

Corporate Control and Executive Selection*

Francesco Lippi

University of Sassari and EIEF

f.lippi@uniss.it

Fabiano Schivardi

University of Cagliari and EIEF

fschivardi@unica.it

September 2011

Abstract

In firms with concentrated ownership the controlling shareholder may pursue non-monetary private returns, such as electoral goals in a firm controlled by politicians or family prestige in family firms. We use a simple theoretical model to analyze how this mechanism affects the selection of executives and, through this, the firm's productivity compared to a benchmark where the owner only cares about the value of the firm. We discuss identification and derive two structural estimates of the model, based on different sample moments. The estimates, based on a matched employer-employee dataset of Italian firms, suggest that private returns are larger in family and government controlled firms than in firms controlled by a conglomerate or by a foreign entity. The resulting distortion in executive selection can account for TFP differentials between control types of up to 10%.

Key Words: corporate governance, private returns, TFP.

JEL classification numbers: D2, G32, L2

*We thank Fernando Alvarez, Gadi Barlevy, Marco Bassetto, Jeff Butler, Jeff Campbell, Hugo Hopenhayn, Mariacristina De Nardi, Giuseppe Moscarini, Tommaso Nannicini, Marco Pagano, Ariel Pakes, Mark Roberts, Bernard Salanié, Yossi Spiegel, Andy Skrzypacz, Luke Taylor, Nicholas Trachter, Francis Vella and seminar participants at EIEF, University of Sassari, Tor Vergata University, Catholic University in Milan, University of Bologna, the SED 2009 meetings, the CEPR-Bank of Italy conference "Corporate Governance, Capital Structure and Firm Performance", the Federal Reserve Bank of Chicago.

1 Introduction

This paper studies executive selection in firms with concentrated ownership, a control structure that the recent corporate finance literature has shown to be very diffused around the world (Laporta, Lopez-De-Silanes, and Shleifer, 1999; Faccio and Lang, 2002). Unlike public corporations, where the separation between ownership and control naturally puts agency issues at center stage, our hypothesis is that in firms with concentrated ownership the controlling shareholder may pursue non-monetary private returns, such as electoral goals in a firm controlled by politicians or family prestige in a firm controlled by an individual.¹ We use a simple model to analyze how this assumption affects the selection of executives and reduces the firm productivity compared to a benchmark where the owner only cares about the value of the firm. Two structural estimates of the model, based on different sample moments, indicate that this mechanism is quantitatively important and allow us to compute, via counterfactuals, the productivity losses that are due to the private returns.

We focus on a specific form of *owners'* private benefits: we assume that owners may derive utility not only from profits but also from employing executives with whom they have developed personal ties. Personal ties and repeated interaction can facilitate the delivery of non monetary payoffs that are typically not verifiable in court and therefore cannot be part of the employment contract. For example the owner of a family business might enjoy a compliant entourage and/or a group of executives who pursue family prestige, possibly at the expense of the value of the firm.² A politician (the “owner” of a government controlled firm) might want executives who serve his political interests, resulting in, e.g., hiring workers in his constituency. Finally, diverting resources at the expenses of minority shareholders or plain wrongdoing requires obliging executives, as some recent corporate scandals have shown. If owners value personal ties, they might do so at the expenses of ability, thus distorting the process of executive selection with respect to a situation where only the value of the firm matters.

¹Consistent with the hypothesis that control commands a premium, Dyck and Zingales (2004) provide cross-country evidence that controlling blocks are sold on average at a 14% premium, up to 65% in certain countries—see below for more details. The importance of private benefits of control is also stressed by Moskowitz and Vissing-Jørgensen (2002), who show that in the US average returns of privately held firms are dominated by the market portfolio. They conclude that owners of private firms must be obtaining some form of non-monetary return.

²Becker was the first to stress the importance of non-pure consumption components of preferences for individual decision making, e.g. in his classic analysis of discrimination (Becker, 1971). In the introduction to the book collecting his contributions on this topic, he states that “Men and women want respect, recognition, prestige, acceptance, and power from their family, friends, peers, and others” (Becker, 1998, p. 12).

We consider a simple partial-equilibrium, infinite horizon, economy in which the firm owner chooses the firm’s executives. The problem solved by the firm owner is reminiscent of the optimal search problem first studied by [McCall \(1970\)](#) and [Weitzman \(1979\)](#). We assume that average managerial ability determines the firm TFP.³ Executives are characterized by two random variables: their ability (productivity) x and their relationship value r , whose distributions are known to the owner. Both variables are specific to the executive-firm match, and the owner only learns their realization after an in-office trial period for the executive. Upon learning x and r the owner decides whether to give tenure to the executive, who in this case turns “senior”, or to replace him with a “junior” one. We assume that relationship-building takes time, so that only senior executives may deliver the private returns from the personal relationship. In particular, the expected value of r is zero for a junior executive, while $\mathbb{E}(r) = qR$ for a senior, where q denotes the ex-ante probability that the senior executive delivers a valuable relationship and R is the value of the relationship. The key decision for the firm owner is whether to retain the executive in the firm after learning the value of $s = x + r$, or to replace the executive with a junior one. Once tenured, executives in office die with an exogenous probability ρ . A standard model where the owner maximizes the value of the firm is obtained as a special case assuming that $R = 0$.

We solve the model analytically and show that the optimal policy has a threshold form: the owner retains an executive if and only if the sum of the ability and the relationship value is above a cutoff value. We then characterize the steady state predictions of this model and develop several comparative statics exercise. The firm’s productivity is maximized when the owner derives no utility from personal ties ($R = 0$) and it declines monotonically with R . This is because, as R increases, the owner’s tenure decisions are based less on ability and more on personal ties, thus weakening the selection effect on managerial ability. Similarly, the owners who attach more importance to relationships (higher R) will on average retain a larger share of senior executives, as some low-ability executives with whom they have developed a personal relationship will not be fired.

We use our model to discuss what dimensions of the data allow the model parameters to be identified. For a given distribution of ability $G(x)$ and hazard ρ the model provides a mapping between the fundamental parameters q, R and two firm-level observables: the productivity X and the share of senior executives ϕ . Our objective is to use this mapping to quantify the effects of

³This is akin to [Lucas \(1978\)](#), where TFP depends on the *entrepreneurial* ability.

different control types, modeled as different values of R , on X, ϕ . A natural interpretation of R is that it provides a measure of the importance that different owner types assign to private returns, measured in units of the firm-level productivity. In our model the heterogeneity in the value of personal ties R gives rise to persistent cross-sectional differences in firms' productivity, in accordance with the findings of a vast empirical literature (see [Syverson, 2010](#), for a recent survey).

The estimates are based on a sample of Italian manufacturing firms for which we have detailed information on the firms' characteristics, including the complete work history of their executives. We construct TFP using the [Olley and Pakes \(1996\)](#) procedure and define senior executives those who have been with the firm for at least five years. To ensure that our results are not driven by time trends or sectoral composition potentially correlated with the ownership types, all our empirical exercises control for such effects via a full set of year and sectoral dummies. The data also classify the controlling shareholder into the four broad control types: individual/family, the government, a conglomerate, or a foreign institution. Different control types might assign different importance to personal ties. In the estimation we therefore allow the importance of private returns, R , to vary across control types. The first empirical exercise is based on the correlation between productivity and the share of senior executives across firms *of a given control type*. The model predicts that, for a given owner-type, this correlation measures the strength of the selection effect on managerial ability, defined as the difference in the average ability of senior and junior executives. Intuitively, if selection based on talent (i.e., x) is important for the firm owner, then observing a high-retention rate by a firm (i.e., a large fraction of senior executives) signals the high quality of these managers. On the contrary, when quality plays a smaller role in the selection decision, because the owner also cares about personal ties, the correlation between the share of senior executives and the firm's productivity decreases. We prove that this correlation can be estimated by an OLS regression. We find that selection is weaker in government and family firms, and that private benefits generate substantial losses in productivity.

One potential criticism of these estimates is that different ownership types might differ along several dimensions, in addition to executive selection. This is not a problem for the OLS estimates, which include ownership dummies and are therefore robust to potential unobserved factors affecting the average share of senior executives and average productivity for a given control type. Still, omitted variables and unobserved heterogeneity might give rise to the correlation we observe within

type. For example, one might conjecture that, within family firms, older firms might be both more productive and employ more senior executives. We address this important criticism in two ways. First we argue, and show formally, that the omitted variable bias might explain the correlation within type, but not the *differences across types*. It is precisely the variation of the estimated coefficients across types that is predicted by our model and confirmed by the data. Second, we perform a series of additional regressions in which we utilize more controls, including firm age and size and various indicators of the characteristics of the workforce. The estimation of the parameters of interest is hardly affected, even when these additional controls are interacted with ownership dummies. Finally, we check that the results are robust to a series of additional modifications, such as changing the length of time used to classify “senior” executives and using alternative measures of profitability – rather than productivity – as performance indicators. The results appear robust.

The OLS only estimate a subset of the structural parameters. We then move on to a structural estimation of all the model parameters. We again net out of the distribution of individual ability $G(x)$ sectoral and time effects. We explain the residual variation in X, ϕ by variation in the nature of the control type, i.e., R , and one value of q common to all control-types. We show that given the distribution of abilities $G(x)$ there is a one-to-one mapping between X, ϕ and the two structural parameters q, R . We invert this mapping to estimate the private benefits of control. The estimates of the model’s fundamental parameters q, R and the ability distribution $G(x)$ are used to quantify the relevance of private returns and the related efficiency losses for each of the four ownership types. We find that the executive selection is distorted with respect to the benchmark of no private values ($R = 0$) in all control types. The distortion is smallest for conglomerate and foreign controlled firms. Given the model parameters, we can run counterfactual experiments by changing the importance of private benefits to compute the effects on productivity. The importance of private benefits accounts for a decrease in average managerial ability (i.e., firm productivity) of around 6% in family firms and of 10% in government firms, as compared to conglomerate and foreign controlled firms. According to our estimates, this happens because controlling shareholders of family and government firms select executives almost exclusively on the basis of personal ties: they tend to keep all the executives with whom they developed a relationship, independent of ability, and fire all the others. This mechanism inhibits the selection effect of managerial ability. When compared to the theoretical benchmark of no private benefits, productivity losses are in the order of 7% for conglomerates, 8% for foreigners,

13% for family and 17% for Government firms.

As a final robustness check we compare the OLS and the structural estimates. This comparison is informative because the two sets of estimates use different sources of data variability to identify the parameters: within-type correlation in ϕ, X for the OLS and variation in average sample moments across control types for the structural estimates. Nothing in the estimation procedure imposed that the estimation results should yield quantitatively similar results in these two sets of estimates. The fact that the estimates do produce quantitatively similar effects is supportive of our hypothesis.

The paper is organized as follows. The next section discusses the connection with the literature, in particular with [Albuquerque and Schroth \(2010\)](#), [Bandiera et al. \(2009\)](#) and [Taylor \(2010b\)](#), who study related problems. [Section 3](#) develops a model to study how the presence of private returns affects the selection of executives and average productivity within a firm. [Section 4](#) describes the matched employer-employee data and the classification of the various types of corporate control. In [Section 5](#) we study the mapping between the model and the data, discussing the identification of the structural model parameters. [Section 6](#) presents a test of the model hypothesis that exploits a model-based regression analysis. A subsection discusses several robustness checks and alternative hypothesis. [Section 7](#) presents structural estimates of the model parameters, which are used to quantify the “costs” of private benefits in terms of foregone productivity by means of simple counterfactual analysis. [Section 8](#) concludes.

2 Related literature

The idea that private benefits play a central role in shaping firms’ performance is central to the recent corporate governance literature ([La Porta et al., 2000](#)). [Dyck and Zingales \(2004\)](#) empirically estimate the value of private benefits of control using the difference between the price per share of a transaction involving a controlling block and the price on the stock market before that transaction. They find large values of private benefits of control. In particular, at 37% the value for Italy is the second highest in a sample of 39 countries. Compared to this literature, we focus on a very specific channel through which the private benefits arise: the relationship between the owner and her executives. Moreover, we use the model’s predictions on productivity and executive seniority distribution to estimate the value of such a relationship, rather than referring to stock market data.

In our model the inefficient selection derives since the firm owner may assign value to the personal relationship with the management. The fact that personal ties between firms' high-ranked stakeholders (large shareholders, board members, top managers) is detrimental for firm performance finds support in recent literature. [Bandiera, Barankay, and Rasul \(2008\)](#) study the effects of social connections among managers and workers on performance. Using a field experiment, they show that managers favor workers they are socially connected to, possibly at the expense of the firm's performance. [Kramarz and Thesmar \(2006\)](#) study the effects of social networks on the composition of firms' board of directors and performance in France. They find that networks, defined in terms of school of graduation, influence the board composition; moreover, firms with a higher share of directors from the same network have a worse performance.

Our work also relates to the vast body of literature that documents that firms of very different productivity levels coexist even within narrowly defined markets (see, e.g., [Bartelsmann and Doms, 2000](#); [Syverson, 2010](#)). As in [Lucas \(1978\)](#), in our model dispersion in firm productivity derives from the underlying dispersion in managerial ability, subject to a cutoff level dictated by the selection effect.⁴ Differently from Lucas, in our model owners are willing to accept a low return on their investment because they derive other types of returns, which weakens the selection effect and increases the cross sectional dispersion in firms productivity. Empirically, we find this to be more relevant for family firms, in line with a growing literature on empirical work practices and performance in family firms (see, for example, [Moskowitz and Vissing-Jørgensen, 2002](#); [Bloom and Van Reenen, 2007](#); [Bandiera et al., 2009](#); [Michelacci and Schivardi, 2010](#)).

In terms of managerial turnover, [Volpin \(2002\)](#) studies top executive turnover in Italian listed firms. Consistent with our findings, family controlled firms tend to have lower turnover rates than foreign controlled firms.⁵ Executive compensation, promotion policies and turnover are subject to a growing and heterogeneous body of model-based empirical analysis, using, among others, assignment ([Gabaix and Landier, 2008](#); [Terviö, 2008](#)) or moral hazard models ([Gayle, Golan, and Miller, 2009](#)). Compared to this literature, we do not explicitly formalize the market for executives, but focus on

⁴Heterogeneity in an underlying unobservable firm characteristic has become the standard way to model productivity dispersion at the firm level both in IO ([Jovanovic, 1982](#); [Hopenhayn, 1992](#); [Ericson and Pakes, 1995](#)) and in trade ([Bernard et al., 2003](#); [Melitz, 2003](#)).

