CEO Behavior and Firm Performance*

Oriana Bandiera
London School of Economics

Stephen Hansen
Universitat Pompeu Fabra

Andrea Prat
Columbia University

Raffaella Sadun
Harvard University

March 25, 2016

****PRELIMINARY AND INCOMPLETE****

Abstract

We measure the behavior of over 1,100 CEOs in six countries (Brazil, France, Germany, India, UK and US) using a new methodology that combines (i) a survey measuring each activity undertaken in a random work-week by the executives and (ii) a machine learning algorithm that projects these high dimensional data onto one CEO behavior index. A simple firm-CEO matching model yields the null hypothesis that, in absence of matching frictions, CEO behavior is uncorrelated with firm performance. Combining the CEO behavior index with firm level accounting data we reject this null. We find a large and significant correlation between CEO behavior and firm performance, which appears only gradually over time after the CEO is appointed and is stronger in emerging economies. Our results suggest that CEO-firm matching frictions may account for a sizable fraction of the cross country productivity differential observed in the data.

*This project was funded by Columbia Business School, Harvard Business School and the Kauffman Foundation. We are grateful to Wouter Dessein, Bob Gibbons, Rebecca Henderson, Ben Hermalin, Amit Khandelwal, Antoinette Schoar, Steve Tadelis and seminar participants at Bocconi, Chicago, Columbia, Harvard Business School, Politecnico di Milano, Stanford Management Conference, Tel Aviv, Toronto, Uppsala, Warwick for useful suggestions.

1 Introduction

The impact of CEOs on firm performance is at the core of many economic debates. The conventional wisdom, backed by a growing body of empirical evidence (Bertrand and Schoar 2003, Bennedsen et al. 2007, Kaplan et al. 2012) is that the identity of the CEO matters for firm performance. But what do CEOs actually do? Do different CEOs behave differently? and (why) does it matter?

In this paper we develop a new methodology to measure CEO behavior in large samples combining (i) a survey that systematically codifies CEO diaries at the activity level and (ii) a machine learning algorithm that projects the several dimensions of CEO behavior onto a low-dimensional behavior index. We use this data to study the correlation between CEO behavior and firm performance within the framework of a simple firm-CEO matching model.

Our survey methodology is inspired by the classic study of CEO behavior of Mintzberg (1973), which was based on data collected by shadowing five CEOs over the course of one week. We scale-up this methodology by focusing on CEOs' diaries rather than shadowing individuals directly. This approach allows us to collect detailed and comparable data on the behavior of 1,114 CEOs of manufacturing firms in six countries: Brazil, France, Germany, India, UK and the US.

The survey gathered information on all the activities the CEOs undertook each day over the course of one randomly selected week using a team of forty enumerators who phoned the CEOs or their PAs every day. For each activity, we collected information on five features: its nature (e.g. meeting, site visits, public event, etc.), the planning horizon, the number of participants, the number of different functions and the type of participants (i.e firm employees vs outsiders) and their function (e.g. finance, marketing, clients, suppliers, etc.). Overall, we collected data on 42,233 activities of different lengths, covering an average of 50 working hours per CEO.[1] Each of these activities is characterized by one of the 4,253 combinations of the five features described above.

We use an unsupervised Bayesian machine learning algorithm, Latent Dirichlet Allocation (Blei et al. 2003) to project this high-dimensional feature space onto a lower-dimensional behavior space in a non-subjective fashion.[2] We begin by estimating behaviors common to all CEOs as probability vectors over the activity feature set described above. We then estimate a CEO-specific behavior index as the distribution over the prototype behaviors - namely we allow, but do not force, each CEO to have a different mix

[1]In earlier work (Bandiera et al. (2013) we use the same data to measure the CEOs' labor supply and assess whether and how it depends on family ownership.

[2]The typical application of LDA is to natural language, where it is widely cited (for an example in economics, see Hansen et al. 2014). It is less commonly used for survey data, but in principle it is able to also usefully reduce the dimensionality of any dataset of counts.

of behaviors.

In our baseline specification we estimate two "pure" behaviors, and a uni-dimensional behavioral index that ranges from 0 (for CEOs who follow a pure behavior 0) to 1 (for CEOs who follow a pure behavior 1). The two behaviors differ considerably: the feature combinations that are most frequent in behavior 0 are least frequent in behavior 1, and vice versa. Low values of the CEO behavior index are associated with direct monitoring of production, less planning and one-on-one meetings with outsiders alone. In contrast, high values are associated with CEOs participating in larger meetings that involve high level functions both inside and outside the firm, and that are planned in advance.

While the diary data reveal that different CEOs behave differently on all dimensions, there is no theoretical reason to expect either type of behavior to lead to higher/lower performance for all firms, or to be more/less costly for all CEOs. To the contrary, the fact that different behaviors coexist suggests that they might be best responses to different circumstances faced by the firm. Indeed, in our data we find that on average CEOs with high values of the behavior index are more likely to be found in larger firms and in industries with more complex production processes (Autor et al. 2003).

However, performance differentials related to CEO behavior may still arise in the presence of significant matching frictions, i.e. if CEOs with different behavioral patterns are not optimally matched with the specific needs of the firms they run. To illustrate this point, we develop a model of firm-CEO matching model with two types of firms and two types of CEOs. Firm type determines which CEO behavior is most productive given its specific features, while CEO type determines the cost of adopting a certain behavior. The pool of potential CEOs is larger than the pool of firms seeking a CEO, and one type of CEO is relatively more abundant than the other, i.e. its share is higher than the share of firms who seek CEOs of that type. We allow for two types of frictions in the market for CEOs. First, the screening technology is imperfect and it cannot always correctly identify the actual CEO's type. Second, after hiring the CEO, the firm can offer him incentives to adopt the right behavior, but these are limited due to poor governance or labor laws that make dismissals costly.

The model specifies that, if frictions are small, all firms will hire CEOs of the right type, and that these will adopt the behavior that is optimal for the firm. Therefore, in these circumstances the correlation between CEO behavior and firm performance will be zero, which is the null hypothesis we test in the data. In alternative, frictions are large enough that, in equilibrium, some of the abundant type CEOs will match with the wrong type of firms, so that they would in some instances (the mismatched firms) be associated with lower productivity. In this case, we would observe a positive correlation between the behavior adopted by the scarce CEO type and firm performance.

Guided by the model, we then combine our estimated CEO behavior index with firm level accounting data. Using the set of 831 firms (75% of the CEO sample) for which accounting data is available, we find that high values of the CEO behavior index are significantly correlated with firm productivity. A standard deviation increase in the CEO behavior index is associated with a 0.10 log points increase in productivity, which is about 12% of the increase associated with a standard deviation increase in capital. In light of the model, these results imply that matching frictions are sufficiently large to create some mismatches between firms and CEOs, and that the (unobserved) CEO type that leads to high values of the CEO behavior index is relatively scarce in the population.

This interpretation relies on the identifying assumption, transparent in the model, that firm traits that determine which CEO behavior is optimal are orthogonal to unobservable determinants of firm productivity. In other words, conditional on size, capital and industry, the potential productivity of firms that need low index CEOs and those that need high index CEOs is the same. If this assumption fails, the fact that a firm hires a low index CEO might just reflect firm traits that lead to low productivity. We test this identifying assumption using accounting data for the period before the current CEO was appointed. We find that *before* current CEO is hired, the productivity of firms that currently hire low index CEOs is the same as that of firms that currently hire high index CEOs. This rules out the possibility that the results are driven by time invariant firm traits that correlate with productivity and the type of CEOs that firm hire. Furthermore, we find that the correlation between CEOs behavior and firm performance only materializes four years after the CEO appointment. This helps us address the concern that the results may be driven by time varying firm traits, namely that firms hire low index CEOs after their productivity starts declining (in which case we should see an effect the year before the CEO appointment) or right at the same time as productivity starts declining (in which case we should see the correlation between the CEO index and productivity starting in the same year in which the CEO is appointed).

Next, we exploit the cross-regional variation in regional GDP to proxy for differences in the severity of matching frictions across regions and test the prediction that the quality of the match should be higher and the correlation between behavior and performance should be lower when matching frictions are less severe. First, we allow the correlation between CEO behavior and firm size to vary with the level of regional development within country and find the correlation between CEO behavior and firm size to be stronger in richer regions. Second, we also allow the correlation between CEO behavior and firm performance to vary with the level of regional development within country. In line with the predictions of the matching model, we find that this correlation is weaker in richer regions, and we cannot reject the null that the correlation equals zero in the richest regions in our sample. Taken together, these findings provide further support to the hypothesis that the correlation between CEO behavior and firm performance is driven by matching frictions in the market for CEOs rather than unobservable firm characteristics. Moreover, the findings also cast doubt on the alternative hypothesis that one type of behavior is always better for all the firms. If it were so, we would find that behavior to be positively correlated with firm performance regardless of the severity of the frictions (i.e. across all countries in our sample).

The final part of the analysis brings the model to the data to back out the share of mismatched firms-CEOs pairs and calibrate the parameters that cause the mismatch, and the extent to which matching frictions may be able to account for productivity differences across countries.

To the best of our knowledge, ours is the only study that collects time use data to measure CEO behavior in large samples to study its link to firm performance. The management literature contains some examples of time use analyses but on much smaller samples and for managers on lower rungs of the hierarchy.[3] In economics, our findings are complementary to the literature that studies the correlation between CEO traits and firm performance. Malmendier and Tate (2005) and Malmendier and Tate (2009) focus on overconfidence; they find that this is correlated with higher investment–cash flow sensitivity and mergers that destroy value. Kaplan et al. (2012) and Kaplan and Sorensen (2016) have detailed data on skills and personality traits of several CEOs candidates; they show the CEOs mostly differ along three dimensions: managerial talent, execution skills and interpersonal skills. Of these, only talent and execution skills correlate with firm performance but interpersonal skills increase the likelihood that the candidate is hired. This is consistent with our assumption that screening is imperfect and firms can end up hiring the wrong CEOs. Our methodology is complementary to Mullins and Schoar (2013) who use self-reported survey questions to measure the management style and values of 800 CEOs in emerging economies. Their focus however differs as they aim to explain variation in style and values rather than the link with performance. Finally, this paper is complementary to a growing literature documenting the role of basic management processes on firm performance (Bloom and Van Reenen (2007) and Bloom et al. (2016)). The relationship between CEO behavior and firm performance that we identify is of the same order of magnitude as the effect of management practices. Furthermore, for a subset of our firms we have both CEO behavior data and management scores (measured at middle managerial levels) and we are able to check that both variables retain independent explanatory power, thus suggesting that these might reflect two distinct channels through which managerial activity influences firm performance.

[3]The largest shadowing exercise on top executives known to us –Kotter (1999) –includes 15 general managers, not CEOs. The largest time use study of managerial personnel we are aware of is Luthans (1988), which covers 44 mostly middle managers. Some professional surveys ask large numbers of CEOs general questions about their aggregate time use (e.g. McKinsey 2013), but they do not collect detailed calendar information.

The paper is organized as follows. Section 2 describes the data and the machine learning algorithm that yields the CEO behavior index. Section 3 presents the matching model, which is then used to inform the empirical analysis in section 4. Section 5 investigates the extent to which matching frictions vary across regions, while section 6 calibrates the model to quantify the share of mismatches and their consequences for performance differentials across countries. Section 7 concludes.

2 Measuring CEO Behavior

2.1 Sample

The survey covers CEOs in six of the world's ten largest economies: Brazil, France, Germany, India, the United Kingdom and the United States. For comparability, we chose to focus on established market economies and opted for a balance between high and middle-low income countries. While titles may differ across countries (e.g. Managing Director in the UK) we always interview the highest-ranking authority in charge of the organization who has executive powers and reports to the board of directors. For brevity we refer to them as CEOs in what follows.

Our sampling frame was drawn from ORBIS, a data set that contains firm level accounting data for more than 30 million firms around the world. In line with other studies (Bloom et al. 2016), the sample is restricted to manufacturing to be able to more reliably compare performance across firms. Among firms in this sector we selected those with available sales and employment data, yielding 11,500 potential sample firms. We could find CEOs contact details for 7,744 firms and of these 1,217 later resulted not to be eligible.[4] The final number of eligible firms was thus 6,527 in 32 two-digit SIC industries. We randomly assigned these to different enumerators to call to seek the CEOs' participation, and we managed to interview the CEOs of 1,114 of them[5] - a 17% response rate. This figure is at the higher end of response rates for CEO surveys, which range between 9% and 16% (Graham et al 2011). Our final sample thus comprises of 1,114 CEOs, of which 282 are in Brazil, 115 in France, 125 in Germany, 356 in India, 87 in the UK and 149 in the US.

Table AXX shows that sample firms have on average slightly lower log sales (coefficient 0.071, standard error 0.011) but we do not find any significant selection effect on performance variables, such as labor productivity (sales over employees) and return on capital employed (ROCE).

[4]The reasons for non eligibility included recent bankruptcy or the company's not being in manufacturing. 310 of the 1217 could not be contacted before the project ended.

[5]1,131 CEOs agreed to participate but 17 dropped out before the end of the data collection week for

Table 1: Summary Statistics

Variable	Mean	Median	Standard Deviation	Observations
A. CEOs Traits				
CEO age	50.93	52.00	8.45	1107
CEO gender	0.96	1.00	0.19	1114
CEO has college degree	0.92	1.00	0.27	1114
CEO has MBA	0.55	1.00	0.50	1114
CEO has studied abroad	0.48	0.00	0.50	1114
CEO tenure in post	10.29	7.00	9.55	1110
CEO tenure in firm	17.10	16.00	11.58	1108
CEO belongs to the owning family	0.41	0.00	0.49	1114
B. Firms Traits				
Employment	1275.5	300.0	6497.7	1114
Sales ('000 $)	205407.2	36803.2	1417493.0	831
Capital ('000 $)	76436.6	9340.2	494726.3	613
Materials ('000 $)	142859.8	22198.5	1403714.0	378
Profits per employee ('000 $)	24.8	10.1	32.2	516
Tobin's q	0.8	0.7	0.5	296
C. Industry Traits				
Task abstraction	2.51	2.35	0.72	1050
Capital Intensity	4.26	4.12	0.65	1046
Homogeneous Product	0.67	0.83	0.38	1009
D. Regional Traits				
Log Regional Income per Capita	9.36	9.48	1.08	1111

Notes: "Task abstraction" is an industry metric drawn from Autor et al. (2003) with higher values denoting a higher intensity of abstract tasks in production. "Capital intensity" denotes the average industry level value of capital over labour, built from the NBER manufacturing database (aggregated between 2000 and 2010). "Homogeneous product" is an industry dummy drawn from Rauch (1999). Log regional income per capita at the regional level in current purchasing-power-parity (PPP) dollars is drawn from Gennaioli et al. (2013).

Table 1, Panel A and B shows descriptive statistics on the sample CEOs and their firms. Sample CEOs are 52 years old on average, nearly all (96%) are male and have a college degree (92%). About half of them have an MBA and a similar share has studied abroad. The average tenure is 10 years, with a standard deviation of 9.6; the heterogeneity is mostly due to the distinction between family and professional CEOs as the former have much longer tenures.[6]

personal reasons.

[6]In our sample 57% of the firms are owned by a family, 23% by disperse shareholders, 9% by private individuals, and 7% by private equity. Ownership data is collected in interviews with the CEOs and independently checked using several Internet sources (e.g. The Economic Times of India, Bloomberg,

2.2 The Executive Time Use Survey

2.2.1 Data collection

The data were collected by a team of enumerators through daily phone calls with the personal assistant (PA) of the CEO, or with the CEO himself (43% of the cases), over a week randomly chosen by us.[7] On day one of this week (typically a Monday), the enumerator called in the morning and gathered detailed information on all the activities planned in the CEO diary for the day. The enumerator then called again in the evening, to gather information on the actual activities undertaken by the CEO (including those that were not originally included in the planned agenda), and the activities planned for the following day. On subsequent days, the enumerator called in the evening, again to collect data on the actual activities undertaken during the day, and the planned schedule for the next day.[8]

The survey collects information on all the activities lasting longer than 15 minutes in the order they happened during the day. Figure A.1 shows a screen-shot of the survey tool.[9] For each activity we collect information on the following features: (1) type (e.g. meeting, public event, etc.); (2) duration (15m, 30m, etc.); (3) planning (planned or unplanned); (4) number of participants (one, more than one); (5) functions of participants, divided between employees of the firms or "insiders" (finance, marketing, etc.) and "outsiders" (clients, banks, etc.).

