European Bank
for Reconstruction and Development

CEO-Firm Matches and Productivity in 42 Countries

Amanda Dahlstrand, Dávid László, Helena Schweiger, Oriana Bandiera, Andrea Prat and Raffaella Sadun

Abstract

Firms are key to economic development, and CEOs are key to firm productivity. Are firms in countries at varying stages of development led by the right CEOs, and if not, why? We develop a parsimonious measure of CEO time use that allows us to differentiate CEOs into "leaders" and "managers" in a survey of 4,800 manufacturing firms across 42 countries, with income per capita ranging from USD 4,000 to 45,000. We find that poorer countries have fewer leaders and relate this to training opportunities. Even when suitable leaders are available, they often do not lead the firms that would benefit the most, resulting in mismatches that can cause up to a 20 per cent loss in productivity for the mismatched firms. The findings imply that policies that address the causes of mismatch could significantly enhance growth without additional resources.

Keywords: CEO-firm match, productivity,

JEL Classification Number: J24, M12, O47

Contact details: Helena Schweiger, European Bank for Reconstruction and Development, 5 Bank St, London, E14 4BG, UK. Email: schweigh@ebrd.com.

Amanda Dahlstrand is an Assistant Professor at the Economics Department of University of Zurich. Dávid László is a Pre-Doctoral Research Assistant at the London School of Economics. Helena Schweiger is a Lead Research Economist at the EBRD. Oriana Bandiera is the Sir Anthony Atkinson Professor of Economics at the London School of Economics. Andrea Prat is the Richard Paul Richman Professor of Business at Columbia Business School and Professor of Economics at the Department of Economics, Columbia University. Raffaella Sadun is the Charles E. Wilson Professor of Business Administration at Harvard Business School.

We thank Ralph De Haas, Cevat Giray Aksoy, and participants at the CEPR IMO-ENT conference, 2024 ASSA Annual Meeting, University of Zurich, and Office of the Chief Economist at the EBRD for helpful comments. Pablo García-Guzmán provided excellent research assistance with collecting and processing the data on higher education institutions. The views expressed in this paper are our own and not necessarily those of the EBRD. Amanda Dahlstrand acknowledges funding from the University Research Priority Program "Equality of Opportunity" at the University of Zurich.

The EBRD Working Papers intend to stimulate and inform the debate about the economic transformation of the regions in which the EBRD operates. The views presented are those of the authors and not necessarily of the EBRD.

Working Paper No. 300 **Prepared in January 2025**

CEO-Firm Matches and Productivity in 42 Countries

Amanda Dahlstrand[*], Dávid László[†], Helena Schweiger[‡], Oriana Bandiera[§], Andrea Prat[¶], and Raffaella Sadun[‖**]

January 2025

Abstract

Firms are key to economic development, and CEOs are key to firm productivity. Are firms in countries at varying stages of development led by the right CEOs, and if not, why? We develop a parsimonious measure of CEO time use that allows us to differentiate CEOs into "leaders" and "managers" in a survey of 4,800 manufacturing firms across 42 countries, with income per capita ranging from USD 4,000 to 45,000. We find that poorer countries have fewer leaders and relate this to training opportunities. Even when suitable leaders are available, they often do not lead the firms that would benefit the most, resulting in mismatches that can cause up to a 20% loss in productivity for the mismatched firms. The findings imply that policies that address the causes of mismatch could significantly enhance growth without additional resources.
JEL codes: J24, M12, O47

[*]University of Zurich. Email: amanda.dahlstrand@econ.uzh.ch.
[†]London School of Economics. Email: d.laszlo@lse.ac.uk.
[‡]European Bank for Reconstruction and Development. Email: schweigh@ebrd.com.
[§]London School of Economics. Email: o.bandiera@lse.ac.uk.
[¶]Columbia University. Email: andrea.prat@columbia.edu.
[‖]Harvard University. Email: rsadun@hbs.edu.

[**]We thank Ralph De Haas, Cevat Giray Aksoy, and participants at the CEPR IMO-ENT conference, 2024 ASSA Annual Meeting, University of Zurich, and Office of the Chief Economist at the EBRD for helpful comments. Pablo García-Guzmán provided excellent research assistance with collecting and processing the data on higher education institutions. The views expressed in this paper are our own and not necessarily those of the EBRD. Amanda Dahlstrand acknowledges funding from the University Research Priority Program "Equality of Opportunity" at the University of Zurich.