⁵He also finds that the sensitivity of turnover to performance, an indicator of the quality of the governance, is not significantly different in the two groups. However, his sample has a small number of foreign controlled firms and in fact his main analysis is focussed on family controlled firms only.

the owner’s decision to confirm or replace incumbent executives.

Three recent papers are closely related to ours. [Albuquerque and Schroth \(2010\)](#) use a structural model to estimate the private benefits of control in negotiated block transactions. They find evidence consistent with the hypothesis that those benefits are large. [Taylor \(2010b\)](#) builds and estimates a structural model of CEO turnover with learning about managerial ability and costly turnover. He finds that only very high turnover costs can rationalize the low turnover rate observed in the data.⁶ He interprets this result in terms of CEO entrenchment and weak governance.⁷ Compared to his paper we propose a different, possibly complementary, reason for inefficient turnover: owners trade off efficiency for the private benefits of the personal relationships with their executives. We also use data on all top executives, rather than CEOs alone, and exploit different ownership structures to estimate our model parameters. [Bandiera et al. \(2009\)](#) analyze the role of incentive schemes in family and non family firms, assuming that family firms pursue private benefits of control. They show that family firms rely less on performance-based compensation schemes and attract more risk averse and less able managers. The model predictions are supported by reduced-form regressions. We see this paper and ours as complementary. In fact, [Bandiera et al. \(2009\)](#) focus on the optimal compensation scheme, an issue that is ignored in our model. On the other hand, we introduce learning about managerial ability in a dynamic setting, so that we can study turnover and seniority composition. Moreover, we provide direct structural estimates of the importance of private benefits by control type, supplying supporting evidence for a central assumption of their model.

3 A model of executive tenure and firm productivity

We model the decision problem of a firm owner in charge of selecting the executives who run the firm. Our aim is to use the model to organize the empirical analysis. Given that we will structurally estimate the model parameters, we keep it as simple as possible. In particular, we focus on the owner’s selection problem and completely put aside the market for executives.⁸ A firm employs

⁶See also [Taylor \(2010a\)](#) for a model of learning about CEOs’ ability and wage dynamics.

⁷[Garrett and Pavan \(2010\)](#) study managerial turnover when the quality of the match between a firm and its top managers changes stochastically over time and is privately observed by the managers. They show that the optimal retention decision becomes more permissive over time, offering an alternative explanation for what would look like CEO entrenchment.

⁸Our conjecture, to be verified in future work, is that the effects we identify would also hold both in a competitive market for executives and in a search framework. In fact, as long as the private benefits create a surplus, they should affect executive selection in the same way as in our framework, independently from the surplus splitting rule.

n executives, depending on its size (not modeled here). Each executive is characterized by an ability level x_i . We assume that ability is a shifter of the production function, $\bar{x}F(K, L)$, where $\bar{x} = \frac{1}{n} \sum_i x_i$ is the average managerial ability, K the capital stock and L labor. This assumption has two consequences. First, it implies that we will be able to measure average managerial ability by firm level TFP, which we will use in the empirical section. Second, the fact that overall TFP is additive in individual ability implies that we can study the problem of the owner for each single executive in isolation from the others, as we exclude spillovers in ability among them. This assumption simplifies the analysis in the model that we present next.

3.1 The model

The problem describes executive selection by the firm owner (the principal henceforth). The executives are hired at the junior level and become senior - and eligible for tenure - after one period. We think of this period as the time during which executive's quality is learned by the principal. An executive's quality is characterized by two independent exogenous variables: his productivity x , a non-negative random variable with continuous and differentiable CDF $G(x)$ with $G(0) = 0$ and expected value $\mu = \int_0^\infty x dG(x)$, and his relationship value r , a non-negative random variable that is identically zero for a junior executive and equals zero with probability $1 - q$ or R with probability q for a senior executive. The personal relationship is valuable because it facilitates the delivery of non-monetary payoffs, that cannot be explicitly included in an employment contract. For example, a politician might value executives who serve his political interests in government controlled firms, by hiring workers in his constituency. The owner of a family business might enjoy a compliant entourage and/or a group of executives who pursue the prestige of the family. The rationale for the value of relationships to mature only for senior executives is that relationships take time to be developed.⁹

We assume that upon hiring a (junior) executive the principal observes neither x nor r , but only knows their distribution. At the end of the first period the principal learns the value of the executive's productivity, a realization of x , and the value of his relationship, the realization of r

⁹Of course, personal relationships might develop before the match forms, for example if an owner hires friends or relatives. In this case, exactly the same logic would apply to "connected" (friends or relatives) vs. "unconnected" executives. Unfortunately, in the data we have no way of detecting this type of relationships, while we observe seniority. The two channels of personal ties are not mutually exclusive and might both be at play.

(either 0 or $R > 0$). It is assumed that both the executive relationship value and productivity are specific to an executive-firm match, so that if an executive moves to a new firm both his x and r are unknown to the new principal. After learning the realizations of x and r the principal decides whether to keep the executive in office (i.e., give him tenure) or to replace him with a junior one (i.e., fire the incumbent executive). It is convenient to define a new random variable $s \equiv x + r$. Using that x and r are independent, the CDF is

$$F(s) = q G(\max(s - R, 0)) + (1 - q) G(s) \quad \forall s > 0 \quad (1)$$

so that the probability that a senior executive with $r + x \geq s$ is observed is $1 - F(s)$.

If appointed, the (senior) executive stays one period with the firm and then dies with an exogenous constant hazard ρ , so that the expected office tenure of a senior executive is $1/\rho$. When a senior executive dies the principal replaces her with a junior one.

The per-period return for the risk-neutral principal is given by the realizations of $s_t = x_t + r_t$, his utility is given by the expected present value of the sum of these realizations: $v \equiv \sum_{t=0}^{\infty} \beta^t s_t$, where β is a time discount. The principal cares about the executive productivity and his/her relationship value, and decides whether or not to fire an executive after observing the realization of both variables at the end of the first period. When a junior executive is in office at the beginning of period t there is no further decision to be taken for the principal, and the expected value for the principal is

$$v_y = \mu + \beta \mathbb{E}_{\tilde{s}} \max\{ v_y, v_o(\tilde{s}) \} \quad (2)$$

where expectations are taken with respect to the next-period realization of the executive value \tilde{s} and $v_o(\tilde{s})$ denotes the value of a senior executive with *known* value \tilde{s} . This value is

$$v_o(\tilde{s}) = \max\{ v_y, \tilde{s} + \beta [\rho v_y + (1 - \rho) v_o(\tilde{s})] \}. \quad (3)$$

where the value function $v_o(\tilde{s})$ is continuous and increasing in \tilde{s} .

The optimal policy follows a threshold rule: the principal fires the senior executive if $s < s^*$, i.e., if the value of $s = r + x$, learned when the executive becomes senior, is below the threshold

s^* .¹⁰ To be clear about the condition that pins down the optimal threshold value s^* , it is useful to introduce two more pieces of notation. Let $v_{y,\bar{s}}$ denote the conditional value of a junior executive under a generic policy threshold: \bar{s} . Likewise, let us define $v_{o,\bar{s}}(\tilde{s})$ as the conditional value of a senior executive of type $\tilde{s} \geq \bar{s}$ under a generic policy threshold \bar{s} . Obviously \bar{s} enters both (conditional) value functions because it determines the senior executives that will be fired, i.e., all those for which $s < \bar{s}$. We can now state the condition that defines the *optimal* value of the threshold s^* as the smallest value of s that leaves the firm indifferent between keeping the senior executive or appointing a junior one, namely

$$v_{o,s^*}(s^*) = v_{y,s^*} . \quad (4)$$

Using [equation \(3\)](#) and the optimal threshold s^* it is straightforward to compute the *expected* value of a senior executive *conditional on being in office* as

$$\begin{aligned} v_o \equiv \mathbb{E}_s(v_o(s)|s \geq s^*) &= \int_{s^*}^{\infty} s \frac{dF(s)}{1 - F(s^*)} + \beta [\rho v_y + (1 - \rho)v_o] \\ &= \frac{1}{1 - \beta(1 - \rho)} \left(\int_{s^*}^{\infty} s \frac{dF(s)}{1 - F(s^*)} + \beta \rho v_y \right) . \end{aligned} \quad (5)$$

Given s^* , the expected value of a junior executive can be rewritten as

$$v_y = \mu + \beta [F(s^*) v_y + (1 - F(s^*)) v_o] \quad (6)$$

Using [equation \(6\)](#) and the expression for v_o in [equation \(5\)](#) gives a closed form equation for v_y as a function of s^* :

$$v_y = \frac{\mu(1 - \beta(1 - \rho)) + \beta \int_{s^*}^{\infty} s dF(s)}{(1 - \beta)[1 + \beta(\rho - F(s^*))]} . \quad (7)$$

Using [equation \(3\)](#) to write the value of a senior executive of type s^* as

$$v_o(s^*) = \frac{1}{1 - \beta(1 - \rho)} (s^* + \beta \rho v_y)$$

and replacing this expression into [equation \(4\)](#) gives the optimality condition

$$s^* = (1 - \beta)v_y . \quad (8)$$

¹⁰Like in [McCall \(1970\)](#) model, the proof relies on the fact that the functions $v_o(\tilde{s})$ and v_y cross only once.

Using [equation \(8\)](#) and the expression for v_y in [\(7\)](#) gives one equation in one unknown for s^* :

$$H(s^*, R) \equiv s^* [1 + \beta(\rho - F(s^*))] - \mu(1 - \beta(1 - \rho)) - \beta \int_{s^*}^{\infty} s dF(s) = 0 . \quad (9)$$

This leads us to:

PROPOSITION 1. Given the primitives $\beta, \rho, G(\cdot), q$, there exists a unique optimal threshold $s^*(R)$.

Moreover:

(i) $s^*(0) > \mu$

(ii) $s^*(R)$ satisfies:

$$0 < \frac{\partial s^*(R)}{\partial R} = q\beta \frac{1 - G(s^* - R)}{1 + \beta(1 - F(s^*))} < q\beta < 1 \quad (10)$$

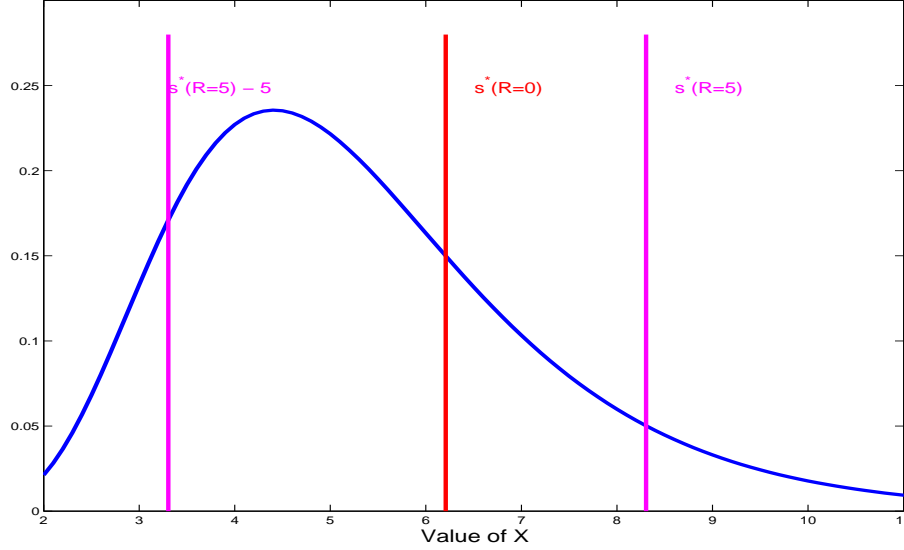
Proof. See [Appendix A](#).

The proposition states that when $R \equiv 0$ (relationships bring no value to the principal) the optimal threshold $s^*(0) > \mu$. Hence the senior executive retains office only if he is sufficiently above the expected value of a junior μ . That is because the appointment of a junior, and the possibility of future replacement, gives the policy of appointing a junior a positive option value. The fact that productivity x is learned after one period induces a selection whereby senior executives who retain office are more productive than the average junior executive. This is shown in [Figure 1](#) where the optimal threshold for the $R = 0$ case lies to the right of μ (the mean of the ability distribution).

The second part of the proposition characterizes how the optimal threshold s^* varies with R . The larger the importance of the non-monetary returns to the principal (as measured by a higher value of R), the greater is the value of the threshold s^* . This has two contrasting effects on the productivity of the executives who get tenure: the fact that $\partial s^*/\partial R < 1$ implies that, as R increases, the productivity threshold for the executives who develop a relationship (i.e., those with $r = R$) falls, since $s^* - R$ is decreasing in R . On the other hand, the threshold for the executives who do not develop a valuable relationship (i.e., those with $r = 0$) increases: these executives must compensate for their lack of “relationship” value with a higher productivity, such that $x \geq s^*$. As shown in [Figure 1](#), the ability thresholds for executives with and without relationship value move apart as R increases.

We now turn to the model prediction concerning the seniority composition of the firm’s execu-

Figure 1: Example of selection thresholds for $R = 0$ and $R = 5$



Note: The figure uses the following parameters: $\beta = 0.98$ (per year), $\rho = 0.11$ (per year), $q = 0.75$. $G(\cdot)$ is lognormal with log mean $\lambda_m = 1.6$, and log std $\lambda_\sigma = 0.36$ which imply $\mu = 5.3$.

tives in a steady state. The fraction of senior executives in office, ϕ , follows the law of motion

$$\phi_t = \phi_{t-1}(1 - \rho) + (1 - \phi_{t-1})(1 - F(s^*)) ,$$

so that the steady state fraction of senior executives is

$$\phi(s^*) = \frac{1}{1 + \frac{\rho}{1 - F(s^*)}} \in (0, 1) \quad (11)$$

which is decreasing in ρ . Mechanically, a lower hazard rate increases the fraction of senior executives. It is also immediate that ϕ is decreasing in $F(s^*)$. To study how ϕ depends on R we need to compute the total derivative of $F(s^*)$, since changes in R affect the CDF directly and also affect the threshold s^* . Note that

$$\frac{dF(s^*)}{dR} = [q g(s^* - R) + (1 - q) g(s^*)] \frac{\partial s^*}{\partial R} - q g(s^* - R). \quad (12)$$

Recalling that $\frac{\partial s^*}{\partial R} < q$ (see [Proposition 1](#)) shows that the derivative is negative at $R = 0$, which means that at $R = 0$ the share of senior executives is increasing in R . Intuitively, when $R > 0$

the appeal of senior executives increases because, all other things equal, their expected return is increased by the expected value of relationships, qR . However, the effect of R on ϕ cannot be signed in general when $R > 0$. The reason is that an increase in R has two opposing effects. On the one hand it *lowers* the threshold $s^* - R$ for that fraction (q) of senior executives who display valuable relationships ($r = R$). This increases ϕ . On the other hand a higher s^* raises the acceptance threshold for the senior executive with no relationship capital ($r = 0$). This reduces ϕ . The final effect thus depends on the features of the distribution of x and r . For instance, for sufficiently small values of q the relationship between ϕ and R is non-monotone. An example is displayed in the left panel of [Figure 2](#).