Overall we collect data on 42,233 activities of different duration, equivalent to 225,721 15-minute blocks. The average CEO thus has 202 15-minute activities, adding up to 50 hours per week.

2.2.2 Feature description and combinations

In 57,216 times blocks (25.3% of total time), CEOs are either working alone or sending emails; in 21,895 (9.7%) they are engaged in personal or family time; and in 18,950 (8.4%)

etc.), information provided on the company website and supplemental phone interviews. We define a firm to be owned by an entity if this controls more than 25.01% of the shares; if no single entity owns at least 25.01% of the share the firm is labeled as "Dispersed shareholder".

[7]The data collection methodology discussed in this section is an evolution of the approach followed in Bandiera et al. (2012) to collect data on the agenda of 100 Italian CEOs. While the data collection of the Italian data was outsourced to a private firm, the data collection described in this paper was internally managed from beginning to end. Due to this basic methodological difference and other changes introduced after the Italian data was collected (e.g. the vector of features used to characterize every activity) we decided not to combine the two samples.

[8]For 70% of the CEOs in our sample, the work week consisted of 5 days. The remaining 30% of the CEOs also reported to work during the weekend (21% for 6 days and 9% for 7 days). Analysts were instructed to call the CEO after the weekend to retrieve data on Saturdays and Sundays. On the last day of the data collection, the analysts also interviewed the CEO to validate the activity data (if collected through his PA) and to collect information on the characteristics of the CEO and of the firm.

[9]The survey tool can also be found online on `www.executivetimeuse.org`.

they are traveling. In the remaining 127,660 time blocks (56.6% of total time), CEOs spend time with at least one other person. In the baseline analysis we only consider these latter interactive activities because they are the ones for which we can measure with precision the vector of specific attributes (e.g. planning, number of participants) which, as described below, are used to derive classifications of CEO behavior. Since this approach may potentially eliminate useful information, we also include a series of robustness checks showing that the main results are not sensitive to this choice.

Table 2, panel A shows the share of time devoted to different options within features for the interactive activities. Thus, within the "type" feature, the most frequent entry is "meeting", which accounts for 74.1% of time. Table 2, panel B shows the time the average CEO spends with different functions. Perhaps unsurprisingly given that we are working with a sample of manufacturing firms, the average CEO is most likely to spend time with employees involved in production. CEOs also spend more time with inside than outside functions. Functions are not mutually exclusive, and CEOs can spend time with more than one function in a single activity; in 39.5% of activities there is more than one function present.

While Table 2 shows average behavior, the data features substantial heterogeneity across CEOs. For example, while the average CEO spends 75% of his or her time in planned activities, the 25th and 75th percentiles are 64% and 91%, respectively. The corresponding percentiles for time spent with production functions is 19% and 51%.

In order to fully describe each 15-minute block of CEO time, we combine all the features into a single overall variable. More specifically, we define each block of time according to the five distinct features described above (type of activity, duration, planning horizon, number of participants, type of functions involved). Using this approach, we obtain 4,253 unique combinations in the data.[10] Examples of such combinations are:

1. Meeting; Duration of 1 hour or more; Planned; Two or more participants; With production

2. Meeting; Duration of 30 minutes max; Unplanned; One participant; With marketing

3. Meeting; Duration of 1 hour or more; Unplanned; Two or more participants; With marketing and production

4. Public Event; Duration of 1 hour or more; Planned; Two or more participants; With clients, suppliers and competitors

The most frequent, associated with 3,620 15-minute time blocks, is example (1) above.

[10] In all cases, the value of the first four features is unique, while the value of the last feature—the functions present in the activity—is a set that contains one or more elements.

Table 2: Average Time Shares for all CEOs

(a) Distribution of time within features

Type		**Duration**		**Planned**		**Participants**	
value	share	value	share	value	share	value	share
meeting	0.741	1hr+	0.642	planned	0.754	size2+	0.62
business_meal	0.07	1hr	0.198	unplanned	0.244	size1	0.362
phone_call	0.06	30m	0.138	missing	0.002	missing	0.018
site_visit	0.059	15m	0.022				
conference_call	0.033						
public_event	0.02						
workrelated_leisure	0.011						
video_conference	0.005						
other	0.0						

(b) Distribution of time across functions

Inside Functions		**Outside Functions**	
function	share	function	share
production	0.354	clients	0.108
mkting	0.224	suppliers	0.069
finance	0.173	others	0.059
hr	0.082	associations	0.036
groupcom	0.081	consultants	0.035
bunits	0.055	govoff	0.023
other	0.049	compts	0.02
board	0.043	banks	0.018
admin	0.042	lawyers	0.015
cao	0.036	pemployee	0.015
coo	0.03	investors	0.014
strategy	0.022		
legal	0.018		

Notes: The top table shows the amount of time the average CEO spends on different options within features for the 127,660 interactive 15-minute unit of time in the data. The bottom table shows the amount of time the average CEO spends with different functions. Since there are typically multiple functions in a single activity, these shares sum to more than one.

2.3 Estimating Behavior: Machine Learning

Knowing which aspects of CEO behavior to focus on *ex ante* is complicated by the fact that there is no theory to guide our choice. At the same time, the dimensionality of the feature combination space is too high to work with using conventional econometric models. We instead adopt a machine-learning approach that allows us to use *all* dimensions of variation across CEOs to describe a low-dimensional set of behaviors. The particular algorithm we use is Latent Dirichlet Allocation, or LDA(Blei et al. 2003).

One key advantage of LDA over simpler dimensionality-reduction techniques like principal components analysis (PCA) or k-means clustering is that it is a so-called "generative" model that provides a complete probabilistic description of time use patterns linked to statistical parameters. In this sense, LDA is akin to structural estimation in econometrics. In contrast, PCA performs an eigenvalue decomposition of the variance-covariance matrix, while k-means solves for centroids with the smallest squared distance from the observations. Neither procedure estimates the parameters of a statistical model, which can make interpreting their output difficult. Moreover, we believe that LDA can provide a basic framework for embedding more complex time use patterns in future work.

Suppose all CEOs have F possible ways of organizing any given unit of time, and let x_f be a particular way of organizing time. In our baseline case, $F = 654$ and x_f is a combination of the values of the five features described above. We refer to $X \equiv \{x_1, \ldots, x_F\}$ as the *activity feature set*.

A management *behavior* k is a probability distribution $\boldsymbol{\beta}_k$ over X that is common to all CEOs. That is, every CEO who adopts management behavior k draws elements from the activity feature set according to the same distribution $\boldsymbol{\beta}_k$. The fth element of $\boldsymbol{\beta}_k$ $\left(\beta_k^f\right)$ gives the probability of generating x_f when adopting behavior k. All behaviors are potentially associated with all elements of X ($\boldsymbol{\beta}_k \gg 0$ is compatible with the definition of behavior), but some can be associated with some behaviors more than others ($\beta_k^f \neq \beta_{k'}^f$ in general when $k \neq k'$). We assume there are K behaviors in the data, and discuss its value in section 2.3.1 below.

Describing time use in terms of combinations of individual features allows for arbitrary covariance patterns among features. For example, rather than simply estimating the probability that behavior k plans activities or not, we estimate a behavior-specific probability for planning larger versus smaller meetings, for activities with many versus fewer functions, etc. Later, we look at the marginal distributions over each individual feature separately, but we emphasize that no independence assumptions between individual features are built into the estimation procedure.

We associate to each CEO i a distribution over behaviors $\boldsymbol{\theta}_i$, a K-dimensional probability vector whose kth element θ_i^k gives the probability that CEO i adopts behavior k when organizing a unit of time. This allows CEOs to potentially mix different behaviors over the course of a workday or workweek, albeit with a tendency to adopt some behaviors more than others.[11] As $\boldsymbol{\theta}_i$ describes CEO i's time use on a low-dimensional space of behaviors, we refer to it as a *behavioral index*.

For each of the $N = 1,114$ CEOs, we observe T_i distinct units of managerial time, each with an associated $y_{i,t} \in X$. $y_{i,t}$ is generated in two steps:

1. First, CEO i draws a behavior associated with the t-th unit of time from $\boldsymbol{\theta}_i$. Denote this behavior $z_{i,t}$.

2. Second, given the assignment of behavior $z_{i,t}$, draw an activity feature $y_{i,t}$ from the activity feature set according to the distribution $\boldsymbol{\beta}_{z_{i,t}}$.

The probability of observing $y_{i,t}$ given the parameters $\boldsymbol{\beta} \equiv (\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_K)$ and $\boldsymbol{\theta} \equiv (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_N)$ is

$$\Pr[\, y_{i,t} \mid \boldsymbol{\beta}, \boldsymbol{\theta} \,] = \sum_{z_{i,t}} \Pr[\, y_{i,t}, z_{i,t} \mid \boldsymbol{\beta}, \boldsymbol{\theta} \,] = \sum_{z_{i,t}} \Pr\left[\, y_{i,t} \,\middle|\, \boldsymbol{\beta}_{z_{i,t}} \right] \Pr[\, z_{i,t} \mid \boldsymbol{\theta}_i \,]. \tag{1}$$

By independence, the probability of all the observed data is $\prod_i \prod_t \Pr[\, y_{i,t} \mid \boldsymbol{\beta}, \boldsymbol{\theta} \,]$.

The independence assumption of time blocks within a CEO may appear strong since one might imagine that CEOs' behavior is persistent across a day or week. However our goal in this initial application of machine learning methods in the economics of management literature is to understand overall patterns of time use for each CEO rather than issues such as the evolution of behavior over time, or other more complex dependencies. These are of course interesting, but outside the scope of the paper.[12]

While in principle one can attempt to estimate $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ via maximum likelihood, in practice this problem is intractable. Instead, LDA uses Bayesian inference and places Dirichlet priors on each of the $\boldsymbol{\beta}_k$ and $\boldsymbol{\theta}_i$ terms. The Dirichlet distribution of dimensionality M is defined on the $M-1$-simplex, and provides a flexible means of modeling the probability of the weights for multinomial or categorical distributions. (The Dirichlet with $M = 2$ corresponds to the beta distribution). Symmetric Dirichlet distributions are parameterized by a scalar α. When $\alpha = 1$, the Dirichlet places uniform probability on all elements of simplex; when $\alpha < 1$ it places more weight on the corners of the simplex and so generates multinomial weights that tend to have a few large values and many small values; and when $\alpha > 1$ it places more weight on the center of the simplex and so generates multinomial weights that are similar in magnitude. Hereafter we let α denote the parameter associated with the symmetric Dirichlet prior on the behavioral indices, and η the parameter associated with the prior on the behaviors.

[11] Note that this model nests a simpler alternative in which CEOs adopt a single behavior k for all their time use in the sense that θ_i^k can be arbitrarily close to 1.

[12] The independence assumption taken literally also implies that units of time within the same activity are independent. We explain in section 2.4 why we treat the unit of analysis as a time block rather than an activity.

2.3.1 Model selection

There are three model parameters whose values we impose: the number of behaviors K, and the hyperparameters α and η. Choosing the dimensionality of the latent space in unsupervised learning—in our case K—is a persistent challenge in this literature. We estimate a model with $K = 2$ as a baseline. This is the minimal model size that admits heterogeneity between CEOs, and captures the main distinctions among CEOs in the data regarding time use. In the appendix we explore the model with $K = 3$. When $K = 2$, the CEO behavioral index can be summarized by a scalar $\theta_i \equiv \theta_i^1$, or the probability that CEO i adopts behavior 1. [13]

As for the hyperparameters, we set $\alpha = 1$, which corresponds to a uniform prior on each CEO's behavioral index. We also set $\eta = 0.1$. As noted above, this means the prior on the β_k terms places more weight on probability vectors that have their mass concentrated on a limited number of elements of the activity feature set. In other words, we set the prior so that behaviors feature some combinations prominently, but put little weight on many others.

2.4 Inference Algorithm

Exact posterior inference for LDA is intractable due to the high dimensionality of the model. In a model with K styles there are $K^{98,347}$ possible realizations of the latent variables—each block of time in the data can take any of K values. Enumerating all these events to compute the posterior distribution is computationally infeasible. One must therefore use approximate posterior inference algorithms, and we follow the Markov Chain Monte Carlo approach of Griffiths and Steyvers (2004) to sample the behaviors associated with each unit of time.[14] The general idea of MCMC estimation is to randomly seed the model with initial values for the behaviors associated to time units $z_{i,t}$, perform initial sampling iterations while the Markov Chain "burns in" to its stationary distribution, and then draw samples every nth iteration thereafter. The gap between sample draws is called a thinning interval, and is introduced to reduce autocorrelation between samples. The samples are then averaged to form estimates as in Monte Carlo simulations.

The specific procedure we adopt is:[15]

[13] An alternative approach would be to apply statistical criteria to choose K such as cross-validation or marginal likelihood methods (see Taddy 2012 for further discussion). Preliminary analysis indicates that, for our data, the resulting K is larger than 50. While such a model might predict feature combinations better than our baseline, interpreting its results would be very challenging. In a natural language context, Chang et al. (2009) also show a tension between predictability and interpretability in the choice of K, with larger values favoring the former and smaller values the latter.

[14] We use a collapsed Gibbs sampling procedure that integrates out the $\boldsymbol{\beta}_k$ and $\boldsymbol{\theta}_i$ terms from the posterior distribution, and samples just the latent assignment variables $z_{i,t}$ described above. For a more technical discussion, see Heinrich (2009) or the appendix of Hansen et al. (2014).

[15] We run five chains beginning from five different seeds, and select the one for analysis that has the

1. Randomly allocate to each time block a style drawn uniformly from $\{1, \ldots, K\}$.
2. For each time block in sequence, draw a new style using multinomial sampling. The probability that block t for CEO i is assigned to style k is increasing in:
   (a) The number of other blocks for CEO i that are currently assigned to k.
   (b) The number of other occurrences of the feature combination $y_{i,t}$ in the entire dataset that is currently assigned to k.
3. Repeat step 2 5,000 times as a burn in phase.
4. Repeat step 2 5,000 more times, and store every 50th sample.

Steps 2a and 2b mean that feature combinations that regularly co-occur in CEOs' time use will be grouped together to form behaviors. Also, step 2a means that feature combinations within individual CEOs will tend to be concentrated rather than spread across behaviors.

Many combinations are rare: there are 183 combinations that appear in just one time block, and 430 that appear in two. Since inference in LDA relies on co-occurrence, the assignment of such rare combinations to behaviors is noisy. For this reason, we drop any combination that is not present in at least 30 CEOs' time use. This leaves 654 combinations and 98,347 time blocks in the baseline analysis. Tables A.1 in appendix show average CEO time shares across features on this subsample, which are very similar to those of the whole sample reported in table.[16]

For each draw in step 4, the estimate $\widehat{\theta}_i^k$ is proportional to the total number of time units of CEO i allocated to behavior k plus the prior α, and the estimate $\widehat{\beta}_k^f$ is proportional to the total number of times x_f is allocated to behavior k plus the prior η. We then average these estimates across all draws, to form the final objects we analyze in the paper.

To make the inference procedure more concrete, consider a simplified dataset with three CEOs and an activity feature set $X = \{\text{unplanned,planned}\} \times \{\text{size1, size2+}\}$. Table 3 tabulates the number of time blocks of each CEO according to their value of x_f and their allocation across two behaviors—which we denote B0 and B1—at different points in a Markov chain. The row sums within each value of $x_f \in X$ represent the total number of time blocks of a CEO associated to x_f. CEO A's time is dominated by planned activities with two or more people (162 out of 168 time blocks have x_f = size2+planned); CEO B's time is dominated by unplanned activities; while CEO C has a broader distribution of time use across feature combinations.

best goodness-of-fit across the draws we take after burn in.