1 Introduction

The separation of ownership and control has facilitated firm growth to such an extent that large corporations now account for the majority of both employment and output in most countries. Unsurprisingly, CEOs at their helm receive considerable attention from scholars and practitioners alike. However, much of this research focuses on CEOs in high-income countries. In contrast, we know much less about CEOs in countries at other stages of development despite the fact that the success of firms in these countries could play a critical role in driving economic development. We focus on middle-income countries, as CEOs play a smaller role in the lowest-income countries, where the informal sector and small firms represent a large share of the economy (LaPorta and Shleifer, 2014).

Does the behavior of CEOs in middle-income countries mirror that of their counterparts in high-income countries? Does it match firms' needs, or do market frictions create systematic mismatches that reduce productivity? There is reason to believe that matching is worse in middle-income countries in this arguably most important labor market, as evidence suggests that the match of worker skills to jobs is also poorer in these regions (Bandiera, Kotia, Lindenlaub, Moser, and Prat, 2024). Yet, answering these questions is not straightforward. Middle-income countries are a much more varied set – in terms of institutions, culture, and economic development – than their richer counterparts, and tools to measure CEOs' behavior and their match with firm needs are both expensive and difficult to implement at scale.

In this paper, we propose a method to measure both CEOs' behavior and firms' need for that behavior using a firm survey. We embed our new survey questions in the European Bank for Reconstruction and Development (EBRD), European Investment Bank (EIB), and World Bank (WB) Enterprise Survey (henceforth "ES"), covering 4,800 manufacturing firms in 42 countries. Seventy percent of these countries are middle-income countries, and they host 796 million people (10% of the world population). The countries also span a wide range of economic development stages: GDP per capita in the poorest country in our sample, Tajikistan, is 1/12 that of the richest, Malta. In addition to economic development, these countries have a wide range of institutional features, from ex-communist countries to some of the first members of the European Union. Together, these features allow us to study how CEO behavior and CEO-firm matching varies across stages of development.

We build on Bandiera, Hansen, Prat, and Sadun (2020), from which we derive the distinction between two important types of CEOs regarding how they spend their work hours (their time use). We use granular diary data to show that the behavioral types ("managers" and "leaders") uncovered by the latent Dirichlet allocation (LDA) learning process in Bandiera et al. (2020) can be approximated by a small number of questions. We add these questions

to the firm survey in 42 countries and classify CEOs into managers, whose focus is within the firm, particularly on monitoring processes and implementing firm policies, and leaders, who coordinate between internal and external constituencies of the firm, providing organizational alignment.

To measure firm demand for managers and leaders in middle-income countries, we use the correlations between firm-level characteristics and CEO types in high-income countries – where the structural estimates of Bandiera et al. (2020) suggest that the share of mismatched firms is low - to predict which types of CEOs firms in poorer countries should hire, were they to use the same matching algorithm as their counterparts in richer countries. Having assigned each CEO their closest behavioral type and each firm their best CEO fit, we can compute, for each country, the share of mismatched firm-CEO pairs and assess whether this is due to a specific type being in excess demand and/or whether mismatches occur at the firm level, even though aggregate demand and supply of CEO types are balanced.

We have three key findings. First, using our new, simple survey instrument, we find that the distinction between managers and leaders is correlated with firm productivity. The productivity gap between the average firm with a leader CEO and the average firm with a manager CEO is 12%, with leader firms having higher productivity in terms of revenue per employee. Reassuringly, the correlation is of a similar magnitude to that in the Bandiera et al. (2020) study (which uses full CEO diary data collection in 6 countries). Moreover, mirroring the results from those 6 countries, the share of leader CEOs increases with GDP per capita across the 42 countries in our sample.

Second, we find that the overall productivity advantage of leaders stems from the demand for leader CEOs outstripping supply, rather than from leader CEOs being better suited to every firm. Indeed, a firm that we predict needs a leader but hires a manager CEO has 20% lower productivity than a firm that needs a leader and hires a leader CEO. Vice versa, a firm that needs a manager but hires a leader CEO has 15% lower productivity than a firm that needs a leader and hires a leader CEO. Accounting for the total number of firms in the mismatched categories and the revenue of these firms suggests that eliminating the mismatch would be associated with a 9% increase in productivity across the entire sample. Finally, we find that around half of the aggregate correlation between mismatch and productivity is due to an excess demand for leader CEOs, while the other half is due to mismatch conditional on demand equal to supply.