We now analyze how changes in R affect the firm's average productivity in the steady state. Let X denote the mean productivity of the firm, given by the weighted average of the expected productivity of the junior and senior executives:

$$X(s^*) = \mathbb{E}_{r,x}(x) = \mu + \phi(s^*) [X_o(s^*) - \mu] \quad (13)$$

where some algebra shows that the senior executives' average productivity is

$$X_o(s^*) = \frac{q \int_{s^*-R}^{\infty} x dG(x) + (1-q) \int_{s^*}^{\infty} x dG(x)}{1 - F(s^*)} . \quad (14)$$

This leads us to:

PROPOSITION 2. Let $\beta \rightarrow 1$, then the steady state firm productivity $X(s^*)$ is:

- (i) maximal under the policy $s^*(R = 0)$, with $\frac{\partial X}{\partial R} \Big|_{R=0} = 0$
- (ii) decreasing in R : $\frac{\partial X}{\partial R} \Big|_{R>0} < 0$

Proof. See [Appendix A](#).

The proposition shows that the mean productivity of a firm is maximized when the firm only cares about ability, i.e., under the policy $s^*(R = 0)$. Any policy $s^*(R)$ with $R > 0$ induces on average a lower firm productivity. Moreover, the proposition shows that X is monotone decreasing in R . This will be useful in the discussion of the parameters identification below. The assumption that $\beta \rightarrow 1$ simplifies the derivation and is useful to interpret the mean X as a cross section average.¹¹

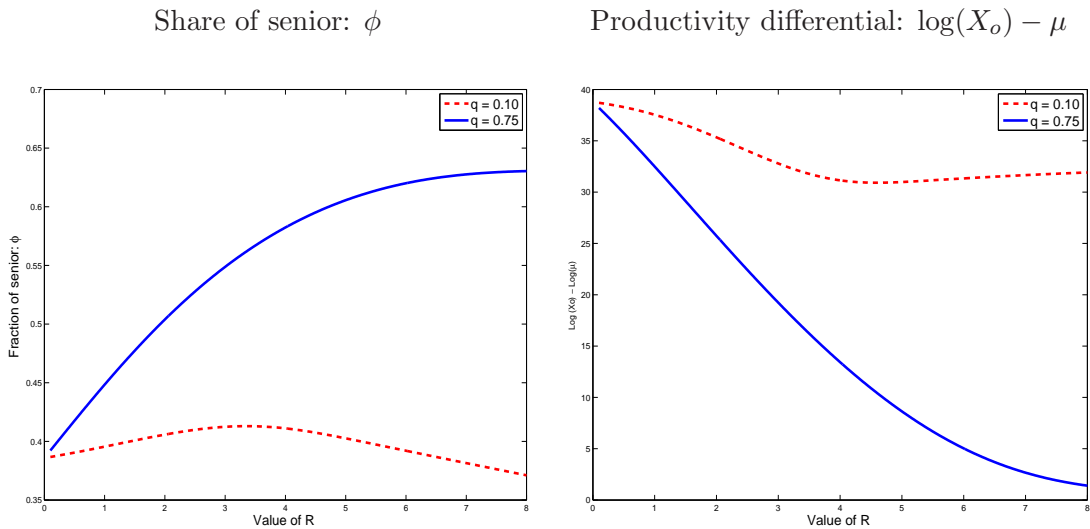
¹¹The numerical analysis of the model for $\beta \in (0.85, 1)$ (per year) gives very similar results. The relationship

The proposition also establishes that the derivative of X with respect to R is zero at $R = 0$. This result, and the fact that the share of senior executives is increasing at $R = 0$ (discussed above), implies that the productivity differential between senior and junior executives, $X_o - \mu$, is decreasing in R at $R = 0$, as can be seen from [equation \(14\)](#):

$$\left. \frac{\partial \{X_o(s^*) - \mu\}}{\partial R} \right|_{R=0} = \frac{g(s^*) \left(q - \frac{\partial s^*}{\partial R} \right) \left[s^* - \frac{\int_{s^*}^{\infty} x dG(x)}{1-G(s^*)} \right]}{1-G(s^*)} < 0 \quad (15)$$

The intuition behind this pattern is simple: as R increases the owner selects less on ability, so that the senior executives become more similar to the unselected pool of junior executives. The right panel of [Figure 2](#) shows that this pattern holds for a wide set of parameter values and, in particular, for the parameters that are in a (broad) neighborhood of our structural estimates (thick line). The figure also shows that in the parametrization with the high value of $q = 0.75$, both ϕ and $X_o - \mu$ become flat functions of R for $R \cong 8$: at this point owners are basically already firing all executives with $r = 0$ and keeping all those with $r = 8$, so that further increases in R do not influence the selection process anymore. In other words, X and ϕ asymptote to constant values as R grows large. We will come back to this observation when we comment on the results of our estimates.

Figure 2: Share of Senior executives and Senior-Junior differential as R varies



Note: The figure uses the following parameters: $\beta = 0.98$ (per year), $\rho = 0.11$ (per year), $q = 0.75$. $G(\cdot)$ is lognormal with log mean $\lambda_m = 1.6$, and log std $\lambda_\sigma = 0.36$ which imply $\mu = 5.3$.

between X and R is always decreasing.

4 Data description

In this section we describe the main features of our data, referring to [Appendix B](#) for more details. The data match a large sample of executives with a sample of Italian firms. The executives represent approximately the top 2% of the firm’s employment. The firm data are drawn from the Bank of Italy’s annual survey of manufacturing firms (INVIND), an open panel of around 1,200 firms per year representative of manufacturing firms with at least 50 employees. It contains detailed information on firms’ characteristics, including industrial sector, year of creation, number of employees, value of shipments, value of exports and investment. It also reports sampling weights to replicate the universe of firms with at least 50 employees. We completed the dataset with balance-sheet data collected by the Company Accounts Data Service (CADS) since 1982, from which it was possible to reconstruct the capital series, using the perpetual inventory method.

Our measure of productivity is TFP. We assume that production takes place with a Cobb-Douglas production function of the form:

$$Y_{i,t} = \text{TFP}_{i,t} K_{i,t}^{\beta} L_{i,t}^{\alpha} \quad (16)$$

where Y is value added, K is capital, L is labor, and i, t are firm and year indices, respectively. TFP depends on average managerial ability X and, possibly, on other additional observable and unobservable characteristics $W_{i,t}$, such as the industrial sector, the firm size and time effects:

$$\text{TFP}_{i,t} = \left(\frac{1}{n_{i,t}} \sum_{j=1}^{n_{i,t}} X_j \right) e^{W_{i,t} + \epsilon_{i,t}} \quad (17)$$

where $n_{i,t}$ is the number of executives in firm i at t , X_j is the ability of executive $j = 1, 2, \dots, n_{i,t}$, and $\epsilon_{i,t}$ is an iid shock unobserved to the firm or, more simply, measurement error in TFP. We estimate TFP using the [Olley and Pakes \(1996\)](#) approach. The procedure is briefly described in [Appendix B](#); full details are in [Cingano and Schivardi \(2004\)](#).

The survey contains several questions regarding the controlling shareholder. The most relevant for our purpose is “What is the nature of the controlling shareholder?”, from which we construct an indicator that groups firms into one of four control categories (see [Appendix B](#) for the details): 1) individual or family; 2) government (local or central or other government controlled entities); 3)

conglomerate, that is, firms belonging to an industrial conglomerate; 4) institution, such as banks and insurance companies, and foreign owners. We expect these different types of ownership to be characterized by different degrees of relevance of personal relationships. For instance owners of family business are likely to derive utility from controlling the firm above and beyond the pure monetary returns. Part of these returns might come from a compliant entourage and/or a group of executives who pursue the prestige of the family. A politician (the “owner” of a government controlled firm) might want executives who serve his political interests and might care little about how efficiently the firm is run. We therefore expect these types of firms to be characterized by positive values of R . Firms controlled by other entities, such as a foreign institution or a conglomerate, are instead more likely to put weight on pure monetary returns.¹² Independently from these presumptions, in the estimation exercise we will not impose any restriction on the values of R and will let the data speak.

Table 1 reports summary statistics for the firm data used in the regression analysis both for the total sample and by control type. For the total sample, on average, firms have value added of 30 million euros (at 1995 prices) and employ 691 workers of which 13 executives. The average ratio of executives to total workforce is 2.6%. Around 41% of firms are classified as medium-high and high-tech according to the OECD 2003 system and 3/4 are located in the north. Clear differences emerge according to the control type. Family firms are substantially smaller than the average (11 million euros and less than 300 employees) and specialize in more traditional activities. Importantly, they have a lower TFP level, followed by government controlled firms, while foreign firms have the highest TFP.

The executives’ data are taken from the Social Security Institute (Inps), which was asked to provide the complete work histories of *all* workers who were ever employed in an INVIND firm over the period 1981-1997. Workers are classified as blue collar (*operai*), white collar (*impiegati*) and executives (*dirigenti*). The data on workers include age, gender, area where the employee works, occupational status, annual gross earnings, number of weeks worked and the firm identifier. We only use workers classified as executives. In our preferred specification an executive turns senior

¹²We lump institutional and foreign owners together because both ownership types are not likely to be identifiable with a single individual, so that from our perspective it makes sense to assume a common R . Moreover, these two types by themselves have substantially fewer observations than family or conglomerate firms (see **Table 1**), making inference less reliable. We have experimented with five categories, distinguishing between foreign and institutions, finding similar (although less precise) results.

Table 1: Descriptive statistics: firms' characteristics, by Control type

	V.A.	Empl.	# Exec.	% Exec.	TFP	% High Tech	% North	N. obs.
All firms								
Mean	30.0	692	13.3	0.026	2.41	0.41	0.74	7,773
S.D.	127.3	3,299	29.1	0.021	0.51	0.49	0.44	
Family								
Mean	11.2	281	5.7	0.024	2.33	0.33	0.73	2,906
S.D.	18.0	420	9.8	0.016	0.46	0.47	0.44	
Conglomerate								
Mean	44.5	1024	15.0	0.026	2.44	0.40	0.82	2,390
S.D.	214.2	5,637	27.5	0.025	0.54	0.49	0.39	
Government								
Mean	40.6	1,013	21.8	0.022	2.38	0.47	0.51	687
S.D.	87.4	2,076	57.5	0.021	0.61	0.50	0.50	
Foreign								
Mean	37.0	791	19.8	0.030	2.53	0.52	0.75	1,790
S.D.	69.1	1,563	32.9	0.022	0.48	0.50	0.44	

NOTE: V.A. is value added (in millions of 1995 euros), # Exec. is the number of executives, % Exec. is the share of executives over the total number of employees, TFP is the log of total factor productivity, High Tech is the share of firms classified as medium-high and high tech according to the OECD classification system (OECD, 2003), North is the share of firms located in the North, N. obs. is the number of firm-year observations.

after five years of tenure. **Table 2** reports the statistics on executives' characteristics for the total sample and by control type. For the total sample, average gross weekly earnings at 1995 constant prices are 1,236 euros, the share of executives that have been with the firm at least five years is 0.57 and at least seven years 0.45. Executives are on average 46.5 years old and 96% are male. Family controlled firms pay lower wages to their executives and have a higher share of senior executives (62%). Executives' characteristics at conglomerate controlled firms are fairly similar to the overall ones. Government controlled firms employ older and almost exclusively male executives. Finally, foreign control firms pay their executives more, while their executives' characteristics resemble the average in terms of the tenure, age and gender composition.

Table 2: Descriptive statistics: executives' characteristics, by Control type

	Wage	ϕ_5	ϕ_7	Age	Male
All firms					
Mean	1,236	0.57	0.45	46.5	0.96
S.D.	330	0.30	0.31	4.6	0.13
Family					
Mean	1,130	0.62	0.51	46.0	0.94
S.D.	288	0.33	0.33	5.3	0.17
Conglomerate					
Mean	1,294	0.53	0.41	46.6	0.97
S.D.	321	0.28	0.28	4.0	0.09
Government					
Mean	1,298	0.54	0.42	47.7	0.99
S.D.	349	0.27	0.27	4.6	0.03
Foreign					
Mean	1,309	0.54	0.43	46.9	0.96
S.D.	361	0.29	0.29	4.1	0.12

NOTE: Wage (gross, per week) is in 1995 euros. ϕ_5 the share of executives with at least 5 years of seniority, ϕ_7 with at least 7 years. Age is the average executives' age and Male is the share of male executives.

5 Identification

This section discusses the mapping between the model and the data and, in particular, the data variability that identifies the model's parameters. We first show that the model yields a restriction

that has a natural interpretation in terms of an OLS regression run over a cross section of firms *of a given ownership type*. This regression identifies a subset of the model's parameters, while allowing to control for several unobserved variables that are not accounted for by our theory. We then show that more structural parameters can be identified by comparing the average X, ϕ values *across ownership-types*. These two sets of estimates exploit completely different dimensions of data variability. The results therefore can be compared to gain some insights on the robustness of our findings.

The model yields a simple prediction on the productivity differential between the senior and the junior executives within a given ownership-type. Fix R, q and consider a set of firms drawn from a given model parametrization. Firm i employs n_i executives. Those firms differ with respect to the quality of executives, which depends on the realizations of $x + r$ for each executive. The econometrician observes the firm's productivity X_i and the fraction of senior executives ϕ_i where the mean productivity of firm i is given by $X_i = \frac{\sum_{j=1}^{n_i} x_{i,j}}{n_i}$ and the fraction of senior executives is $\phi_i = \frac{\sum_{j=1}^{n_i} \mathcal{I}_{i,j}}{n_i}$ where $\mathcal{I}_{i,j}$ is an indicator function equal to 1 if executive j in firm i is senior. Let X_o and $X_y = \mu$ be the large sample conditional productivity of the incumbent senior and junior executives, respectively. Obviously if n_i is not very large the average productivities may differ from X_o, X_y due to sampling variability. Using [equation \(13\)](#) we establish the following proposition:

PROPOSITION 3. The productivity of firm i can be written as:

$$X_i = \mu + (X_o - \mu) \phi_i + \varepsilon_i \tag{18}$$

where $\mathbb{E}\{\varepsilon_i\} = 0$ and $\mathbb{E}\{\phi_i \varepsilon_i\} = 0$.