[16] For robustness, we have also kept combinations present in 15 and, alternatively, 45 CEOs' time use, and find very similar results (see Table A4)

Table 3: Example of MCMC Estimation of Allocation of Time Blocks to Behaviors

(a) Random Seed

	size1unplanned		size1planned		size2+unplanned		size2+planned		
CEO	B0	B1	B0	B1	B0	B1	B0	B1	$\widehat{\theta}_i$
A	0	0	1	3	0	2	82	80	0.506
B	9	4	1	0	5	4	12	19	0.5
C	35	43	0	0	38	30	0	0	0.5
	0.24	0.254	0.011	0.017	0.235	0.195	0.513	0.535	
	$\widehat{\beta}_0^1$	$\widehat{\beta}_1^1$	$\widehat{\beta}_0^2$	$\widehat{\beta}_1^2$	$\widehat{\beta}_0^3$	$\widehat{\beta}_1^3$	$\widehat{\beta}_0^4$	$\widehat{\beta}_1^4$	

(b) Iteration 2

	size1unplanned		size1planned		size2+unplanned		size2+planned		
CEO	B0	B1	B0	B1	B0	B1	B0	B1	$\widehat{\theta}_i$
A	0	0	4	0	2	0	35	127	0.753
B	10	3	1	0	5	4	4	27	0.625
C	73	5	0	0	63	5	0	0	0.074
	0.421	0.047	0.026	0.001	0.355	0.053	0.198	0.899	
	$\widehat{\beta}_0^1$	$\widehat{\beta}_1^1$	$\widehat{\beta}_0^2$	$\widehat{\beta}_1^2$	$\widehat{\beta}_0^3$	$\widehat{\beta}_1^3$	$\widehat{\beta}_0^4$	$\widehat{\beta}_1^4$	

(c) Iteration 5

	size1unplanned		size1planned		size2+unplanned		size2+planned		
CEO	B0	B1	B0	B1	B0	B1	B0	B1	$\widehat{\theta}_i$
A	0	0	0	4	2	0	0	162	0.982
B	13	0	0	1	9	0	0	31	0.589
C	78	0	0	0	68	0	0	0	0.007
	0.535	0.001	0.001	0.026	0.464	0.001	0.001	0.973	
	$\widehat{\beta}_0^1$	$\widehat{\beta}_1^1$	$\widehat{\beta}_0^2$	$\widehat{\beta}_1^2$	$\widehat{\beta}_0^3$	$\widehat{\beta}_1^3$	$\widehat{\beta}_0^4$	$\widehat{\beta}_1^4$	

Notes: This table shows the allocation of three CEO's time use to behaviors at different points in an example Markov chain. The algorithm samples each unit of time into one of two behaviors, from which we derive estimates of the behavioral index $\widehat{\theta}_i$ and behaviors $\widehat{\boldsymbol{\beta}}_0$ and $\widehat{\boldsymbol{\beta}}_1$. In this simple example, the chain converges within a few iterations.

Table 3 represents the random seed from which sampling begins. Since behavior assignments are drawn uniformly, each CEO's time is split roughly evenly between behaviors. The last column shows the behavioral indices derived from these assignments, which is around 0.5 for all CEOs. The last row shows the estimated probability that each x_f appears in each behavior, which begins around the empirical frequency of x_f in the overall sample.

As sampling proceeds from the random seed, time units are re-allocated between behaviors. size1unplanned and size2+unplanned activities begin to be pulled into B0, while size1planned and size2+planned activities are pulled into B1. As this happens, A's behavioral index moves towards one, C's moves towards zero, and B's remains around 0.5. This shows the importance of allowing CEOs to mix behaviors, as forcing B into one of the two behaviors would not capture the full heterogeneity of his or her time use.

In such a small dataset, the chain converges quickly and by the fifth iteration stabilizes. The only time units whose assignments vary substantially in further sampling are the two that CEO A spends in size2+unplanned activities. This combination is both strongly associated with B0—which favors sampling its value to 0—and present in a CEO's time use that is strongly associated to B1—which favors sampling its value to 1. Averaging over numerous draws accounts for this uncertainty.

2.5 Estimation Results

2.5.1 Behaviors

The first two objects of interest are the behaviors $\boldsymbol{\beta}_0$ and $\boldsymbol{\beta}_1$. A first question is the extent to which the algorithm identifies behavioral differences in the data. To answer it, we construct Figure 1. First, we reorder the elements of the activity feature set according to their probability in $\widehat{\boldsymbol{\beta}}_0$. Second, we plot the estimated probabilities of each element of X in both behaviors. There is a clear overall pattern in which the combinations most associated with behavior 0 have low probability in behavior 1 and vice versa. In other words, behaviors are indeed sharply characterized.

Since the elements of X are combinations of features, interpreting the raw estimated probabilities associated to behaviors is rather difficult. Instead we compute marginal distributions over separate, individual features. For example, from the 654 elements of $\widehat{\boldsymbol{\beta}}_0$ and $\widehat{\boldsymbol{\beta}}_1$ one can compute a two-element marginal distribution over the "planned" feature in each behavior. Figure 2 displays the ratios of all the marginal distributions that we compute.[17] A value of 1 for the ratio indicates that both behaviors placed the same probability on the feature category; a value greater than (less than) 1 indicates a

[17]We only report feature categories for which at least one of the two estimated behaviors has more than 0.05 probability in its marginal distribution.

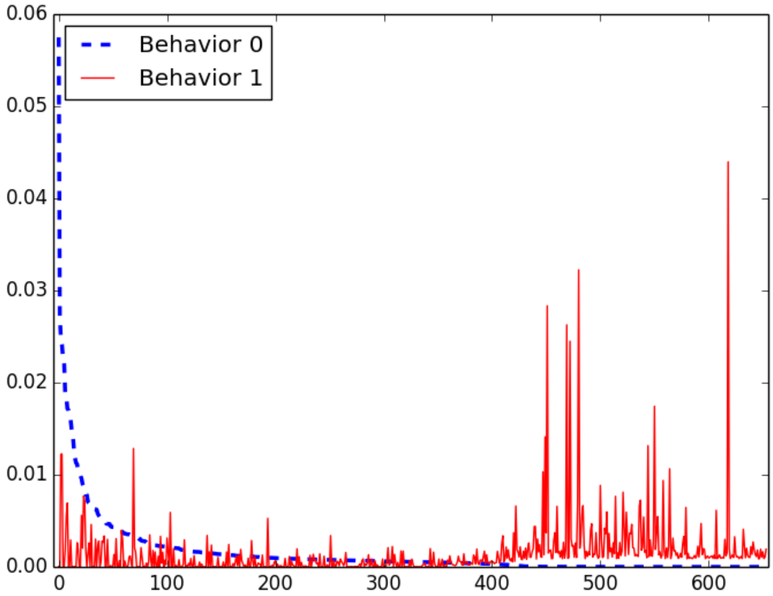


Figure 1: Probability of Feature Combinations

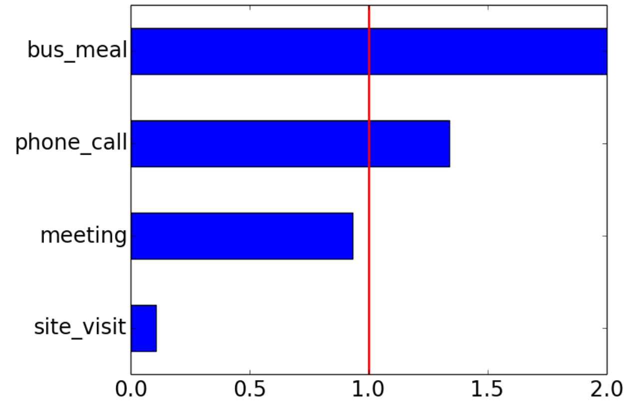
Notes: This plots the probability of different elements of the activity feature set in behaviors 0 and 1. The 654 elements of X are ordered left to right according to their probability in behavior 0.

higher (lower) probability for behavior 1. Finally, where bars extend to the edges of the figure, we have truncated the ratio for visual coherence.

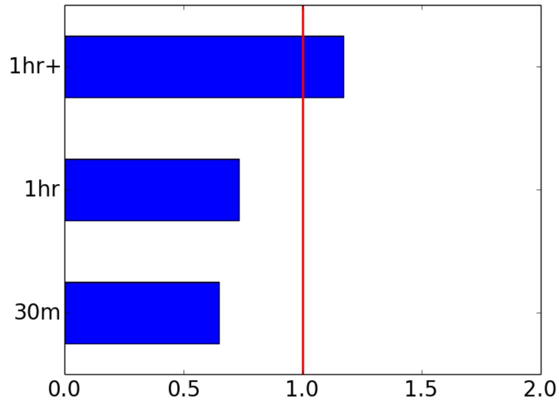
For activity types the most prominent distinction is site visits, which is ten times more likely in behavior 0. Another notable difference is for business meals, which behavior 1 is over twice as likely to generate. Less prominent differences exist for phone calls, which are 34% more likely in behavior 1 and meetings, which are 7% more likely in behavior 0.

For meeting duration, behavior 0 is clearly more associated with shorter activities, with 30-minute durations 54% more likely and 1-hr durations 36% more likely. In contrast, behavior 1 is 17% more likely to generate activities that last more than one hour. Behavior 1 is also more likely to engage in planned activities (17% more likely); activities with two or more participants (14% more likely); and especially activities with two or more functions (50% more likely).

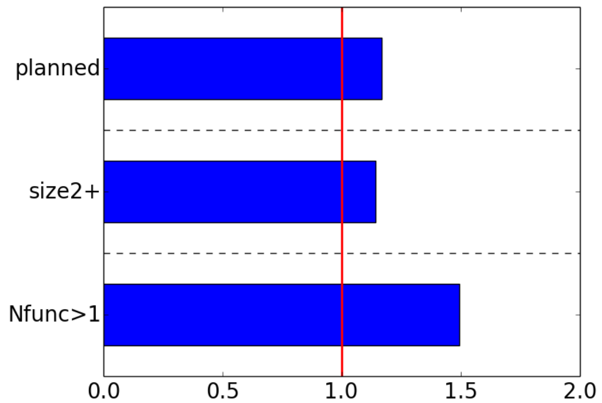
The remaining differences we explore are time spent with functions. While both behaviors spend time in activities with only inside functions in equal amounts, behavior 1 is twice as likely to spend time with both inside and outside functions together, and behavior 0 is twice as likely to spend time with only outside functions. Very stark differences emerge in time spent with specific inside functions. Behavior 1 is over ten times



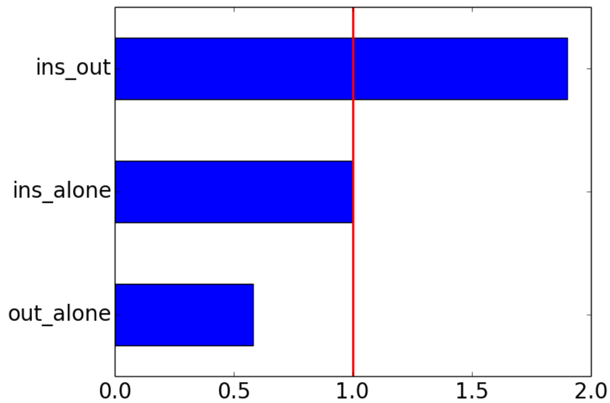
(a) Activity Type



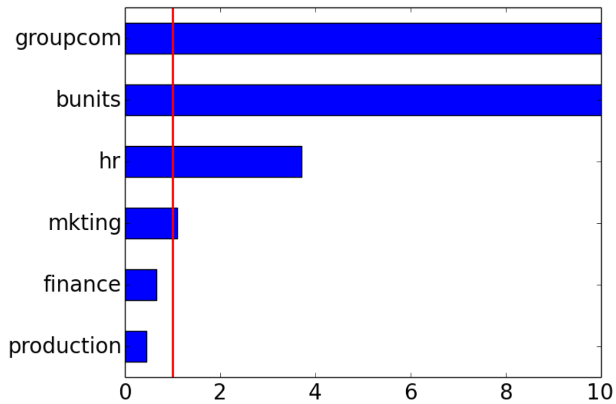
(b) Duration



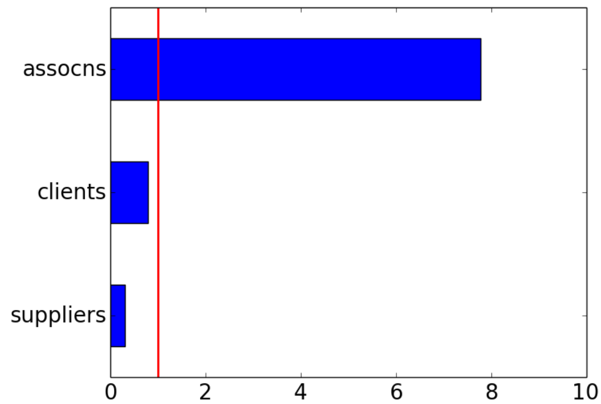
(c) Planning; Size; Number of Functions



(d) Inside vs Outside Time



(e) Inside Functions



(f) Outside Functions

Figure 2: Ratios of Marginal Distributions (Behav1/Behav0)

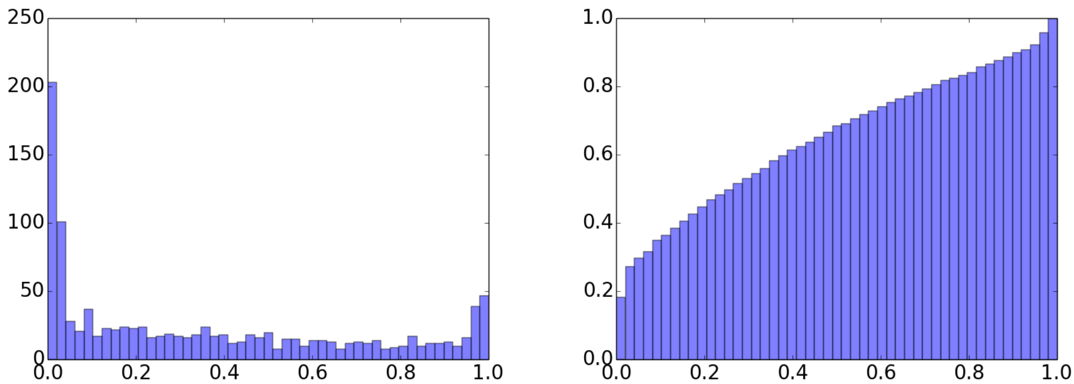
Notes: We generate these figures in two steps. First, we create marginal distributions for each behavior along several dimensions. Then, for each category that has more than 5 per cent probability in either behavior, we report the probability of the category in behavior 1 over the probability in behavior 0. The third panel represents three separate marginal distributions. Each has two categories, so we report the ratio for only one.

as likely to spend time in activities with commercial-group and business-unit functions, and nearly four times as likely to spend time with the human-resource function. On the other hand, behavior 0 is over twice as likely to engage in activities with production. Smaller differences exist for finance (50% more likely in behavior 0) and marketing (10% more likely in behavior 1) functions. In terms of outside functions, behavior 0 is over three times as likely to spend time with suppliers and 25% more likely to spend time with clients, while behavior 1 is almost eight times more likely to attend trade associations.

In summary, an overall pattern arises in which behavior 0 engages in short, small, production-oriented activities and behavior 1 engages in long, planned activities that combine numerous functions, especially high-level insiders.

2.5.2 The CEO Behavior Index

The two behaviors we estimate represent extremes. As discussed above, individual CEOs generate time use according to the behavioral index θ_i that gives the probability that any specific time block's feature combination is drawn from behavior 1. Figure 3 plots both the frequency and cumulative distributions of θ_i in our sample.



