# Corporate control and executive selection

Francesco Lippi
University of Sassari and EIEF

Fabiano Schivardi
LUISS and EIEF

In firms with concentrated ownership the controlling shareholder may pursue nonmonetary 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 data set 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 total factor productivity differentials between control types of up to 10%.

Keywords. Corporate governance, private returns, total factor productivity.

JEL classification. D2, G32, L2.

## 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 (La Porta, 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 nonmonetary private returns such as electoral goals in a firm controlled by politicians or family prestige in a firm

controlled by an individual.[1] 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 nonmonetary 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] If owners value personal ties, they might do so at the expenses of managerial ability, thus distorting the process of executive selection with respect to a situation where the owner only aims to maximize the value of the firm. We formalize this mechanism and propose two distinct sets of structural estimates to quantify its effects on firms' productivity.

We consider a simple partial-equilibrium, infinite horizon economy in which the firm owner chooses the firm's executives. We assume that average managerial ability determines the firm total factor productivity (TFP).[3] Executives are heterogeneous along two dimensions: their ability (productivity) and their relationship value. The owner only learns these values after an in-office trial period for the executive. Upon learning ability and relationship value, 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 private returns from the personal relationship. The key decision for the firm owner is whether to retain the executive in the firm after the trial period or to replace her with a junior one. Once tenured, executives in office die with an exogenous probability.

The model is formulated as an optimal sequential search problem in the spirit of McCall (1970) and Weitzman (1979). We solve it analytically and derive three key predictions that form the basis for our empirical investigation. First, the firm's productivity declines monotonically with the importance assigned to private benefits. This is because the owner's tenure decisions are based less on ability and more on personal ties as the importance of private benefits increases. Second, the owners who attach more importance to relationships 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. Finally, we show that the cross-firm correlation between productivity and the share of senior executives measures the strength of the selection effect on managerial ability, defined as the difference in the average ability of senior and junior executives.

---

[1]Consistent with the hypothesis that control commands a premium, Dyck and Zingales (2004) provided 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 showed that in the United States, average returns of privately held firms are dominated by the market portfolio. They concluded that owners of private firms must be obtaining some form of nonmonetary return.

[2]Becker was the first to stress the importance of nonpure consumption components of preferences for individual decision making, for example, in his classic analysis of discrimination (Becker (1971)). In the introduction to the book that collects his contributions on this topic, he stated that "Men and women want respect, recognition, prestige, acceptance, and power from their family, friends, peers, and others" (Becker (1998, p. 12)).

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

Intuitively, if selection is based on talent only, then a large fraction of senior executives signals that the firm found a talented pool of executives and is, therefore, highly productive. Instead, when the owner cares mostly about personal ties, a large share of senior executives is not very informative about the firm's productivity, because their ability played little role in the selection process. We prove that this correlation can be estimated by an ordinary least squares (OLS) regression of TFP on the share of senior executives.

The empirical analysis is 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 as those who have been with the firm for at least 5 years. The data classify the controlling shareholder into four broad control types: individual/family, the government, a conglomerate, or a foreign/institutional owner. We estimate the model by allowing the importance of private returns to vary across control types (conditional on additional covariates, such as time and sectoral effects).

Our first exercise is based on measuring the selection effect via an OLS regression of TFP on the share of senior executives. We find that selection is weaker in government and family firms, and that private benefits generate substantial losses in productivity. Note that these estimates are not based on differences in average productivity and seniority across types, as we include ownership dummies. Still, omitted variables and unobserved heterogeneity might give rise to the correlation we observe within type. For example, one might conjecture that older firms might both be 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 large series of robustness checks, such as adding more controls and using profitability—rather than productivity—as a performance indicator.

We then move on to a structural estimation of all the model parameters. We show that there is a one-to-one mapping between the average productivity and seniority measured in the data, and the model parameters that capture the importance of private benefits. We find that the executive selection is distorted with respect to the benchmark of no private benefits in all control types. The distortion is smallest for conglomerate and foreign-controlled firms. Private benefits account for a decrease in average firm productivity of around 6% in family firms and of 10% in government firms, as compared to conglomerate- and foreign-controlled firms. Compared to the theoretical benchmark of no private benefits, productivity losses are on the order of 7% for conglomerates, 8% for foreign firms, 13% for family firms, and 17% for government firms.

We finally 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 productivity and seniority for the OLS, and variation in average sample moments across control types for the structural estimates. Nothing in the estimation procedure imposes that the two sets of estimates should be consistent. The fact that they do produce quantitatively similar results offers further support for the mechanism we propose.

Two important caveats should be kept in mind. First, our empirical analysis is based on a set of joint predictions: on the degree of correlation between TFP and seniority *within* an ownership type, as well as on the correlation of average TFP and average seniority *across* ownership types. These patterns are explained in terms of differences in the degree of private benefits. Of course, firms with different ownership types differ along other aspects. Arguably, one can think of alternative mechanisms that can explain each specific pattern we find in the data. However, the advantage of a formal model is that it supplies an explicit and coherent framework that accounts for all these facts simultaneously and allows us to assess their consistency, as well as their impact on productivity via counterfactuals, in a quantitative way. Second, we cannot draw normative implications from our analysis, as private benefits do enter the owners' utility function. We only stress that they come at a cost in terms of productivity.

The paper is organized as follows. The next section discusses the connection with the literature, in particular with Albuquerque and Schroth (2010), Bandiera, Guiso, Prat, and Sadun (forthcoming), and Taylor (2010), who studied 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 hypotheses that exploits a model-based regression analysis. A subsection discusses several robustness checks and alternative hypotheses. 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, Lopez-de-Silanes, Shleifer, and Vishny (2000). Dyck and Zingales (2004) empirically estimated the value of private benefits of control using the difference between the price per share of a transaction that involved a controlling block and the price on the stock market before that transaction. They found large values of private benefits of control. In particular, at 37%, the value of private benefits in 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 occurs because the firm owner assigns value to the personal relationship with the executives. 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 (2009) studied the effects of social connections among managers and workers on performance. Using a field experiment, they showed that managers favor workers

who they are socially connected to, possibly at the expense of the firm's performance. Kramarz and Thesmar (2013) studied the effects of social networks on the composition of firms' boards of directors and performance in France. They found 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. Battistin, Graziano, and Parigi (2012) studied the effects of connections of top executives in Italian banks on bank performance and turnover. Consistent with our results, they concluded that connections reduce turnover and worsen performance, particularly in local banks.

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 (2011)). 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. (forthcoming), Michelacci and Schivardi (2013)).

In terms of managerial turnover, Volpin (2002) studied 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 the owner's decision to confirm or replace incumbent executives.

Three recent papers are closely related to ours. Albuquerque and Schroth (2010) used a structural model to estimate the private benefits of control in negotiated block transactions. They found evidence consistent with the hypothesis that those benefits are large. Taylor (2010) built and estimated a structural model of chief executive officer (CEO) turnover with learning about managerial ability and costly turnover. He found that only very high turnover costs can rationalize the low turnover rate observed in the data.[6] He interpreted this result in terms of CEO entrenchment and weak governance.[7]

---

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

[5]He also found 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.

[6]See also Taylor (2013) for a model of learning about CEOs' ability and wage dynamics.

[7]Garrett and Pavan (2012) studied 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 showed

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. (forthcoming) analyzed the role of incentive schemes in family and nonfamily firms, assuming that family firms pursue private benefits of control. They showed 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 their paper and ours as complementary. In fact, Bandiera et al. (forthcoming) focussed 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 $n$ 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, $K$ is the capital stock, and $L$ is 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 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 greatly simplifies the analysis in the model that we present next. We also assume that, conditional on $\bar{x}$, $K$ and $L$ are chosen optimally so that the profit function can be written only as a function of $\bar{x}$, and capital and labor can be ignored in what follows without loss of generality.

### 3.1 *The model*

The problem describes executive selection by the firm owner. 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 owner. An

---

that the optimal retention decision becomes more permissive over time, offering an alternative explanation for what would look like CEO entrenchment.

[8]Our 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.

executive's quality is characterized by two independent exogenous variables: his productivity $x$, a nonnegative random variable with continuous and differentiable cumulative distribution function (CDF) $G(x)$ with $G(0) = 0$ and expected value $\mu = \int_0^\infty x\, dG(x)$, and his relationship value $r$, a nonnegative random variable that is identically zero for a junior executive and equals zero with probability $1 - q$ or equals $R$ with probability $q$ for a senior executive. The personal relationship is valuable because it facilitates the delivery of nonmonetary payoffs, which 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 develop.[9]

We assume that on hiring a (junior) executive, the owner observes neither $x$ nor $r$, but only knows their distribution. At the end of the first period, the owner learns the value of the executive's productivity—a realization of $x$—and the value of his relationship—the realization of $r$ (either 0 or $R > 0$). It is assumed that both the executive relationship value and the 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 owner. After learning the realization of $x$ and $r$, the owner 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

$$F(s) = qG\big(\max(s - R, 0)\big) + (1 - q)G(s) \quad \forall s > 0, \tag{1}$$

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 $\rho$, so that the expected office tenure of a senior executive is $1/\rho$. When a senior executive dies, the owner replaces her with a junior one.

The per-period return for the risk-neutral owner 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 $\beta$ is a time discount. The owner 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 owner, and the expected value for the owner is

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

---

[9]Of 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) versus "unconnected" executives. Unfortunately, in the data, we have no way to detect this type of relationships, while we observe seniority. The two channels of personal ties are not mutually exclusive and might both be at play.

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

$$v_o(\tilde{s}) = \max\{v_y, \tilde{s} + \beta[\rho v_y + (1 - \rho)v_o(\tilde{s})]\}, \tag{2}$$

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

The optimal policy follows a threshold rule: the owner fires the senior executive if $s < s^*$, that is, 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 who will be fired, that is, all those for whom $s < \bar{s}$. 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

$$v_{o,s^*}(s^*) = v_{y,s^*}. \tag{3}$$

Appendix A shows how equation (3) can be used to obtain an analytical characterization of $s^*$ that is useful for the comparative statics analysis below. This characterization leads us to state our first proposition.