Proof. See [Appendix A.3](#).

A key result from this proposition is that deviations ε_i about the (large sample or unconditional) mean values are uncorrelated with the share of senior executives ϕ_i . Intuitively, this property holds since an increase (or a decrease) in the quota of senior executives ϕ_i about its unconditional mean ϕ does not contain any information on the innovation ε_i , i.e., the amount by which the productivity of the senior (junior) executive exceeds the selection threshold s^* in firm i . From a statistical point of view this result is an immediate corollary of the properties of the conditional mean. The proposition

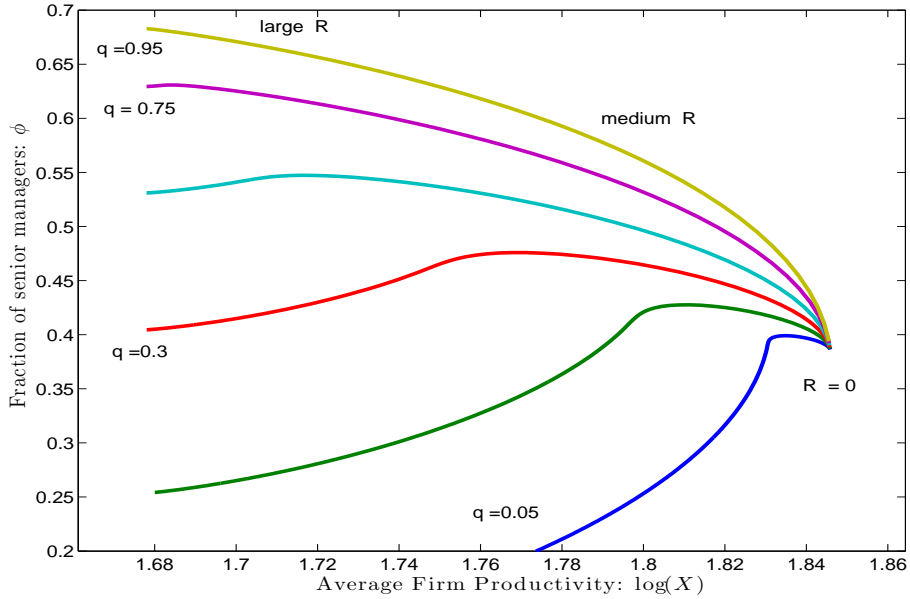
implies that the productivity differential $X_o - \mu$ can be estimated with an OLS regression of X_i on ϕ_i . The intuition is the following. When selection is weak (i.e., R is large), two firms with different shares of senior executives differ little in productivity, since on average senior executives are not much more productive than junior executives. This implies that the correlation between X and ϕ is low. If R is low, the selection mechanism is effective, senior executives are on average more productive than junior ones and differences in ϕ will go together with substantial differences in X , yielding a high correlation between X and ϕ .

Another prediction of the model concerns a comparison across ownership-types. Consider the firm productivity X and the fraction of senior executives ϕ . For a given vector of model primitives $\beta, \rho, G(\cdot)$, **Figure 3** shows that for each admissible (i.e., model generated) observable pair X, ϕ there is at most one pair of parameter values R, q that can produce it.¹³ Each line in the figure is indexed by one value of q . Increasing q shifts the locus upward: a higher probability of maturing a valuable relationship increases the likelihood of being tenured and hence ϕ . Notice that all lines depart from the same point in the X, ϕ plane, which corresponds to $R = 0$. This is the productivity maximizing situation, obtained when relationships have no value. Starting from this point, an increase in R moves the model outcomes along one line (indexed by q) from right to left. We know this because, as shown in **Proposition 2**, X is decreasing in R . Moreover, as discussed above, the effect of an increase in R on ϕ is, in general, not monotone. This explains why the lines that correspond to low values of q are hump-shaped. The important point of this figure is that those lines never cross, so that given any point in the space spanned by the model, one can invert it and retrieve the values of q and R that produced it.

Compared to this structural estimate, the regression of **Proposition 3** only identifies the productivity differential, $X_o - \mu$, and not the levels of R, q . The two approaches to estimating the model that we discussed above are useful for a number of reasons. First, they allow us to assess one key prediction of the model using different moments from the data: while the structural estimates use the average values of X, ϕ of a given control-type to back out the R, q parameters across types, the OLS regression exploits the partial correlation coefficient between X and ϕ across firms for a given R, q . The latter estimate does not depend on the average X, ϕ for a given control type, but

¹³**Section 7** discusses how we pin down the parameters that are not estimated within the structural estimation routine: β, ρ and the standard deviation of the lognormal distribution $G(\cdot)$.

Figure 3: Productivity and Seniority: space spanned by the model



Note: The figure uses the following parameters: $\beta = 0.98$ (per year), $\rho = 0.11$ (per year). $G(\cdot)$ is lognormal with log mean $\lambda_m = 1.6$, and log std $\lambda_\sigma = 0.36$ which imply that $\log(\mu) = 1.67$.

on how X and ϕ co-vary across firms of a given type. This implies that the OLS estimates are robust to potential differences in either TFP or seniority structure that, on average, affect all firms in a control type equally. For example, one might argue that, due to career considerations, foreign controlled firms are uniformly more appealing to junior executives than the other types. This would affect the structural estimates through changes in the average ϕ across control types not related to R , but would not bias the OLS estimates. A second important feature of the OLS estimates is that they do not require one to pin down all the structural parameters. In particular, they are independent of (and, of course, do not supply any information on) $\beta, q, \rho, G(\cdot)$. They therefore offer a test of robustness of the structural estimates with respect to the values of the auxiliary parameters they hinge upon. Finally, within a regression framework it is easy to perform robustness analysis, something that we exploit in the next section.

6 Model-based OLS regressions

This section presents various estimates of the model predictions discussed in [Proposition 3](#) and explores their robustness.

6.1 Basic framework and results

Equation (18) establishes a relationship between the share of senior executives and firm level TFP that we use to construct an OLS based estimate of the productivity differential between senior and junior executives. By taking the log of both sides and applying a first order Taylor series approximation around μ , we obtain:

$$\log X_i = \gamma_0 + \gamma_1 \phi_i + \eta_i \quad (19)$$

where $\gamma_0 = (1 + \log \mu)$ and $\gamma_1 = \frac{X_o - \mu}{\mu}$ and where, as shown in **Proposition 3**, η_i is uncorrelated with ϕ_i . This equation shows that, in a regression of log TFP on the share of senior executives, the coefficient γ_1 measures the percentage difference in average ability between senior and junior executives. We bring **equation (19)** to the data using the following specification:

$$\log TFP_{i,t} = \gamma_0 + \gamma_1 \phi_{i,t} + \gamma_{fam} D_{fam} \cdot \phi_{i,t} + \gamma_{gov} D_{gov} \cdot \phi_{i,t} + \gamma_{for} D_{for} \cdot \phi_{i,t} + \gamma_5 \mathbf{W}_{i,t} + \epsilon_{i,t} \quad (20)$$

where the dummies D_k with $k = fam, gov, for$ are control status dummies equal to one for family, government and foreign controlled firms, respectively, and $\mathbf{W}_{i,t}$ is a vector of controls that includes year dummies, 2-digit sector dummies and control-status dummies, that account for potential unobserved heterogeneity across firms with different control types. The coefficient γ_1 measures the percentage difference in the average ability of senior and junior executives in conglomerate controlled firms, the reference group in this specification. Under our assumption that the productivity of junior executives is the same across groups, the coefficients γ_k measure the difference in the average ability of senior executives for the corresponding control type with respect to the conglomerate firms: $\gamma_k = X_o(R_k) - X_o(R_{cong})$. To obtain population consistent estimates we weight observations with population weights, available from the INVIND survey, unless otherwise specified.

The estimates reported in Column [1] of **Table 3** show that the relationship between productivity and the share of senior executives is positive for conglomerate-controlled firms: the coefficient is 0.11 with a standard error of 0.040. The positive sign of this coefficient is consistent with the selection hypothesis, i.e., with the assumption that senior executives are on average more productive than junior ones. To give a sense of the size of the effect, the productivity of a firm with a share of

executives that is one standard deviation above the mean (see Table 2) is higher by about 3%. The estimated coefficient for the foreign-controlled firms, γ_{for} is not statistically different from zero, hence we cannot reject the hypothesis that the selection in foreign-controlled firms is similar to the one in conglomerate-controlled firms at conventional levels of significance. Instead, selection appears significantly smaller in family firms, where senior executives are on average 17% less productive than in conglomerate controlled firms. Finally, we obtain a very large negative coefficient for government firms (-0.47). This value implies a negative selection in such firms, i.e., that junior executives are more efficient than senior ones, an outcome that our model cannot predict (the worst selection we can have in the model makes the productivity of the senior executives equal to the one of the junior). This indicates that government firms display some features that the model cannot match, an issue that will also arise in the structural estimates of Section 7. Altogether, the sign and magnitude of the effect is indicative of very poor selection in government controlled firms.

Table 3: TFP and share of senior executive relationship, by Control type

	Dependent variable: log TFP					
	[1]	[2]	[3]	[4]	[5]	[6]
ϕ	0.111*** (0.040)	0.022 (0.034)	0.081* (0.044)	0.089** (0.040)	0.113** (0.044)	0.125*** (0.045)
$\phi \cdot \text{Foreign}$	-0.035 (0.063)	0.012 (0.049)	0.020 (0.069)	-0.053 (0.063)	0.017 (0.067)	-0.024 (0.069)
$\phi \cdot \text{Family}$	-0.175*** (0.047)	-0.096** (0.040)	-0.170*** (0.052)	-0.127** (0.049)	-0.107** (0.051)	-0.129** (0.052)
$\phi \cdot \text{Government}$	-0.469*** (0.088)	-0.316*** (0.074)	-0.417*** (0.105)	-0.370*** (0.089)	-0.344*** (0.109)	-0.288** (0.121)
Observations	5,875	5,943	4,897	6,840	5,136	5,136
R-squared	0.47	0.49	0.45	0.47	0.53	0.54

Note: ϕ is the share of senior executives who have been with the firm at least 5 years in columns [1],[2],[5],[6], at least 7 years in column [3] and at least 3 years in column [4]. All regressions are weighted with sampling weights with the exception of column [2], which is unweighted. All regressions include control type dummies, year dummies, 2-digit sector dummies. Column [5] also includes firm size (log of the number of employees), firm age (log), the average age of the workforce (log), the share of executives, of white collar and of male workers as a fraction of the total workforce. Column [6] includes the same additional controls as in column [5] interacted with ownership dummies. Robust standard errors in parenthesis. Significance levels for the null hypothesis of a zero coefficient are labelled as follows: * is 10%, ** is 5%, *** is 1%.

As a first robustness check, in column [2] we do not weight observations. In this case, the results

are somehow weaker: the coefficient on ϕ is positive but statistically insignificant. The interaction terms for family and government are negative and significant, again pointing to weaker selection in these firms. Another important issue is length of time assumed to become senior, 5 years in the baseline regressions. It is important to check to what extent our results depend on this choice. To do so, in Column [3] we use a 7-year based definition of seniority and in column [4] a 3-year definition. We find no substantial differences with respect to the basic specification.

6.2 Additional robustness checks and alternative hypothesis

This section discusses potential criticisms of the regressions reported above to further assess their robustness.

A first criticism of the regression analysis may concern an omitted variable bias. Our stylized theoretical model excludes other potential determinants of seniority and productivity. For example, a firm with a good human resource (HR) department might be more productive and better at retaining senior managers. Formally, let Z_i measure the quality of the HR department and assume that the correct regression equation is

$$\log X_i = \gamma_0 + \gamma_1\phi_i + \gamma_2Z_i + \eta_i$$

Then, the estimated coefficient on seniority using (19) would be equal to $\hat{\gamma}_1 = \gamma_1 + \gamma_2 \frac{\text{cov}(\phi_i, Z_i)}{\text{var}(\phi_i)}$. The omitted variable hypothesis might thus offer an alternative explanation of the correlation between seniority (ϕ) and productivity (X) that we found within each ownership type. But notice that while this bias might explain the presence of a positive correlation between productivity and seniority, more assumptions are necessary to challenge our main finding, namely that $\hat{\gamma}_1$ differs across control types in the way predicted by our theory. In particular, to reproduce our finding that the conditional correlation between X and ϕ is high in firms under conglomerate-control and small in family-firms requires to assume much more than an omitted variable, namely that either γ_2 or $\text{cov}(\phi_i, Z_i)/\text{var}(\phi_i)$ differs systematically across ownership types. Following up on the above example, one would need a theory that explains why the quality of the HR department has a *different impact* on productivity and/or on the seniority structure in, e.g., family with respect to conglomerate firms. To challenge our identification mechanism one needs an alternative theory of why the effect of an omitted variable

varies across control types. We were unable to identify any obvious alternative explanation in the literature.

Despite this important theoretical objection, we further explore the role of omitted variables empirically, as our database contains a rich set of firm characteristics, particularly on the workforce composition, owing to the matched employer-employee nature of the data. Rather than trying to propose and dismiss a specific hypothesis, we select a set of potential determinants of productivity and include them in the regression. In column [5] of [Table 3](#) we report the results when including as additional controls firm size (log of the number of employees), firm age, the average age of the workforce, the share of executives, of white collar and of male workers as a fraction of the total workforce. To save on space, we report the coefficients on this additional controls in [Table A-9](#) in the Appendix. Even with this very rich set of controls, the pattern that emerges from the data is unchanged: the share of senior executives is positively correlated with TFP in conglomerate and foreign controlled firms, while significantly lower in the other two types.

A potential criticism of the specification of column [5] is that we are giving the share of senior executives a better chance to affect the results than to the other controls, since the coefficient of the seniority share varies by control type while those of the additional controls do not. We relax this restriction in column [6], where all the additional regressors listed above are interacted with the control type dummies. Again, the results are hardly affected. Interestingly, the interaction between the additional controls and the ownership type dummies are almost all non significantly different from zero ([Table A-9](#), in the Appendix), suggesting that the differential response we find for the share of senior managers is not a general feature of the data.

One might still argue that only exogenous variation in ϕ can conclusively dismiss omitted variable bias (as well as any other endogeneity concern). We argue that this is not the case. The estimates of the productivity differential, $X_o - \mu$, are based on the *model predicted correlation* between the share of senior managers and productivity and do not reflect a generic form of causation from seniority to productivity. In fact, changes in ϕ not attributable to the selection mechanism will not identify $X_o - \mu$. For example, an exogenous increase in the share of senior managers derived from a tightening of labor market regulation that makes firing more costly would reduce the selection and weaken the seniority-productivity correlation.¹⁴ In other words, the interpretation of our estimate is the correct

¹⁴Note that in Italy, over the period considered, executives can be fired at will, so this issue does not arise.

one under our maintained assumption that the theoretical model is the data generating process. In this sense the estimates are structural, and the OLS correlation, as opposed to instrumental variable estimation, is the proper way to measure selection consistently with our model.