We next examine the factors underlying our main results, focusing on what prevents the market from producing more leader-type CEOs in certain countries and what factors contribute to mismatches within countries conditional on demand equaling supply. We first find that country-level characteristics explain only a small portion of the variation in which

firms are mismatched or have leader-type CEOs. Therefore, we investigate local market- and firm-level characteristics. In particular, as research suggests human capital is essential to explain development on a sub-national level (Gennaioli, LaPorta, de Silanes, and Shleifer, 2013), we explore the possibility that the low supply of leaders is due to a lack of local training opportunities. As a starting point, we collect data on location and type of higher education institutions (HEIs) offering business education with the World Higher Education Database (WHED) as a starting point, and find that the availability of business degrees close to firms is associated with a higher supply of leaders.

To shed more light on the mechanism and assess its robustness, we differentiate business education by expected quality. If it is indeed the business education that increases leader supply, we should observe a greater effect from higher-quality HEIs. While school rankings do not cover the institutions in our sample, we propose an alternative approach to evaluating quality. In many middle-income economies, a surge in private HEIs offering business education began in 1990. These newer HEIs, unconnected to established universities, appear to provide lower-quality education. We find that business education at the more established HEIs is associated with a higher supply of leaders in firms nearby, while business education at HEIs founded in the booming higher education market post-1990 are unassociated with any higher supply.

We further investigate reasons for mismatches within countries. When making a hiring decision, a firm must be aware of the CEO type it needs to select the appropriate candidate. We hypothesize that the ideal and actual CEO types are more likely to match within firms that participate in international trade or are exposed to best practices through foreign ownership or the presence a board. We do indeed find that openness to international trade, the existence of a supervisory board and foreign ownershipare significantly positively associated with a firm having a correctly matched CEO type. This aligns with the hypothesis that influence from the global market can contribute to less mismatch.

This paper contributes to the literature on CEO management style (Bertrand and Schoar, 2003; Kaplan, Klebanov, and Sorensen, 2012; Bandiera et al., 2020). In that literature, there are at least two different approaches to measuring CEOs effects on firms: inferring it from fixed effects (Bertrand and Schoar, 2003), versus directly measuring CEO behavior through shadowing (Bandiera et al., 2020). While the former (fixed effects) is silent about the mechanisms through which CEOs affect firm performance, the latter (shadowing) is often expensive to implement at scale. In this paper, we take a different approach and demonstrate that important components of CEO behavior can be captured with a parsimonious survey instrument that can be deployed at scale. In addition to this methodological difference, we complement the earlier literature through providing evidence of horizontal differentiation of

CEO types across many markets.

The paper also relates to the broader literature on management, firm organization, and productivity (overview in Scur, Sadun, Van Reenen, Lemos, and Bloom, 2021). In the equilibrium of the competitive assignment model of Gabaix and Landier (2008), higher-talent CEOs run larger companies. In that model, there is only one dimension of managerial talent, while we study potential horizontal differentiation. Our paper's primary contribution to this literature is the focus on the matching of CEOs to firms in countries across different levels of development. An important contemporary paper is by Hjort, Malmberg, and Schoellman (2024), who show that the cost of hiring middle managers is a major barrier to growth in lower-income countries. While Hjort et al. (2024) focus on middle managers and their cost, we focus on CEOs and their behavior. Together, both papers show that obstacles to hiring the right managers at different levels can negatively impact productivity in middle- and lower-income countries.

Finally, we build on the literature relating factor misallocation to productivity differences across and within countries (Restuccia and Rogerson, 2008; Hsieh and Klenow, 2009). We contribute by examining misallocation in arguably one of the most important labor markets – that for CEOs. Like Hsieh and Klenow (2009), we focus on misallocation across rather than within firms. We explore a novel potential source (mismatch in managerial input) of total factor productivity differences across countries, and a mechanism for how human capital affects development through firms, building on Gennaioli et al. (2013) and Valero and Van Reenen (2019).

The rest of the paper is organized as follows. Section 2 introduces the data and main variables. Section 3 presents descriptive results of how the frequency of CEO types varies across countries and how CEO type correlates with firm performance. Section 4 presents a simple methodology for understanding the extent of mismatch between CEOs and firms. Section 5 presents results on the level of mismatch across different countries, the relationship between CEO-firm matching and productivity, and predictors of mismatch. Section 6 concludes.