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

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

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

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

See Appendix B for the proofs of all propositions.

The proposition states that when $R \equiv 0$ (relationships bring no value to the owner), the optimal threshold $s^*(0) > \mu$. Hence the senior executive retains office only if he is sufficiently above the expected value of a junior $\mu$. This 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 $\mu$ (the mean of the ability distribution).

---

[10]As in the McCall (1970) model, the proof relies on the fact that the functions $v_o(\tilde{s})$ and $v_y$ cross only once.
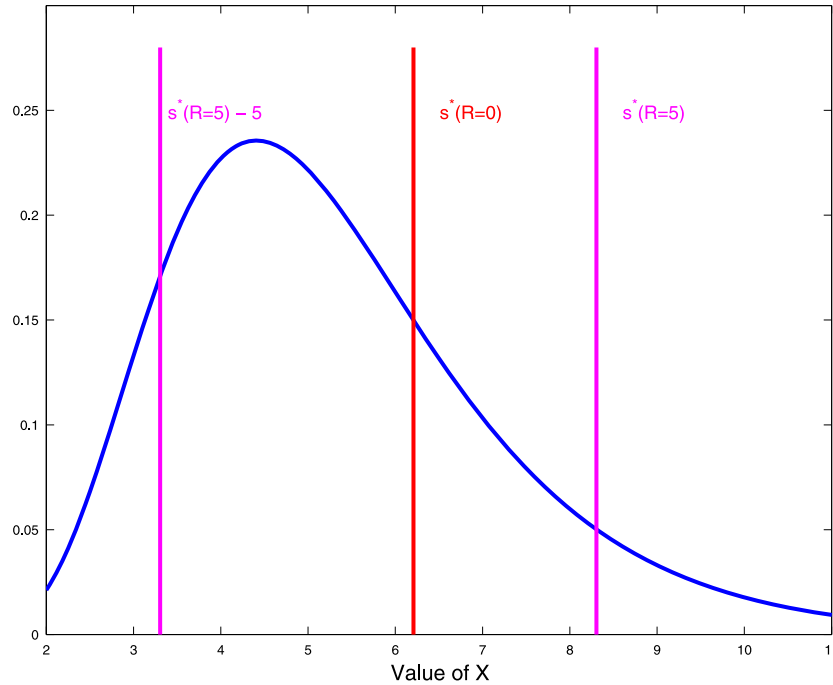
FIGURE 1. Example of selection thresholds for $R = 0$ and $R = 5$. *Note*: The figure uses the parameters $\beta = 0.98$ (per year), $\rho = 0.11$ (per year), and $q = 0.75$; $G(\cdot)$ is log normal with log mean $\lambda_m = 1.6$ and log std $\lambda_\sigma = 0.36$, which imply $\mu = 5.3$.

The second part of the proposition characterizes how the optimal threshold $s^*$ varies with $R$. The larger is the importance of the nonmonetary returns to the owner (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, respectively given by $s^* - R$ and $s^*$, move apart as $R$ increases.

We now turn to the model prediction concerning the seniority composition of the firm's executives in a steady state. The fraction of senior executives in office, $\phi$, follows the law of motion

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

so that the steady state fraction of senior executives is

$$\phi(s^*) = \frac{1}{1 + \dfrac{\rho}{1 - F(s^*)}} \in (0, 1), \tag{4}$$

which is decreasing in $\rho$. Mechanically, a lower hazard rate increases the fraction of senior executives. It is also immediate that $\phi$ is decreasing in $F(s^*)$. To study how $\phi$ 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

$$\frac{\mathrm{d}F(s^*)}{\mathrm{d}R} = \left[qg(s^* - R) + (1-q)g(s^*)\right]\frac{\partial s^*}{\partial R} - qg(s^* - R). \tag{5}$$

Recall 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 $\phi$ cannot be signed, in general, when $R > 0$, because 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 $\phi$. On the other hand, a higher $s^*$ raises the acceptance threshold for the senior executive with no relationship capital ($r = 0$); this reduces $\phi$. The final effect thus depends on the features of the distribution of $x$ and $r$. In fact, when $q$ is sufficiently small and $R$ is sufficiently large, the option value of searching for an executive who delivers $R$ is high enough to induce a low retention rate.[11] An example is displayed in the left panel of Figure 2.
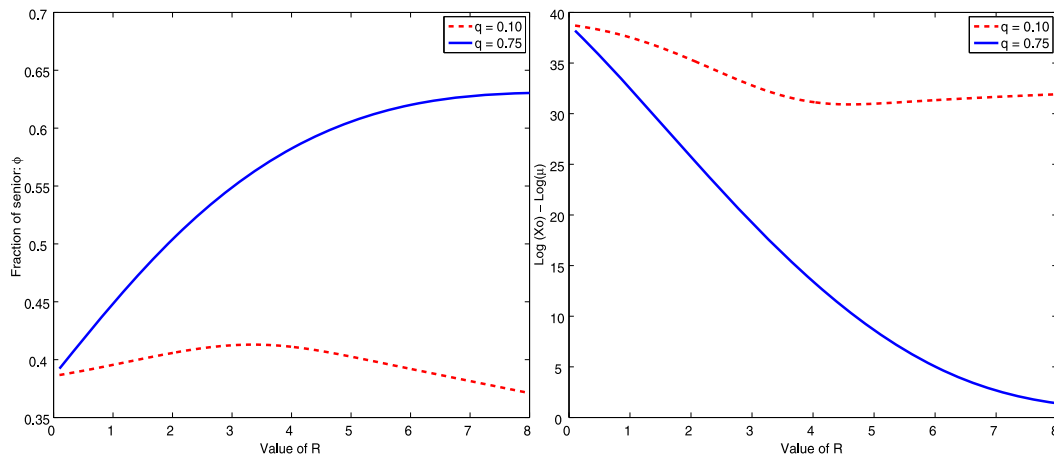


FIGURE 2. Share of senior executives and senior–junior differential as $R$ varies. *Note*: The figure uses the parameters $\beta = 0.98$ (per year), $\rho = 0.11$ (per year), and $q = 0.75$; $G(\cdot)$ is log normal with log mean $\lambda_m = 1.6$ and log std $\lambda_\sigma = 0.36$, which imply $\mu = 5.3$.

---

[11]It is immediate to show that a decrease in $q$ accompanied by a corresponding increase in $R$, to keep the expected value of $r$ fixed, leads to an increase in the variance of $r$. This, in turn, increases the option value of searching for an executive who delivers the relationship value.

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,

$$X(s^*) = \mathbb{E}_{r,x}(x) = \mu + \phi(s^*)\big[X_o(s^*) - \mu\big], \tag{6}$$

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

$$X_o(s^*) = \frac{q \displaystyle\int_{s^*-R}^{\infty} x\,dG(x) + (1-q)\displaystyle\int_{s^*}^{\infty} x\,dG(x)}{1 - F(s^*)}. \tag{7}$$

This leads us to the next proposition.

PROPOSITION 2. *Let $\beta \to 1$. 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 $R$, $\frac{\partial X}{\partial R}\big|_{R>0} < 0$.*

Proposition 2 shows that the mean productivity of a firm is maximized when the firm only cares about ability, that is, 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 $R$.[12] This will be useful in the discussion of the parameters identification below. The assumption that $\beta \to 1$ simplifies the derivation and is useful to interpret the mean $X$ as a cross section average.[13] 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 (7):

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

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

---

[12]One might argue that in a competitive equilibrium, firms whose owners have a large $R$ should not be able to survive, as they are less efficient than small $R$ firms. This is not necessarily the case. An owner who enjoys a private benefit might be willing to accept a return on assets below the market return, as she is trading off monetary for nonmonetary returns. Of course, she would still need to meet the no-bankruptcy constraint, which is, however, less stringent than matching the market return. This implies that in equilibrium, firms with different $R$'s—and, therefore, different productivity levels—can coexist. As discussed in the literature review, there is evidence that the return on privately held firms is dominated by the market portfolio (Moskowitz and Vissing-Jørgensen (2002)).

[13]The numerical analysis of the model for $\beta \in (0.85, 1)$ (per year) gives very similar results. The relationship between $X$ and $R$ is always decreasing.

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 $\phi$ 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 $R$ no longer influence the selection process. In other words, $X$ and $\phi$ asymptote to constant values as $R$ grows large. We will come back to this observation when we comment on the results of our estimates.

We conclude this section with a brief discussion of our modeling assumptions about the distribution of abilities. For the sake of simplicity, the model introduced an asymmetry between ability and the relationship value, assuming that only the latter grows with seniority. We notice that this is not a necessary assumption for Proposition 2(ii) to hold; what is necessary for this result, even if we allowed ability to evolve with seniority, is that such evolution is uncorrelated with $R$. Intuitively even if executives become more productive when turning senior (on average), it would still be the case that owners who are more interested in $R$ will select more based on private benefits and less on productivity. Of course, this result would break down if owners who care more about private benefits are also attached to executives whose productivity increases relatively faster with seniority than owners who are not interested in private benefits. This would be the case if such owners, for example, invest more in executive training. Note, however, that in such cases, the relationship between $R$ and productivity could even become positive, an implication we will be able to test (and dismiss) empirically.

## 4. Data description

In this section, we describe the main features of our data, referring to Appendix D for more details. The data match a large sample of executives with a sample of Italian firms. The executives represent approximately 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 1200 firms per year that is 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 data set 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. All these manufacturing firms belong to the private sector and are, therefore, subject to the same legislation, particularly in terms of labor laws.

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

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

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,

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

where $n_{i,t}$ is the number of executives in firm $i$ at $t$, $X_j$ is the ability of executive $j = 1, 2, \ldots, n_{i,t}$, and $\epsilon_{i,t}$ is an independent and identically distributed (i.i.d.) 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 D; full details are given 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 D for the details): (i) individual or family; (ii) government (local or central or other government controlled entities); (iii) conglomerate, that is, firms belonging to an industrial conglomerate; (iv) 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 characterize 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.[14] 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 are 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 Organization for Economic Cooperation and Development (OECD) (2003) system, and 75% 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.

