5	42.2	$\lambda = 0.495$	43.3	57.3	70.6	84.1
Skill mismatch (%)									
$\lambda = 0.759$	37.0	32.3	26.6	20.4					
$\lambda = 0.637$	37.2	30.0	22.2	14.6					
$\lambda = 0.556$	33.9	24.1	14.8	6.5					
$\lambda = 0.495$	29.0	17.5	7.7	0					

TABLE 3. Law enforcement versus cost of skill acquisition

Table (3) summarizes the most relevant statistics. The share of skilled employees naturally decreases with the rising cost of skill acquisition. Managerial work predictably drops as it is relatively skill-intensive. The positive correlation between the managerial

share and the share of skilled employees - keeping law enforcement fixed - can be directly compared with the empirical findings in Section 2. The qualitative effect is therefore identical to a drop in λ , though the force exerted by contractual frictions is quantitatively stronger. Another noticeable aspect is that the covariant effect from λ and ψ is subdued. This is true for most of the other statistics as well. One moment where there is a stronger covariant effect is GDP. The sole increase in ψ - at the benchmark level of expropriation - has a minor effect on GDP, presumably because there is significantly less skill mismatch. At the lowest level of λ , however, a high cost of skill acquisition is particularly pernicious for aggregate income because attenuation through reduced skill mismatch reaches its limit. In other words, countries with very expensive skill acquisition may see their GDP drop disproportionately with poor law enforcement. One final concern is to consider statistics that allow to disentangle contractual frictions from the cost of skill acquisition. The model, after all, predicts very similar qualitative effects associated with either. In quantitative terms, however, there are some differences. The most important one is the size of the largest firm (or more generally employment in the largest firms). Expropriation drastically reduces employment in such firms while the cost of skills generates a more muted response.

6.2.2. TFP

Statistic	$A = 1.000$	$A = 0.284$	$A = 0.153$	$A = 0.106$		$A = 1.000$	$A = 0.284$	$A = 0.153$	$A = 0.106$
Managerial workers (%)					GDP (benchmark=1)				
$\lambda = 0.759$	35.0	28.7	23.2	16.8	$\lambda = 0.759$	1.000	0.250	0.100	0.050
$\lambda = 0.637$	32.1	27.2	21.6	15.2	$\lambda = 0.637$	0.881	0.213	0.081	0.042
$\lambda = 0.556$	30.3	26.1	20.9	14.6	$\lambda = 0.556$	0.820	0.192	0.074	0.039
$\lambda = 0.495$	29.0	25.3	20.5	14.2	$\lambda = 0.495$	0.773	0.180	0.070	0.038
Avg. employer firm size					Own-account workers (%)				
$\lambda = 0.759$	20.5	10.3	7.8	9.4	$\lambda = 0.759$	7.1	4.8	0	0
$\lambda = 0.637$	10.9	6.6	6.3	8.3	$\lambda = 0.637$	9.6	5.6	0	0
$\lambda = 0.556$	8.3	5.5	5.9	8.1	$\lambda = 0.556$	12.1	7.0	0	0
$\lambda = 0.495$	7.0	5.1	5.7	7.9	$\lambda = 0.495$	14.3	8.2	0	0
Max. layers					Largest firm ('000)				
$\lambda = 0.759$	6	6	6	6	$\lambda = 0.759$	360	289	204	147
$\lambda = 0.637$	5	5	5	5	$\lambda = 0.637$	163	112	67	45
$\lambda = 0.556$	5	5	4	4	$\lambda = 0.556$	102	59	24	18
$\lambda = 0.495$	5	4	4	4	$\lambda = 0.495$	60	28	17	12
Premium b_u (%)					Premium b_s (%)				
$\lambda = 0.759$	19.4	6.0	7.6	8.2	$\lambda = 0.759$	56.5	107.5	64.0	37.8
$\lambda = 0.637$	19.5	7.6	7.9	7.3	$\lambda = 0.637$	62.9	90.4	40.0	19.4
$\lambda = 0.556$	16.2	7.7	7.0	6.1	$\lambda = 0.556$	63.1	69.5	25.3	11.8
$\lambda = 0.495$	13.1	7.4	6.1	5.2	$\lambda = 0.495$	62.1	56.1	18.8	8.2
Elast. wage (%)					Elast. productivity (%)				
$\lambda = 0.759$	3.34	4.75	5.33	5.69	$\lambda = 0.759$	2.31	2.95	3.30	3.21
$\lambda = 0.637$	5.53	7.13	8.03	8.93	$\lambda = 0.637$	3.26	4.19	4.78	4.82
$\lambda = 0.556$	7.25	8.85	10.04	11.02	$\lambda = 0.556$	4.41	5.22	5.01	5.33
$\lambda = 0.495$	8.41	10.08	11.44	12.36	$\lambda = 0.495$	4.17	5.21	6.21	6.49
Skilled empl. (%)					Skill premium (%)				
$\lambda = 0.759$	53.1	39.5	17.4	7.7	$\lambda = 0.759$	31	164	521	1115
$\lambda = 0.637$	49.7	31.6	11.4	5.0	$\lambda = 0.637$	36	202	647	1382
$\lambda = 0.556$	47.5	26.2	8.9	4.0	$\lambda = 0.556$	40	230	721	1489
$\lambda = 0.495$	45.9	22.9	7.7	3.6	$\lambda = 0.495$	43	248	761	1548
Skill mismatch (%)									
$\lambda = 0.759$	37.0	0	0	0					
$\lambda = 0.637$	37.2	0	0	0					
$\lambda = 0.556$	33.9	0	0	0					
$\lambda = 0.495$	29.0	0	0	0					