2 Data and measurement

Our analysis relies on combining information on CEO time use, firm characteristics, and performance from the EBRD-EIB-WB Enterprise Survey (*"Survey data"*) with data on CEO behavior from CEO diary shadowing in Bandiera et al. (2020) ("Diary data"). We discuss both sources briefly below. Appendix Table A.1 presents all the countries included in our sample (ordered by GDP per capita), with the number of observations (firms) per country.

It also indicates which countries belong to each of the two datasets.

2.1 Diary data for high-income countries

Bandiera et al. (2020) developed a new shadowing method to collect high-frequency, high-dimensional diary data for CEOs. Specifically, they collected data on 1,114 CEOs of manufacturing firms in four high-income countries (France, Germany, the United Kingdom, and the United States with 476 firms) and two middle-income countries (Brazil and India with 638 firms) by interviewing the CEOs or their personal assistants twice a day during one workweek. They applied an unsupervised machine learning approach (LDA by Blei, Ng, and Jordan, 2001) to identify a small number of primitive behavioral types that are most likely to generate the observed behavioral patterns in the Diary data. The authors show that two behavioral types can sufficiently explain most of the variance and label the two "pure" behavioral types as "leader" and the "manager," as defined by Kotter (1990a). Leadership requires communication across a broad variety of stakeholders to create organizational alignment; the leader diary is characterized by high-level multifunction meetings. Management, on the other hand, consists of monitoring and implementation tasks and creating systems that enable precise and efficient execution of plans; the manager diary is characterized by many individual meetings with core internal firm functions. Bandiera et al. (2020) found that firms that hire leaders tend to have higher productivity and profitability, both across firms in a given year and within a firm over time (starting a few years after a leader-type CEO is appointed).

In this paper, we use the Diary data for two separate purposes. First, we use them to identify survey questions to cost-effectively collect information on CEO behavior across a wide range of countries, as discussed in the next subsection. Second, they allow us to build a benchmark for matching in high-income countries using data from France, Germany, the UK, and the US that we can then compare to matching patterns in middle-income countries, as discussed in Section 2.3.2.[1]

2.2 Collecting CEO behavior data in middle-income countries

While less expensive than in-person shadowing, the Bandiera et al. (2020) survey method is still prohibitively costly to carry out across a large number of countries. To expand the data collection cost-effectively across middle-income countries, we develop a parsimonious survey

[1]Several variables that were collected in the ES are not available for Brazil and India. In Appendix G.4, we show that our results are robust to including data from both countries (which were in the original Bandiera et al., 2020 sample) in the sample of middle-income countries whenever these variables are available.

instrument and add it to the ES.

2.2.1 Development of the simple survey tool

We leveraged the Diary data to select five time use variables that were most predictive of the continuous leader index constructed with the LDA algorithm in Bandiera et al. (2020) and could be included as questions in the ES (see Appendix C.1 for a description of the approach followed to identify the variables). The decision to focus on a handful of variables was based on the pragmatic consideration that we could not add more than five additional variables in the ES due to its overall length. These questions (listed in Appendix C.2) measure the number of times that the CEO engages in the following activities in a typical workweek: (i) meetings with C-suite executives, board members, business unit managers, or managers from a parent company; (ii) meetings with suppliers; (iii) meetings with employees involved in production activities; (iv) meetings involving more than one participant; and (v) meetings lasting longer than one hour. The probability of pure leader behavior increases in the frequency of all the above except for meetings with suppliers (ii) and production employees (iii). The last two activities are negatively correlated with the probability of pure leader behavior.

For these questions, respondents were given five response options, which are mapped to a scale from 1 to 5, with larger numbers corresponding to higher intensities. Appendix Figure C.1 describes the exact response options and the distribution of answers. Given that we added the questions to an existing large-scale survey that was soon to be conducted by the WB, EIB, and EBRD in 42 countries, there was unfortunately no possibility to conduct a cognitive testing of the new survey questions. The standard pilot of the entire ES questionnaire was, however, conducted in all countries and focused on checking that the translation was correct and the questions were appropriate for the local environment. Striking this balance between relatively cheap large-scale surveying and precision of the questions meant that one question did not measure the aspect that the Diary data dimension measured. Appendix C.3 details how we validated the questions and the reasons for excluding the question on meetings lasting longer than one hour.