(a) Frequency Distribution **(b)** Cumulative Distribution

Figure 3: CEO Behavior Index Distributions

Notes: The left-hand side plot displays the number of CEOs with behavioral indices in each of 50 bins that divide the space $[0, 1]$ evenly. The right-hand side plot displays the cumulative percentage of CEOs with behavioral indices lying in these bins.

Many CEOs are estimated to be mainly associated with one behavior: 316 have a behavioral index less than 0.05 and 94 have an index greater than 0.95. As Figure 3 shows, though, away from these extremes the distribution of the index is essentially uniform, and the bulk of CEOs draw their time use from both behaviors. This again highlights the value of using a mixed-membership model that allows CEOs to be associated with both estimated behaviors. Finally, we calculate the estimated time shares the average CEO spends within different categories for each feature displayed in Table 2 from the marginal distributions computed in Figure 2 and the estimated behavioral indices displayed in Figure 1. Table A.2 in appendix contains the results, which track very closely the actual time shares computed on the subsample used in estimation contained in Table 2. This provides assurance that the differences between behaviors that LDA uncovers are consistent with the raw time-use data.

2.5.3 Correlations with Firm Characteristics

Since the CEO behavior index is estimated using solely time use information over a single week of activity, a possible concern is that the variation we observe in the data may entirely be driven by high-frequency noise. To address this concern, we provide evidence on the match between CEO behavior and firm characteristics by estimating basic conditional correlations between CEO behavior and firm characteristics, that is:

$$\theta_{ifs} = \alpha + \beta x_f + \gamma a_s + \mathbf{Z}_i \delta + \varepsilon_{ifs} \tag{2}$$

where θ_{ifs} is the behavior index of CEO i in firm f in sector s, $\mathbf{x}_f$ denotes firm characteristics, a_s denotes industry features, and $\mathbf{Z}_i$ is a vector of CEO characteristics. We include country dummies throughout the analysis and cluster the standard errors at the industry level to account for errors correlation within industry, due to the fact that all firms in the same industry might have common needs for a given CEO behavior. We also include in all regressions a set of controls to take into account for factors that may have affected the informational content of the time use data collected across CEOs.[18] Since the behavior index is meant to represent "typical" CEO behavior, regardless of the specific week in which the data was collected, all regressions in this table and throughout the analysis are weighted by a score (ranging between 1 and 10) attributed by the CEO to the survey week to denote its level of representativeness.

The results of this analysis are shown in Table 4. Recall that low values of the behavior index are associated with direct involvement in production and one-on-one unplanned meetings, while high value of the index are associated with large, planned meetings with high level executives. We find that the behavior index is positively associated with firm size, as proxied by the log of the number of total employees (column 1), thus

[18]The noise controls included throughout the analysis are: a dummy to denote whether the data was collected through the PA (rather than the CEO himself), a reliability score attributed by the analyst at the end of the week of data collection, a set of dummies to denote the specific week in which the data was collected and a dummy to denote whether the CEO formally reported to another manager (this was the case in 6% of the sample).

Table 4: CEO-Firm Match

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable			CEO behavior index			
log(employment)	0.055***	0.053***	0.055***	0.056***	0.051***	0.053***
	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)
COO Dummy		0.066***	0.062***	0.062***	0.064***	0.057**
		(0.021)	(0.021)	(0.021)	(0.021)	(0.025)
task abstraction (industry)			0.032**	0.035**	0.028**	
			(0.012)	(0.013)	(0.012)	
capital intensity (industry)				-0.006		
				(0.018)		
homogeneous product (industry)				-0.031		
				(0.030)		
log(CEO tenure)					-0.025***	-0.022**
					(0.009)	(0.010)
CEO has an MBA					0.053**	0.065**
					(0.022)	(0.026)
Adjusted R-squared	0.242	0.248	0.252	0.251	0.262	0.264
Observations	1114	1114	1114	1114	1114	1114
Controls:						
Country	y	y	y	y	y	y
Noise	y	y	y	y	y	y
Industry						y

Notes: *** (**) (*) denotes significance at the 1%, 5% and 10% level, respectively.The COO dummy takes value one if the firm has an officer with the COO title, "task abstraction" is an industry metric drawn from Autor et al. (2003), with higher values denoting a higher intensity of asbtract tasks in production. "Capital intensity" denotes the average industry level value of capital over labour, built from the NBER manufacturing database (aggregated between 2000 and 2010). "Homogeneous product" is an industry dummy drawn from Rauch (1999). "Log CEO tenure" is the log of 1+number of years CEO is in office, "CEO has an MBA"is a dummy taking value one is the CEO has attained an MBA degree or equivalent postgraduate qualification. Noise controls are a full set of dummies to denote the week in the year in which the data was collected, a reliability score assigned by the interviewer at the end of the survey week and a dummy taking value one if the data was collected through the PA of the CEO, rather than the CEO himself. Country dummies are included in all columns. Industry controls are 2 digit SIC dummies. All columns weighted by the week representativeness score assigned by the CEO at the end of the interview week. Errors clustered at the 2 digit SIC level.

suggesting a greater demand for structured and multilateral interactions in larger firms. Conditional on firm size, the index is also higher in firms in which a COO position exists (column 2), suggesting that engagement in coordinating activities (or conversely, the lack of engagement in production tasks) is complementary with the ability to delegate operational tasks to other executives.

In column (3) we investigate the extent to which the CEO index is correlated with industry characteristics, conditional on firm size and organizational structure. We focus in particular on the "task abstraction" index, an industry metric developed by Autor et al. (2003) which captures the intensity in production of "abstract analytical and managerial tasks, which may require creativity, hypothesis formation, problem solving or persuasion" (Autor 2013).[19] We use this metric to evaluate whether the time spent by CEOs in structured and multilateral activities (as captured by higher values of the CEO index) systematically correlates with an industry level metric of complexity in production. We find that a standard deviation change in the task abstraction index is correlated with a 0.023 change in the CEO behavior index (significant at the 5%). Column (4) shows that this result is robust to controlling for other industry characteristics, such as capital intensity (derived from the NBER manufacturing database) and degree of product differentiation (using the Rauch 1999 classification). We also experimented with other industry metrics capturing more indirectly the intensity of complex tasks in production, namely industry level measures of intensity of R&D spending and intensity of human capital in production with very similar results (see Table A3 for details). Overall, these results show that industry characteristics predict some of the observed heterogeneity in CEO behavior.

In column (5) we investigate the extent to which the match between CEO behavior and firm and industry characteristics examined above are accounted for by observable CEO characteristics. We find that CEOs with formal managerial training (as captured by a dummy denoting whether they have an MBA degree of equivalent postgraduate qualification) are associated with higher values of the CEO index, while the index tends to be significantly lower in CEOs with longer tenure (as measured by the log of (1+ years as CEO)). However, the inclusion of these basic CEO characteristics hardly changes the magnitude and significance of the correlations with firm size, organizational structure and task abstraction.[20] We conclude this analysis by showing the robustness of the results to

[19]This variable and the other industry controls described in this section are only available for the US. In the analysis we project the data on all other countries, under the assumption that similar sectors would display similar characteristics across the other countries in our sample. This is clearly a strong assumption, which is likely to introduce measurement error in the estimates. To the extent that we find a relationship between the CEO behavior index and these industry measures, they should thus be considered a lower bound of the magnitude of the true relationship.

[20]This statement is true also when we consider an extended vector of CEO characteristics, as shown in Table AX.

the inclusion of industry (SIC2) fixed effects (which are collinear with the task abstraction index). The significance and magnitude of the coefficients is unchanged when we include industry dummies.

The fact that our behavior index correlates with firm and industry characteristics in a somewhat predictable way (e.g. the time spent in coordinative activities increases with firm size and complexity) reassures us about the quality of the coarse classification implemented by the machine learning approach. Furthermore, this finding is also indicative of a certain degree of matching between CEOs and firms. What is still unclear, however, is the extent to which the match between CEOs and firms is indeed optimal, or whether efficiency gains could be achieved by reallocating CEO behaviors across firms. In the next section we present a simple matching model to guide the empirical analysis of this specific question.

3 CEO Behavior and Firm Performance: Theory

In this section we develop a model of that specifies the conditions under which the cross-sectional analysis of CEOs behavior and firm performance may reveal the presence of frictions in the matching of firms and CEOs. This minimalistic CEO-firm matching model is based on two assumptions. First, both CEOs and firms have "types" . The type of a firm determines which CEO behavior makes it most productive and the type of the CEO determines how willing or able she is to adopt a certain behavior. Moreover, one type of CEO may be relatively more abundant than the other type, in the sense that its share is higher than the share of firms who seek CEOs of that type. Second, there are frictions in the market for CEOs. On the hiring side, the firm's screening technology is imperfect and it cannot always correctly identify the CEO's type. On the dismissal side, firing a CEO may be a lengthy process. This second assumption makes our story different from existing theories of manager-firm matching, where the matching process is frictionless and the resulting allocation of managerial talent achieves productive efficiency (Gabaix and Landier (2008), Tervio (2008), Bandiera et al. (2015)).

The main result of the model is that there may be mismatch in equilibrium. Some prospective CEOs who belong to the more abundant type will "pass" as CEOs of the scarce type. After they are hired, they will behave in a way that is suboptimal for their firms. The firms they run will have lower productivity. Because of this mismatch, abundant-type CEOs will underperform on average scarce-type CEOs. Moreover, as the mismatch is based on screening errors, once one conditions on CEO type, observable attributes of CEOs have no predictive value on firm performance.

The model has a natural dynamic extension where the effect of CEO behavior on is gradual. It takes time for a newly hired CEO to affect the performance of the firm. This leads to predictions on the shape of the performance residual of firms as a function of CEO tenure and CEO type, which will guide our empirical analysis.

3.1 Static Version

There are two possible CEO behaviors: $x = 0$ and $x = 1$. Once a CEO is hired, he decides how he is going to manage the firm that hired him. CEOs come in two types. Type 0 prefers behavior 0 to behavior 1. Namely, he incurs a cost of 0 if he selects behavior 0 and cost of c, which we normalize to one, if he selects behavior 1. Type 1 is the converse: he incurs a cost of 0 if he selects behavior 1 and cost of c if he selects behavior 0. The cost of choosing a certain behavior can be interpreted as coming from the preferences of the CEO (he finds one behavior more enjoyable)) or his skill set (he finds one behavior less costly to implement).

Firms too have types. A type-0 firm is more productive if the CEO chooses $x = 0$. Namely, the firm's output is $R = 1$ if the CEO chooses $x = 0$ and $R = 0$ if the CEO chooses behavior 1. A type-1 firm is the converse.

All firms offer the same linear compensation scheme

$$w(R) = \bar{w} + \beta R,$$

where $\bar{w}$ is a is a fixed part, and $\beta \geq 0$ is a parameter that can be interpreted directly as the performance-related part of CEO compensation or indirectly as how likely it is that a CEO is retained as a function of his performance (in this interpretation the CEO receives a fixed per-period wage but he is more likely to be terminated early if firm performance is low).[21]

The total utility of the CEO is equal to compensation less behavior cost. After a CEO is hired, she chooses her behavior. If the CEO is hired by a firm with the same type, she will obviously choose the behavior that is preferred by both parties. The interesting case is when the CEO type and the firm type differ. If $\beta > 1$, the CEO will adapt to the firm's desired behavior, produce an output of 1, and receive a total payoff of $\bar{w} + \beta - 1$. If instead $\beta < 1$, the CEO will choose her preferred behavior, produce output $R = 0$ and receive a payoff $\bar{w}$. We think of β as a measure of governance. A higher β makes CEO behavior more aligned with the firm's interests.

Now that we know what happens once a match is formed, let us turn our attention to the matching process. There are a mass 1 of firms. A proportion ϕ of them are of type 1, the remainder are of type 0. The pool of potential CEOs is larger than the pool of firms seeking a CEO. There is a mass $m >> 1$ of potential CEOs. Without loss of generality,

[21] We assume that CEO compensation is not directly dependent on CEO behavior or CEO type. If it were, we would be in a frictionless environment.

assume that a proportion $\gamma \leq \phi$ of CEOs are of type 1. The remainder are of type 0. From now on, we refer to type 1 as the *scarce* CEO type and type 0 as the *abundant* CEO type. We emphasize that scarcity is relative to the share of firm types. So, it may be the case that the scarce type is actually more numerous than the abundant type.

The market for CEOs works as follows. In the beginning, every prospective CEO sends his application to a centralized CEO job market. The applicant indicates whether he wishes to work for a firm of type 0 or a firm of type 1. All the applications are in a large pool. Each firm begins by downloading an application meant for its type. Each download costs k to the firm.[22] If the application is of the wrong type, deception is detected with probability $\rho \in [0, 1]$, where $\rho = 1$ denotes perfect screening and $\rho = 0$ represents no screening.[23]

Potential CEOs maximize their expected payoff, which is equal to the probability they are hired times the payoff if they are hired. Firms maximize their profit less the screening cost (given by the number of downloaded application multiplied by k).[24]

We can show:

Proposition 1 *Assume that the screening process is sufficiently unreliable, governance is sufficiently poor, and one CEO type is sufficiently abundant.*[25] *Then, in equilibrium:*

- *All scarce-type CEOs are correctly matched;*
- *Some abundant-type CEOs are mismatched;*
- *The average productivity of firms run by abundant-type CEOs is lower than that of firms run by scarce-type CEOs.*

Proof. We verify that the situation described in the proposition corresponds to a Bayesian equilibrium. First note, that if $\beta > 1$, all CEOs will choose the behavior that is optimal for the firm that hires them. This means that CEO behavior only depends on firm type. Therefore, in what follows we assume that governance is sufficiently poor, so $\beta < 1$.

In that case, when a CEO is hired, her utility is $\bar{w} + \beta$ if she works for a firm of the same type and $\bar{w}$ if she works for a firm of a different type. To simplify notation,

[22] We can allow firms to mis-represent their type. In equilibrium, they will report their type truthfully.

[23] We assume that would-be-CEOs know their own type before they apply to firms. It is easy to see that our mismatch result would hold a fortiori if prospective applicants had limited or no knowledge of their own type.

[24] We assume that k is sufficiently low that a firm would not hire the first applicant independently of her type.

[25] Formally, this is given by the conditions: $\beta < 1$ and

$$\rho < \frac{\phi - \gamma}{\phi - \gamma\phi}.$$

normalize $\bar{w}+\beta$. Hence the utility of a correctly matched CEO is one and the utility of a mismatched CEO is

$$b \equiv \frac{\bar{w}}{\bar{w}+\beta}.$$

Note that b is a measure of the quality of governance, with $b=1$, being the worst level of governance.

A type-0 firm faces an abundant supply of type-0 CEOs. As all the applications it receives come from type-0 CEOs, the firm will simply hire the first applicant. A type-1 firm instead may receive applications from both CEO types. If c is sufficiently low, the optimal policy consists in waiting for the first candidate with $s=1$ and hire him.

We now consider CEOs. Suppose that all type-1 CEOs apply to type-1 firms and type-0 CEOs apply to type-1 firms with probability z and to type-0 firms with probability $1-z$.

If a type-0 CEO applies to a type-0 firm, he will get a job if and only if his application is downloaded. The mass of type-0 firms is $1-\phi$. The mass of type-0 CEOs applying to type-0 firms is $(1-\gamma)(1-z)\,m$. The probability the CEO is hired is

$$P_0 = \frac{1-\phi}{(1-\gamma)(1-z)\,m}.$$

If instead a type-0 CEO applies to a type-1 firm, he will get a job if and only if his application is considered and the firm does not detect deception. Computing the first probability requires an additional step, because some firms consider more than one application before they find an application which passes the screening process.