TABLE 4. Law enforcement versus TFP

The variations in TFP are summarized in Table (4). These changes are large and the associated wages w very low so that the necessary skill premium skyrockets because

of the assumed expenditure cost of skill accumulation. As a result there are very few skilled workers, few managerial workers and no mismatch in skills. Also, the share of own-account workers goes to zero. The reason for that is the presence of the operating cost κ . As wages and TFP fall in tandem, the threshold $\underline{z}$ does not change much. Yet the drop in wages also implies that individuals with a high enough efficiency to overcome the operating cost all become employers, lowering the average size of employer firms. In that sense, the share of own account workers is one important variable that qualitatively distinguishes a reduction in contract enforcement from a reduction in TFP. As for GDP changes, note that at the intermediate level of $A = 0.284$ the proportional losses due to expropriation are the largest, while the ones at the lower end of the TFP spectrum are comparable to those at the benchmark.

7. CONCLUSION

This paper explores the effect of law enforcement on firm organization, selection, skill accumulation and ultimately GDP per capita. A drop in law enforcement generating plausible changes in the aggregate managerial share is shown to replicate many features characteristic of poor countries, i.e. those at income levels of between 5 and 25 percent of the U.S. The modeling choice is arguably conservative in that firms - in conjunction with the general equilibrium feedback - do find means to circumvent contractual frictions to some extent. Skill-mismatch, for instance, decreases as contractual frictions start to bite significantly. Still, the attendant productivity losses turn out to be sizeable. The simplicity of the main building blocks - the generalized Lucas span of control production function and the structure of the dynamic game - allows for many exciting extensions in future work. Examples include the firm's limited commitment in its dealings with third parties, e.g. via market for credit or intermediate inputs; tenure progression and accumulation of experience; dynamic changes in firm efficiency. There is some hope that the study of frictions affecting the internal firm organization will yield further advances in the quest for the ultimate sources of aggregate productivity differences across countries.

8. APPENDIX

8.1. Solution to the firm's problem

The first-order conditions with respect to the various forms of employees in (10), conditional on $L \geq 1$, imply the following system of equations:

$$n_u = [1 - B(1 - \lambda)(L - 1)]\phi\mu(1 - \theta)\frac{r(z)}{w_u}, \quad (13)$$

$$n_s = \frac{1 - \mu}{\mu} \frac{w_u}{w_s} n_u, \quad (14)$$

$$m_{u,l} = (\gamma\theta)^l \frac{1}{1 + b_u} n_u, \forall l \in \{1, 2, \dots, L - 1\}, \quad (15)$$