2.2.2 Data collection

The ES, conducted face-to-face with firm managers, aims to measure the quality of the local business environment in terms of, for example, infrastructure, labor, and business-government relations. The survey also collects various firm characteristics and performance measures (for more details on the ES methodology, see Appendix B). Eligible firms were

selected using stratified random sampling. A total of 28,778 firms were interviewed between October 2018 and March 2020 in 42 countries across Europe, Central Asia, and the Middle East and North Africa, primarily covering the pre-pandemic period. Of these, 8,899 reported to have at least 50 employees during the screening interview and were thus eligible to answer the CEO time use questions.[2]

Of the 8,899 eligible firms (6,011 manufacturing and 2,888 services firms), our final sample contains 4,817 manufacturing firms with non-missing values for all the required variables – referred to as the Survey data. We restrict our analysis to manufacturing firms to maximize the comparability of our data, as the Diary data from Bandiera et al. (2020) does not include service firms. Appendix Tables D.2 and D.1 provide details on each sample's restrictions and how many observations are dropped as a result.

2.3 Constructing analysis samples and variables

2.3.1 The CEO time use index

We use the individual time use data collected in the ES survey to build a CEO time use index. We first calculate the average of the scores of the four time use questions, and then reverse the scale for the questions on the frequency of meetings with suppliers and production/sales employees, as these activities are negatively correlated with leader type (Bandiera et al., 2020). Appendix C.1 provides further details. The index takes values from 1 to 5, with larger values corresponding to a CEO more closely resembling the leader type.[3] Our simple leader index recovers a large portion of the variation in the continuous Bandiera et al. (2020) CEO behavior measure constructed with the LDA using the full diary shadowing data rather than just our four questions. The two indices exhibit a 49% correlation across the full Diary data.

[2]Fifty-two interviews in Armenia and 44 in Spain were conducted on the phone (1.1% of the Survey data). In Spain, 2 of the interviews were delayed and not completed until January 2022. In our results, we control for whether the interview was completed over the phone (which mostly happened during the pandemic) and for the fiscal year the interview referred to. Weights adjusting for representativeness are used in all country-level results when the regression or figure includes only Survey data countries. Similar weights are not available for the Diary data countries.

[3]While this is our preferred definition, we check robustness to using an alternative z-score definition. For this, we first normalize the scores for each question over the full sample so that they have a sample mean of 0 and a standard deviation of 1. We then add up the first four questions' values, which we normalize again to create an overall CEO time use index. This definition makes the index less dependent on the arbitrary 1–5 scoring of the time use questions but at the cost of the scoring becoming contingent on the sample, over which the normalization is done. Using this alternative definition, the results are qualitatively unchanged from the results presented in the paper.

2.3.2 Training and testing samples

The structural estimates of Bandiera et al. (2020) suggest that the share of firms with mismatched CEO types is likely to be low (approximately 5%) in the high-income Diary data countries. Based on this, we use a set of high-income countries as benchmarks for studying whether matching can be improved in the middle-income countries. Our hypothesis is that CEO behavior types are more misallocated in middle-income countries (building on evidence of misallocation in middle-income countries cited in the introduction) and productivity would be higher if CEO-firm allocation resembled higher-income country patterns. If we find that firms in middle-income countries not matched similarly to the benchmark have lower productivity than those matched according to prediction, it would support our hypothesis.

However, we do not solely rely on the high-income Diary data countries for this purpose, for two reasons. First, we want to make sure that any differences between the Diary and Survey data are not artificially created by the different data collection methods. Second, we want to increase power by including more than the 7.6% of firms (395 firms in the Diary data out of the total 5,212 firms in the pooled samples; see Table A.1) in the prediction.

Therefore, we rank the countries in the Survey and Diary data by GDP per capita and include all the high-income countries in the Diary data, as well as the Survey data countries with at least 90% of the GDP per capita of the lowest-income country in the Diary data (France). Using this rule, our high-income country *training sample* includes seven countries: the US, the UK, Germany, France, Malta, Italy, and Cyprus. The remaining countries in the Survey data are in the *testing sample*.