---

[14]We 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.

|  | VA | Empl. | No. Exec. | % Exec. | TFP | % High Tech | % North | No. Obs. | No. Firms |
|---|---|---|---|---|---|---|---|---|---|
| All firms |  |  |  |  |  |  |  |  |  |
| Mean | 30.0 | 692 | 13.3 | 0.026 | 2.41 | 0.41 | 0.74 | 7773 | 802 |
| S.D. | 127.3 | 3299 | 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 | 2906 | 349 |
| 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 | 2390 | 300 |
| S.D. | 214.2 | 5637 | 27.5 | 0.025 | 0.54 | 0.49 | 0.39 |  |  |
| Government |  |  |  |  |  |  |  |  |  |
| Mean | 40.6 | 1013 | 21.8 | 0.022 | 2.38 | 0.47 | 0.51 | 687 | 99 |
| S.D. | 87.4 | 2076 | 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 | 1790 | 274 |
| S.D. | 69.1 | 1563 | 32.9 | 0.022 | 0.48 | 0.50 | 0.44 |  |  |

*Note*: The notation VA is value added (in millions of 1995 euros), No. 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, No. Obs. is the number of firm-year observations, and No. Firms is the number of unique firms observations.

The executives' data are taken from the Istituto Nazionale della Previdenza Sociale (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 after 5 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 1236 euros, and the share of executives who have been with the firm at least 5 years is 0.57 and at least 7 years is 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-controlled firms pay their executives more, while their executives' characteristics resemble the average in terms of the tenure, age, and gender composition.

## 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

TABLE 2. Descriptive statistics: executives' characteristics by control type.

|  | Wage | $\phi_5$ | $\phi_7$ | Age | % Male |
|---|---|---|---|---|---|
| **All Firms** | | | | | |
| Mean | 1236 | 0.57 | 0.45 | 46.5 | 0.96 |
| S.D. | 330 | 0.30 | 0.31 | 4.6 | 0.13 |
| **Family** | | | | | |
| Mean | 1130 | 0.62 | 0.51 | 46.0 | 0.94 |
| S.D. | 288 | 0.33 | 0.33 | 5.3 | 0.17 |
| **Conglomerate** | | | | | |
| Mean | 1294 | 0.53 | 0.41 | 46.6 | 0.97 |
| S.D. | 321 | 0.28 | 0.28 | 4.0 | 0.09 |
| **Government** | | | | | |
| Mean | 1298 | 0.54 | 0.42 | 47.7 | 0.99 |
| S.D. | 349 | 0.27 | 0.27 | 4.6 | 0.03 |
| **Foreign** | | | | | |
| Mean | 1309 | 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; $\phi_5$ is the share of executives with at least 5 years of seniority, and $\phi_7$ is the share with at least 7 years. Age is the average executives' age and % Male is the share of male executives.

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 $X$, $\phi$ 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 consistency 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 $\phi_i$, 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. If $n_i$ is not large, the average productivity will differ from $X_o$, $X_y$ due to sampling variability. Using equation (6), we establish the following proposition.

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

$$X_i = \mu + (X_o - \mu)\phi_i + \varepsilon_i, \tag{8}$$

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

A key result from this proposition is that deviations $\varepsilon_i$ about the (large sample or unconditional) mean values are uncorrelated with the share of senior executives $\phi_i$. Intuitively, this property holds since an increase (or a decrease) in the quota of senior executives $\phi_i$ about its unconditional mean $\phi$ does not contain any information on the innovation $\varepsilon_i$, that is, 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 $\phi_i$. 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 $\phi$ is low. If $R$ is low, the selection mechanism is effective, senior executives are, on average, more productive than junior ones, and differences in $\phi$ will go together with substantial differences in $X$, yielding a high correlation between $X$ and $\phi$.

Another prediction of the model concerns a comparison across ownership types. Consider the firm productivity $X$ and the fraction of senior executives $\phi$. For a given vector of model primitives $\beta$, $\rho$, $G(\cdot)$, Figure 3 shows that for each admissible (i.e., model generated) observable pair $X$, $\phi$, there is at most one pair of parameter values $R$, $q$ that can produce it. 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 $\phi$. Notice that all lines depart from the same point in the $X$, $\phi$ 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
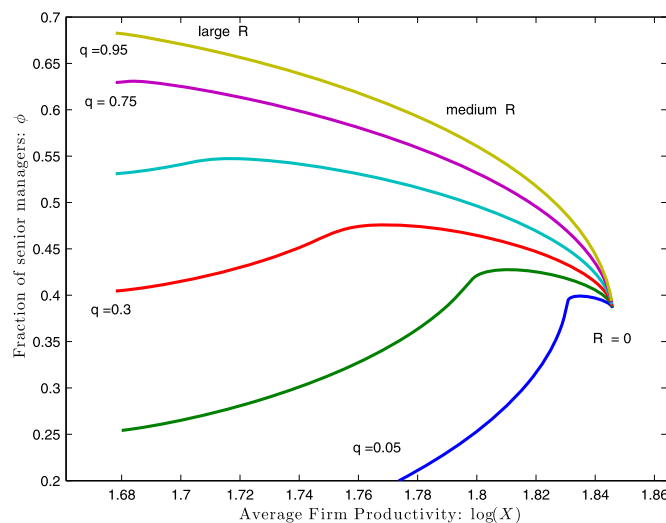


FIGURE 3. Productivity and seniority: space spanned by the model. *Note*: The figure uses the parameters $\beta = 0.98$ (per year) and $\rho = 0.11$ (per year); $G(\cdot)$ is log normal with log mean $\lambda_m = 1.6$ and log std $\lambda_\sigma = 0.36$, which imply that $\log(\mu) = 1.67$.

this because, as shown in Proposition 2, $X$ is decreasing in $R$. Moreover, as discussed above, the effect of an increase in $R$ on $\phi$ 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 $X$, $\phi$ 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 $\phi$ across firms for a given $R$, $q$. The latter estimate does not depend on the average $X$, $\phi$ for a given control type, but on how $X$ and $\phi$ covary 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 $\phi$ across control types unrelated 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) $\beta$, $q$, $\rho$, and $G(\cdot)$. They therefore offer a test of robustness of the structural estimates with respect to the values of the auxiliary parameters they hinge on. 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. After presenting the baseline estimates, we analyze their robustness and discuss alternative hypotheses for interpreting the data.

### 6.1 *Basic framework and results*

Equation (8) 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 expansion around $X_i = \mu$, we obtain

$$\log X_i = \gamma_0 + \gamma_1 \phi_i + \eta_i, \tag{9}$$

with $\gamma_0 = \log \mu$, $\gamma_1 = \frac{X_o - \mu}{\mu}$, and $\eta_i = \frac{1}{\mu} \varepsilon_i$, and where, as shown in Proposition 3, $\varepsilon_i$ is uncorrelated with $\phi_i$. This equation shows that in a regression of log TFP on the share of senior executives, the coefficient $\gamma_1$ measures the percentage difference in average

ability between senior and junior executives. We bring equation (9) to the data using the specification

$$\log \mathrm{TFP}_{i,t} = \gamma_0 + \gamma_1 \phi_{i,t} + \gamma_{\mathrm{fam}} D_{\mathrm{fam}} \cdot \phi_{i,t} + \gamma_{\mathrm{gov}} D_{\mathrm{gov}} \cdot \phi_{i,t} + \gamma_{\mathrm{for}} D_{\mathrm{for}} \cdot \phi_{i,t}$$
$$+ \gamma_5 \mathbf{W}_{i,t} + \eta_{i,t},$$

where the dummies $D_k$ with $k = (\mathrm{fam}, \mathrm{gov}, \mathrm{for})$ are control status dummies equal to 1 for family-, government-, and foreign-controlled firms, respectively, and $\mathbf{W}_{i,t}$ is a vector of controls that includes year dummies, two-digit sector dummies, and control-status dummies that account for potential unobserved heterogeneity across firms with different control types. The coefficient $\gamma_1$ 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 $\gamma_k$ measure the difference in the average ability of senior executives for the corresponding control type with respect to the conglomerate firms: $\gamma_k = X_o(R_k) - X_o(R_{\mathrm{cong}})$. 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 0.04. The positive sign of this coefficient is consistent with the selection hypothesis, that is, with the assumption that senior executives are, on average, more productive than junior ones. To give a sense of the size

TABLE 3. TFP and share of senior executive relationship by control type.

| | Dependent Variable: log TFP | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $\phi$ | 0.11*** | 0.02 | 0.08* | 0.09** | 0.11** | 0.12*** |
| | (0.04) | (0.03) | (0.04) | (0.04) | (0.04) | (0.04) |
| $\phi \cdot$ foreign | −0.03 | 0.01 | 0.02 | −0.05 | 0.02 | −0.02 |
| | (0.06) | (0.05) | (0.07) | (0.06) | (0.07) | (0.07) |
| $\phi \cdot$ family | −0.17*** | −0.10** | −0.17*** | −0.13** | −0.11** | −0.13** |
| | (0.05) | (0.04) | (0.05) | (0.05) | (0.05) | (0.05) |
| $\phi \cdot$ government | −0.47*** | −0.32*** | −0.42*** | −0.37*** | −0.34*** | −0.29** |
| | (0.09) | (0.07) | (0.11) | (0.09) | (0.11) | (0.12) |
| Observations | 5875 | 5943 | 4897 | 6840 | 5136 | 5136 |
| $R$-squared | 0.47 | 0.49 | 0.45 | 0.47 | 0.53 | 0.54 |

*Note*: The variable $\phi$ is the share of senior executives who have been with the firm at least 5 years in columns (1), (2), (5), and (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, and two-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), and the share of executives, of white collar workers, 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 are given in parentheses. Significance levels for the null hypothesis of a zero coefficient are labelled with asterisks: * is 10%, ** is 5%, and *** is 1%.

of the effect, the productivity of a firm with a share of executives that is 1 standard deviation above the mean (see Table 2) is higher by about 3%. The estimated coefficient for the foreign-controlled firms, $\gamma_{\text{for}}$ is not statistically different from zero; hence, we cannot reject the hypothesis that the selection in foreign-controlled firms is similar to that in conglomerate-controlled firms at conventional levels of significance. Instead, selection appears to be 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, that is, 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 that 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.