$$m_{s,l} = \frac{(1 - \gamma^l \mu)\theta^l}{\mu} \frac{w_u}{(1 + b_s)w_s} n_u, \forall l \in \{1, 2, \dots, L - 1\}, \quad (16)$$

as well as $r(z) = R^{1-\phi}y^\phi(z)$ with $y(z)$ defined in (8). Collecting terms:

$$\begin{aligned} \pi_{L \geq 1}^*(z) = & [1 - \phi(1 - \theta^L)] \left\{ R^{1-\phi} z^\phi \alpha^{\phi\theta^L} [\phi(1 - \theta)]^{\phi(1-\theta^L)} [1 - B(1 - \lambda)(L - 1)] \right. \\ & \left. [(1 + b_u)^\mu (1 + b_s)^{1-\mu} Q(L)]^{\phi(1-\theta)} \right\}^{\frac{1}{1-\phi(1-\theta^L)}} - \kappa \end{aligned} \quad (17)$$

where

$$Q(L) \equiv \prod_{l=0}^{L-1} \theta^{l\theta^l} \left[\frac{\gamma^l \mu}{(1+b_u)w_u} \right]^{\mu(\gamma\theta)^l} \left[\frac{1-\gamma^l \mu}{(1+b_s)w_s} \right]^{(1-\gamma^l \mu)\theta^l}. \quad (18)$$

8.2. Proofs

8.2.1. Proposition 1

Workers of type $i \in \{u, s\}$ enter the labor market expecting value $V_{i,h} = q_{i,0}V_{i,0} + \int \sum_{l=1}^{L(z)-1} q_{i,l}(z)V_{i,l}(z)dz$. Replacing $V_{i,l}(z)$ from (11) gives $V_{i,h} = q_{i,0}[w_i + \beta(1-\delta_a)V_{i,h}] + \int \sum_{l=1}^{L(z)-1} q_{i,l}(z)((1-\lambda)/[\beta(1-\Delta)]r(z)/[m_{u,l}(z) + m_{s,l}(z)] + V_{i,h})dz$. Since $1 - q_0 = \int \sum_{l=1}^{L(z)-1} q_{i,l}(z)dz$ and $(1+b_i)w_i \equiv [1 - \beta(1-\delta_a)]V_{i,h}$ simplification yields $b_i = 1/w_i + (1-\lambda)/[\beta(1-\Delta)](1/q_{i,0}) \int \sum_{l=1}^{L(z)-1} q_{i,l}(z)r(z)/[m_{u,l}(z) + m_{s,l}(z)]dz$. Finally, we need to establish the probability measures q . The probability of acceding to a managerial job $q_{i,l \geq 1}(z)$, is the ratio between the numerator of the mass of such job openings and the denominator of searching job candidates. The mass of such managerial jobs in the economy is $m_{i,l}(z)dF(z)$. In each period, however, only a fraction of those become available, namely $\delta_a + (1-\delta_a)\delta_m = \Delta$. Employees exiting the economy as well as those not exiting but displaced due to exogenous separation (including due to the probability of the entrepreneur, i.e. the firm, exiting) make managerial posts available. The denominator includes all job searchers with skills $i \in \{u, s\}$, i.e. newly entering individuals opting for employment, surviving production employees in search of better opportunities, and surviving but displaced managerial employees. This denominator also holds for $q_{i,0}$. Its numerator is the mass of production job openings, which is $(1-a) \int \mathbb{1}_{L \geq 1}(z)n_u(z)dF(z)$ for unskilled workers and $\int \mathbb{1}_{L \geq 1}(z)n_s(z)dF(z) + a \int \mathbb{1}_{L \geq 1}(z)n_u(z)dF(z)$ for skilled workers. In other words, all production employee jobs are available in each period as none of the previous occupants is keen on staying. Also, the demand and supply of workers of a particular skill level are not necessarily identical as skilled workers may perform tasks requiring no skills ($a > 0$) when the period wage for these two activities is equal. Applying the respective values of q gives the required result.

8.2.2. Proposition 2

Case 1: For $\lambda < 1$ it is obvious from (17) that as $L \rightarrow \infty$ the term $1 - B(1-\lambda)(L-1)$ turns negative, so this cannot be profit-maximizing. The optimal level of $L(z)$ must be finite.