Tables 1 and 2 provide descriptive statistics for the training and testing samples, respectively, showing wide variation in sales and employee count in both samples.[4] [,5] Training sample firms are larger, both in terms of sales and number of employees, and are more productive on average. In the training sample, 12% of firms are listed and 50% are family-owned, and they have a mean age of 56 years. In the testing sample, 15% of firms are listed and 40% are family-owned, and the mean firm age is 26 years. In the testing sample, we have additional variables on firms' local labor market and local education environment.

[4]In the Diary data, firm performance information was sourced from Orbis for several years around the collection of the time use data. In the analysis, we use averaged values of available data points to decrease the potential influence of yearly fluctuations. As show in Appendix G.2, our results are robust to using the data available closest to the time use data collection. For 98% of firms, this is at most one year away.

[5]The baseline sample—which conditions on the availability of the predictor variables used in the analysis presented in the main paper—consists of 636 firms from the training sample and 4,576 firms from the testing sample. In robustness checks, we sometimes use an alternative set of predictors. In these sections, we change the sample so that the all predictor variables included in a particular section have no missing values (see Appendix D.1 for details). Some firm characteristics and education variables, not used for the prediction, have a few missing observations in the baseline sample. As these are less central to the analysis, we condition for their missingness them rather than drop them. This is discussed in Appendix D.2.

The final panel in both tables shows CEO characteristics, which are not strikingly different between the two samples. The CEOs in the training sample have on average 18 years of experience at the firm (not necessarily as CEOs), and 6% are women. The top managers in the testing sample have on average 23 years of experience in the sector and 14 years of experience at the firm, while 11% of them are women. Figure 1 presents the distributions of the CEO time use index in the training and testing samples, showing slightly higher values (indicating more leader-like behavior) in the training sample.

Table 1 Training sample summary statistics

Variable Name	N	Median	Mean	Std. Dev.	Min	Max
Predictors						
Sales (million USD)	636	41.30	228.99	1423.63	0.14	29,913.40
Number of employees	636	159.79	1,009.51	8,478.62	2.00	186,920.00
Log (labor productivity)	636	12.34	12.28	0.91	7.56	16.02
Listed firm (indicator)	636	0.00	0.12	0.33	0.00	1.00
Family-owned firm (indicator)	636	0.00	0.50	0.50	0.00	1.00
Establishment age	636	44.00	55.56	43.93	1.00	273.00
Predictors in robustness checks						
Family-managed firm (indicator)	635	0.00	0.26	0.44	0.00	1.00
Multi-establishment firm (indicator)	625	0.00	0.47	0.50	0.00	1.00
CEO characteristics						
Top manager female (indicator)	636	0.00	0.06	0.24	0.00	1.00
Top manager's years of experience at the firm	545	16.00	17.57	12.43	0.00	60.00

Notes: This table provides summary statistics for the training sample variables. Table A.1 presents all the countries included in our sample with the number of observations (firms) per country and an indication of the sample they belong to. *N* indicates the number of observations with non-missing observations; summary statistics are calculated on this sample. Predictor variables, when used, are required to be non-missing for all observations. Sample restrictions, discussed in Appendix D.1, result in 636 training sample firms being included in the baseline training sample. For Survey data firms, the number of permanent, full-time employees is reported. *Sources:* EBRD-EIB-WB Enterprise Survey and Bureau van Dijk's Orbis.

Table 2 Testing sample summary statistics

	N	Median	Mean	Std. Dev.	Min	Max
Predictors						
Sales (million USD)	4,576	5.95	26.26	180.99	0.001	7,356.25