As a final set of robustness checks, we have experimented with the measure of performance. We have estimated the production function directly, rather than using the two-step procedure that first estimates TFP and then relates it to the share of senior executives. The results are reported and discussed in [Appendix B](#). They are aligned with those of [Table 3](#). Moreover, we have also experimented with a profit-based measure of performance. In fact, one might argue that profits is the correct indicator of managerial ability, and that efficiency and profitability might not be simply one-to-one ([Foster, Haltiwanger, and Syverson, 2008](#)). In the OLS framework we can directly test if the patterns that emerge for productivity also hold for profitability measures. In [Table 4](#) we use return on assets (ROA) as the dependent variable. In this case, the coefficient on the share of senior executives can be interpreted as the difference in the average contribution to ROA of senior and junior executives. The patterns we find are exactly the same as those emerging when the productivity measure is used, if anything stronger. In particular, profitability is positively related to the share of senior executives in conglomerate controlled firms.¹⁵ The interaction for foreign firms is not significantly different from zero, while it is negative and significant for family and government firms. Again, for the latter the effect is very strong, indicating negative selection in such firms. This is fully consistent with the results obtained when performance is measured by TFP. In particular, in family and government firms the selection effect for senior executives is absent compared to the other control types. Moreover, the results are robust to all the additional checks performed for TFP, reported in columns [2]-[6]. Similar results are obtained when using return on equity (not reported for brevity). This shows that our results are robust with respect to different performance measures.

To conclude, the OLS estimates indicate that there are substantial differences in the effectiveness of the selection process of executives across different types of owners. We have argued that such differences are not likely to be due to omitted variable bias or to one specific performance indicator. We now turn to the structural estimation exercise, which will allow us to check if these conclusions are confirmed and to assess the effects of the private benefits from the personal relationships on

¹⁵To get a sense of the size of the estimated effect, considering a firm whose share of senior executives increases by one standard deviation above the mean (see [Table 2](#)) would increase ROA by 1 percentage point (the median ROA is 7.8, the mean 8.6).

Table 4: ROA and share of senior executives, by control type

	Dependent variable: ROA					
	[1]	[2]	[3]	[4]	[5]	[6]
ϕ	3.40*** (0.752)	2.93*** (0.658)	2.30*** (0.852)	3.22*** (0.741)	3.952*** (0.863)	4.013*** (0.863)
$\phi \cdot \text{Foreign}$	-1.57 (1.377)	0.24 (1.068)	-0.92 (1.352)	-0.71 (1.360)	0.323 (1.599)	0.344 (1.645)
$\phi \cdot \text{Family}$	-4.48*** (0.951)	-3.69*** (0.781)	-3.32*** (1.070)	-4.25*** (0.964)	-3.885*** (1.047)	-4.052*** (1.051)
$\phi \cdot \text{Government}$	-7.50*** (1.612)	-3.73** (1.500)	-4.71** (1.837)	-7.22*** (1.669)	-5.817*** (2.108)	-5.953** (2.315)
Observations	5,875	5,943	4,897	6,840	5,136	5,136
R-squared	0.08	0.09	0.08	0.08	0.12	0.13

Note: ROA is return on assets, in percentage units. ϕ is the share of senior executives who have been with the firm at least 5 years in columns [1],[2],[5],[6], at least 7 years in column [3] and at least 3 years in column [4]. All regressions are weighted with sampling weights with the exception of column [2], which is unweighted. All regressions include control type dummies, year dummies, 2-digit sector dummies. Column [5] also includes firm size (log of the number of employees), firm age (log), the average age of the workforce (log), the share of executives, of white collar and of male workers as a fraction of the total workforce. Column [6] includes the same additional controls as in column [5] interacted with ownership dummies. Robust standard errors in parenthesis. Significance levels for the null hypothesis of a zero coefficient are labelled as follows: * is 10%, ** is 5%, *** is 1%.

firms' productivity through counterfactual exercises.

7 Estimates of the structural parameters

This section describes the estimation of the model structural parameters. We begin by discussing the assumptions needed for the estimation of the model parameters using firm level observations on TFP and seniority of the executives, i.e., what we see as empirical measures of X and ϕ . The exercise assumes that the data are generated by the model and observed with classical measurement error. The estimation is developed under the assumption of observed heterogeneity, as some structural parameters are linked to observable characteristics of the firm, along the lines of [Alvarez and Lippi \(2009\)](#). Assuming that the distribution of productivity $G(x)$ is lognormal, the model is characterized by six fundamental parameters: the discount factor β , the hazard rate ρ , the lognormal parameters $\lambda_m, \lambda_\sigma$ (log mean and log standard deviation, respectively), the probability of developing a relationship, q , and the value of the relationship R . Five parameters, namely β ,

ρ , q , λ_σ and λ_m , are assumed common to all firms. The parameter R is assumed to vary with one observable characteristic of the firm: the control type. Given a parametrization, the model uniquely determines the values of X, ϕ to be observed in the data. Differences between datapoints with identical observables (e.g. two firms with the same control type) are accounted for by classical measurement error. Next, we fill in the details that relate to the data used in the estimation, and describe the estimation algorithm and the results.

Our parsimonious structural estimate concerns six parameters: $\theta_p \in \Theta_{6,1}$, $p = 1, 2, \dots, 6$, where θ_6 gives the probability of developing a relationship $q = \frac{\theta_6}{1+\theta_6}$. The firm-level observations vary across 14 years (index t), 13 two-digit sectors (index τ), and four control types described in the previous section (index κ). We assume that R varies across firms according to the nature of the controlling shareholder, with $R_\kappa = \theta_\kappa$, $\kappa = 1, 2, 3, 4$ for the firms under, respectively, family (θ_1), conglomerate (θ_2), government (θ_3), and foreign (θ_4) control. The technological parameter $\lambda_m = \theta_5$ is related to the mean TFP of the firm. In the data TFP has a clear time component, as well as a sectoral one, that are ignored by the simple structure of our model. Thus, before turning to the structural estimation, we normalize the TFP data by removing common time and sector effects. Our measure of X for firm i in year t and sector τ is thus given by

$$\log X_{i,t,\tau} \equiv \log TFP_{i,t,\tau} - \mathbf{a}_1 \cdot \mathbf{year}_{i,t} - \mathbf{a}_2 \cdot \mathbf{sect}_{i,\tau} - a_3 \cdot Z_{i,t,\tau} \quad (21)$$

where \mathbf{a}_1 and \mathbf{a}_2 are the vector of coefficients from an OLS regression of TFP on 13 year and 12 two digit sector dummies. Below we also consider a specification that controls for the effect of firm size $Z_{i,t,\tau}$ (log employment) on TFP.

To reduce the computational burden some parameters are pinned down outside the estimation routine. We calibrate the time discount β to an annual value of 0.98, as standard in the literature. The hazard rate of senior executives, ρ , is computed from the survival function of senior executives, that is with at least five years of seniority, using the [Kaplan and Meier \(1958\)](#) estimator on the individual data. We estimate $\rho = 0.11$ per year, which implies that the expected tenure of senior executives is approximately 10 years. The variance of the talent distribution is computed using the junior executives' compensation data. The idea is that junior executives' compensation, being

independent from the private benefits R , reflects on average the executive's individual ability.¹⁶ We therefore regress the log wage on age, age squared, firm size (log of the employees) and dummies for years, sectors, control type and seniority (from 1 to 5) and take the standard deviation of the residuals as our measure of the standard deviation of the ability distribution. We find a value $\lambda_\sigma = 0.36$, which varies very little with respect to changes in the set of controls. The sensitivity of the structural estimates to the values of λ_σ is discussed at the end of this section.

These assumptions imply that, after removing time and sectoral differences, all firms in a group –indexed by the control type $\kappa = 1, 2, 3, 4$ – are expected to have the same X and ϕ . For each firm i in group κ there are two observables $y_{i,\kappa}^j$, $j = 1, 2$. We assume that the variable $y_{i,\kappa}^j$ is measured with error ε_i^j that is normal, with zero mean, independent across variables, groups and observations. Inspection of the raw data suggests that measurement error is multiplicative in levels for TFP, X , and additive for the share of senior executives, ϕ .¹⁷ Hence the ML estimates use the following observables $y_{i,\kappa}^1 = \log X_{i,\kappa}$ and $y_{i,\kappa}^2 = \phi_{i,\kappa}$. The measurement error variances σ_j^2 , $j = 1, 2$ is assumed common across groups, and is computed as the variance of the residuals of an OLS regression of $\log X$ and ϕ on year and sector dummies. This gives $\sigma_{\log X}^2 = 0.35$ and $\sigma_\phi^2 = 0.29$.

Let $f^j(\Theta, \kappa)$ be the model prediction for the j^{th} variable in group κ under the parameter setting Θ . The observation for the corresponding variable for firm i in group κ is

$$y_{i,\kappa}^j = f^j(\Theta, \kappa) + \varepsilon_i^j.$$

Let Y be the vector of observations and n_κ be the number of firms i in group κ . Define the objective function F as

$$F(\Theta; Y) \equiv \sum_{\kappa=1}^4 \sum_{j=1}^2 \left(\frac{n_\kappa}{\sigma_j^2} \right) \left(\frac{\sum_{i=1}^{n_\kappa} y_{i,\kappa}^j}{n_\kappa} - f^j(\Theta, \kappa) \right)^2 \quad (22)$$

Appendix C shows that the likelihood function is related to the objective function by:

$$\log L(\Theta; Y) = -\frac{1}{2} \sum_{\kappa=1}^4 \sum_{j=1}^2 n_\kappa (1 + \log(2\pi\sigma_j^2)) - \frac{1}{2} F(\Theta; Y)$$

¹⁶In the model, ability is revealed abruptly when an executive turns senior. In theory, therefore, all junior executives should be paid the same wage, as the owner has no information on ability. Of course, in reality the process of learning about executives skills is more gradual, so that the wage over the junior period does convey some information about individual ability.

¹⁷This statement is based on an analysis of the deviations of X and ϕ (in levels and in logs) from the mean of each groups. Details are available from the authors upon request.

We estimate the six structural parameters in Θ by minimizing (22). At each iteration, the algorithm solves the model for each of the four groups and computes the objective function under a given parametrization. Since each group has two observables there is a total of eight moments to be fitted using 6 parameters, hence the model is over-identified with two degrees of freedom. The formulas for the score and the information matrix used for the inference are derived in [Appendix C](#).

Table 5: Structural estimates of model parameters

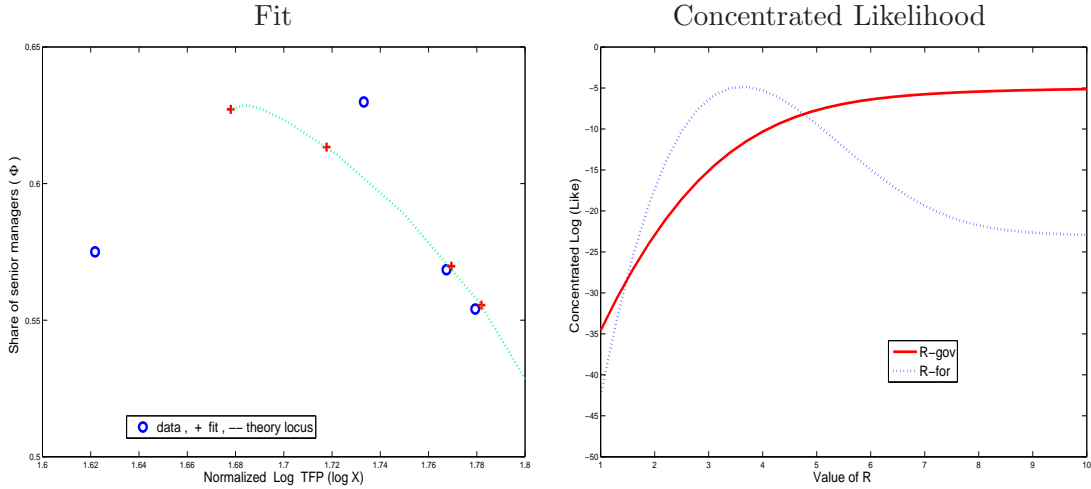
	q (θ_6)	R_{fam} (θ_1)	R_{cong} (θ_2)	R_{govt} (θ_3)	R_{forgn} (θ_4)	λ_m (θ_5)
A: Baseline	0.74 (5.6)	5.6 (10.1)	3.2 (12.5)	41.4 (0.0)	3.6 (10.0)	1.61 (173)
B: With firm-size	0.72 (5.7)	5.1 (10.3)	3.4 (11.8)	86.9 (0.0)	3.7 (8.9)	1.61 (166)

Note: t-statistics in parenthesis. The estimation assumes $\sigma_{\log X}^2 = 0.35$ and $\sigma_\phi^2 = 0.29$. The parameters $\beta, \rho, \lambda_\sigma$ are fixed from an auxiliary estimation (see the main text). The measure for productivity $\log X$ is the firm level (log) TFP net of the common components due to year effects and sector effects, as from [equation \(21\)](#) (see the main text); in the specification B firm-size effects (as measured by the the (log) number of employees) are also controlled for.

The structural estimates of the model parameters are reported in [Table 5](#). The estimated value of $q = 0.74$ indicates that approximately 3/4 of the executives develop a relationship. The value of R varies substantially across control types, but all types enjoy some degree of private benefits. This is consistent with the findings of [Taylor \(2010b\)](#), according to which CEO entrenchment is substantial even in US listed firms, where family and government firms play a minor role. The importance of the relationships is lowest for firms belonging to a conglomerate (3.2), followed by foreign controlled firms (3.6), family (5.6), and government (41.4). Given that the estimated unconditional mean level of TFP is around 7, the estimated values for R show that the non-monetary characteristics of the executives (i.e., their relationship value) are quantitatively important in the selection process. The values are all strongly statistically significant but for the government case. The reason for the lack of significance for the coefficient of government-controlled firms is revealed by the analysis of the likelihood function in [Figure 4](#).