As a first robustness check, in column (2) we do not weight observations. In this case, the results are somehow weaker: the coefficient on $\phi$ 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 the length of time assumed necessary to become senior, which is 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), we use a 3-year definition. We find no substantial differences with respect to the basic specification.

### 6.2 *Additional robustness checks and alternative hypotheses*

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

$$\log X_i = \gamma_0 + \gamma_1 \phi_i + \gamma_2 Z_i + \eta_i.$$

Then the estimated coefficient on seniority using (9) 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 ($\phi$) 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 $\phi$ is high in firms under conglomerate control and is small

in family firms requires us to assume much more than an omitted variable, namely that either $\gamma_2$ or $\mathrm{cov}(\phi_i, Z_i)/\mathrm{var}(\phi_i)$ 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, for example, 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 data base 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, and the share of executives, of white collar workers, and of male workers as a fraction of the total workforce. To save on space, we report the coefficients on these additional controls in Table 9 in Appendix D. 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 it is 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 insignificantly different from zero (Table 9 in Appendix D), 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 $\phi$ 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 $\phi$ that are 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.[15] In other words, the interpretation of our estimate is the correct 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.

---

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

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 D. They are aligned with those of Table 3. We have also used profit-based measures of performance, as profits might capture different dimensions of managerial ability.[16] The results using return on assets (ROA) are reported in Table 4. 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 that emerge when the productivity measure is used; if anything, they are stronger. In particular, profitability is positively related to the share of senior executives in conglomerate-controlled firms.[17] 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

TABLE 4. ROA and share of senior executives by control type.

| | Dependent Variable: ROA | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $\phi$ | 3.40*** | 2.93*** | 2.30*** | 3.22*** | 3.95*** | 4.01*** |
| | (0.75) | (0.66) | (0.85) | (0.74) | (0.86) | (0.86) |
| $\phi \cdot$ foreign | −1.57 | 0.24 | −0.92 | −0.71 | 0.32 | 0.34 |
| | (1.38) | (1.07) | (1.35) | (1.36) | (1.60) | (1.65) |
| $\phi \cdot$ family | −4.48*** | −3.69*** | −3.32*** | −4.25*** | −3.88*** | −4.05*** |
| | (0.95) | (0.78) | (1.07) | (0.96) | (1.05) | (1.05) |
| $\phi \cdot$ government | −7.50*** | −3.73** | −4.71** | −7.22*** | −5.82*** | −5.95** |
| | (1.61) | (1.50) | (1.84) | (1.67) | (2.11) | (2.32) |
| Observations | 5875 | 5943 | 4897 | 6840 | 5136 | 5136 |
| $R$-squared | 0.08 | 0.09 | 0.08 | 0.08 | 0.12 | 0.13 |

*Note*: ROA is return on assets, in percentage units; $\phi$ is the share of senior executives who have been with the firm at least 5 years in columns (1), (2), (5), and (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, and two-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), and the share of executives, of white collar workers, 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 given in parentheses. Significance levels for the null hypothesis of a zero coefficient are labelled by asterisks: * is 10%, ** is 5%, and *** is 1%.

[16]Foster, Haltiwanger, and Syverson (2008) showed that efficiency and profitability might not be simply one-to-one. Sraer and Thesmar (2007) showed that family firms in France pay lower wages in exchange for more job security. Owners of family firms might, therefore, attract lower quality, higher risk-averse executives and pay them less. In principle, family firms might, therefore, be both less efficient and more profitable.

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

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 personal relationships on firms' productivity through counterfactual exercises.

## 7. Estimates of the structural parameters

This section presents the structural estimation of the model parameters. We begin by discussing the assumptions needed for identification using firm-level observations on TFP and seniority of the executives, which we see as empirical measures of $X_i$ and $\phi_i$, that is, the observed productivity and share of senior executives in firm $i$. A third observable used for the estimation is the variance of the regression residual in equation (9), which we label $\text{var}(\eta_i)$.

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). The distribution of productivity $G(x)$ is assumed to be log normal, so that the model is characterized by six fundamental parameters: the discount factor $\beta$, the hazard rate $\rho$, the log normal parameters $\lambda_m$ and $\lambda_\sigma$ (log mean and log standard deviation, respectively), the probability of developing a relationship, $q$, and the value of the relationship $R$. Five parameters, namely $\beta$, $\rho$, $q$, $\lambda_\sigma$, and $\lambda_m$, are assumed to be common to all firms. The parameter $R$ is assumed to vary with one observable characteristic of the firm: the control type. Given a parametrization (i.e., the vector $\beta$, $\rho$, $q$, $\lambda_\sigma$, $\lambda_m$, $R$), the model uniquely determines the values of $X$, $\phi$, and $\text{var}(\eta)$ to be observed in the data. In the estimation procedure, the differences between data points 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 estimates concern seven parameters: $\theta_p \in \Theta_{7,1}$, $p = 1, 2, \ldots, 7$, described next. The firm-level observations vary across 14 years (index $t$), 13 two-digit sectors (index $\tau$), and four control types described in the previous section (index $\kappa$). We assume that $R$ varies across firms according to the nature of the controlling stakeholder, with $R_\kappa = \theta_\kappa$, $\kappa = 1, 2, 3, 4$, for the firms with control type family ($\theta_1$), conglomerate ($\theta_2$), government ($\theta_3$), and foreign ($\theta_4$), respectively. The parameter $\lambda_m = \theta_5$ measures the mean (log) TFP of the firm. In the data, TFP has a clear time component as well as a sectoral one, which are ignored by the simple structure of our model. Thus, before turning to the structural estimation, we normalize the TFP data by removing com-

mon time and sector effects. Our measure of $X$ for firm $i$ in year $t$ and sector $\tau$ is thus given by

$$\log X_{i,t,\tau} \equiv \log \mathrm{TFP}_{i,t,\tau} - \mathbf{a}_1 \cdot \mathbf{year}_{i,t} - \mathbf{a}_2 \cdot \mathbf{sect}_{i,\tau} - a_3 \cdot Z_{i,t,\tau}, \tag{10}$$

where $\mathbf{a}_1$ and $\mathbf{a}_2$ are the vectors of coefficients from an OLS regression of TFP on 13 year and 12 two-digit sector dummies. We also consider a specification that controls for the effect of firm size $Z_{i,t,\tau}$ (log employment) on TFP. The parameter $\theta_6$ measures the probability of developing a relationship $q = \frac{\theta_6}{1+\theta_6}$. The parameter $\theta_7$ measures the standard deviation of the ability distribution (see Appendix C for a discussion of the mapping between this parameter and the observable regression squared error $\mathrm{var}(\eta_i)$).

To reduce the computational burden, two parameters are pinned down outside the estimation routine. The time discount $\beta$ is calibrated to an annual value of 0.98, as is standard in the literature. The hazard rate of senior executives, $\rho$, is computed from the survival function of senior executives, that is, with at least 5 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.

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 $\phi$. For each firm $i$ in group $\kappa$, there are two observables $y_{i,\kappa}^j$, $j = 1, 2$. We assume that the variable $y_{i,\kappa}^j$ is measured with error $e_i^j$ 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, $\phi$.[18] Hence the maximum likelihood (ML) estimates use the observables $y_{i,\kappa}^1 = \log X_{i,\kappa}$ and $y_{i,\kappa}^2 = \phi_{i,\kappa}$. The measurement error variance $\sigma_j^2$, $j = 1, 2$, is assumed common across groups, and is computed as the variance of the residuals of an OLS regression of $\log X$ and $\phi$ on year and sector dummies. This gives $\sigma_{\log X}^2 = 0.35$ and $\sigma_\phi^2 = 0.29$. Finally, for each control group $\kappa$, we use the variance of the regression residual in equation (9) as the third observable, $y_\kappa^3 = \mathrm{var}(\eta)$, to be fitted by the model for each group.

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

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

Let $Y$ be the vector of observations and let $n_\kappa$ be the number of firms $i$ in group $\kappa$. Define the objective function $F$ as

$$F(\Theta; Y) \equiv \sum_{\kappa=1}^{4} \sum_{j=1}^{3} \left( \frac{n_\kappa}{\sigma_j^2} \right) \left( \frac{1}{n_\kappa} \sum_{i=1}^{n_\kappa} y_{i,\kappa}^j - f^j(\Theta, \kappa) \right)^2. \tag{11}$$

---

[18]This statement is based on an analysis of the deviations of $X$ and $\phi$ (in levels and in logs) from the mean of each groups. Details are available from the authors on request.

TABLE 5. Structural estimates of model parameters.

|                   | $q$  | $R_{\text{fam}}$ | $R_{\text{cong}}$ | $R_{\text{gov}}$ | $R_{\text{for}}$ | $\lambda_m$ | $\lambda_\sigma$ |
|-------------------|------|---------|---------|--------|---------|--------|--------|
| A. Baseline       | 0.78 | 5.5     | 3.3     | 6.0    | 3.5     | 1.59   | 0.39   |
|                   | (2.5)| (8.6)   | (10.1)  | (6.3)  | (10.3)  | (76.9) | (30.9) |
| B. With firm size | 0.76 | 5.1     | 3.4     | 6.9    | 3.5     | 1.59   | 0.40   |
|                   | (2.5)| (8.8)   | (10.2)  | (4.3)  | (9.9)   | (74.4) | (30.5) |

*Note*: The $t$-statistics are given in parentheses. The estimation uses the estimated measurement errors $\sigma_{\log X}^2 = 0.35$ and $\sigma_\phi^2 = 0.29$. The parameters $\beta$ and $\rho$ are calibrated (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 (10) (see the main text); in specification B, firm-size effects (as measured by the (log) number of employees) are also controlled for.