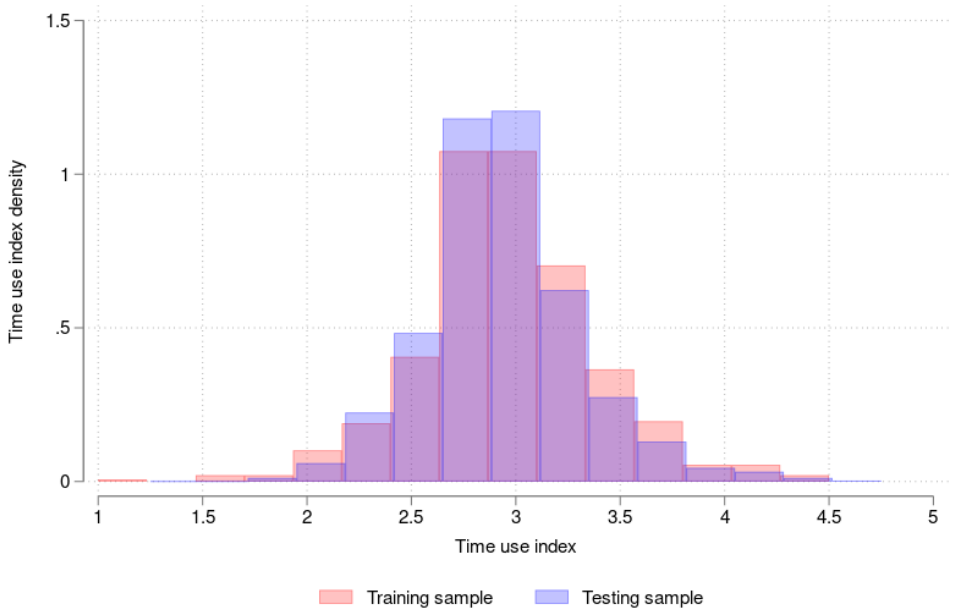
Continued on next page

Table 2 – continued from previous page

	N	Median	Mean	Std. Dev.	Min	Max
Number of permanent, full-time employees	4,576	130.00	246.23	500.72	5.00	20,000.00
Log (labor productivity)	4,576	10.66	10.49	1.55	0.95	17.91
Listed firm (indicator)	4,576	0.00	0.15	0.35	0.00	1.00
Family-owned firm (indicator)	4,576	0.00	0.40	0.49	0.00	1.00
Establishment age	4,576	22.00	26.46	20.20	1.00	202.00
Predictors in robustness checks						
Number of employees, screener	4,576	132.00	256.88	553.76	50.00	20,000.00
Family-managed firm (indicator)	4,543	0.00	0.17	0.38	0.00	1.00
Multi-establishment firm (indicator)	4,576	0.00	0.17	0.38	0.00	1.00
Firm characteristics						
Exporter or importer status (indicator)	4,481	1.00	0.81	0.39	0.00	1.00
At least 25% foreign ownership (indicator)	4,550	0.00	0.16	0.37	0.00	1.00
Has a board of directors or supervisory board (indicator)	4,572	1.00	0.62	0.49	0.00	1.00
Credit constrained (indicator)	4,576	0.00	0.20	0.40	0.00	1.00
Sole proprietorship (indicator)	4,539	0.00	0.09	0.28	0.00	1.00
Local labor market proxies						
Capital/main business city (indicator)	4,576	0.00	0.13	0.34	0.00	1.00
Population $>$ 1 million (indicator)	4,576	0.00	0.20	0.40	0.00	1.00
250,000 $<$ Population $\leqslant$ 1 million (indicator)	4,576	0.00	0.15	0.35	0.00	1.00
50,000 $<$ Population $\leqslant$ 250,000 (indicator)	4,576	0.00	0.19	0.39	0.00	1.00
Population $<$ 50,000 (indicator)	4,576	0.00	0.46	0.50	0.00	1.00
Education variables						
Min. distance to business degree HEIs (km)	4,576	11.38	32.29	56.01	0.05	723.65
Min. distance to non-business degree HEIs (km)	4,576	19.03	48.22	111.09	0.06	1,618.83
Min. distance to established business degree HEIs (km)	4,576	20.12	50.26	85.08	0.11	863.91
Min. distance to new business degree HEIs (km)	4,527	22.33	67.33	167.87	0.05	2,250.01
CEO characteristics						
Female top manager (indicator)	4,573	0.00	0.11	0.32	0.00	1.00
Top manager's years of experience at firm	2,070	13.00	14.26	6.22	1.00	55.00
Top manager's years of experience in sector	4,501	21.00	22.68	11.77	1.00	70.00

Notes: This table provides summary statistics for the testing sample variables. Table A.1 presents all the countries included in our sample with the number of observations (firms) per country and an indication of the sample they belong to. *N* indicates the number of observations with non-missing observations; summary statistics are calculated on this sample. Predictor variables are required to be non-missing for all observations. Sample restrictions, discussed in Appendix D.1, result in 4,576 testing sample firms being included in the baseline testing sample. When non-predictor variables are missing for these observations, they are imputed. Imputation of missing values for firm characteristics and education variables is discussed in Appendix D.2. *Sources:* EBRD-EIB-WB Enterprise Survey, World Higher Education Database, and authors' calculations.