Case 2: $\bar{z} < \infty$. Without loss of generality, let $\lambda = 1$ (implying $b_u = b_s = 0$). We consider two sub-cases. The first is $\phi < 1$ and the proof is by contradiction. Optimally employing infinite layers implies that $\forall L, \exists k = \{1, 2, \dots\}$ such that as $k \rightarrow \infty$, $\lim_{k \rightarrow \infty} \pi_{L+k}^*(z) \geq \pi_L^*(z)$. From (17) we have that $\pi_{L+k}^*(z) \geq \pi_L^*(z)$ implies

$$z \geq \phi(1-\theta) \left(\frac{\alpha}{R} \right)^{\frac{1-\phi}{\phi}} \left[\frac{1}{Q(L)} \right]^{1-\theta} \left[\left(\frac{1-\phi(1-\theta^L)}{1-\phi(1-\theta^{L+k})} \right)^{\frac{1-\phi(1-\theta^{L+k})}{\phi}} \left(\frac{1}{T(L,k)} \right)^{1-\theta} \right]^{\frac{1-\phi(1-\theta^L)}{\phi\theta L(1-\theta^k)}}$$

where

$$T(L, k) \equiv \prod_{l=L}^{L+k-1} \theta^{l\theta^l} \left[\frac{\gamma^l \mu}{w_u} \right]^{\mu(\gamma\theta)^l} \left[\frac{1-\gamma^l \mu}{w_s} \right]^{(1-\gamma^l \mu)\theta^l}.$$

Consider the case where $\phi < 1$ and let $k \rightarrow \infty$. The condition becomes

$$z \geq \phi(1 - \theta) \left(\frac{\alpha}{R} \right)^{\frac{1-\phi}{\phi}} \left[\frac{1}{Q(L)} \right]^{1-\theta} \left(1 + \frac{\phi}{1-\phi} \theta^L \right)^{\frac{1-\phi}{\phi} \left(1 + \frac{1-\phi}{\phi} \frac{1}{\theta^L} \right)} \left(\frac{1}{\lim_{k \rightarrow \infty} T(L, k)} \right)^{(1-\theta) \left(1 + \frac{1-\phi}{\phi} \frac{1}{\theta^L} \right)}.$$

Finally, we have that as $L \rightarrow \infty$ the right side tends to infinity as well. Notice that $\lim_{L \rightarrow \infty} Q(L)$ converges to a finite number while $\lim_{L, k \rightarrow \infty} T(L, k)$ converges to 1. The second to last term, however, tends to infinity as the increase of the exponent dominates the drop in the argument beneath (which is strictly larger than 1). Since no such z exists, this concludes the proof for the sub-case $\phi < 1$.

For $\phi = 1$, notice that as $L \rightarrow \infty$, profits tend to $-\kappa$ from the first term in (17). In that case, it is strictly optimal to choose $L(z) = 0$ which gives $z\alpha - \kappa > \kappa$. It follows that $\forall \phi \in [0, 1]$ and $\bar{z} < \infty$, the optimum number of layers $L^*(z)$ is finite.

To show that $L^*(z)$ is increasing in z consider some L such that $\forall k \in \{1, 2, \dots\}$, $\pi_L^*(z) \geq \pi_{L+k}^*(z)$. If $L \geq 1$, then from (17) we have that $\pi_L^*(z) \geq \pi_{L+k}^*(z) \Leftrightarrow z \leq D(L \geq 1, k)$ where $D(L \geq 1, k)$ is some number independent of z . Clearly, any $\tilde{z} < z$ also satisfies that condition. Hence, if $L = L^*(z) \geq 1$ then there exists no $\tilde{L} > L$ such that $\tilde{L} = L^*(\tilde{z})$. Now consider the case where $L^*(z) = 0$ such that $\forall k \in \{1, 2, \dots\}$, $\pi_0^*(z) \geq \pi_k^*(z)$. We have $\pi_0^*(z) = R^{1-\phi} z^\phi \alpha^\phi \geq \pi_k^*(z)$ where $\pi_k^*(z)$ is taken from (17). Again this implies that $z \leq D(L = 0, k)$ where $D(L = 0, k)$ is some number independent of z , and any $\tilde{z} < z$ satisfies that condition as well.