Appendix E shows that the likelihood function is related to the objective function by

$$\log L(\Theta; Y) = -\frac{1}{2} \sum_{\kappa=1}^{4} \sum_{j=1}^{3} n_\kappa \big(1 + \log\big(2\pi\sigma_j^2\big)\big) - \frac{1}{2} F(\Theta; Y).$$

We estimate the structural parameters in $\Theta$ by minimizing (11). At each iteration, the algorithm solves the model for each of the four groups and computes the objective function under the candidate parametrization. Since each group has three observables, there is a total of 12 moments to be fitted using 7 parameters; hence, the model is overidentified with 5 degrees of freedom. The formulas for the score and the information matrix used for the inference are derived in Appendix E.

The estimates of the model structural parameters are reported in Table 5. The estimated value of $q = 0.78$ indicates that approximately 75% of the executives develop a relationship. The value of $R$ varies substantially across control types, but all types enjoy some degree of private benefits: $R$ is always significantly different from zero. This is consistent with the findings of Taylor (2010), according to which CEO entrenchment is substantial even in U.S. listed firms, where family and government firms play a minor role. The importance of the relationships is lowest for firms that belong to a conglomerate (3.3), followed by foreign-controlled firms (3.5), family (5.5), and government (6.0). Given that the estimated unconditional mean level of TFP is around 7, the estimated values for $R$ show that the nonmonetary characteristics of the executives (i.e., their relationship value) are quantitatively important in the selection process.

To help with the interpretation of the structural estimates, Table 6 computes selected statistics produced by the model in the steady state at the estimated parameters. The first column solves the model for the (counterfactual) case in which the firm's owner gives no value to relationships in executive selection ($R = 0$) and, hence, $s^* = x^*$ for all senior executives. In this case, senior executives are confirmed if their ability $x$ is above $x^* = 6.2$, which occurs in around 25% of cases. The (log) average ability of senior executives is 2.09, almost 30% higher than the unconditional average ability of junior executives. On average, around 38% of executives are senior in this "frictionless" case. These figures can be compared to those obtained for family firms (second column of the table), for which $s^* = 8.6$ and $x^* = s^* - R = 3.1$ (the latter is the cutoff ability of executives

TABLE 6. Model predictions under the baseline estimates.

|  | Control Type | | | | |
|---|---|---|---|---|---|
|  | $R = 0$ | Family | Conglomerate | Government | Foreign |
| $R$ | 0 | 5.5 | 3.3 | 6.0 | 3.5 |
| $s^*$ | 6.2 | 8.6 | 7.4 | 8.9 | 7.5 |
| $x^*$ | 6.2 | 3.1 | 4.2 | 2.9 | 4.1 |
| $\log \mu$ | 1.67 | 1.67 | 1.67 | 1.67 | 1.67 |
| $\log X$ | 1.85 | 1.72 | 1.79 | 1.71 | 1.78 |
| $\log X_o$ | 2.09 | 1.75 | 1.87 | 1.73 | 1.86 |
| $\log X_o \vert R = 0$ | 2.09 | 2.34 | 2.22 | 2.37 | 2.23 |
| $\log X_o \vert R > 0$ | 2.09 | 1.73 | 1.85 | 1.72 | 1.83 |
| $\phi$ | 0.38 | 0.61 | 0.55 | 0.62 | 0.56 |
| Fired | 0.73 | 0.30 | 0.45 | 0.27 | 0.43 |
| Fired$\vert R = 0$ | 0.72 | 0.92 | 0.86 | 0.93 | 0.86 |
| Fired$\vert R > 0$ | 0.72 | 0.11 | 0.34 | 0.08 | 0.31 |

*Note*: 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, $X_o$ is the average managerial ability of the senior executives, $\log X_o \vert R = 0$ is the average ability of the senior who did not develop a relationship, and $\log X_o \vert R > 0$ is the average ability of executives who developed a relationship. Fired is the probability that a junior executive is replaced when turning senior.

who have developed a relationship with the owner). It is apparent that selection is much weaker in family-controlled firms: the senior executives' average ability ($X_o$) is only 8% higher than that of junior executives ($\mu$). An executive who develops a relationship has a 92% chance of being tenured. This probability drops to 11% 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. The productivity of the senior executives in conglomerate firms is about 20% higher than the productivity of junior executives, which indicates a much greater selection than in family firms. For government firms, the estimates imply that selection occurs almost exclusively on the basis of developing a relationship: 93% of executives with $r = R$ are retained; 92% of those with $r = 0$ are fired. As a consequence, the average ability of senior executives is only slightly higher than that of junior executives as the selection effect is inhibited.

The structural estimates allow us to develop a simple counterfactual exercise: setting the private benefits to zero implies a productivity gain of 6% in conglomerates, 7% in foreign firms, 13% in family firms, and 14% in government firms. If we take conglomerate-owned firms, which record the lowest value private benefits, as a benchmark, productivity is around 6% lower in family firms and 10% lower in government firms due to weaker selection. In light of the model and consistent with much anecdotal evidence, this happens because family- and government-controlled firms select executives by putting a large emphasis on nonmonetary values, independent of executive ability, thus reducing the productivity-enhancing effect of managerial selection.

An important robustness check consists of 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 OLS results (taken from the first

TABLE 7. Comparison between the structural and the OLS estimates.

|  | $\log X_{o,\text{cong}} - \log \mu$ | $\Delta X_{o,\text{for}}$ | $\Delta X_{o,\text{fam}}$ | $\Delta X_{o,\text{gov}}$ |
|---|---|---|---|---|
| OLS estimates | 0.11 | −0.03 | −0.17 | −0.47 |
| Structural estimates | 0.20 | −0.01 | −0.12 | −0.14 |

*Note*: The first row reports the OLS regression estimates from column (1) of Table 3. The first column represents the difference in senior–junior ability in conglomerate firms. The second, the third, and the fourth columns report the difference in the senior executives' average ability for each control type with respect to conglomerates: $\Delta X_{o,i} \equiv \log X_{o,i} - \log X_{o,\text{cong}}$, where $i = (\text{for}, \text{fam}, \text{gov})$.

column of Table 3) and compare them with the corresponding values implied by the estimates of Table 6. First, both estimation methods imply a positive effect of selection on the senior–junior productivity differential for conglomerate firms: 0.11 in the OLS and 0.20 in the structural estimates. The other columns report the difference in the ability of the senior managers for the relevant control type as compared to conglomerate firms, the most efficient type. For foreign firms, the difference is −0.01, a value that is close to that produced by the OLS estimates (−0.03). For family firms, both estimates show that selection is substantially weaker than in conglomerate-controlled firms. Although the OLS estimate is larger in absolute value than the structural estimate (−0.17 versus −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, that is, that junior executives are more efficient than senior ones. This is a result that the structural model cannot deliver; in fact, the structural estimates are close to the lowest possible selection effect, with −0.14, so that the average ability of senior and junior managers is almost the same. All in all, it is remarkable that the two sets of estimates give comparable results, although they are based on totally different dimensions of data variability.

We have further explored the robustness of our results. We saw in Table 1 that average firm size differs across control types. We reestimated equation (10) with the inclusion of firm size among the determinants of TFP and 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 firm size. All the quantitative predictions are virtually identical to those obtained under the baseline specification. A second robustness check relates to the estimate of the standard deviation of the ability distribution. In a previous version of this paper, we used additional information on executive compensation to estimate the standard deviation of the ability distribution outside the structural estimation routine (see Lippi and Schivardi (2010) for details). Again, the results we obtained were very similar to those reported here. All in all, our results have proven to be remarkably robust and consistent across specifications and estimation methods.

## 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 back out the structural parameters of the model; in particular, to infer the value of the personal relationship with executives enjoyed by the firm's owner. We estimated the model using matched employer–employee data for a sample of Italian firms and found that the nonmonetary objectives are quantitatively important. In particular, the value of personal relationships is highest in the firms under government and family control, and is smallest in conglomerate- or foreign-owned firms. From a quantitative point of view, those differences account for up to an $8\%$ differential in the firms' TFP. These results are robust to several controls and estimation methods.

One important question is what mechanisms could mitigate the inefficiency in executive selection that 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.

### Appendix A: The threshold rule in closed form

This appendix characterizes the optimal threshold $s^*$ as the unique solution of one equation in one unknown, and provides closed form expressions for the value functions expressed in terms of $s^*$. Using equation (2) and the optimal threshold $s^*$, it is straightforward to compute the *expected* value of a senior executive *conditional on being in office* as

$$
\begin{aligned}
v_o \equiv \mathbb{E}_s\big(v_o(s)|s \geq s^*\big) &= \int_{s^*}^{\infty} s\frac{dF(s)}{1-F(s^*)} + \beta\big[\rho v_y + (1-\rho)v_o\big] \\
&= \frac{1}{1-\beta(1-\rho)}\bigg(\int_{s^*}^{\infty} s\frac{dF(s)}{1-F(s^*)} + \beta\rho v_y\bigg).
\end{aligned}
\tag{12}
$$

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

$$
v_y = \mu + \beta\big[F\big(s^*\big)v_y + \big(1-F\big(s^*\big)\big)v_o\big].
\tag{13}
$$

Using equation (13) and the expression for $v_o$ in equation (12) gives a closed form equation for $v_y$ as a function of $s^*$:

$$
v_y = \frac{\mu(1-\beta(1-\rho)) + \beta\displaystyle\int_{s^*}^{\infty} s\,dF(s)}{(1-\beta)[1+\beta(\rho-F(s^*))]}.
\tag{14}
$$

Using equation (2) to write the value of a senior executive of type $s^*$ as

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

and replacing this expression into equation (3) gives the optimality condition

$$
s^* = (1-\beta)v_y.
\tag{15}
$$

Using equation (15) and the expression for $v_y$ in (14) gives one equation in one unknown for $s^*$:

$$H(s^*, R) \equiv s^*[1 + \beta(\rho - F(s^*))] - \mu(1 - \beta(1 - \rho)) - \beta \int_{s^*}^{\infty} s\, dF(s) = 0. \qquad (16)$$

## Appendix B: Proofs

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

Proof of Proposition 1. Simple algebra shows that $H(s^*, R)$ is continuous in $s^*$, that $H(0, R) < 0$, and that the first order derivative with respect to (w.r.t.) $s^*$ is positive, $H_{s^*}(s^*, R) = 1 + \beta(\rho - F(s^*)) > 0$, and in the limit, $\lim_{s^* \to \infty} H_{s^*}(s^*, R) > 0$. Hence, there exists one and only one $s^* > 0$ that solves equation (16).