The probability that a type-1 firm application is accepted if it is considered is:

$$H = \frac{(1-\gamma)\,z\,(1-\rho)+\gamma}{(1-\gamma)\,z+\gamma}.$$

The mass of applications that are downloaded by type-1 firms is therefore:

$$\phi\left(1+(1-H)+(1-H)^2+...\right) = \phi\frac{1}{H}.$$

Given that the mass of applicants to type-1 firms is $m\,((1-\gamma)\,z+\gamma)$, the probability that an application is considered is

$$\frac{\phi}{m\,(\gamma+(1-\gamma)\,z)\,H} = \frac{\phi}{m\,((1-\gamma)\,z\,(1-\rho)+\gamma)}$$

The probability that a type-0 applicant passes the screening process is $1-\rho$. Thus, the probability that a type-0 applicant is hired by a type-1 firm is

$$P_1 = \frac{(1-\rho)\,\phi}{m\left((1-\gamma)\,z\,(1-\rho)+\gamma\right)}.$$

In the equilibrium under consideration a type-0 CEO must be indifferent between applying to the two types of firms. As the benefit of being hired by a same-type firm is one, while the benefit of being hired by a type-1 firm is b, the indifference condition is $P_0 = bP_1$, which yields:

$$\frac{1-\phi}{(1-\gamma)\,(1-z)} = \frac{(1-\rho)\,\phi b}{\left((1-\gamma)\,z\,(1-\rho)+\gamma\right)},$$

yielding

$$z = \frac{(1-\gamma)\,(1-\rho)\,\phi b-(1-\phi)\,\gamma}{(1-\phi+\phi b)\,(1-\gamma)\,(1-\rho)}.$$

The solution of z will be positive – meaning that some 0-types will apply to 1-firms – if

$$\rho < 1-\frac{(1-\phi)\,\gamma}{(1-\gamma)\,\phi b},$$

which is satisfied as long as ρ is not too high, b is not too low, and γ is sufficiently smaller than ϕ. For instance, the combination of $\rho = 0$, $b = 1$, and $\phi > \gamma$ would work.

Type-1 CEOs always produce 1, while the average productivity of a type-0 CEO is equal to the probability that he is matched with a type 0 firm, which is

$$\frac{1-z}{1-z+z\,(1-\rho)}.$$

By replacing z, we find the average productivity of a type-0 CEO:

$$\frac{(1-\phi)\left((1-\gamma)\,(1-\rho)+\gamma\right)}{(1-\phi)\,(1-\gamma)\,(1-\rho)+(1-\phi)\,\gamma+\left((1-\gamma)\,(1-\rho)\,\phi b-(1-\phi)\,\gamma\right)(1-\rho)},$$

which is smaller than one whenever $\rho < 1$. ■

The intuition for this result is as follows. If all abundant-type CEOs applied to their firm type, they would have a low probability of being hired and they would prefer to apply to the other firm type and try to pass as scarce-type CEO. In order for this to be true, it must be that the share of abundant types is sufficiently larger than the share of scarce types and that the risk that they are screened out is not too large. If this is the case, then in equilibrium some abundant-type CEOs will apply to the wrong firm type up to the point where the chance of getting a job is equalized under the two strategies. The application strategy of CEOs means that all scarce-type CEOs are matched to the right firm and will produce high performance, while some abundant-type CEOs are matched to the wrong firm and will produce low performance. The average performance of abundant-type CEOs is therefore lower.

Note that three conditions are required for the Proposition to hold: imperfect screening, imperfect governance, and a degree of unbalance between firm types and CEO types. If at least one of the three conditions fail, then in equilibrium we should observe no correlation between CEO type and firm performance. Let us review all three possible violations. If there is no scarce CEO type ($\gamma = \phi$), a CEO has no reason to apply to a firm of a different type. If screening is perfect ($\rho = 1$), a CEO who applies to a firm of the other time is always caught (and hence he won't do it). If governance is good ($\beta > 1$), a CEO who is hired by a firm of the other type will always behave in the firm's ideal way (and hence there will either be no detectable effect or CEOs will only applied to firms of their type).

Some remarks are in order. First, under Proposition 1, the economy under consideration does not achieve productive efficiency. As the overall pool of scarce-type CEOs is assumed to be sufficient to cover all firms that prefer that CEO type ($m >> 1$), it would be possible to give all firms their preferred type and thus increase overall production.[26]

Second, one can consider the extreme case where there are no 0-type firms: $\phi = 1$. This is no longer a matching problem. No firm wants the abundant-type CEO. Everybody applied to type-1 firms.

Corollary 1 *In the extreme case where there are no 0-type firms, all abundant-type CEOs apply to the wrong firm type. A share $\gamma(1-\rho)$ of firms is run by the wrong type of CEO. All employed abundant-type CEO underperform.*

Third, one can tweak the model by assuming that some CEOs have observable attributes that make them more or less likely to be one type of CEO. For instance, assume that the share of CEOs with an MBA degree is μ_0 in the abundant type and μ_1 in the scarce type, with $\mu_1 > \mu_0$.

If type-1 firms used the presence of an MBA degree to screen applicants, then only abundant-type CEOs with an MBA will apply, but that would make having an MBA a "negative" signal. In equilibrium it must be that the abundant-type CEOs who apply to type-1 firms have the same share of MBA degrees as scarce-type CEOs.

Corollary 2 *If CEOs have observable attributes that are correlated to their type, in equilibrium firms do not use those attributes to screen out abundant-type CEOs.*

[26]If side transfers were feasible, this would also be a Pareto-improvement as a type-1 CEO matched with a type-0 firm generates a higher bilateral surplus than a type-0 CEO matched with a type-1 firm, and the new firm-CEO pair could therefore compensate the now unemployed type-0 CEO for her job loss.

Note that the corollary does not imply that the presence of visible CEO attributes is inconsequential. Type signals may make it harder for abundant-type applicants to pretend to be scarce-type applicants, which in turn reduces CEO type mismatch and improves firm productivity.[27]

3.2 Dynamic Version

We now explore the dynamic implication of our CEO-firm matching model. In particular, we ask a question that will be useful for the empirical analysis when we will address the role of time invariant firm characteristics is the selection of CEOs. Suppose that we know the behavior of the current CEO, but not the type of the firm and the behavior of the previous CEO. What can we say about the evolution of firm performance over time?

The starting premise is that the influence of the behavior of the CEO on the performance of her firm is not immediate. As in the model of Halac and Prat (2014), it takes time for a corporate leader to change the existing management practice and to affect the company's culture.[28]

Let us assume that the conditions for Proposition 1 are satisfied. There are two types of CEOs ($c \in \{0, 1\}$) and two types of firms ($f \in \{0, 1\}$). We assume that the abundant CEO type is $c = 0$. The performance of a firm is $w_f + x_{fc}$, where $x_{fc} = 1$ if the firm type and the CEO type match ($f = c$) and $x_{fc} = 0$ if there is a mismatch ($f \neq c$)x_{fc}, and the term w_f indicate that the two firm types may have different baseline productivities.

Let us consider a firm whose CEO is replaced at time 0. Let x_{fc}^{old} and x_{fc}^{new} denote the match quality of the previous CEO and the current CEO, respectively. The performance of the firm at time $t < 0$ was determined uniquely by the performance of the old CEO (thus assuming that he had been in the job sufficiently long). The performance at $t \geq 0$ is given by

$$y_t = w_f + (1 - \alpha_t)\, x_{fc}^{\text{old}} + \alpha_t x_{fc}^{\text{new}},$$

where α_t is increasing and s-shaped in t. Namely,$\alpha_0 = 0$,$\alpha_t' > 0$, $\lim_{t\to 0+} \alpha_t' = 0$, $\lim_{t\to\infty} \alpha_t = 1$, and $\alpha"_t > 0$ if t is low and $\alpha"_t < 0$ if t is high. As time passes, the company's performance is determined more and more by the type of the new CEO as his tenure increases. The s-shaped assumption captures the idea that the effect of a new CEO is limited in the beginning, it increases with time, but then it reaches a stable plateau.

Consider a large sample of firms. Suppose we observe the type of the current CEO, but we do not observe the type of the previous CEO, nor the type of the firm. What can we say about them?

[27] If the number of MBAs increases so much that it eliminates the incentive for abundant-type CEOs to apply to type-1 firms, then the Corollary is no longer applicable.

[28] Bloom et al. (2016) estimate adjustment costs in managerial capital of similar magnitude to the ones estimated for physical capital.

If the current CEO belongs to the scarce type, we know for sure that the firm has type-1. The previous CEO was the scarce type too with probability π and the abundant-type with probability $1-\pi$.[29]

Focus on performance growth, taking $t = 0$ as the baseline year: $\Delta y_t = y_t - y_0$. If the current CEO belongs to the scarce type, we have

$$\Delta y_t \left(c^{\text{new}} = 1\right) = \begin{cases} 0 & \text{if } t < 0 \\ \left(\left(1-\alpha_t\right) E\left[x_{fc}^{\text{old}} | x_{fc}^{\text{new}} = 1\right] + \alpha_t\right) - E\left[x_{fc}^{\text{old}} | x_{fc}^{\text{new}} = 1\right] & \text{if } t > 0 \end{cases}$$

but note that $E\left[x_{fc}^{\text{old}} | x_{fc}^{\text{new}} = 1\right] = \pi < 1$. Therefore,

$$\Delta y_t \left(c^{\text{new}} = 1\right) = \begin{cases} 0 & \text{if } t < 0 \\ \alpha_t \left(1-\pi\right) & \text{if } t > 0 \end{cases},$$

which implies that average performance growth in a sample of firms run by scarce-type CEOs was flat before the new CEO was hired and becomes increasing and s-shaped thereafter.

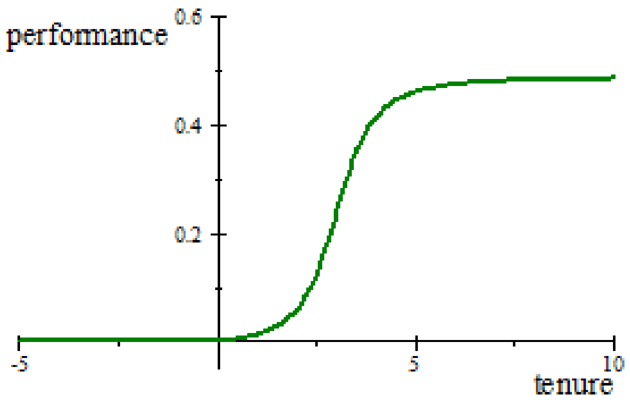


Figure 4: Average performance of a set of firms managed by scarce-type CEOs by years of CEO tenure.

Figure 4 depicts $\Delta y_t \left(c^{\text{new}} = 1\right)$ under the assumption that α_t is a sigmoid function ($\alpha_t = t/\sqrt{1+t^2}$) and $\pi = \frac{1}{2}$. The average effect of having a scarce-type CEO is positive, gradual, and s-shaped. This result implies that if we observe a set of firms run by scarce type CEOs who were all hired at the same date, we should predict that the average

[29]This probability is given in equilibrium by

$$\pi = \frac{\gamma}{\gamma + (1-\gamma) z},$$

where

$$z = \frac{(1-\gamma)(1-\rho)\phi - \gamma(1-\phi)}{(1-\gamma)(1-\rho)}.$$

performance of those firms is constant before the CEOs are hired and increasing and s-shaped afterwards.

If instead we consider a sample of firms run by abundant-type CEOs, a specular argument applies: we would observe that the average performance decreases after the current CEO is hired and follows a similarly s-shaped curve.

4 CEO Behavior and Firm Performance: Evidence

Guided by the model, we now combine our CEO behavior data with accounting data to test the null hypothesis of zero correlation between CEO behavior and firm performance.[30] Our main measure of performance is the value of sales (in constant 2010 dollars) controlling for log(employees) - hence, a measure of labor productivity - since this is available for the largest number of firms and countries. Conditional on data availability, we also test the relationship between the CEO index and sales controlling also for capital and materials (so closer to a TFP specification), profits per employee and Tobin's q.

We start by examining the correlation between the CEO behavior index and firm performance in the years in which the CEO is in office. To avoid overweighting CEOs based on their tenure, we use for all firms at most 5 of the most recent years pre-dating the survey year (2011 for India and 2013 for the rest of the countries) and average all inputs across this time period (so that we end up with one observation per firm). Our baseline specification is a production function of the form:

$$y_{ifts} = \alpha\theta_i + \beta^E e_{ft} + \beta^K k_{ft} + \beta^M m_{ft} + \mathbf{Z}_i\gamma + \zeta_t + \eta_s + \varepsilon_{ifts} \quad (3)$$

where y_{ifts} is the performance of firm f, led by CEO i, in year t and sector s, θ_i is the behavior index of CEO i, e_{ft}, k_{ft}, and m_{ft} denote, respectively, the natural logarithm of the number of firm employees and, when available, capital and materials. $\mathbf{Z}_c$ is a vector of CEO characteristics (MBA dummy and log(1+years as CEO)), ζ_t and η_s are year and SIC2 sector fixed effects, respectively. We include country by year dummies throughout, as well the set of noise controls described above. We cluster the standard errors at the industry level throughout the table and weight observations according to the self-reported week representativeness, as discussed above.[31]

[30] Data on firm performance was extracted from ORBIS. We were able to gather at least one year of sales and employment data in the period in which the sampled CEO was in office for 831 of the 1,114 firm with time use data. Of these: 29 did not report sales information at all; 128 were dropped in cleaning, 126 had data that referred only to years in which the CEO was not in office, or outside the 5 year window pre-dating the survey. The data covers the time period 2003-2013 (this is the maximum number of years of data which can be retrieved from Orbis). See the data Appendix for more details.

[31] Since the data is aggregated into a single average, year dummies are set as the average year for which the performance data is available. The results discussed in this section are robust to using multiple years and clustering the standard errors at the firm level instead of using averages.

Table 5: CEO Behavior and Firm Performance

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable		Log(sales)			Profits/Emp	Tobin's Q
CEO behavior index	0.369***	0.310**	0.326**	0.324**	10.945**	0.209
	(0.102)	(0.122)	(0.156)	(0.162)	(5.222)	(0.236)
log(employment)	0.908***	0.486***	0.325***	0.332***	-0.055	
	(0.038)	(0.063)	(0.105)	(0.106)	(0.112)	
log(capital)		0.435***	0.191***	0.193***		
		(0.036)	(0.054)	(0.052)		
log(materials)			0.449***	0.438***		
			(0.080)	(0.074)		
COO Dummy				0.178		
				(0.142)		
log(CEO tenure)				-0.120*		
				(0.070)		
CEO has an MBA				-0.045		
				(0.096)		
Adjusted R-squared	0.727	0.824	0.895	0.897	0.530	0.082
Number of firms	831	613	378	378	516	296
Underlying number of Observations	2554	1804	1190	1190	1552	1125
Sample	all	with k	with k & m	with k & m	all	listed
Controls:						
Industry	y	y	y	y	y	y
Country by year	y	y	y	y	y	y
Noise	y	y	y	y	y	y

Notes: *** (**) (*) denotes significance at the 1%, 5% and 10% level, respectively. We include at most 5 years of data for each firm and build a simple average across output and all inputs over this period. The sample in Column 1 includes all firms with at least one year with both sales and employment data. Columns 2, 3 and 4 restrict the sample to firms with additional data on capital (column 2) and capital and materials (columns 2 and 3). The sample in column 6 is restricted to listed firms. "Firm size" is the log of total employment in the firm, "Log CEO tenure" is the log of 1+number of years CEO is in office, "CEO has an MBA" is a dummy taking value one is the CEO has attained an MBA degree or equivalent postgraduate qualification. Noise controls are a full set of dummies to denote the week in the year in which the data was collected, a reliability score assigned by the interviewer at the end of the survey week and a dummy taking value one if the data was collected through the PA of the CEO, rather than the CEO himself. Country by year dummies are included in all columns. Industry controls are 2 digit SIC dummies. All columns weighted by the week representativeness score assigned by the CEO at the end of the interview week. Errors clustered at the 2 digit SIC level.