The left panel of the picture shows that the observations to be fitted for the government control group are outside the space that can be spanned by the model in the X dimension. The TFP level for government-controlled firms is below μ , the unconditional mean value of X . In attempting to fit such a low value the model uses a high value of R (which reduces X by [Proposition 2](#)). Notice

Figure 4: Fit of estimates



Note: This figure uses the parameters reported in the baseline estimate of Table 5.

that the model has an asymptote: once R is so large that the selection threshold $s^* - R$ hits zero, then we have that $X \cong \mu$. There is no selection because senior executives are retained in office if and only if $r = R$, independently from their ability, and their average ability equals that of an unselected pool of junior executives. This is the lowest possible value achievable by the model. This effect is clearly seen from the analysis of the concentrated likelihood function reported in the right panel of the figure.¹⁸ The picture clearly shows that in the estimation of the government $R = \theta_3$ the concentrated likelihood function (the thick line) becomes virtually flat for values of $\theta_3 \geq 8$. The null hypothesis that $\theta_3 = 0$ is strongly rejected at the 1 per cent confidence level, as the large differences in the value of the log likelihood on the vertical axis indicate. This evidence shows that the value of R for the government controlled firms is such that selection is completely absent, so that the ML estimate is not able to identify a single value of R in the range $(8, +\infty)$. This is in stark contrast with, e.g., the estimate for R in foreign controlled firms, where the concentrated likelihood is single peaked (dashed line) around 3.6 (the ML estimate of Table 5).

To understand the implications of the estimates Table 6 reports some statistics on the model's steady state produced using the estimates of Table 5. The first column solves the model for the (counterfactual) case in which the firm's principal gives no value to relationships in executive selec-

¹⁸This is computed by evaluating the likelihood function along the parameter of interest (θ_3 for the government R) freezing all other parameters at their estimated values reported in Table 5.

Table 6: Model predictions under the benchmark estimates

	Control type				
	$R = 0$	Family	Conglomerate	Government	Foreign
R	0	5.58	3.21	41.4	3.62
s^*	6.21	8.61	7.39	29.2	7.59
x^*	6.21	3.03	4.18	0	3.96
$\log \mu$	1.68	1.68	1.68	1.68	1.68
$\log X$	1.85	1.72	1.78	1.68	1.77
$\log X_o$	2.07	1.74	1.86	1.68	1.83
$\log X_o R = 0$	2.07	2.32	2.2	3.55	2.22
$\log X_o R > 0$	2.07	1.72	1.83	1.68	1.81
ϕ	0.39	0.61	0.56	0.63	0.57
Fired	0.72	0.30	0.45	0.26	0.41
Fired $ R = 0$	0.72	0.93	0.86	1	0.87
Fired $ R > 0$	0.72	0.080	0.31	0	0.26

The table reports key statistics for the steady state of our model solved using the benchmark estimates of Table 5. $\log X$ is the (log) average managerial ability, $\log X_o$ is the average managerial ability of the senior executives, $\log X_o|R = 0$ is the average ability of the senior that did not develop a relationship, $\log X_o|R > 0$ for those that did develop a relationship. Fired is the probability that a junior executive is replaced when turning senior.

tion ($R = 0$) and hence $s^* = x^*$ for all senior executives. In this case, senior executives are confirmed if their ability x is above $x^* = 6.21$, which occurs in around 25% of cases. The (log) average ability of senior executives is 2.07, almost 30 log points higher than the unconditional average ability of junior ones. On average, around 39% of executives are senior. These figures can be compared to the ones obtained for family firms (second column of the table), for which $s^* = 8.61$ and $s^* - R = 3.01$ (the latter is the cutoff ability of executives that have developed a relationship with the owner). It is apparent that selection is much weaker in family owned firms: the senior executives' average ability (X_o) is only 6% higher than that of junior executives (μ). An executive who develops a relationship has a 92% chance of being tenured. This probability drops to 7% for a senior executive who does not develop a relationship. For conglomerate and foreign controlled firms, the situation is intermediate between the $R = 0$ case and the family-firm case. For government firms, instead, the estimates imply that selection occurs exclusively on the basis of developing a relationship: all executives with $r = R$ are retained, all others are fired. As a consequence, the average ability of senior executives is identical to that of junior executives: the selection effect is completely inhibited.

The structural estimates allow us to analyze simple counterfactual exercises by comparing the

effects of private benefits across the different ownership types. Setting the private benefits to zero implies a productivity gain of 7% in conglomerates, 8% in foreigners, 13% in family and 17% in Government firms. If we take conglomerate-owned firms as a benchmark, who record the lowest value private benefits, productivity is around 6% lower in family firms and 10% lower in government firms due to weaker selection. Again, this happens because controlling shareholders of family and government firms select executives almost exclusively on the basis of personal ties, independent of ability, thus reducing the productivity enhancing effect of managerial selection.

An important robustness check consists in comparing the structural and the OLS estimates. As argued above, the OLS estimates measure the difference in the average ability of senior and junior managers: $X_o - \mu$. It is immediate to compute this statistic for the structural estimates as well. In [Table 7](#) we report the estimates of the first column of [Table 3](#) and the corresponding values obtained from the structural estimates, derived from [Table 6](#). First, both estimates imply a positive effect of selection on the senior-junior productivity differential: 0.11 in the OLS and 0.18 in the structural. The other columns report the difference in the ability of the senior managers for the relevant control type as compared to the conglomerate, the most efficient type. For foreign firms, the difference is -0.03 for both estimations. For family firms, again both estimates show that selection is substantially weaker than for conglomerate-controlled firms. Although the OLS estimate is larger in absolute value than the structural estimate (-0.17 vs. -0.12), with a standard error of 0.05 in the OLS estimates, we cannot reject the hypothesis that the two values are the same at conventional levels of significance. Finally, the OLS give a very large negative coefficient for government firms (-0.47). This value implies a negative selection in such firms, i.e., that junior executives are more efficient than senior ones. This is a result that the structural model cannot deliver: in fact, the structural estimate imply the lowest possible selection effect: with -0.18 , the average ability of senior and junior managers is the same. All in all, it is remarkable that the two sets of estimates give comparable results, although based on totally different dimensions of data variability.

We have checked the robustness of our estimates along several dimensions. First, we saw in [Table 1](#) that average firm size differs across control types. We therefore re-estimated [equation \(21\)](#) also including firm size among the determinants of TFP. We estimated the model with this measure of ability. The results are reported in the lower panel of [Table 5](#) and are similar to those without

Table 7: Comparison between the structural and the OLS estimates

	$\log X_{o,\text{con}} - \log \mu$	$\Delta X_{o,\text{for}}$	$\Delta X_{o,\text{fam}}$	$\Delta X_{o,\text{gov}}$
OLS estimates	0.11	-0.03	-0.17	-0.47
Structural estimates	0.18	-0.03	-0.12	-0.18

Note: the first row reports the OLS regression estimates from Column [1] of Table 3. The first column represents the the difference in the senior-junior ability in conglomerate firms. Columns [2]-[4] report the difference in the senior executives' average ability for each control type with respect to Conglomerates: $\Delta X_{o,i} \equiv \log X_{o,i} - \log X_{o,\text{con}}$, $i = \text{for, fam, gov}$.

firm size. Another important parameter relates to the standard deviation of managerial ability. To check the sensitivity of the results to this parameter, Table 8 reports the results of two alternative estimation exercises that use two different values of λ_σ , equal respectively to 0.5 and 1.5 times the value used in the baseline estimates. First, the table shows that the estimated patterns for R are robust. In all exercises government and family controlled firms have the largest values of R , while the enterprises controlled by conglomerates are the most efficient (smallest R). Moreover, the effects of selection are more important the higher is the dispersion of ability. The differences in the ability of junior and senior executives increase with λ_σ (conditional on performing some selection). Naturally, when ability is very dispersed selection is very effective in increasing productivity. Second, a higher value of λ_σ increases the level of the estimated R .¹⁹ The reason is simple: increasing the dispersion level enhances the effect of selection. Thus an increase in the variance of ability – keeping the mean constant– allows the principal to achieve a much better average productivity. A larger value of R is thus necessary to offset this force and keep the mean productivity predicted by the model aligned with the data. Finally, the share of senior executives changes only marginally for the different values of λ_σ . This shows that changes in λ_σ give rise to qualitatively similar results and that the quantitative differences can be easily understood within the logic of the model.

¹⁹The exception is for family firms, where R is higher in the low λ_σ case. This is because the estimate in this case has reached the asymptote, as was the case for the government firms in the baseline estimates. This is apparent from the fact that selection is totally absent for family firms (as well as for government firms) for this parametrization.

Table 8: Estimation results: sensitiveness with respect to λ_σ

	Control type				
	$R = 0$	Family	Conglomerate	Government	Foreign
	$\lambda_\sigma = .18$				
R	0	16.2	1.59	13.3	1.88
s^*	6	14.8	6.59	13.2	6.73
x^*	6	0	5.01	0	4.85
$\log \mu$	1.72	1.72	1.72	1.72	1.72
$\log X$	1.8	1.72	1.77	1.72	1.77
$\log X_o$	1.91	1.72	1.82	1.72	1.8
$\log X_o R = 0$	1.91	2.74	1.98	2.63	2
$\log X_o R > 0$	1.91	1.72	1.8	1.72	1.79
ϕ	0.41	0.62	0.55	0.62	0.57
Fired	0.69	0.28	0.45	0.28	0.42
Fired $ R = 0$	0.69	1	0.85	1	0.87
Fired $ R > 0$	0.69	0	0.31	0	0.25
	$\lambda_\sigma = .54$				
R	0	7.28	5.1	116	5.54
s^*	6.51	9.44	8.33	70.9	8.54
x^*	6.51	2.16	3.23	0	3
$\log \mu$	1.65	1.65	1.65	1.65	1.65
$\log X$	1.9	1.72	1.79	1.65	1.77
$\log X_o$	2.23	1.76	1.88	1.65	1.85
$\log X_o R = 0$	2.23	2.52	2.42	4.54	2.44
$\log X_o R > 0$	2.23	1.71	1.83	1.65	1.8
ϕ	0.36	0.61	0.56	0.62	0.57
Fired	0.75	0.32	0.44	0.28	0.41
Fired $ R = 0$	0.75	0.91	0.87	1	0.88
Fired $ R > 0$	0.75	0.09	0.27	0	0.22

The table reports the key statistics from the model estimated under two values of λ_σ . The choice of the auxiliary parameters is the same as in Table 5. $\log X$ is the average managerial ability, X_o is the average managerial ability of the senior executives, $\log X_o|R = 0$ is the average ability of the senior that did not develop a relationship, $\log X_o|R > 0$ for those that did develop a relationship. Fired is the probability that a junior executive is replaced when turning senior.

8 Concluding remarks

We formulated a model of executive selection in which the firm's owner cares about managerial ability and, in addition, derives a private benefit from developing a personal relationship with the executives. The theory yields joint predictions on two observables: the firm's average productivity

and the share of senior executives in the firm. Compared to an owner who is only interested in ability, the selection of executives in the case of multiple objectives reduces the productivity of the firm and the rate at which executives leave the company. These predictions can be “inverted” to infer the structural parameters of the model, in particular to learn how important is the personal relationship value of executives enjoyed by the firm’s owner.

A structural estimation of the model, based on matched employer-employee data in a sample of Italian manufacturing firms, shows that the non-monetary objectives appear quantitatively important in accounting for the data. In particular, the value of personal relationship is highest in the firms under government control, and smallest in conglomerate or foreign-owned firms. From a quantitative point of view, those differences account for a 10% differential in the firms’ total factor productivity. These results are robust to several controls and estimation methods.

One important question is what mechanisms could mitigate the inefficiency in executive selection we identify. We expect that competition in the product markets and contestability of control would reduce the extent of such inefficiency. We plan to explore this in future work.

References

- Albuquerque, R. and E. Schroth. 2010. “Quantifying private benefits of control from a structural model of block trades.” *Journal of Financial Economics* 96:33–55.
- Alvarez, F. E. and F. Lippi. 2009. “Financial Innovation and the Transactions Demand for Cash.” *Econometrica* 77 (2):363–402.
- Bandiera, O., I. Barankay, and I. Rasul. 2008. “Social Connections and Incentives in the Workplace: Evidence From Personnel Data.” *Econometrica* Forthcoming.
- Bandiera, O., L. Guiso, A. Prat, and R. Sadun. 2009. “Matching Firms, Managers, and Incentives.” Mimeo, LSE.
- Bartelsmann, E. and M. Doms. 2000. “Understanding productivity: Lessons from longitudinal micro-data.” *Journal of Economic Literature* 38:569–595.
- Becker, Gary S. 1971. *The Economics of Discrimination (Economic Research Studies)*. University of Chicago Press.
- Becker, G.S. 1998. *Accounting for tastes*. Harvard University Press, Cambridge, Ma.
- Bernard, A.B., J. Eaton, J.B. Jensen, and S. Kortum. 2003. “Plants and productivity in international trade.” *The American Economic Review* 93:1268–1290.
- Bloom, N. and J. Van Reenen. 2007. “Measuring and Explaining Management Practices Across Firms and Countries.” *Quarterly Journal of Economics* 122:1351–1408.
- Cingano, F. and F. Schivardi. 2004. “Identifying the sources of local productivity growth.” *Journal of the European Economic Association* 2:720–742.
- Dyck, A. and Luigi Zingales. 2004. “Private Benefits of Control: An International Comparison.” *Journal of Finance* 59:537–600.
- Ericson, R. and A. Pakes. 1995. “Markov-Perfect Industry Dynamics: A Framework for Empirical Work.” *Review of Economic Studies* 62:53–82.
- Faccio, M. and L.H.P. Lang. 2002. “The Ultimate Ownership of Western European Corporations.” *Journal of Financial Economics* 65:365–395.
- Foster, L., J. Haltiwanger, and C. Syverson. 2008. “Reallocation, firm turnover, and efficiency: Selection on productivity or profitability?” *American Economic Review* 98:394–425.
- Gabaix, X. and A. Landier. 2008. “Why Has CEO Pay Increased So Much?” *Quarterly Journal of Economics* 123:49–100.
- Garrett, D.F. and A. Pavan. 2010. “Managerial Turnover in a Changing World.” Mimeo, Northwestern University.
- Gayle, G., L. Golan, and R. Miller. 2009. “Promotion, Turnover and Compensation in the Executive Market.” Mimeo, Carnegie-Mellon university.
- Hopenhayn, H. 1992. “Entry, Exit, and Firm Dynamics in Long Run Equilibrium.” *Econometrica* 60:1127–1150.