  We now show that the implicit function $s^*(R)$ is increasing in $R$. Applying the implicit function theorem to equation (16) gives

$$\frac{\partial s^*}{\partial R} = \frac{(1 - \beta)\dfrac{\partial v_y}{\partial R}}{1 - (1 - \beta)\dfrac{\partial v_y}{\partial s^*}}. \qquad (17)$$

Let us use expression (14) to compute

$$\frac{\partial v_y}{\partial s^*} = \left( -\beta s^* f(s^*)[(1 - \beta)(1 + \beta(\rho - F(s^*)))] \right.$$

$$\left. + \beta(1 - \beta)f(s^*)\left[\mu(1 - \beta(1 - \rho)) + \beta \int_{s^*}^{\infty} s\, dF(s)\right] \right)$$

$$\Big/ [(1 - \beta)(1 + \beta(\rho - F(s^*)))]^2$$

$$= \beta f(s^*) \frac{-s^* + (1 - \beta)\dfrac{\mu(1 - \beta(1 - \rho)) + \beta \displaystyle\int_{s^*}^{\infty} s\, dF(s)}{(1 - \beta)(1 + \beta(1 - F(s^*)))}}{(1 - \beta)(1 + \beta(1 - F(s^*)))}$$

$$= \beta f(s^*) \frac{-s^* + (1 - \beta)v_y}{(1 - \beta)(1 + \beta(1 - F(s^*)))}.$$

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 (14) as

$$\int_{s^*}^{\infty} s\, dF(s) = q \int_{s^*-R}^{\infty} (x + R)\, dG(x) + (1 - q) \int_{s^*}^{\infty} x\, dG(x)$$

$$= q\left( R(1 - G(s^* - R)) + \int_{s^*-R}^{\infty} x\, dG(x) \right) + (1 - q) \int_{s^*}^{\infty} x\, dG(x).$$

Using this expression in (14) and taking the derivative w.r.t. $R$ yields

$$(1 - \beta)\frac{\partial v_y}{\partial R}$$

$$= \beta \frac{\dfrac{\partial \displaystyle\int_{s^*}^{\infty} s\, dF(s)}{\partial R} + (1 - \beta)v_y \dfrac{\partial F(s^*)}{\partial R}}{1 + \beta(1 - F(s^*))}$$

$$= \beta q \frac{[1 - G(s^* - R) + Rg(s^* - R) + (s^* - R)g(s^* - R)] - s^* g(s^* - R)}{1 + \beta(1 - F(s^*))}$$

$$= \beta q \frac{1 - G(s^* - R)}{1 + \beta(1 - F(s^*))} \in (0, 1),$$

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 (16) 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 $R \geq 0$.                    □

PROOF OF PROPOSITION 2.  Rewrite the average productivity $X$ defined in equation (6) as

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

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)

$$\frac{\partial X}{\partial R} = \frac{1}{1 + \rho - F(s^*)} \cdot \left[ qRg(s^* - R)\frac{\partial(s^* - R)}{\partial R} \right.$$

$$+ \frac{\partial F(s^*)}{\partial R} \tag{18}$$

$$\left. \times \left( \frac{\rho\mu + q\displaystyle\int_{s^*-R}^{\infty} x\, dG(x) + (1 - q)\displaystyle\int_{s^*}^{\infty} x\, dG(x) - \mu(1 - F(s^*))}{1 + \rho - F(s^*)} - s^* \right) \right].$$

Now use equation (16) with $\beta = 1$ to get the implicit equation for $s^*$:

$$
\begin{aligned}
s^* &= \frac{\rho\mu + \int_{s^*}^{\infty} s\, dF(s)}{1 + \rho - F(s^*)} \\
&= \frac{\rho\mu + q\int_{s^*-R}^{\infty} x\, dG(x) + (1-q)\int_{s^*}^{\infty} x\, dG(x) + qR(1 - G(s^* - R))}{1 + \rho - F(s^*)}.
\end{aligned}
$$

Replacing the expression on the right hand side for $s^*$ into equation (18) and using the expression for $\frac{\partial F(s^*)}{\partial R}$ computed in equation (5) gives (after some rearranging and cancellations)

$$
\begin{aligned}
\frac{\partial X}{\partial R} = \frac{qR}{1 + \rho - F(s^*)} &\left[ g(s^* - R)\left( \frac{1 + \rho - (q + (1-q)G(s^*))}{1 + \rho - F(s^*)} \right) \frac{\partial(s^* - R)}{\partial R} \right. \\
&\left. - \frac{1 - G(s^* - R)}{1 + \rho - F(s^*)}(1 - q)g(s^*)\frac{\partial s^*}{\partial R} \right].
\end{aligned}
\tag{19}
$$

Inspection of equation (19) and the results on the sign of the partial derivatives established in Proposition 1 reveal that the derivative is zero at $R = 0$ and that it is negative at $R > 0$. □

PROOF OF PROPOSITION 3. 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 $\mu$. 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

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

Then

$$
\varepsilon_i \equiv \phi_i \left( \frac{1}{n_{o,i}} \sum_{j=1}^{n_{o,i}} \zeta_{i,j} - \frac{1}{n - n_{o,i}} \sum_{j=1}^{n - n_{o,i}} \xi_{i,j} \right) + \frac{1}{n - n_{o,i}} \sum_{j=1}^{n - n_{o,i}} \xi_{i,j}.
\tag{20}
$$

We show that $\mathrm{cov}(\phi_i, \varepsilon_i) = 0$, so that the OLS regression assumptions are satisfied. Let $n$ be the number of executives in each firm. For notational convenience, let us define $z_i \equiv \frac{1}{n_{o,i}} \sum_{j=1}^{n_{o,i}} \zeta_{i,j}$ and $u_i \equiv \frac{1}{n - n_{o,i}} \sum_{j=1}^{n - n_{o,i}} \xi_{i,j}$ to write

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

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

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

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 senior executive productivity $x_j$ from 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 $\phi_i$ takes the values $(0, \frac{1}{n}, \ldots, \frac{k}{n}, \ldots, 1)$. As productivity realization is 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 $a = 1, 2$)

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

$$= \sum_{k=0}^{n} p(k, n)\left(\frac{k}{n}\right)^a \mathbb{E}_z[z_i|n_{o,i} = k] = \sum_{k=0}^{n} p(k, n)\left(\frac{k}{n}\right)^a \cdot 0 = 0.$$

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 owner and. hence, it cannot be correlated with his decisions about the tenure of the senior executives.                                  $\square$

### APPENDIX C: IDENTIFYING THE VARIANCE OF THE DISTRIBUTION OF ABILITIES

This section shows that the model of Section 3 can be inverted to infer the variance of the distribution of abilities from a moment that is available in the data. Recall that the distribution of abilities $G(\cdot)$ is chosen to be log normal so that it is defined by two parameters: $\lambda_m$ and $\lambda_\sigma$, respectively, the log mean and the log standard deviation of $x$. Now notice that a given vector of model primitives $\beta$, $\rho$, $\lambda_m$, $\lambda_\sigma$, $R$, and $q$, in addition to determining the observable pair $X$, $\phi$, also determines the variance of the selected $x$, which we denote by $\mathrm{var}(\eta_i)$, in accordance with equation (9).

Intuitively, a larger variance of the primitive parameter $\lambda_\sigma$ increases the option value of waiting, thus raising the selection threshold $\bar{\mu}$ and the steady state proportion of senior executives $\phi$ as well as the ex post variance of abilities.

Figure 4 shows how the standard deviation of the distribution of abilities $\mathrm{std}(\eta)$ varies as a function of the dispersion of abilities $\lambda_\sigma$. The figure has three lines, each one corresponding to a different value of $R$. Notice that for a given $R$, a larger volatility of the unfiltered distribution implies a larger variance in the selected sample: $\mathrm{std}(\eta)$. Such monotone curves are invertible and can thus be used to estimate $\lambda_\sigma$.
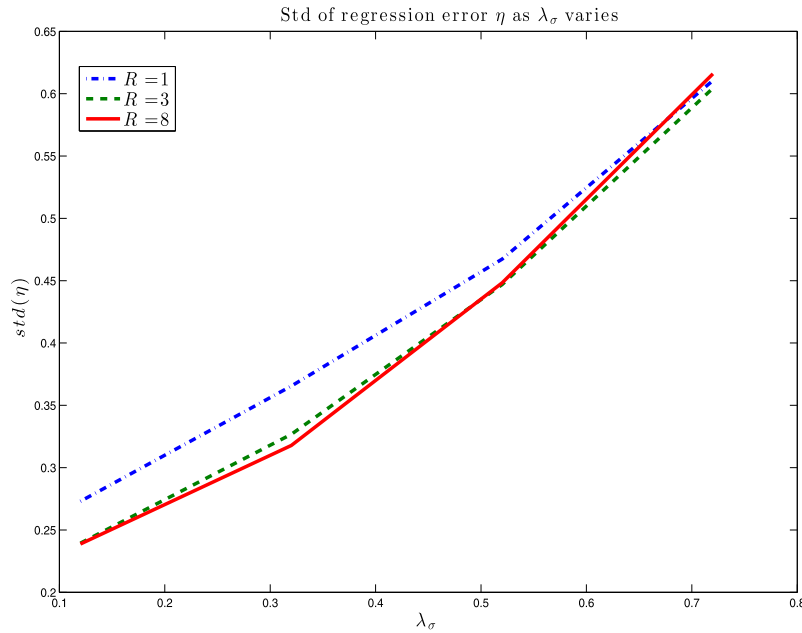
FIGURE 4. Model comparative statics as $\lambda_\sigma$ varies. *Note*: The figure uses the parameters $\beta = 0.98$ (per year), $\rho = 0.11$ (per year), and $q = 0.75$, and $G(\cdot)$ is log normal with log mean $\lambda_m = 1.6$.