8.2.3. Proposition 3

This results directly from combining equation (9) with (13), (15) and (16).

8.2.4. Proposition 4

The share x equals $1 + \sum_{l=1}^{L-1} (m_{u,l} + m_{s,l})$ divided by $1 + n_u + n_s + \sum_{l=1}^{L-1} (m_{u,l} + m_{s,l})$. Using (14) through (16) yields the required share. As for revenue productivity, it equals r divided by $1 + n_u + n_s + \sum_{l=1}^{L-1} (m_{u,l} + m_{s,l})$. Combining (13) through (16) makes the fraction proportional to

$$\frac{1}{[1 - B(1 - \lambda)(L(z) - 1)] \left(\frac{\mu}{n_u} + \mu + (1 - \mu) \frac{w_u}{w_s} + \sum_{l=1}^{L(z)-1} \theta^l \left[\frac{\gamma^l \mu}{1 + b_u} + \frac{(1 - \gamma^l \mu) w_u}{(1 + b_s) w_s} \right] \right)}.$$

Combined with $x(z)$ this yields the required result.

REFERENCES

- Adamopoulos, Tasso and Diego Restuccia**, “The Size Distribution of Farms and International Productivity Differences,” *American Economic Review*, June 2014, 104 (6), 1667–1697.
- Aghion, Philippe and Jean Tirole**, “Formal and Real Authority in Organizations,” *Journal of Political Economy*, February 1997, 105 (1), 1–29.
- , **Nicholas Bloom, and John Van Reenen**, “Incomplete Contracts and the Internal Organization of Firms,” *Journal of Law, Economics and Organization*, August 2014, 30 (1), 37–63.
- Akcigit, Ufuk, Harun Alp, and Michael Peters**, “Lack of Selection and Imperfect Managerial Contracts: Firm Dynamics in Developing Countries,” *Mimeo*, 2014.
- Alder, Simeon**, “In the Wrong Hands: Complementarities, Resource Allocation, and TFP,” *Mimeo*, 2012.
- Alvarez, Fernando and Urban J. Jermann**, “Efficiency, Equilibrium and Asset Pricing with Risk of Default,” *Econometrica*, July 2000, 68 (4), 775–797.
- Amaral, Pedro and Erwan Quintin**, “Limited Enforcement, Financial Intermediation, and Economic Development,” *International Economic Review*, August 2010, 51 (3), 785–811.