- Iranzo, S., F. Schivardi, and E. Tosetti. 2008. "Skill dispersion and Productivity: an Analysis with Employer-Employee Matched Data." *Journal of Labor Economics* 26:247–285.
- Jovanovic, B. 1982. "Selection and the Evolution of Industry." *Econometrica* 50:649–670.
- Kaplan, E.L. and P. Meier. 1958. "Nonparametric estimation from incomplete observations." *Journal of the American Statistical Association* 53:457–481.
- Kramarz, F. and D. Thesmar. 2006. "Social Networks in the Boardroom." CEPR discussion papers 5496.
- La-Porta, R., F. Lopez-De-Silanes, and A. Shleifer. 1999. "Corporate Ownership Around the World." *Journal of Finance* 54:471–517.
- La Porta, R., F. Lopez-de Silanes, A. Shleifer, and R. Vishny. 2000. "Investor protection and corporate governance." *Journal of Financial Economics* 58:3–27.
- Lucas, Robert E., Jr. 1978. "On the Size Distribution of Business Firms." *Bell Journal of Economics* 2:508–523.
- McCall, John J. 1970. "Economics of Information and Job Search." *The Quarterly Journal of Economics* 84 (1):113–26.
- Melitz, M.J. 2003. "The impact of trade on intra-industry reallocations and aggregate industry productivity." *Econometrica* 71:1695–1725.
- Michelacci, C. and F. Schivardi. 2010. "Does idiosyncratic business risk matter?" Mimeo, CEMFI and EIEF.
- Moskowitz, T.J. and A. Vissing-Jørgensen. 2002. "The returns to entrepreneurial investment: A private equity premium puzzle?" *American Economic Review* 92:745–778.
- OECD. 2003. *OECD Science, Technology and Industry Scoreboard, Annex 1*. OECD.
- Olley, S. G. and A. Pakes. 1996. "The Dynamics of Productivity in the Telecommunications Equipment Industry." *Econometrica* 64:1263–1297.
- Syverson, C. 2010. "What Determines Productivity?" *Journal of Economic Literature* Forthcoming.
- Taylor, L.A. 2010a. "CEO Wage Dynamics: Evidence from a Learning Model." Mimeo, Wharton School.
- . 2010b. "Why are CEOs rarely fired? Evidence from structural estimation." *Journal of Finance* Forthcoming.
- Terviö, M. 2008. "The difference that CEOs make: An assignment model approach." *American Economic Review* 98:642–668.
- Volpin, P.F. 2002. "Governance with poor investor protection: Evidence from top executive turnover in Italy." *Journal of Financial Economics* 64:61–90.
- Weitzman, Martin L. 1979. "Optimal Search for the Best Alternative." *Econometrica* 47 (3):641–654.

A Proofs

This appendix provides the proofs of the three propositions in the paper. The arguments are based on standard analysis and probability notions.

A.1 Proof of Proposition 1.

Proof. Simple algebra shows that $H(s^*, R)$: is continuous in s^* , that $H(0, R) < 0$, and that the first order derivative w.r.t. s^* is positive, $H_{s^*}(s^*, R) = 1 + \beta(\rho - F(s^*)) > 0$, and in the limit $\lim_{s^* \rightarrow \infty} H_{s^*}(s^*, R) > 0$. Hence there exist one and only one $s^* > 0$ that solves equation (9).

We now show that the implicit function $s^*(R)$ is increasing in R . Applying the implicit function theorem to equation (9) gives

$$\frac{\partial s^*}{\partial R} = \frac{(1 - \beta) \frac{\partial v_y}{\partial R}}{1 - (1 - \beta) \frac{\partial v_y}{\partial s^*}} \quad (\text{A-23})$$

Let us use expression (7) to compute:

$$\begin{aligned} \frac{\partial v_y}{\partial s^*} &= \frac{-\beta s^* f(s^*) [(1 - \beta)(1 + \beta(\rho - F(s^*))) + \beta(1 - \beta)f(s^*) [\mu(1 - \beta(1 - \rho)) + \beta \int_{s^*}^{\infty} s dF(s)]]}{[(1 - \beta)(1 + \beta(\rho - F(s^*)))]^2} \\ &= \beta f(s^*) \frac{-s^* + (1 - \beta) \frac{\mu(1 - \beta(1 - \rho)) + \beta \int_{s^*}^{\infty} s dF(s)}{(1 - \beta)(1 + \beta(1 - F(s^*)))}}{(1 - \beta)(1 + \beta(1 - F(s^*)))} = \beta f(s^*) \frac{-s^* + (1 - \beta)v_y}{(1 - \beta)(1 + \beta(1 - F(s^*)))}. \end{aligned}$$

Using that at the optimum $(1 - \beta)v_y = s^*$ gives $\frac{\partial v_y}{\partial s^*} = 0$. Hence $\frac{\partial s^*}{\partial R} = (1 - \beta) \frac{\partial v_y}{\partial R}$. Next, we show that $0 < (1 - \beta) \frac{\partial v_y}{\partial R} < 1$. Rewrite the integral term in the numerator of (7) as

$$\begin{aligned} \int_{s^*}^{\infty} s dF(s) &= q \int_{s^* - R}^{\infty} (x + R) dG(x) + (1 - q) \int_{s^*}^{\infty} x dG(x) \\ &= q \left(R(1 - G(s^* - R)) + \int_{s^* - R}^{\infty} x dG(x) \right) + (1 - q) \int_{s^*}^{\infty} x dG(x) \end{aligned}$$

Using this expression in (7) and taking the derivative w.r.t. R yields:

$$\begin{aligned} (1 - \beta) \frac{\partial v_y}{\partial R} &= \beta \frac{\frac{\partial \int_{s^*}^{\infty} s dF(s)}{\partial R} + (1 - \beta) v_y \frac{\partial F(s^*)}{\partial R}}{1 + \beta(1 - F(s^*))} \\ &= \beta q \frac{[1 - G(s^* - R) + Rg(s^* - R) + (s^* - R)g(s^* - R)] - s^*g(s^* - R)}{1 + \beta(1 - F(s^*))} \\ &= \beta q \frac{1 - G(s^* - R)}{1 + \beta(1 - F(s^*))} \in (0, 1) \end{aligned}$$

where the second equality uses $s^* = (1 - \beta)v_y$.

Finally we show that $s^* > \mu$. Note that for $R = 0$ equation (1) gives $F(z) = G(z)$. Using equation (9) to evaluate and $H(s^*, R)$ at $R = 0$ gives

$$H(s^*, 0) = (s^* - \mu)(1 + \beta\rho) + \beta \left(\mu - \int_{s^*}^{\infty} z dG(z) - s^* G(s^*) \right)$$

Simple algebra shows that at $s^* = \mu$ we have that $H(\mu, 0) < 0$ (since $\mu - \int_{\mu}^{\infty} z dG(z) < 0$). Using that $H(s^*, R)$ is increasing in s^* implies that $s^*(0) > \mu$. Using that $s^*(R)$ is increasing in R , implies

that $s^*(R) > \mu$ for any $R \geq 0$. ■

A.2 Proof of Proposition 2.

Proof. Rewrite the average productivity X defined in equation (13) as

$$X = \mu + \frac{q \int_{s^*-R}^{\infty} x dG(x) + (1-q) \int_{s^*}^{\infty} x dG(x) - \mu(1-F(s^*))}{1 + \rho - F(s^*)}$$

The parameter R enters this expression directly and via s^* . Taking the first order derivative with respect to R , accounting for both direct and indirect effects, gives (after some algebra and collecting terms):

$$\begin{aligned} \frac{\partial X}{\partial R} = & \frac{1}{1 + \rho - F(s^*)} \cdot \left[q R g(s^* - R) \frac{\partial(s^* - R)}{\partial R} + \right. \\ & \left. \frac{\partial F(s^*)}{\partial R} \left(\frac{\rho \mu + q \int_{s^*-R}^{\infty} x dG(x) + (1-q) \int_{s^*}^{\infty} x dG(x) - \mu(1-F(s^*))}{1 + \rho - F(s^*)} - s^* \right) \right] \end{aligned} \quad (\text{A-24})$$

Now use equation (9) with $\beta = 1$ to get the following implicit equation for s^*

$$s^* = \frac{\rho \mu + \int_{s^*}^{\infty} s dF(s)}{1 + \rho - F(s^*)} = \frac{\rho \mu + q \int_{s^*-R}^{\infty} x dG(x) + (1-q) \int_{s^*}^{\infty} x dG(x) + q R (1 - G(s^* - R))}{1 + \rho - F(s^*)}$$

Replacing the expression on the right hand side for s^* into equation (A-24) and using the expression for $\frac{\partial F(s^*)}{\partial R}$ computed in equation (12) gives (after some rearranging and cancellations)

$$\begin{aligned} \frac{\partial X}{\partial R} = & \frac{q R}{1 + \rho - F(s^*)} \left[g(s^* - R) \left(\frac{1 + \rho - (q + (1-q)G(s^*))}{1 + \rho - F(s^*)} \right) \frac{\partial(s^* - R)}{\partial R} \right. \\ & \left. - \frac{1 - G(s^* - R)}{1 + \rho - F(s^*)} (1-q)g(s^*) \frac{\partial s^*}{\partial R} \right] \end{aligned} \quad (\text{A-25})$$

Inspection of equation (A-25), and the results on the sign of the partial derivatives established in Proposition 1, reveals that the derivative is zero at $R = 0$, and that it is negative at $R > 0$. ■

A.3 Proof of Proposition 3.

Proof. Define $\zeta_{i,j}$ as the deviation of a senior executive j from the productivity of senior executives, X_o . Analogously, let $\xi_{i,j}$ be the deviation of a junior executive j from μ . Naturally the expected value of those deviations is zero. In a small sample of size n , the average productivity of junior and senior incumbent executives in firm i can be written as:

$$X_{o,i} = X_o + \frac{\sum_{j=1}^{n_{o,i}} \zeta_{i,j}}{n_{o,i}}, \quad X_{y,i} = \mu + \frac{\sum_{j=1}^{n-n_{o,i}} \xi_{i,j}}{n-n_{o,i}}$$

$$\text{Then, } \varepsilon_i \equiv \phi_i \left(\frac{\sum_{j=1}^{n_{o,i}} \zeta_{i,j}}{n_{o,i}} - \frac{\sum_{j=1}^{n-n_{o,i}} \xi_{i,j}}{n-n_{o,i}} \right) + \frac{\sum_{j=1}^{n-n_{o,i}} \xi_{i,j}}{n-n_{o,i}}.$$

We show that $cov(\phi_i, \varepsilon_i) = 0$, so that the OLS regression assumptions are satisfied. Let n be

the number of executives in each firm. For notational convenience let us define $z_i \equiv \frac{\sum_{j=1}^{n_{o,i}} \zeta_{i,j}}{n_{o,i}}$ and $u_i \equiv \frac{\sum_{j=1}^{n-n_{o,i}} \xi_{i,j}}{n-n_{o,i}}$, to write:

$$\text{cov}(\phi_i, \varepsilon_i) = \mathbb{E}[\phi_i^2(z_i - u_i) + \phi_i u_i] - \mathbb{E}(\phi_i) \mathbb{E}[\phi_i(z_i - u_i) + u_i]$$

The key of the proof is that the conditional expectation $\mathbb{E}(z_i | n_{o,i} = k) = 0$, for all $k = 0, 1, \dots, n$. To see this note that, for a given k :

$$\mathbb{E}\left[\frac{\sum_{j=1}^k \zeta_{i,j}}{k} \mid n_{o,i} = k\right] = \frac{1}{k} \sum_{j=1}^k \mathbb{E}(\zeta_{i,j} | x_{i,j} + r_{i,j} > s^*) = 0$$

This holds since $\mathbb{E}(\zeta_{i,j} | x_{i,j} + r_{i,j} > s^*) = 0$ for each j . Recall that $\zeta_{i,j}$ is the deviation of a senior executive productivity x_j from the senior executive unconditional productivity, X_o . It is immediate that conditioning on the information that an executive is tenured does not provide any information on how much above (or below) the average tenured executives' level (X_o) he is.

Recall that ϕ_i takes the values $(0, \frac{1}{n}, \dots, \frac{k}{n}, \dots, 1)$. As productivity realization are independent across executives, the probability of each $\phi_i = \frac{k}{n}$ outcome is $\Pr(\frac{k}{n}) \equiv p(k, n)$ from a binomial distribution. Then (for $a = 1, 2$)

$$\begin{aligned} \mathbb{E}_{\phi, z}(\phi_i^a z_i) &= \mathbb{E}_{\phi}[\mathbb{E}_z(\phi^a z_i) | \phi_i = \phi] = \sum_{k=0}^n p(k, n) \mathbb{E}_z\left[\left(\frac{k}{n}\right)^a \cdot z_i | n_{o,i} = k\right] \\ &= \sum_{k=0}^n p(k, n) \left(\frac{k}{n}\right)^a \mathbb{E}_z[z_i | n_{o,i} = k] = \sum_{k=0}^n p(k, n) \left(\frac{k}{n}\right)^a \cdot 0 = 0 \end{aligned}$$

The same logic shows that $\mathbb{E}_{\phi, u}(u_i \phi_i^a) = 0$ for $a = 1, 2$. This is immediate as the productivity of the junior is not observed by the principal, hence it cannot be correlated with his decisions about the tenure of the senior executives. ■

B Data and OLS regressions details

The INVIND survey is based on a questionnaire comprised of a fixed and monographic section that changes from year to year, used to investigate in-depth specific aspects of firms' activity. In 1992 a large section was devoted to corporate control. The determination of the nature of the controlling shareholders begins with that year. Among other things, the questionnaire asked about each firm's main shareholder, distinguishing between 10 different categories. Since 1992, the questions on control structure have been included every year. Starting in 1996, the categories have been reduced to five: 1) individual or family; 2) government (local or central or other publicly controlled entities); 3) conglomerate; 4) institution (financial or not); 5) foreign owner. We collapse the last two categories into one and map the previous classification into these four groups. Before 1992 the nature of the controlling shareholder was not investigated. However, in 1992 the firm was asked the year of the most recent change in control. We extend the control variable of 1992 back to the year of the most recent control change. Moreover, if a firm has a certain controller type in year t and the same in year t' , and some missing values in the year in between, we assume that the control has remained of the same type for all the period $[t, t']$. Note that there might be some cases of misclassification, in particular among firms that are classified as not controlled by an individual. For example, a foreign entity controlling a resident firm might in turn be controlled by a resident

that uses the offshore firm for taxation purposes. The same holds true for firms that report an institution as the controlling shareholder. This would bias the difference in the estimates between family and non family firms downward, because we would be classifying as foreign some family firms (the opposite case is not very likely). This implies that our results can be seen as a lower bound of the difference we find.

The CADS data are used to construct the capital stock using the permanent inventory method. Investment is at book value, adjusted using the appropriate two-digit deflators and depreciation rates, derived from National Accounts published by the National Institute for Statistics. For consistency with the capital data, in the estimation of the production function we take value added and labor from the CADS database. Both the INVIND and the CADS samples are unbalanced, so that not all firms are present in all years.