TABLE 8. Estimated variance of abilities across control types.

|                          | Family | Conglomerate | Government | Foreign |
|--------------------------|--------|--------------|------------|---------|
| Estimated var($\eta_i$)  | 0.30   | 0.37         | 0.39       | 0.43    |
| Number obs.              | 2162   | 1818         | 1393       | 502     |

*Note*: We use the estimates in column (1) of Table 3 to obtain the standard deviation of the residuals of the productivity equation by control type. Given that the estimates use sample weight, we use the same weights to compute the standard deviation.

The empirical proxy of std($\eta$) that we used is taken from the standard deviation of the residual in the regression of equation (9). This is fully consistent with our theory provided that the number of executives in each firm is finite, a reasonable assumption as our firms on average employ about seven executives. Table 8 reports the sample estimates of std($\eta$) that are used as observables in the structural estimation.

## Appendix D: 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: (i) individual or family; (ii) government (local or central or other publicly controlled entities); (iii) conglomerate; (iv) institution (financial or not); (v) 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 the entire 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 that controls a resident firm might in turn be controlled by a resident who 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 nonfamily 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 data base. 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 showed that the unobservable productivity shock can be approximated by a nonparametric 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 detail 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,

TABLE 9. Coefficients of the additional controls in Tables 3, 4, and 10.

| | Dependent Variable | | | | | |
|---|---|---|---|---|---|---|
| | TFP | TFP | log VA | log VA | ROA | ROA |
| Firm age | −0.03*** | −0.05*** | −0.02** | −0.03* | −0.16 | −0.33 |
| | (0.009) | (0.018) | (0.009) | (0.017) | (0.222) | (0.337) |
| Firm age · family | | 0.04** | | 0.03 | | 0.76* |
| | | (0.023) | | (0.021) | | (0.448) |
| Firm age · foreign | | 0.03 | | 0.00 | | −0.45 |
| | | (0.031) | | (0.028) | | (0.750) |
| Firm age · government | | 0.02 | | −0.03 | | −0.12 |
| | | (0.034) | | (0.028) | | (0.518) |
| Work age | −0.48*** | −0.71*** | −0.43*** | −0.65*** | −17.33*** | −17.37*** |
| | (0.076) | (0.138) | (0.072) | (0.133) | (1.655) | (2.789) |
| Work age · family | | 0.23 | | 0.16 | | −0.29 |
| | | (0.172) | | (0.160) | | (3.411) |
| Work age · foreign | | 0.51** | | 0.64*** | | 1.12 |
| | | (0.241) | | (0.242) | | (6.067) |
| Work age · government | | −0.25 | | −0.43 | | −5.78 |
| | | (0.526) | | (0.501) | | (8.890) |
| Share exec | 1.61*** | 1.48 | 2.20*** | 2.41** | −6.27 | −24.65 |
| | (0.507) | (1.027) | (0.470) | (0.952) | (11.678) | (19.825) |
| Share exec · family | | −0.50 | | −0.72 | | 21.48 |
| | | (1.208) | | (1.109) | | (25.353) |
| Share exec · foreign | | 0.82 | | 0.07 | | 5.80 |
| | | (1.555) | | (1.454) | | (33.950) |
| Share exec · government | | 10.05*** | | 7.90*** | | 193.57*** |
| | | (2.915) | | (2.660) | | (63.083) |
| Share white-collar | 0.66*** | 0.79*** | 0.63*** | 0.71*** | 1.80 | 0.51 |
| | (0.054) | (0.099) | (0.048) | (0.096) | (1.348) | (1.797) |
| Share white-collar · family | | −0.16 | | −0.10 | | 2.31 |
| | | (0.114) | | (0.106) | | (2.357) |
| Share white-collar · foreign | | −0.19 | | −0.12 | | 2.50 |
| | | (0.145) | | (0.129) | | (3.494) |
| Share white-collar · government | | −0.25 | | −0.11 | | −3.57 |
| | | (0.167) | | (0.160) | | (3.476) |
| No. workers | 0.02** | 0.01 | 0.74*** | 0.72*** | 0.17 | 0.37 |
| | (0.009) | (0.015) | (0.013) | (0.018) | (0.215) | (0.301) |
| No. workers · family | | −0.02 | | 0.01 | | −0.29 |
| | | (0.020) | | (0.020) | | (0.431) |
| No. workers · foreign | | 0.04 | | 0.04 | | 0.09 |
| | | (0.025) | | (0.024) | | (0.600) |
| No. workers · government | | −0.02 | | −0.05 | | −1.29* |
| | | (0.036) | | (0.033) | | (0.730) |

(*Continues*)

TABLE 9. *Continued.*

| | TFP | TFP | log VA | log VA | ROA | ROA |
|---|---|---|---|---|---|---|
| | | | *Dependent Variable* | | | |
| Share male | 0.07** | 0.05 | 0.17*** | 0.15** | −0.97 | −1.34 |
| | (0.038) | (0.061) | (0.038) | (0.058) | (0.906) | (1.417) |
| Share male · family | | −0.01 | | 0.03 | | 1.38 |
| | | (0.065) | | (0.061) | | (1.498) |
| Share male · foreign | | 0.05 | | −0.08 | | −2.27 |
| | | (0.110) | | (0.099) | | (2.731) |
| Share male · government | | 0.51*** | | 0.43*** | | −0.15 |
| | | (0.184) | | (0.150) | | (3.160) |

*Note*: The table reports the coefficients on the additional controls in the last two columns of Tables 3, 4 and 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 10. Value added and share of senior executives by control type.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| | | | Dependent Variable: log *Value Added* | | | | |
| $\phi$ | 0.131*** | 0.127*** | 0.056* | 0.102** | 0.109*** | 0.129*** | 0.135*** |
| | (0.039) | (0.039) | (0.032) | (0.043) | (0.039) | (0.042) | (0.042) |
| $\phi \cdot$ foreign | −0.088 | −0.060 | −0.030 | 0.003 | −0.072 | −0.010 | −0.048 |
| | (0.061) | (0.060) | (0.046) | (0.066) | (0.061) | (0.063) | (0.065) |
| $\phi \cdot$ family | −0.204*** | −0.186*** | −0.115*** | −0.179*** | −0.137*** | −0.124*** | −0.135*** |
| | (0.046) | (0.045) | (0.038) | (0.050) | (0.046) | (0.047) | (0.048) |
| $\phi \cdot$ government | −0.471*** | −0.442*** | −0.304*** | −0.370*** | −0.347*** | −0.309*** | −0.246** |
| | (0.084) | (0.081) | (0.072) | (0.098) | (0.085) | (0.091) | (0.107) |

*Note*: The variable $\phi$ 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 third degree polynomial in capital and investment to control for the unobserved productivity shock. All regressions include control type dummies, year dummies, and two-digit sector dummies. Robust standard errors are given in parentheses. Significance levels for the null hypothesis of a zero coefficient are labelled by asterisks: * is 10%, ** is 5%, and *** is 1%.

we again follow Olley and Pakes, and include in the regression a third degreeg polynomial series in $i$ and $k$ and their interactions, which approximate the unobserved productivity shock.[19] In Table 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 col-

[19]Note 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.

umn (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) uses 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.

## Appendix E: Derivation of the likelihood function

The likelihood for a sample of observations $Y$ under the parametrization $\Theta$ is

$$L(\Theta; Y) = \prod_{\kappa=1}^{K} \prod_{j=1}^{3} \prod_{i=1}^{n_\kappa} \frac{1}{(2\pi\sigma_j^2)^{1/2}} \exp\left(-\frac{1}{2}\left[\frac{y_{i,\kappa}^j - f^j(\Theta, \kappa)}{\sigma_j}\right]^2\right)$$

$$= \prod_{\kappa=1}^{K} \prod_{j=1}^{3} (2\pi\sigma_j^2)^{-n_\kappa/2} \prod_{i=1}^{n_\kappa} \exp\left(-\frac{1}{2}\left[\frac{y_{i,\kappa}^j - f^j(\Theta, \kappa)}{\sigma_j}\right]^2\right),$$

where $K$ is the number of "groups" in the model (four control types in our case). This gives

$$\log L(\Theta; Y) = -\frac{1}{2}\sum_{\kappa=1}^{K}\sum_{j=1}^{3} n_\kappa \log(2\pi\sigma_j^2) - \frac{1}{2}\sum_{\kappa=1}^{K}\sum_{j=1}^{3}\sum_{i=1}^{n_\kappa}\left[\frac{y_{i,\kappa}^j - f^j(\Theta, \kappa)}{\sigma_j}\right]^2. \tag{21}$$

For all observables $j$ and for each group $\kappa$ (hence omitting the $j, \kappa$ subindices),

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

Replacing this expression in equation (21), we can rewrite the likelihood function by minimizing the distance between the theoretical value $f(\Theta, k)$ and the sample average $\bar{y}_\kappa^j$ for each variable $j$, or

$$\log L(\Theta; Y) = -\frac{1}{2}\sum_{\kappa=1}^{K}\sum_{j=1}^{3} n_\kappa \ln(2\pi\sigma_j^2)$$

$$-\frac{1}{2}\sum_{\kappa=1}^{K}\sum_{j=1}^{3}\left[n_\kappa^j + \frac{n_\kappa^j}{\sigma_j^2}(f^j(\Theta, \kappa) - \bar{y}_\kappa^j)^2\right]. \tag{22}$$

The measurement error for variable $j$ (common for all group $\kappa$) is $\sigma_j^2 \equiv \text{var}(y^j) = \sum_{i=1}^{n_\kappa} \frac{1}{n_\kappa}(y_{i,\kappa}^j - \bar{y}_\kappa^j)^2$.