Column 1, Table 5 shows the estimates of Equation (3) controlling for firm size, country by year and industry fixed effects, and noise controls. The estimate of α is positive and precise at the 1% level (coefficient 0.369, standard error 0.107). Next, in column (2) we add as control the log of capital, which is positive and statistically significant. While including the capital variable restricts the sample to 613 firms, the coefficient on the CEO behavior index remains large and statistically significant (coefficient 0.310, standard error 0.115). In column (3) we examine an even smaller sample (378 firms) for which we have at least one year of capital and materials to look at a specification closer to TFP. Even in this case, while the other inputs are statistically significant and of expected magnitudes, their inclusion does not change substantially the magnitude and the significance of the CEO behavior index (coefficient 0.326, standard error 0.154).[32]

In column (4) we test whether the correlation simply proxies for other observable firm and CEO characteristics, rather than behavior per se. To do so, we add as controls the CEO variables examined in Table 4 (MBA dummy and log of CEO tenure), as well as a dummy to denote firms with a formal COO position. Including these variables hardly changes the magnitude of the CEO behavior index (coefficient 0.324, standard error 0.152), and the variables themselves are not significant at standard significance levels. This last finding is consistent with Corollary 2 of our model. If firms are using an observable CEO trait to select among candidates, then in equilibrium that trait cannot predict the probability that the CEO they hire is mismatched (if it did, it would mean the firm has not used that information optimally).

To assess the magnitude of the coefficient of the CEO behavior index, consider the results shown in Column (2), where we control for capital and employment. Given the coefficient of 0.310, a one standard deviation increase in the CEO behavior index is associated with a 0.10 log points higher log sales. This magnitude is about 2/3 of the effect of a one standard deviation change in management practices on firm performance (0.15, estimated in Bloom et al. 2016) and about 12% of the effect of a one standard deviation increase in capital (taking the coefficient of 0.435 times the in sample standard deviation of log capital of 1.88). Table A4 shows that the main productivity results are robust to alternative specifications and measurements of the CEO behavior index.

Columns (5) and (6) analyze the correlation between CEO behavior and two measures of firms profitability: profits per employee and Tobin's q. This allows us to assess whether CEOs capture all the extra rent they generate, or whether firms profit from being matched with the scarce type CEO. The results are consistent with the latter interpretation: the correlation between the CEO index and profits per employee is positive and precisely

[32] We also experimented with the same specification on the subsample of firms with capital and material data which are also listed on stock market to check whether the results could be driven by input mismeasurement in private firms. The coefficient on the CEO behavior index is of even larger magnitude (0.511) and significant at the 5% level (standard error 0.233) even in the smaller sample of 261 firms.

estimated, while the results on the Tobin's q are positive but below standard significance levels (probably because this variable is only available for 25% of the sample firms). The magnitudes are also large: a one standard deviation increase in the CEO behavior index is associated with an increase of $3,400 in profits per employee and with a 0.06 increase in Tobin's q.[33]

In light of the model, the results in Table 5 imply that frictions are sufficiently large to create some mismatches between firms and CEOs, and that the (unobserved) CEO type associated with higher values of the CEO behavior index is relatively scarce in the population. This interpretation relies on the identifying assumption, transparent in the model, that conditional on factor inputs and sector of activity, the unobserved productivity of firms that need low index CEOs and those that need high index CEOs is the same. If this assumption fails, the fact that a firm hires a low index CEO might just reflect unobservable firm traits that lead to low productivity. We study the empirical validity of this assumption in the next section.

4.1 A Placebo Test of the Identifying Assumption

In this section we provide evidence on the validity of the identifying assumption. The intuition behind the test is the following. If the identifying assumption fails, then CEO behavior index simply captures unobserved performance differentials across firms choosing a specific type of CEO. In this case, we would expect to see a correlation between the CEO behavior index and firm performance even in the years *before* the individual we survey is appointed as CEO of the firm. In contrast, if the identifying assumption holds, the performance differential arising from the scarce CEO type would arise only in the years in which the specific CEO is in office. In this case, conditional performance would be the same across different firms in expectation.

To implement this test, we look at the evolution of the performance differential estimated in Table 5 before and after the CEO included in our time use sample is appointed. This is shown in Table 6. We start the analysis on the sample of 613 firms with available sales, employment and capital data examined in Table 5, column (2).[34] This sample

[33]Another way to look at this issue is to compare the magnitude of the relationship between the CEO behavior index and profits to the magnitude of the relationship between the CEO behavior index and CEO pay. We are able to make this comparison for a subsample of 196 firms with publicly available compensation data. Over this subsample, we find that a standard deviation change in the CEO behavior index is associated with an increase in profits per employee of $4,900 (which using the median number of employees in the subsample would correspond to $2,686,000 increase in total profit) and an increase in annual CEO compensation of $33,960. This broadly confirms the finding that the increase in firm performance associated with higher values of the CEO behavior index is not fully appropriated by the CEO in the form of rents.

[34]The results described in this Table are robust to the use of the larger sample of 813 firms with only sales and employment data (Table 3, column 1) Focusing on the subsample of firms with sales, employment and capital allows us to retain a larger sample size and at the same control for at least one

Table 6: CEO behavior and Firm Performance—Tenure Regressions

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable			Log(sales)			
CEO behavior index	0.251***	0.006	-0.017	0.079		
	(0.085)	(0.143)	(0.146)	(0.121)		
log(employment)	0.597***	0.596***	0.596***	0.597***	0.784***	0.800***
	(0.039)	(0.039)	(0.041)	(0.041)	(0.104)	(0.248)
log(capital)	0.316***	0.315***	0.315***	0.315***	0.123***	0.142**
	(0.034)	(0.034)	(0.037)	(0.037)	(0.043)	(0.060)
after CEO appointment		-0.298***				
		(0.090)				
CEO behavior*after CEO appointment		0.314**				
		(0.154)				
year 0-3 after CEO appointment			-0.183**			
			(0.074)			
CEO behavior*year 0-3 after CEO appointment			0.166			
			(0.121)			
year 4-6 after CEO appointment			-0.371***	-0.250***	-0.054	0.027
			(0.102)	(0.080)	(0.048)	(0.127)
CEO behavior*year 4-over after CEO appointment			0.407**	0.297**	0.140**	0.235**
			(0.169)	(0.144)	(0.067)	(0.117)
Adjusted R-squared	0.870	0.871	0.871	0.871	0.983	0.986
Observations	2589	2589	2589	2589	624	202
Number of firms	613	613	613	613	101	101
Sample	all	all	all	all	balanced	balanced & collapsed
Controls:						
Industry	y	y	y	y	y	y
Country by year	y	y	y	y	y	y
Noise	y	y	y	y	y	y
Firm fixed effects	n	n	n	n	y	y
Cluster	firm*period	firm*period	firm*period	firm*period	firm*period	firm*period
Test CEO behavior+CEO behavior*year 0 over (p-value)		0.00				
Test CEO behavior*year 0-3=CEO behavior*year 4-over			0.09			
Test CEO behavior+CEO behavior*year 0-3 over (p-value)			0.22			
Test CEO behavior+CEO behavior*year 4 over (p-value)			0.00	0.00		

Notes: *** (**) (*) denotes significance at the 1%, 5% and 10% level, respectively. We include all available years with information on sales, employment and capital, including up to 5 years prior to the CEO appointment. The sample in columns 5 and 6 is restricted to firms with obervations in both the before and after appointment period and include firm level fixed effects. Column 6 also uses output and input data averages across the two subperiods (instead of using individual years). Noise controls are a full set of dummies to denote the week in the year in which the data was collected, a reliability score assigned by the interviewer at the end of the survey week and a dummy taking value one if the data was collected through the PA of the CEO, rather than the CEO himself. Country by year dummies are included in all columns. Industry controls are 2 digit SIC dummies. All columns weighted by the week representativeness score assigned by the CEO at the end of the interview week. Errors clustered by firm and before/after period includes 2,589 observations, of which 540 relative to years pre-dating the CEO appointment, and the rest relative to years in which the CEO was in office. We start from the estimation of the same specification shown in Table 5, column (2) in this larger sample comprising before and after appointment data, forcing the correlation between the CEO behavior index and firm performance to be the same regardless of whether the CEO for whom the behavior index is computed is in office. Column (1) shows that the correlation between the CEO behavior index and firm performance in this larger sample is similar to the results shown in Table 5, column (2) (0.251, standard error 0.085). [35]

We then allow for the correlation to vary according to the tenure of the CEO in office by estimating:

$$y_{ifts} = \rho\theta_i + \delta A_t \theta_i + \beta^E e_{ft} + \beta^K k_{ft} + \mathbf{Z}_i \gamma + \zeta_t + \eta_s + \varepsilon_{ifts} \tag{4}$$

where the CEO appointment occurs at $t = 0$ and $t \in (-5, +60)$, $A_t = 1$ for $t > 0$, and all other variables and controls are defined above. The identifying assumption holds if $\rho = 0$ and $\delta > 0$, namely that the correlation between CEO behavior and firm performance only materializes after the CEO has been appointed, because it reflects mismatch rather than time invariant firm traits that are correlated with productivity. Column (2) provides evidence that support this identifying assumption: the correlation between the CEO behavior index and firm performance pre-appointment (θ) is equal to 0.006 and we cannot reject the null that it is equal to zero (standard error 0.143), while the correlation after appointment ($\rho + \delta$) is equal to 0.320 (0.006+0.314), significantly different from zero at the 1% level.

The results in column (2) allay the concern that the correlation between CEO behavior and firm performance is driven by time-invariant firm traits that determine its performance. A related concern is that the process is driven by firms' time varying traits, e.g. firms may appoint low index CEOs as a consequence of performance decline. To assess the relevance of this concern, we re-estimate Equation (4) allowing the coefficient of θ_i to vary every year. Figure 5 plots the coefficients estimated in this regression (which includes all the controls discussed above), grouping the years beyond 7 years after the CEO appointment in a single category. The graph shows that the difference between firms that eventually appoint a low index CEO and those that eventually appoint a high index CEO is stable and equal to zero before the appointment of the current CEO. This evidence allays the concern that low index CEOs are appointed in response to a fall in productivity. Moreover, Figure 5 also shows that the correlation between CEO behav-

of the other main factor inputs in production.

[35] To take into account the fact that CEOs may have an effect on firm performance after their appointment, in this table we cluster the standard errors at the firm*period level (i.e. differently for the same firm before and after the CEO appointment). Results are robust to using clustering at the firm level only.

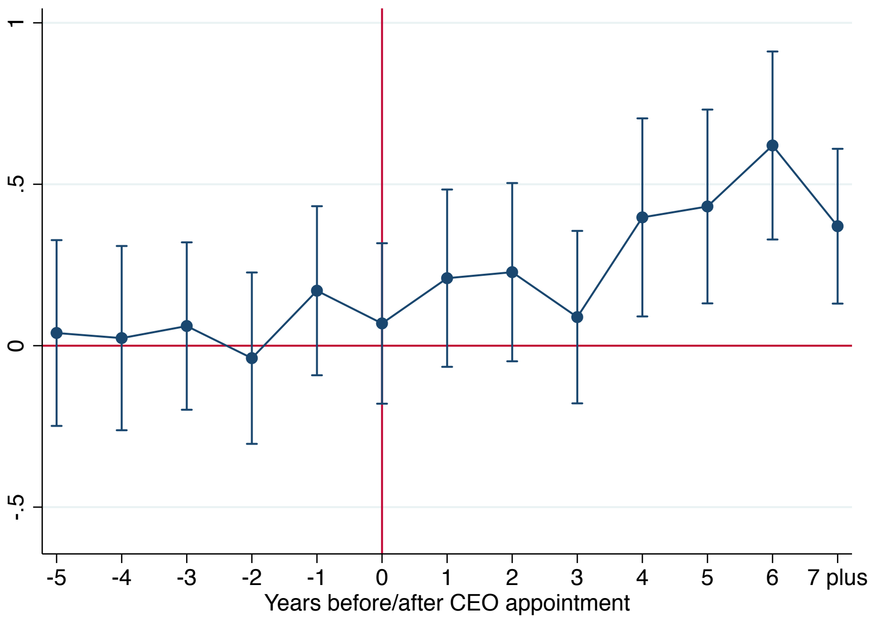


Figure 5: Correlation CEO behavior and TFP before and after the CEO's appointment.

Notes: The figure represents the point estimates and confidence intervals of the coefficients on interactions between the CEO behavior index and a set of dummies measuring the years before and after CEO appointment in the TFP regression of Table 6, columns 1 to 5.

ior and firm performance materializes four years after the CEO appointment, which is consistent with the dynamic version of our CEO-firm matching model (Section 3.2) in which the effect of CEO behavior on performance manifests itself gradually, as it takes time for the CEO's actions to have an impact on the firm. This rules out time varying unobservables driving the positive and significant coefficient of the CEO behavior index, with the exception of the case in which firms can foresee a productivity decline in the future and appoint a low index CEO 4 years beforehand.

The remainder of Table 6 provides robustness checks on these placebo test. Column (3) provides a formal test of the hypothesis that the correlation between CEO behavior and firm performance is the same after year 4 than between years 1 and 3. We reject the null with p-value 0.089. Column (4) re-estimates the regression using as benchmark for the pre-appointment period the window $t \in (-5, +3)$, with very similar results. Column (5) repeats this specification, but on the balanced sample (i.e. the subsample of firms with at least one year of data included in the before and after period), and including a full set of firm fixed effects to exploit *within firm* variation in performance to estimate the coefficient δ. Even in this demanding specification, the coefficient is precisely estimated, although of smaller magnitude (coefficient 0.140, standard error 0.067).[36] Finally, in

[36]In this case the before and after period includes $t \in (-5, +11)$, i.e. focusing on the balanced sample

column (6) we focus again on the balanced sample, but this time averaging the input and output data the data in the two separate periods before and after $t = 4$. The magnitude of the coefficient is even larger (0.235) and precisely estimated (standard error 0.117).

4.2 Robustness Checks

4.2.1 Managers or Management?

So far we have interpreted the CEO index primarily in terms of manager-specific behavior. However, what CEOs do with their time may also reflect broader differences in management processes across firms. For example, the propensity to engage in cross-functional coordination activities (vs. purely operational tasks) captured by higher values of the CEO index may be facilitated by the presence of systematic monitoring systems. To investigate this issue, we matched the CEO behavior index with detailed information on the type of management practices adopted in the firm. The management data was collected using the basic approach of the World Management Survey (Bloom et al. 2016). The survey methodology is based on semi-structured double blind interviews with plant level managers, run independently from the CEO time use survey.[37] To our knowledge, this is the first time that data on middle level management practices and information on CEO behavior is systematically analyzed.[38]

We start by looking at the correlation between the CEO behavior index and the management practices data in a simple specification including country and industry (SIC 1 level, given the smaller sample for which we are able to conduct this analysis) dummies, controls for log firm and plant employment (since the management data is collected at the plant level) and interview noise controls, using the weighting scheme described in previous specifications.[39] Table 7, Column (1) shows that higher values of the CEO behavior index are significantly correlated with a higher management score - a one standard deviation in management is associated with 0.059 increase in the CEO behavior index, or 18% of a

implies that CEOs with very long tenure - for whom the before appointment data is not available in Orbis - are dropped from the analysis. The drop in the magnitude of the coefficient on δ is largely attributable to this compositional change in the data. Estimating column (5) without firm fixed effects yields a coefficient on δ of 0.10.

[37]We collected the majority of the data in the Summer of 2013. A small share of the management data (16 observations out of a total of 191) was collected between 2006 and 2012 in the context of the larger WMS survey waves. We include this data in the analysis only if the CEO was in office at the time in which it was collected, and include wave dummies in all specifications.

[38]Bloom et al. (2016) analyze the correlation between management practices and employees' wage fixed effects and find evidence of sorting of employees with higher fixed effects in better managed firms. The analysis also includes a subsample of top managers, but due to data confidentiality it excludes from the sample highest paid individuals, who are likely to be CEOs.