Data on workers are extensively described in [Iranzo, Schivardi, and Tosetti \(2008\)](#). We cleaned the data by eliminating the records with missing entries on either the firm or the worker identifier, those corresponding to workers younger than 25 (just 171 observations, 0.08% of the total) and those who had worked less than 4 weeks in a year. We also avoided duplication of workers within the same year; when a worker changed employer, we considered only the job at which he had worked the longest.

The main econometric problem in recovering TFP is that inputs are a choice variable and thus are likely to be correlated with unobservables, particularly the productivity shock. This is the classical problem of endogeneity in the estimation of production functions. To deal with it we follow the procedure proposed by [Olley and Pakes \(1996\)](#). Using a standard dynamic programming approach, Olley and Pakes show that the unobservable productivity shock can be approximated by a non-parametric function of the investment and the capital stock. To allow for sectoral heterogeneity in the production function, we estimate it separately at the sectoral level. The estimation procedure, the coefficients, and all the results are described in details in [Cingano and Schivardi \(2004\)](#).

To make sure that our results are not dependent on the TFP measure, we also perform some direct production function estimation exercises. To control for endogeneity we again follow Olley and Pakes and include in the regression a third degree polynomial series in i and k and their interactions, which approximate the unobserved productivity shock.²⁰ In [Table A-10](#) we report a series of exercises analogous to those of [Table 3](#) in the main text. The dependent variable is log value added, the regressors are capital and labor in addition to the share of senior executives interacted with the control dummies, and the control dummies themselves. All the regressions include year and sectoral dummies. As in [Table 3](#), we use sample weights with the exception of column [3]. In column [1] we do simple OLS; the Olley and Pakes controls are introduced in Column [2] and maintained throughout; in Column [3] we do not weight observations. Column [4] uses a 7-year period to become senior and column 5 a 3-year period. Column [6] introduces the additional controls, interacted with ownership dummies in column [7]. Results are similar across specifications; more importantly, they are very much in line with those of [Table 3](#).

²⁰Note that when the nonparametric term in capital and investment is included, the capital coefficient can no longer be interpreted as the parameter of the production function in the first stage of the procedure. However, given that the coefficient on capital is of no particular interest to us, this is inconsequential for our purposes.

Table A-9: Coefficients of the additional controls in Tables 3, 4 and A-10

	Dependent variable:					
	TFP	TFP	log VA	log VA	ROA	ROA
Firm age	-0.03*** (0.009)	-0.05*** (0.018)	-0.02** (0.009)	-0.03* (0.017)	-0.16 (0.222)	-0.33 (0.337)
Firm age·Fam		0.04** (0.023)		0.03 (0.021)		0.76* (0.448)
Firm age·For		0.03 (0.031)		0.00 (0.028)		-0.45 (0.750)
Firm age·Govt		0.02 (0.034)		-0.03 (0.028)		-0.12 (0.518)
Work age	-0.48*** (0.076)	-0.71*** (0.138)	-0.43*** (0.072)	-0.65*** (0.133)	-17.33*** (1.655)	-17.37*** (2.789)
Work age·Fam		0.23 (0.172)		0.16 (0.160)		-0.29 (3.411)
Work age·For		0.51** (0.241)		0.64*** (0.242)		1.12 (6.067)
Work age·Govt		-0.25 (0.526)		-0.43 (0.501)		-5.78 (8.890)
Share exec	1.61*** (0.507)	1.48 (1.027)	2.20*** (0.470)	2.41** (0.952)	-6.27 (11.678)	-24.65 (19.825)
Share exec·Fam		-0.50 (1.208)		-0.72 (1.109)		21.48 (25.353)
Share exec·For		0.82 (1.555)		0.07 (1.454)		5.80 (33.950)
Share exec·Govt		10.05*** (2.915)		7.90*** (2.660)		193.57*** (63.083)
Share WC	0.66*** (0.054)	0.79*** (0.099)	0.63*** (0.048)	0.71*** (0.096)	1.80 (1.348)	0.51 (1.797)
Share WC·Fam		-0.16 (0.114)		-0.10 (0.106)		2.31 (2.357)
Share WC·For		-0.19 (0.145)		-0.12 (0.129)		2.50 (3.494)
Share WC·Govt		-0.25 (0.167)		-0.11 (0.160)		-3.57 (3.476)
N. workers	0.02** (0.009)	0.01 (0.015)	0.74*** (0.013)	0.72*** (0.018)	0.17 (0.215)	0.37 (0.301)
N. workers·Fam		-0.02 (0.020)		0.01 (0.020)		-0.29 (0.431)
N. workers·For		0.04 (0.025)		0.04 (0.024)		0.09 (0.600)
N. workers·Govt		-0.02 (0.036)		-0.05 (0.033)		-1.29* (0.730)
Share male	0.07* (0.038)	0.05 (0.061)	0.17*** (0.038)	0.15** (0.058)	-0.97 (0.906)	-1.34 (1.417)
Share male·Fam		-0.01 (0.065)		0.03 (0.061)		1.38 (1.498)
Share male·For		0.05 (0.110)		-0.08 (0.099)		-2.27 (2.731)
Share male·Govt		0.51*** (0.184)		0.43*** (0.150)		-0.15 (3.160)

Note: the table reports the coefficients on the additional controls in the last two columns of Table 3, 4 and A-10. See the main text for the definition of the variables. Robust standard errors in parenthesis. Significance levels for the null hypothesis of a zero coefficient are labelled as follows: * is 10%, ** is 5%, *** is 1%.

Table A-10: Value added and share of senior executives, by control type

Dependent variable: log value added							
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
ϕ	0.131*** (0.039)	0.127*** (0.039)	0.056* (0.032)	0.102** (0.043)	0.109*** (0.039)	0.129*** (0.042)	0.135*** (0.042)
$\phi \cdot \text{Foreign}$	-0.088 (0.061)	-0.060 (0.060)	-0.030 (0.046)	0.003 (0.066)	-0.072 (0.061)	-0.010 (0.063)	-0.048 (0.065)
$\phi \cdot \text{Family}$	-0.204*** (0.046)	-0.186*** (0.045)	-0.115*** (0.038)	-0.179*** (0.050)	-0.137*** (0.046)	-0.124*** (0.047)	-0.135*** (0.048)
$\phi \cdot \text{Government}$	-0.471*** (0.084)	-0.442*** (0.081)	-0.304*** (0.072)	-0.370*** (0.098)	-0.347*** (0.085)	-0.309*** (0.091)	-0.246** (0.107)

Note: ϕ is the share of senior executives who have been with the firm at least 5 years in columns [1]-[3], at least 7 years in column [4] and at least 3 years in column [5]. All regressions are weighted with sampling weights with the exception of column [3], which is unweighted. Following [Olley and Pakes \(1996\)](#), regressions in columns 2-5 include a 3rd degree polynomial in capital and investment to control for the unobserved productivity shock. All regressions include control type dummies, year dummies, 2-digit sector dummies. Robust standard errors in parenthesis. Significance levels for the null hypothesis of a zero coefficient are labelled as follows: * is 10%, ** is 5%, *** is 1%.

C Derivation of the likelihood function

The likelihood for a sample of observations Y , under the parametrization Θ is

$$\begin{aligned}
L(\Theta; Y) &= \prod_{\kappa=1}^K \prod_{j=1}^2 \prod_{i=1}^{n_\kappa} \frac{1}{(2\pi\sigma_j^2)^{1/2}} \exp\left(-\frac{1}{2} \left[\frac{y_{i,\kappa}^j - f^j(\Theta, \kappa)}{\sigma_j}\right]^2\right) \\
&= \prod_{\kappa=1}^K \prod_{j=1}^2 (2\pi\sigma_j^2)^{-n_\kappa/2} \prod_{i=1}^{n_\kappa} \exp\left(-\frac{1}{2} \left[\frac{y_{i,\kappa}^j - f^j(\Theta, \kappa)}{\sigma_j}\right]^2\right)
\end{aligned}$$

where K is the number of “groups” in the model (4 control types in our case). This is

$$\log L(\Theta; Y) = -\frac{1}{2} \sum_{\kappa=1}^K \sum_{j=1}^2 n_\kappa \log(2\pi\sigma_j^2) - \frac{1}{2} \sum_{\kappa=1}^K \sum_{j=1}^2 \sum_{i=1}^{n_\kappa} \left[\frac{y_{i,\kappa}^j - f^j(\Theta, \kappa)}{\sigma_j}\right]^2 \quad (\text{A-26})$$

For all observable j and for each group κ (hence omitting the j, κ subindices)

$$\begin{aligned}
\sum_{i=1}^{n_\kappa} \left[\frac{y_i - f}{\sigma}\right]^2 &= \frac{n_\kappa}{\sigma^2} \left[\frac{\sum_{i=1}^{n_\kappa} y_i^2}{n_\kappa} + f^2 - 2\frac{\sum_{i=1}^{n_\kappa} y_i}{n_\kappa} f\right] \\
&= \frac{n_\kappa}{\sigma^2} \left[\sigma^2 + \left(\frac{\sum_{i=1}^{n_\kappa} y_i}{n_\kappa}\right)^2 + f^2 - 2\frac{\sum_{i=1}^{n_\kappa} y_i}{n_\kappa} f\right] \\
&= n_\kappa + \frac{n_\kappa}{\sigma^2} (f - \bar{y}_\kappa)^2 \quad \text{where } \bar{y}_\kappa \equiv \frac{\sum_{i=1}^{n_\kappa} y_i}{n_\kappa}
\end{aligned}$$

Replacing this expression in [equation \(A-26\)](#) we can rewrite the likelihood function by minimizing the distance between the theoretical value $f(\Theta, k)$ and the sample average \bar{y}_κ^j for each variable j , or

$$\log L(\Theta; Y) = -\frac{1}{2} \sum_{\kappa=1}^K \sum_{j=1}^2 n_\kappa \ln(2\pi\sigma_j^2) - \frac{1}{2} \sum_{\kappa=1}^K \sum_{j=1}^2 \left[n_\kappa^j + \frac{n_\kappa^j}{\sigma_j^2} (f^j(\Theta, \kappa) - \bar{y}_\kappa^j)^2 \right]. \quad (\text{A-27})$$

The measurement error for variable j (common for all group κ) is

$$\sigma_j^2 \equiv \text{var}(y^j) = \sum_{i=1}^{n_\kappa} \frac{1}{n_\kappa} (y_{i,\kappa}^j - \bar{y}_\kappa^j)^2$$

C.1 Score and Information matrix

Let M be the size of Θ . The n -th element of the score is given by

$$\begin{aligned} \mathbf{s}_n(\Theta; Y) &\equiv \frac{\partial \log L(\Theta; Y)}{\partial \theta_n} = -\frac{1}{2} \frac{\partial F(\Theta; Y)}{\partial \theta_n} \\ &= \sum_{\kappa=1}^K \sum_{j=1}^2 \left(\frac{n_\kappa}{\sigma_j^2} \right) (\bar{y}_\kappa^j - f^j(\Theta, \kappa)) \frac{\partial f^j(\Theta, \kappa)}{\partial \theta_n} \end{aligned}$$

The (n, m) element of the $M \times M$ information matrix $I(\Theta)$ is defined as:

$$I_{n,m}(\Theta) = \mathbb{E} \left[\frac{\partial \log L(\Theta; Y)}{\partial \theta_n} \frac{\partial \log L(\Theta; Y)}{\partial \theta_m} \right] = \mathbb{E}[\mathbf{s}_n(\Theta, Y) \mathbf{s}_m(\Theta, Y)]$$

which in our case becomes

$$\begin{aligned} I_{n,m}(\Theta) &= \mathbb{E} \left\{ \left[\sum_{\kappa=1}^K \sum_{j=1}^2 \left(\frac{n_\kappa}{\sigma_j^2} \right) (\bar{y}_\kappa^j - f^j(\Theta, \kappa)) \frac{\partial f^j(\Theta, \kappa)}{\partial \theta_n} \right] \right. \\ &\quad \left. \left[\sum_{\kappa'=1}^K \sum_{j'=1}^2 \left(\frac{n_{\kappa'}}{\sigma_{j'}^2} \right) (\bar{y}_{\kappa'}^{j'} - f^{j'}(\Theta, \kappa')) \frac{\partial f^{j'}(\Theta, \kappa')}{\partial \theta_m} \right] \right\} \\ &= \sum_{\kappa=1}^K \sum_{j=1}^2 \left(\frac{n_\kappa}{\sigma_j^2} \right) \sum_{\kappa'=1}^K \sum_{j'=1}^2 \left(\frac{n_{\kappa'}}{\sigma_{j'}^2} \right) \\ &\quad \mathbb{E} \left\{ (\bar{y}_\kappa^j - f^j(\Theta, \kappa)) (\bar{y}_{\kappa'}^{j'} - f^{j'}(\Theta, \kappa')) \right\} \frac{\partial f^j(\Theta, \kappa)}{\partial \theta_n} \frac{\partial f^{j'}(\Theta, \kappa')}{\partial \theta_m} \\ &= \sum_{\kappa=1}^K \sum_{j=1}^2 \left(\frac{n_\kappa}{\sigma_j^2} \right) \sum_{\kappa'=1}^K \sum_{j'=1}^2 \left(\frac{n_{\kappa'}}{\sigma_{j'}^2} \right) \\ &\quad \mathbb{E} \left\{ \left(\frac{1}{n_\kappa} \sum_{i=1}^{n_\kappa} \varepsilon_i^j \right) \left(\frac{1}{n_{\kappa'}} \sum_{i=1}^{n_{\kappa'}} \varepsilon_i^{j'} \right) \right\} \frac{\partial f^j(\Theta, \kappa)}{\partial \theta_n} \frac{\partial f^{j'}(\Theta, \kappa')}{\partial \theta_m} \\ &= \sum_{\kappa=1}^K \sum_{j=1}^2 \left(\frac{n_\kappa}{\sigma_j^2} \right) \left(\frac{n_\kappa}{\sigma_j^2} \right) \left\{ \frac{\sigma_j^2}{n_\kappa} \right\} \frac{\partial f^j(\Theta, \kappa)}{\partial \theta_n} \frac{\partial f^{j'}(\Theta, \kappa')}{\partial \theta_m} \\ &= \sum_{\kappa=1}^K \sum_{j=1}^2 \left(\frac{n_\kappa}{\sigma_j^2} \right) \frac{\partial f^j(\Theta, \kappa)}{\partial \theta_n} \frac{\partial f^{j'}(\Theta, \kappa')}{\partial \theta_m} \end{aligned}$$