### E.1 *Score and information matrix*

Let $M$ be the size of $\Theta$. The $n$th element of the score is given by

$$
\mathbf{s}_n(\Theta; Y) \equiv \frac{\partial \log L(\Theta; Y)}{\partial \theta_n} = -\frac{1}{2} \frac{\partial F(\Theta; Y)}{\partial \theta_n}
$$

$$
= \sum_{\kappa=1}^{K} \sum_{j=1}^{3} \left( \frac{n_\kappa}{\sigma_j^2} \right) \left( \bar{y}_\kappa^j - f^j(\Theta, \kappa) \right) \frac{\partial f^j(\Theta, \kappa)}{\partial \theta_n}.
$$

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

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

which in our case becomes

$$
I_{n,m}(\Theta) = \mathbb{E}\Bigg\{ \left[ \sum_{\kappa=1}^{K} \sum_{j=1}^{3} \left( \frac{n_\kappa}{\sigma_j^2} \right) \left( \bar{y}_\kappa^j - f^j(\Theta, \kappa) \right) \frac{\partial f^j(\Theta, \kappa)}{\partial \theta_n} \right]
$$

$$
\times \left[ \sum_{\kappa'=1}^{K} \sum_{j'=1}^{3} \left( \frac{n_{\kappa'}}{\sigma_{j'}^2} \right) \left( \bar{y}_{\kappa'}^{j'} - f^{j'}(\Theta, \kappa) \right) \frac{\partial f^{j'}(\Theta, \kappa)}{\partial \theta_m} \right] \Bigg\}
$$

$$
= \sum_{\kappa=1}^{K} \sum_{j=1}^{3} \left( \frac{n_\kappa}{\sigma_j^2} \right) \sum_{\kappa'=1}^{K} \sum_{j'=1}^{3} \left( \frac{n_{\kappa'}}{\sigma_{j'}^2} \right)
$$

$$
\times \mathbb{E}\big\{ \left( \bar{y}_\kappa^j - f^j(\Theta, \kappa) \right) \left( \bar{y}_{\kappa'}^{j'} - f^{j'}(\Theta, \kappa') \right) \big\} \frac{\partial f^j(\Theta, \kappa)}{\partial \theta_n} \frac{\partial f^{j'}(\Theta, \kappa')}{\partial \theta_m}
$$

$$
= \sum_{\kappa=1}^{K} \sum_{j=1}^{3} \left( \frac{n_\kappa}{\sigma_j^2} \right) \sum_{\kappa'=1}^{K} \sum_{j'=1}^{3} \left( \frac{n_{\kappa'}}{\sigma_{j'}^3} \right)
$$

$$
\times \mathbb{E}\left\{ \left( \frac{1}{n_\kappa} \sum_{i=1}^{n_\kappa} e_i^j \right) \left( \frac{1}{n_{\kappa'}} \sum_{i=1}^{n_{\kappa'}} e_i^{j'} \right) \right\} \frac{\partial f^j(\Theta, \kappa)}{\partial \theta_n} \frac{\partial f^{j'}(\Theta, \kappa')}{\partial \theta_m}
$$

$$
= \sum_{\kappa=1}^{K} \sum_{j=1}^{3} \left( \frac{n_\kappa}{\sigma_j^2} \right) \left( \frac{n_\kappa}{\sigma_j^2} \right) \left\{ \frac{\sigma_j^2}{n_\kappa} \right\} \frac{\partial f^j(\Theta, \kappa)}{\partial \theta_n} \frac{\partial f^{j'}(\Theta, \kappa')}{\partial \theta_m}
$$

$$
= \sum_{\kappa=1}^{K} \sum_{j=1}^{3} \left( \frac{n_\kappa}{\sigma_j^2} \right) \frac{\partial f^j(\Theta, \kappa)}{\partial \theta_n} \frac{\partial f^{j'}(\Theta, \kappa')}{\partial \theta_m}.
$$

### 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. [420, 421]

Alvarez, F. E. and F. Lippi (2009), "Financial innovation and the transactions demand for cash." *Econometrica*, 77, 363–402. [438]

Bandiera, O., I. Barankay, and I. Rasul (2009), "Social connections and incentives in the workplace: Evidence from personnel data." *Econometrica*, 77, 1047–1094. [420]

Bandiera, O., L. Guiso, A. Prat, and R. Sadun (forthcoming), "Matching firms, managers, and incentives." *Journal of Labor Economics*. [420, 421, 422]

Bartelsmann, E. and M. Doms (2000), "Understanding productivity: Lessons from longitudinal microdata." *Journal of Economic Literature*, 38, 569–594. [421]

Battistin, E., C. Graziano, and B. M. Parigi (2012), "Connections and performance in bankers' turnover." *European Economic Review*, 56, 470–487. [421]

Becker, G. S. (1971), *The Economics of Discrimination*. Economic Research Studies. University of Chicago Press, Chicago. [418]

Becker, G. S. (1998), *Accounting for Tastes*. Harvard University Press, Cambridge, Massachusetts. [418]

Bernard, A. B., J. Eaton, J. B. Jensen, and S. Kortum (2003), "Plants and productivity in international trade." *American Economic Review*, 93, 1268–1290. [421]

Bloom, N. and J. Van Reenen (2007), "Measuring and explaining management practices across firms and countries." *Quarterly Journal of Economics*, 122, 1351–1408. [421]

Cingano, F. and F. Schivardi (2004), "Identifying the sources of local productivity growth." *Journal of the European Economic Association*, 2, 720–742. [429, 449]

Dyck, A. and L. Zingales (2004), "Private benefits of control: An international comparison." *Journal of Finance*, 59, 537–600. [418, 420]

Ericson, R. and A. Pakes (1995), "Markov-perfect industry dynamics: A framework for empirical work." *Review of Economic Studies*, 62, 53–82. [421]

Faccio, M. and L. H. P. Lang (2002), "The ultimate ownership of Western European corporations." *Journal of Financial Economics*, 65, 365–395. [417]

Foster, L., J. Haltiwanger, and C. Syverson (2008), "Reallocation, firm turnover, and efficiency: Selection on productivity or profitability?" *American Economic Review*, 98, 394–425. [437]

Gabaix, X. and A. Landier (2008), "Why has CEO pay increased so much?" *Quarterly Journal of Economics*, 123, 49–100. [421]

Garrett, D. F. and A. Pavan (2012), "Managerial turnover in a changing world." *Journal of Political Economy*, 120, 879–925. [421]

Gayle, G., L. Golan, and R. Miller (2009), "Promotion, turnover and compensation in the executive market." Report, Carnegie–Mellon University. [421]

Hopenhayn, H. (1992), "Entry, exit, and firm dynamics in long run equilibrium." *Econometrica*, 60, 1127–1150. [421]

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. [449]

Jovanovic, B. (1982), "Selection and the evolution of industry." *Econometrica*, 50, 649–670. [421]

Kaplan, E. L. and P. Meier (1958), "Nonparametric estimation from incomplete observations." *Journal of the American Statistical Association*, 53, 457–481. [439]

Kramarz, F. and D. Thesmar (2013), "Social networks in the boardroom." *Journal of the European Economic Association*, 11, 780–807. [421]

La Porta, R., F. Lopez-de-Silanes, and A. Shleifer (1999), "Corporate ownership around the world." *Journal of Finance*, 54, 471–517. [417]

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. [420]

Lippi, F. and F. Schivardi (2010), "Corporate control and executive selection." Discussion Paper 8031, CEPR. [442]

Lucas, R. E., Jr. (1978), "On the size distribution of business firms." *Bell Journal of Economics*, 2, 508–523. [418, 421]

McCall, J. J. (1970), "Economics of information and job search." *Quarterly Journal of Economics*, 84, 113–126. [418, 424]

Melitz, M. J. (2003), "The impact of trade on intra-industry reallocations and aggregate industry productivity." *Econometrica*, 71, 1695–1725. [421]

Michelacci, C. and F. Schivardi (2013), "Does idiosyncratic business risk matter?" *Journal of the European Economic Association*, 11, 343–368. [421]

Moskowitz, T. J. and A. Vissing-Jørgensen (2002), "The returns to entrepreneurial investment: A private equity premium puzzle?" *American Economic Review*, 92, 745–778. [418, 421, 427]

OECD (2003), *OECD Science, Technology and Industry Scoreboard*. Annex 1. OECD, Paris. [429, 430]

Olley, S. G. and A. Pakes (1996), "The dynamics of productivity in the telecommunications equipment industry." *Econometrica*, 64, 1263–1297. [419, 429, 449, 451]

Sraer, D. and D. Thesmar (2007), "Performance and behavior of family firms: Evidence from the French stock market." *Journal of the European Economic Association*, 5, 709–751. [437]

Syverson, C. (2011), "What determines productivity?" *Journal of Economic Literature*, 49, 326–365. [421]

Taylor, L. A. (2010), "Why are CEOs rarely fired? Evidence from structural estimation." *Journal of Finance*, 65, 2051–2087. [420, 421, 440]

Taylor, L. A. (2013), "CEO wage dynamics: Evidence from a learning model." *Journal of Financial Economics*, 108, 79–98. [421]

Terviö, M. (2008), "The difference that CEOs make: An assignment model approach." *American Economic Review*, 98, 642–668. [421]

Volpin, P. F. (2002), "Governance with poor investor protection: Evidence from top executive turnover in Italy." *Journal of Financial Economics*, 64, 61–90. [421]

Weitzman, M. L. (1979), "Optimal search for the best alternative." *Econometrica*, 47, 641–654. [418]