[39]Given the limited number of firms in the sample we cannot include a full set of week dummies in the vector of noise controls as in previous specifications. We also include two measures of interview noise drawn from the management interviews, namely a variable denoting the duration of the management interview and the overall reliability of the interview as assessed by the interviewer.

standard deviation. Columns (2) and (3) show that this result is driven primarily by the sections of the management score measuring processes relative to operations, monitoring and targets, rather than people management practices (e.g. use of financial and non financial rewards in managing employees).

Table 7: CEO Behavior, Management, and Firm Performance

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	CEO behavior index			Log(sales)		
CEO behavior index				0.609**		0.614**
				(0.300)		(0.291)
Management (z-score)	0.059**				0.170**	0.172**
	(0.030)				(0.079)	(0.073)
Operations, Monitoring, Targets (z-score)		0.062**				
		(0.030)				
People (zscore)			0.044			
			(0.029)			
log(employment)	0.101***	0.103***	0.101***	0.950***	0.957***	0.910***
	(0.033)	(0.033)	(0.033)	(0.075)	(0.073)	(0.072)
Adjusted R-squared	0.144	0.145	0.133	0.718	0.714	0.733
Number of firms	191	191	191	145	145	145
Controls:						
Industry	y	y	y	y	y	y
Country	y	y	y			
Country by year				y	y	y
Noise	y	y	y	y	y	y
Cluster	Industry	Industry	Industry	Firm	Firm	Firm

We then turn to analyzing both the CEO behavior index and the management variables in the context of the production function of Equation (3).[40] Column (4) shows that the CEO behavior index is positive and statistically significant even in the smaller sample of 145 firms with both management and CEO data (coefficient 0.609, standard error 0.3). Column (5) shows that the management index is also correlated with labor productivity within the same sample (coefficient 0.17, standard error 0.079). Column (6) shows that the two variables retain a similar magnitude and significance level even when both included in the production function regression. The magnitude of the coefficients is also similar: a standard deviation change in the CEO behavior index is associated with an increase of 0.18 log points in sales, versus the 0.17 change implied by a standard deviation change in the management score. To summarize, even if management and the CEO behavior index are positively correlated among each other, they appear to be independently correlated with performance. The latter finding suggests that the positive

[40]We use a labor productivity specification since capital is not available for about a third of the sample of firms with both management and CEO behavior index data. When we do include capital, we end up with a sample of 103 observations and in the specification of column (6) the CEO behavior index remains significant at the 10% level (coefficient 0.57), while the management score drops to 0.07 and is insignificant (standard error 0.72).

relationship between the CEO index and firm performance is not entirely a reflection of the management practices adopted by the CEO when he is in office.

4.2.2 Alternative ways of building the CEO Behavior Index

TBC

4.2.3 Alternative estimation methods of the production function

TBC

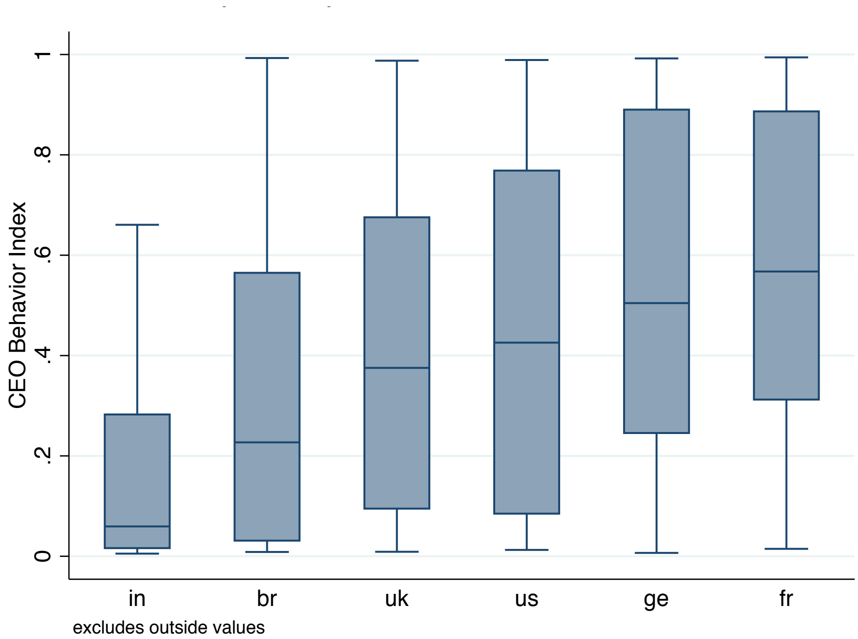
5 CEO-Firm Match along the Development Path

In this section we exploit regional variation in development across and within countries to provide further evidence on how the correlation between CEO behavior and firm performance may be driven by matching frictions in the market for CEOs. The analysis relies on the assumption that frictions are more severe in poorer regions. The reasons underpinning this assumption are manyfold. At the hiring stage, screening might be worse in low income regions because the market for CEOs is less thick and professional headhunting services less common. After hiring, governance might be worse because contract enforcement is less effective in low income regions and courts are slower. To operationalize this idea, we use regional GDP to proxy for the severity of matching frictions and test whether: a) the quality of the match is higher and; b) the correlation between behavior and firm performance is lower when matching frictions are less severe.

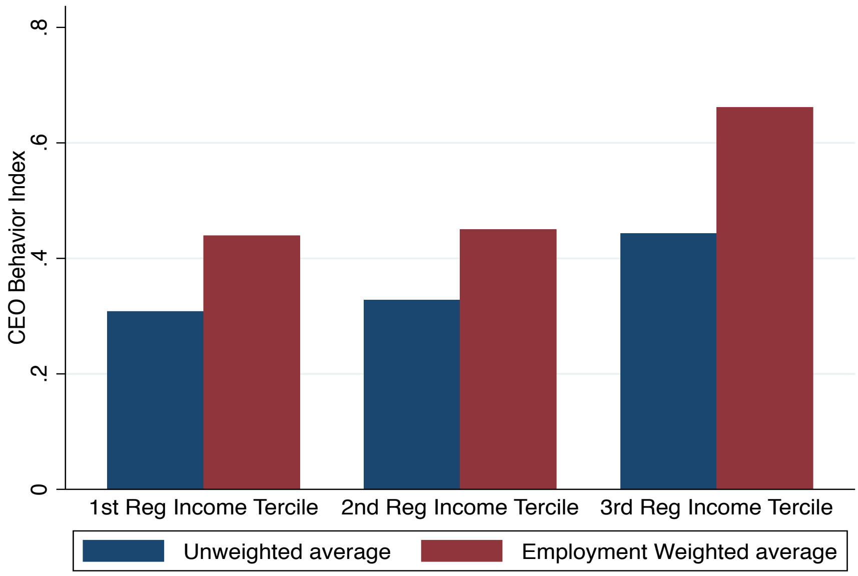
We use within country regional variation in development, using the data on regional income per capita in current purchasing-power-parity (PPP) dollars developed by Gennaioli et al. (2013). The sample firms are located in 121 regions within 6 countries that are at very different stages of the development path. Income per capita in the poorest region in the sample (Uttar Pradesh, India) is \$1,300; in the richest region it is \$143,000 (DC, USA). The median within country range in income per capita is \$23,942 and the median within country standard deviation is \$5,695. The median number of regions within each country is 16.

We start the analysis by showing in Figure 6, Panel A a box plot of the CEO behavior index across countries. Clearly, the median value of the CEO behavioral index is higher in richer countries, and significantly lower in Brazil and India. The graph also shows that there is ample within country variation in each country, although the distribution is more compressed in India.

Figure 6, Panel B shows the values of the CEO behavior index across different terciles of the within country distribution of regional income per capita (i.e. we first normalize



(a) CEO behavior by country



(b) CEO behavior index and regional income per capita

Figure 6: CEO Behavior across Countries and Regions

Notes: The top panel shows the box plot of the CEO behavior index by country of CEO location. Number of observations: India=358; Brazil=280; UK=87; US=149; Germany=125; France=115. Each bar in the bottom panel represents the average of the CEO behavior index by tercile of regional income per capita (expressed in deviations from country means). Within each tercile, the left bar shows raw averages, while the right bar shows employment weighted averages (employment weights computed within each country).

regional income per capita by its country mean, and then look at relatively poorer or richer regions within country). The left bars in the graph refer to raw averages, while right bars report employment weighted averages. Even within countries, richer regions tend to have a higher value of the CEO behavior index. Furthermore, differences between poor and richer regions are larger when we consider employment weighted averages, which is consistent with the idea that, in richer regions, CEOs with high values of the behavior index are more likely to be found in large firms relative to poor regions.

Table 8 exploits within-country, cross-regional variations in income per capita to test whether the quality of the match is higher when frictions are less severe. We estimate:

$$\theta_{ifsr} = \alpha + \beta e_f + \gamma e_f * Y_r + \vartheta a_s + \mathbf{Z}_i \delta + \varepsilon_{ifsr} \tag{5}$$

where θ_{ifsr} is the behavior index of CEO c in firm f, industry s and region r, e_f is log firm employment and Y_r is a measure of regional development that proxies for matching frictions. All specifications include the same set of CEO and noise controls discussed above, as well as SIC2 industry dummies. Throughout the analysis we control for country fixed effects, thereby exploiting the variation in development levels across regions within countries.

Table 8: CEO-Firm Match by Region

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	CEO behavior index			Log(sales)		
log(employment)	0.044***	-0.112	-0.128	0.906***	0.907***	0.923***
	(0.010)	(0.076)	(0.087)	(0.037)	(0.033)	(0.041)
log(employment) * high income country	0.025*					
	(0.015)					
Region income per capita		-0.042			0.169	
		(0.066)			(0.105)	
log(employment) * region income per capita		0.018**	0.020**			
		(0.008)	(0.009)			
CEO behavior index				0.598***	2.469**	2.220**
				(0.162)	(0.990)	(1.100)
CEO behavior* high income country				-0.523***		
				(0.189)		
CEO behavior*region gdp					-0.223**	-0.199*
					(0.101)	(0.114)
Adjusted R-squared	0.265	0.272	0.294	0.730	0.729	0.710
N	1114	1114	1114	831	831	831
Controls:						
COO dummy	y	y	y	y	y	y
CEO (mba & tenure)	y	y	y	n	n	n
Industry	y	y	y	y	y	y
Country	y	y	y	y	y	y
Noise	y	y	y	y	y	y
Region dummies	n	n	y	n	n	y
Cluster	Industry	Region	Region	Firm	Region	Region

If the quality of the match improves as frictions become less severe, we expect $\gamma > 0$. Column (1) proxies for development with a dummy that equals one if the region is located in a high income country (France, Germany, UK and US). In line with the hypothesis that the quality of the match is higher when frictions are less severe, we find that the correlation between firm size and the behavioral index is significantly larger in richer countries. Column (2) uses log regional income per capita to proxy for Y_r, and finds a similar result: the strength of the correlation between the CEO behavior index and firm size increases as regional income per capita increases, indicating that in highly developed regions large firms are more likely to hire CEOs with a high behavior index. Column (3) further probes this correlation with the inclusion of a full set of regional dummies, thus exploiting *within region* variation in CEO behavior and firm size, with remarkably similar results.

Second, we test whether the correlation between CEO behavior and firm performance *decreases* with the level of regional development. Intuitively, if the match between firms and CEOs improves with development, the share of mismatched CEOs should decrease as well, and so should the difference in performance between firms led by different CEO types. To test this idea, we estimate:

$$y_{iftsr} = \alpha\theta_i + \delta\theta_i * Y_r + \beta^E e_{ft} + \mathbf{Z}_i\gamma + \zeta_t + \eta_s + \varepsilon_{iftsr} \tag{6}$$

where all variables are defined above. [41] If the quality of the match improves as frictions become less severe, we expect $\delta < 0$. This is because as fewer CEOs are mismatched, differences in behavior are more likely to reflect optimal responses in firm needs.

Table 8, columns (4) to (6) report the estimates of Equation (6). Column (4) shows that the positive correlation between the CEO behavior index and firm performance shown in 5, column (1) is, in fact, an average of two very different magnitudes: 0.598 in low income countries and 0.075 in high income countries. The difference is precisely estimated at the 1% level. Column (5) defines Y_r as regional income per capita and again finds $\delta < 0$. Figure 7 plots the estimated correlation $\alpha + \delta Y_r$ evaluated at all sample values of Y_r. The correlation between the CEO behavior index and firm performance falls from 0.9 to 0 and it becomes statistically indistinguishable from zero at log(regional income per capita)=10.5, which is at the 75th percentile of the regional income per capita distribution in our sample. Finally, in column (6) we include in the specification a full set of regional dummies, thus estimating $\delta < 0$ relying exclusively on within region variation. The results are robust to the inclusion of the regional dummies, although the significance of the interaction term drops to 10%.

[41] Note that we are now using a labor productivity specification, and thus the larger sample of firms with usable information on sales and employment during at least one year in which the CEO is in office (i.e. the sample of 831 firms used in Table 4, column 1). We do so because imposing the additional requirement of having at least one non missing information on capital needed to estimate a production function closer to a TFP specification would imply the loss of the majority of the Brazilian and the Indian sample, where within country differences in income per capital are larger relative to the rest of developed countries.

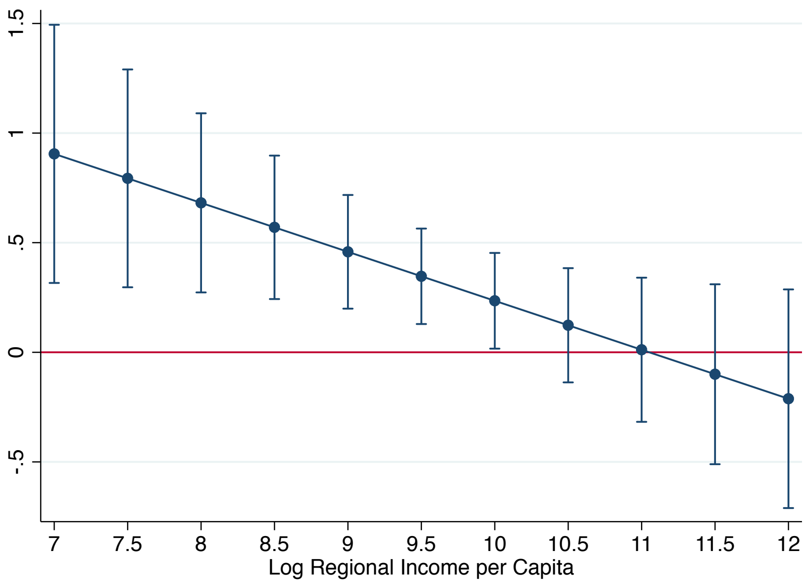


Figure 7: CEO and Firm Performance by Region.

Taken together, the results shown in Table 8 are consistent with the idea that the correlation between CEO behavior and firm performance arises because of matching frictions. Importantly, they rule out a simpler moral hazard story where the high index behavior is always more productive (i.e. there is no demand for low index behavior), since if it were so we would find a positive correlation between firm performance and the index across all regions and countries in our sample.

6 Calibration

TBC

7 Conclusions

This paper combines a new survey methodology with a machine learning algorithm to measure the behavior of CEOs in large samples. We show that CEOs differ in their behavior along a number of dimensions, and that these differences tend to co-vary with observable firm characteristics, such as firm size, organizational structure and industry characteristics. Guided by a simple firm-matching model, we also show evidence of significant matching frictions in the assignment of CEOs to firms, and that these frictions appear to be particularly severe in emerging economies.

While this paper has intentionally taken an agnostic approach to leadership, an obvious next step would be to explore in more detail the precise mechanisms through which different leadership behaviors affect firm performance. The CEO behavior that according to our CEO-firm matching model and our data is scarcer in the population of potential CEOs (and hence produces a better average performance) features a longer planning horizon, larger multi-functional meetings, a focus on higher-level executives and non-production functions. One tentative interpretation is that a CEO that displays this pattern of behavior is a *coordinator*, who delegates operational tasks to high-level executives and spends more of his time making ensuring good communication in the top management team. Within the same interpretation, a CEO that displays the other CEO behavior emerging from the classification exercise is instead a *micromanager*, who tends to intervene directly in operational aspects, who prefers one-on-one meetings with a variety of internal and external constituents, and who puts less emphasis on long term planning.