- Antràs, Pol, Luis Garicano, and Esteban Rossi-Hansberg**, “Offshoring in a Knowledge Economy,” *Quarterly Journal of Economics*, February 2006, 121 (1), 31–77.
- Atkeson, Andrew and Patrick J. Kehoe**, “Modeling and Measuring Organization Capital,” *Journal of Political Economy*, October 2005, 113 (5), 1026–1053.
- Auer, Peter, Janine Berg, and Ibrahim Coulibaly**, “Is a Stable Workforce Good for Productivity?,” *International Labour Review*, September 2005, 144 (3), 319–343.
- Baker, George, Michael Gibbs, and Bengt Holmstrom**, “The Internal Economics of the Firm: Evidence From Personnel Data,” *Quarterly Journal of Economics*, November 1994, 109 (4), 921–955.
- Barro, Robert J. and Jong Wha Lee**, “A New Data Set of Educational Attainment in the World, 1950–2010,” *Journal of Development Economics*, September 2010, 104, 184–198.
- Bartelsman, Eric, John Haltiwanger, and Stefano Scarpetta**, “Cross-Country Differences in Productivity - The Role of Allocation and Selection,” *American Economic Review*, February 2013, 103 (1), 305–334.
- Bloom, Nicholas and John Van Reenen**, “Measuring and Explaining Management Practices across Firms and Countries,” *Quarterly Journal of Economics*, November 2007, 122 (4), 1351–1408.
- and ———, “Why Do Management Practices Differ across Firms and Countries?,” *Journal of Economic Perspectives*, Winter 2010, 24 (1), 203–224.
- , **Aprajit Mahajan, David McKenzie, and John Roberts**, “Why Do Firms in Developing Countries Have Low Productivity?,” *American Economic Review: Papers and Proceedings*, May 2010, 100 (2), 619–623.
- , **Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts**, “Does Management Matter? Evidence from India,” *Quarterly Journal of Economics*, February 2013, 128 (1), 1–51.
- , **Raffaella Sadun, and John Van Reenen**, “The Organization of Firms across Countries,” *Quarterly Journal of Economics*, November 2012, 127 (4), 1663–1705.
- Boehm, Johannes**, “The Impact of Contract Enforcement Costs on Outsourcing and Aggregate Productivity,” *Mimeo*, 2013.
- Buera, Francisco J., Joseph P. Kaboski, and Yongseok Shin**, “Finance and Development: A Tale of Two Sectors,” *American Economic Review*, August 2011, 101 (5), 1964–2002.
- Caliendo, Lorenzo and Esteban Rossi-Hansberg**, “The Impact of Trade on Organization and Productivity,” *Quarterly Journal of Economics*, August 2012, 127 (3), 1393–1467.
- , **Ferdinando Monte, and Esteban Rossi-Hansberg**, “The Anatomy of French Production Hierarchies,” *Journal of Political Economy*, forthcoming.
- Calvo, Guillermo A. and Stanislaw Wellisz**, “Supervision, Loss of Control, and the Optimum Size of the Firm,” *Journal of Political Economy*, October 1978, 86 (5), 943–952.
- and ———, “Hierarchy, Ability, and Income Distribution,” *Journal of Political Economy*, October 1979, 87 (5), 991–1010.
- Caselli, Francesco**, “Accounting for Cross-Country Income Differences,” in Philippe Aghion and Steven Durlauf, eds., *Handbook of Economic Growth*, Vol. 1, Elsevier, 2005, chapter 9, pp. 679–741.
- and **Nicola Gennaioli**, “Dynastic Management,” *Economic Enquiry*, January 2013, 51 (1), 971–996.
- Chen, Cheng**, “Management Quality, Firm Organization, and International Trade,” *Mimeo*, 2014.
- Cole, Harold L., Jeremy Greenwood, and Juan M. Sanchez**, “Why Doesn’t Technology Flow from Rich to Poor Countries?,” *Mimeo*, 2013.
- Dessein, Wouter**, “Authority and Communication in Organizations,” *Review of Economics Studies*, October 2002, 69 (4), 811–838.
- Eeckhout, Jan and Boyan Jovanovic**, “Occupational Choice and Development,” *Journal of Economic Theory*, March 2011, 147 (2), 657–683.
- Erosa, Andrés and Ana Hildago Cabrillana**, “On Finance as a Theory of TFP, Cross-Industry Productivity Differences, and Economic Rents,” *International Economic Review*, May 2008, 49 (2), 437–473.
- García-Santana, Manuel and Josep Pijoan-Mas**, “The Reservation Laws in India and the Misallocation of Production Factors,” *Journal of Monetary Economics*, September 2014, 66, 193–209.
- Garicano, Luis**, “Hierarchies and the Organization of Knowledge in Production,” *Journal of Political Economy*, October 2000, 108 (5), 874–904.
- and **Esteban Rossi-Hansberg**, “Organization and Inequality in a Knowledge Economy,” *Quarterly Journal of Economics*, November 2006, 121 (4), 1383–1435.

- and ———, “Organizing Growth,” *Journal of Economic Theory*, March 2012, 147 (2), 623–656.
- Gennaioli, Nicola, Rafael La Porta, Florencio Lopez de Silanes, and Andrei Shleifer**, “Human Capital and Regional Development,” *Quarterly Journal of Economics*, February 2013, 128 (1), 105–164.
- Gollin, Douglas**, “Nobody’s Business But My Own: Self-Employment and Small Enterprise in Economic Development,” *Journal of Monetary Economics*, March 2008, 55 (2), 219–233.
- Greenwood, Jeremy, Juan M. Sanchez, and Cheng Wang**, “Financing Development: The Role of Information Costs,” *American Economic Review*, September 2010, 100 (4), 1875–1891.
- Guner, Nezih, Gustavo Ventura, and Yi Xu**, “Macroeconomic Implications of Size-Dependent Policies,” *Review of Economic Dynamics*, October 2008, 11 (4), 721–744.
- Hall, Robert E. and Charles I. Jones**, “Why Do Some Countries Produce So Much More Output Per Worker Than Others?,” *Quarterly Journal of Economics*, February 1999, 114 (1), 83–116.
- Hipple, Steven F.**, “Self-Employment in the United States,” *Monthly Labor Review*, September 2010, pp. 17–32.
- Hsieh, Chang-Tai and Peter J. Klenow**, “Misallocation and Manufacturing TFP in China and India,” *Quarterly Journal of Economics*, November 2009, 124 (4), 1403–1448.
- Idson, Todd L. and Walter Y. Oi**, “Workers Are More Productive in Large Firms,” *American Economic Review*, May 1999, 89 (2), 104–108.
- Kehoe, Timothy J. and David K. Levine**, “Debt-Constrained Asset Markets,” *Review of Economic Studies*, October 1993, 60 (4), 865–888.
- Kremer, Michael**, “The O-Ring Theory of Economic Development,” *Quarterly Journal of Economics*, August 1993, 108 (3), 551–575.
- Lucas, Robert E. Jr.**, “On the Size Distribution of Business Firms,” *Bell Journal of Economics*, Autumn 1978, 9 (2), 502–523.
- Luttmer, Erzo G. J.**, “Selection, Growth, and the Size Distribution of Firms,” *Quarterly Review of Economics*, August 2007, 122 (3), 1103–1144.
- Manuelli, Rodolfo and Ananth Seshadri**, “Human Capital and the Wealth of Nations,” *American Economic Review*, September 2014, 104 (9), 2736–2762.
- Midrigan, Virgiliu and Daniel Yi Xu**, “Finance and Misallocation: Evidence from Plant-Level Data,” *American Economic Review*, forthcoming.
- Moll, Benjamin**, “Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation?,” *American Economic Review*, forthcoming.
- Peters, Michael**, “Heterogeneous Mark-Ups, Growth and Endogenous Misallocation,” *Mimeo*, 2013.
- Poschke, Markus**, “The Firm Size Distribution across Countries and Skill-Biased Change in Entrepreneurial Technology,” *Mimeo*, 2014.
- Powell, Michael**, “Productivity and Credibility in Industry Equilibrium,” *Mimeo*, 2013.
- Psacharopoulos, George**, “Returns to Investment in Education: A Global Update,” *World Development*, September 1994, 22 (9), 1325–1343.
- Rajan, Raghuram G. and Luigi Zingales**, “Power in A Theory Of The Firm,” *Quarterly Journal of Economics*, May 1998, 113 (2), 387–432.
- Restuccia, Diego and Richard Rogerson**, “Policy Distortions and Aggregate Productivity with Heterogeneous Plants,” *Review of Economic Dynamics*, October 2008, 11 (4), 707–720.
- and ———, “Misallocation and Productivity,” *Review of Economic Dynamics*, 2013, 16 (1), 1–10.
- Roys, Nicolas and Ananth Seshadri**, “Economic Development and the Organization of Production,” *Mimeo*, 2014.
- Shapiro, Carl and Joseph E. Stiglitz**, “Equilibrium Unemployment as a Worker Discipline Device,” *The American Economic Review*, June 1984, 74 (3), 433–444.
- Troske, Kenneth R.**, “Evidence on the Employer Size-Wage Premium From Worker-Establishment Matched Data,” *The Review of Economics and Statistics*, February 1999, 81 (1), 15–26.
- Tybout, James R.**, “Manufacturing Firms in Developing Countries: How Well Do They Do, and Why?,” *Journal of Economic Literature*, March 2000, 38 (1), 11–44.
- Williamson, Oliver E.**, “Hierarchical Control and Optimum Firm Size,” *Journal of Political Economy*, April 1967, 75 (2), 123–138.