To the best of our knowledge, the coordinator/micromanager dichotomy has not been directly addressed by any of the existing literature on leadership - within and outside economics - although the general idea of leader types is present in recent papers in the economic leadership literature.[42] Future work could utilize information about CEO behavior to inform alternative leadership models. At the same time, it would also be interesting to better explore the connection between our observed behavioral patterns and contributions in the management literature. For example, Kotter (1999) proposes that the key task of a CEO is to align the organization behind a common vision - the emphasis of our Behavior-1 CEOs on large, planned, multi-functional meetings is consistent with an alignment effort.

More generally, a possible next step of this research would be to extend the data collection to the diaries of multiple managerial figures beyond the CEO. This approach would allow us to further explore the importance of managerial interactions and team behavior (Hambrick and Mason 1984), which are now largely absent from our analysis. We leave these topics for further research.

[42]Hermalin (1998, 2007) proposes a rational theory of leadership, whereby the leader possesses private non-verifiable information on the productivity of the venture that she leads. In the dynamic version of the model, the leader can develop a reputation for honestly announcing the true state of the world. In practice, one way of strengthening this reputation is to have formal gatherings where the leader is held accountable for her past announcements. Van den Steen (2010) highlights the importance of shared beliefs in organizations. Shared beliefs lead to more delegation, less monitoring, higher utility, higher execution effort, faster coordination, less influence activities, and more communication. Bolton et al. (2013) propose a model of resoluteness. A resolute leader has a strong, stable vision that makes her credible among her followers. This helps align the followers' incentives and generates higher effort and performance. Finally, Dessein and Santos (2016) explore the interaction between CEO characteristics, CEO attention allocation, and firm behavior: small differences in managerial expertise may be amplified by optimal attention allocation and result in dramatically different firm behavior.

References

Autor, D. H. (2013). The "task approach" to labor markets : an overview. *Journal for Labour Market Research*, 46(3):185–199.

Autor, D. H., Levy, F., and Murnane, R. J. (2003). The Skill Content of Recent Technological Change: An Empirical Exploration. *The Quarterly Journal of Economics*, 118(4):1279–1333.

Bandiera, O., Guiso, L., Prat, A., and Sadun, R. (2012). What Do CEOs Do? CEP Discussion Papers dp1145, Centre for Economic Performance, LSE.

Bandiera, O., Guiso, L., Prat, A., and Sadun, R. (2015). Matching Firms, Managers, and Incentives. *Journal of Labor Economics*, 33(3):623 – 681.

Bandiera, O., Prat, A., and Sadun, R. (2013). Managing the Family Firm: Evidence from CEOs at Work. Harvard Business School Working Papers 14-044, Harvard Business School.

Bennedsen, M., Nielsen, K. M., Perez-Gonzalez, F., and Wolfenzon, D. (2007). Inside the Family Firm: The Role of Families in Succession Decisions and Performance. *The Quarterly Journal of Economics*, 122(2):647–691.

Bertrand, M. and Schoar, A. (2003). Managing with Style: The Effect of Managers on Firm Policies. *The Quarterly Journal of Economics*, 118(4):1169–1208.

Blei, D. M., Ng, A. Y., and Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3:993–1022.

Bloom, N., Sadun, R., and Van Reenen, J. (2016). Management as a Technology? mimeograph.

Bloom, N. and Van Reenen, J. (2007). Measuring and Explaining Management Practices Across Firms and Countries. *The Quarterly Journal of Economics*, 122(4):1351–1408.

Bolton, P., Brunnermeier, M. K., and Veldkamp, L. (2013). Leadership, Coordination, and Corporate Culture. *The Review of Economic Studies*, 80(2):512–537.

Chang, J., Gerrish, S., Boyd-Graber, J. L., and Blei, D. M. (2009). Reading Tea Leaves: How Humans Interpret Topic Models. In *Advances in Neural Information Processing Systems*.

Dessein, W. and Santos, T. (2016). Managerial Style and Attention. mimeograph.

Gabaix, X. and Landier, A. (2008). Why has CEO Pay Increased So Much? *The Quarterly Journal of Economics*, 123(1):49–100.

Gennaioli, N., Porta, R. L., de Silanes, F. L., and Shleifer, A. (2013). Human Capital and Regional Development. *The Quarterly Journal of Economics*, 128(1):105–164.

Griffiths, T. L. and Steyvers, M. (2004). Finding Scientific Topics. *Proceedings of the National Academy of Sciences*, 101(Suppl. 1):5228–5235.

Halac, M. and Prat, A. (2014). Managerial Attention and Worker Engagement. CEPR Discussion Papers 10035, C.E.P.R. Discussion Papers.

Hambrick, D. C. and Mason, P. A. (1984). Upper Echelons: The Organization as a Reflection of Its Top Managers. *The Academy of Management Review*, 9(2):193–206.

Hansen, S., McMahon, M., and Prat, A. (2014). Transparency and Deliberation within the FOMC: a Computational Linguistics Approach. Discussion Paper 9994, CEPR.

Heinrich, G. (2009). Parameter Estimation for Text Analysis. Technical report, vsonix GmbH and University of Leipzig.

Hermalin, B. E. (1998). Toward an Economic Theory of Leadership: Leading by Example. *American Economic Review*, 88(5):1188–1206.

Hermalin, B. E. (2007). Leading for the Long Term. *Journal of Economic Behavior & Organization*, 62(1):1–19.

Kaplan, S. N., Klebanov, M. M., and Sorensen, M. (2012). Which CEO Characteristics and Abilities Matter? *The Journal of Finance*, 67(3):973–1007.

Kaplan, S. N. and Sorensen, M. (2016). Are CEOs Different? Characteristics of Top Managers. mimeograph.

Kotter, J. P. (1999). *John Kotter on What Leaders Really Do*. Harvard Business School Press, Boston.

Luthans, F. (1988). Successful vs. Effective Real Managers. *Academy of Management Executive*, 2(2):127–132.

Malmendier, U. and Tate, G. (2005). CEO Overconfidence and Corporate Investment. *Journal of Finance*, 60(6):2661–2700.

Malmendier, U. and Tate, G. (2009). Superstar CEOs. *The Quarterly Journal of Economics*, 124(4):1593–1638.

Mintzberg, H. (1973). *The Nature of Managerial Work*. Harper & Row., New York.

Mullins, W. and Schoar, A. (2013). How do CEOs see their Role? Management Philosophy and Styles in Family and Non-Family Firms. NBER Working Papers 19395, National Bureau of Economic Research, Inc.

Rauch, J. E. (1999). Networks versus markets in international trade. *Journal of International Economics*, 48(1):7–35.

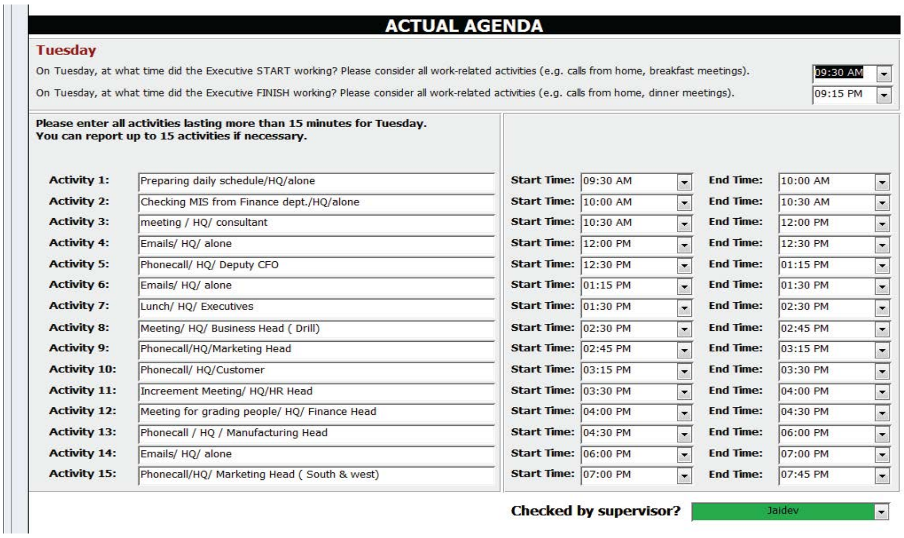
Taddy, M. A. (2012). On estimation and selection for topic models. In *Proceedings of the 15th International Conference on Artificial Intelligence and Statistics (AISTATS)*. JMLR: W&CP 22.

Tervio, M. (2008). The Difference That CEOs Make: An Assignment Model Approach. *American Economic Review*, 98(3):642–68.

Van den Steen, E. (2010). On the origin of shared beliefs (and corporate culture). *RAND Journal of Economics*, 41(4):617–648.

A Data Appendix

A.1 Survey Background



ACTUAL AGENDA

Tuesday

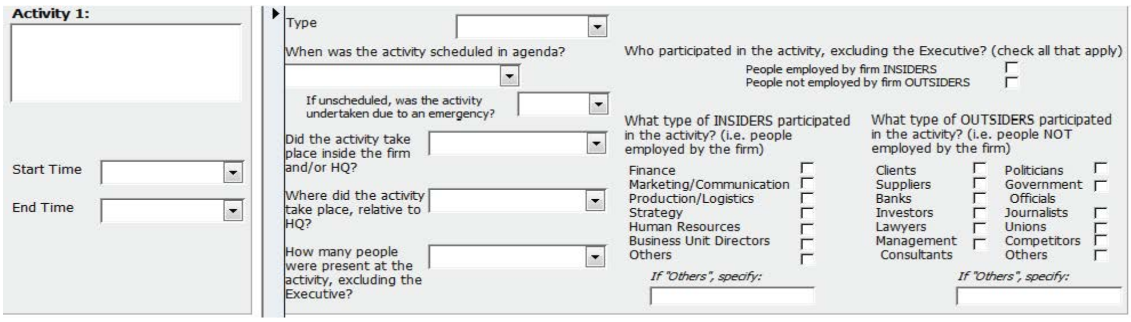
On Tuesday, at what time did the Executive START working? Please consider all work-related activities (e.g. calls from home, breakfast meetings). 09:30 AM

On Tuesday, at what time did the Executive FINISH working? Please consider all work-related activities (e.g. calls from home, dinner meetings). 09:15 PM

Please enter all activities lasting more than 15 minutes for Tuesday.
You can report up to 15 activities if necessary.

Activity 1:	Preparing daily schedule/HQ/alone	Start Time: 09:30 AM	End Time: 10:00 AM
Activity 2:	Checking MIS from Finance dept./HQ/alone	Start Time: 10:00 AM	End Time: 10:30 AM
Activity 3:	meeting / HQ/ consultant	Start Time: 10:30 AM	End Time: 12:00 PM
Activity 4:	Emails/ HQ/ alone	Start Time: 12:00 PM	End Time: 12:30 PM
Activity 5:	Phonecall/ HQ/ Deputy CFO	Start Time: 12:30 PM	End Time: 01:15 PM
Activity 6:	Emails/ HQ/ alone	Start Time: 01:15 PM	End Time: 01:30 PM
Activity 7:	Lunch/ HQ/ Executives	Start Time: 01:30 PM	End Time: 02:30 PM
Activity 8:	Meeting/ HQ/ Business Head (Drill)	Start Time: 02:30 PM	End Time: 02:45 PM
Activity 9:	Phonecall/HQ/Marketing Head	Start Time: 02:45 PM	End Time: 03:15 PM
Activity 10:	Phonecall/ HQ/Customer	Start Time: 03:15 PM	End Time: 03:30 PM
Activity 11:	Increement Meeting/ HQ/HR Head	Start Time: 03:30 PM	End Time: 04:00 PM
Activity 12:	Meeting for grading people/ HQ/ Finance Head	Start Time: 04:00 PM	End Time: 04:30 PM
Activity 13:	Phonecall / HQ / Manufacturing Head	Start Time: 04:30 PM	End Time: 06:00 PM
Activity 14:	Emails/ HQ/ alone	Start Time: 06:00 PM	End Time: 07:00 PM
Activity 15:	Phonecall/HQ/ Marketing Head (South & west)	Start Time: 07:00 PM	End Time: 07:45 PM

Checked by supervisor? Jaidev



Activity 1:

Start Time

End Time

Type

When was the activity scheduled in agenda?

If unscheduled, was the activity undertaken due to an emergency?

Did the activity take place inside the firm and/or HQ?

Where did the activity take place, relative to HQ?

How many people were present at the activity, excluding the Executive?

Who participated in the activity, excluding the Executive? (check all that apply)

People employed by firm INSIDERS

People not employed by firm OUTSIDERS

What type of INSIDERS participated in the activity? (i.e. people employed by the firm)

Finance
Marketing/Communication
Production/Logistics
Strategy
Human Resources
Business Unit Directors
Others

If "Others", specify:

What type of OUTSIDERS participated in the activity? (i.e. people NOT employed by the firm)

Clients
Suppliers
Banks
Investors
Lawyers
Management Consultants
Politicians
Government Officials
Journalists
Unions
Competitors
Others

If "Others", specify:

Figure A.1: Survey Instrument.

A.2 Average CEO Time Shares in Baseline Subsample

Table A.1: Raw Average Time Shares for all CEOs on Estimation Subsample

(a) Distribution of time within features

Type		**Duration**		**Planned**		**Participants**	
value	share	value	share	value	share	value	share
meeting	0.803	1hr+	0.657	planned	0.764	size2+	0.553
site_visit	0.06	1hr	0.188	unplanned	0.236	size1	0.427
phone_call	0.054	30m	0.139			missing	0.019
business_meal	0.049	15m	0.017				
public_event	0.015						
conference_call	0.013						
workrelated_leisure	0.005						
video_conference	0.001						

(b) Distribution of time across functions

Inside Functions		**Outside Functions**	
function	share	function	share
production	0.35	clients	0.103
mkting	0.206	suppliers	0.064
finance	0.147	others	0.05
groupcom	0.073	associations	0.031
hr	0.063	consultants	0.026
bunits	0.042	govoff	0.016
board	0.031	banks	0.013
other	0.029	compts	0.012
admin	0.029	pemployee	0.01
cao	0.023	lawyers	0.008
coo	0.017	investors	0.005
strategy	0.011		
legal	0.008		

Notes: The top table shows the amount of time the average CEO spends on different options within features for the 98,347 15-minute units of time in the baseline estimation exercise excluding rare combinations. The bottom table shows the amount of time the average CEO spends with different functions on the same subsample.

Table A.2: Estimated Average Time Shares for all CEOs on Estimation Subsample

(a) Distribution of time within features

Type		**Duration**		**Planned**		**Participants**	
value	share	value	share	value	share	value	share
meeting	0.801	1hr+	0.687	planned	0.782	size2+	0.573
site_visit	0.062	1hr	0.176	unplanned	0.218	size1	0.411
business_meal	0.053	30m	0.123			missing	0.017
phone_call	0.047	15m	0.014				
public_event	0.017						
conference_call	0.012						
workrelated_leisure	0.006						
video_conference	0.001						

(b) Distribution of time across functions

Inside Functions		**Outside Functions**	
function	share	function	share
production	0.355	clients	0.104
mkting	0.208	suppliers	0.068
finance	0.144	others	0.05
groupcom	0.081	associations	0.033
other	0.077	consultants	0.026
hr	0.062	govoff	0.015
bunits	0.041	compts	0.014
board	0.032	banks	0.013
admin	0.029	pemployee	0.01
cao	0.022	lawyers	0.008
coo	0.015	investors	0.006
strategy	0.01		
legal	0.008		

Notes: The top table shows the estimated amount of time the average CEO spends on different options within features for the baseline estimation exercise. The bottom table shows the estimated amount of time the average CEO spends with different functions on the same subsample. These estimated shares are derived from the marginal distributions computed from the estimated behaviors, and the estimated CEO behavioral indices.

B Additional Results