Figure 1 Distribution of the CEO time use index



Notes: This figure shows the distribution of the CEO time use index of the 4,817 firms in the Survey data and 395 firms in the Diary data, divided into the training and testing samples.

2.3.3 Binary CEO type

To study the possible misallocation of CEO types to firm needs in a concise manner, we create a binary CEO type variable. We refer to CEOs with a CEO time use index above 3 as leaders and to those with a CEO time use index equal to or below 3 as managers (Kotter, 1990b; Bandiera et al., 2020).[6]

[6]We chose the threshold 3 for several reasons. First, since time use questions are scored on a scale from 1 to 5, 3 is the theoretical midpoint of our constructed CEO time use index. Second, it is also the median time use index in the full sample and is close to the mean of time use indices (2.93). Note that when we use the z-score-based time use index, we use 0 as the threshold of leaders versus managers.

Table 3 shows the percentage of leaders observed in the Survey and Diary data, as well as in the training and testing samples. While more than half of the Survey data (53.5%) and two-thirds of the Diary data (66.1%) consist of leaders, they are relatively less common in the middle-income Survey data. When the data are divided into the training and testing samples, the leader shares are more similar (54.0% in testing and 57.4% in training).

Table 3 CEO time use type shares across samples

	CEO time use type				
	Manager		Leader		Total
Sample	No. of firms	%	No. of firms	%	No. of firms
Diary data	134	33.9	261	66.1	395
Survey data	2,242	46.5	2,575	53.5	4,817
Training sample	271	42.6	365	57.4	636
Testing sample	2,105	46.0	2,471	54.0	4,576
Total	2,376	45.6	2,836	54.4	5,212

Notes: This table shows the percentages of leader and manager types across the Diary and Survey data as well as across training and testing samples. Leaders are CEOs with a CEO time use index above 3, while managers are CEOs with a CEO time use index equal to or below 3. Table A.1 presents all the countries included in our sample with the number of observations (firms) per country and an indication of the sample they belong to.

3 Leader frequency across countries and correlation with firm productivity

In this section, we investigate how common leaders are across countries at varying stages of development, and analyze whether on aggregate, leaders are associated with higher productivity in firms.

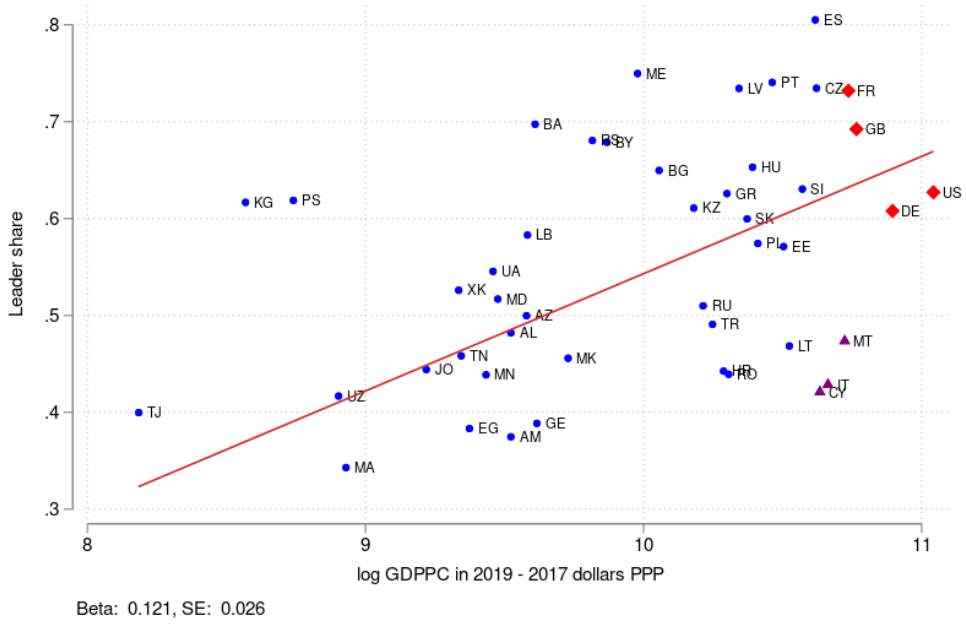
3.1 Leader distribution across countries

Using the binary leader variable, we find that the distribution of leader-type CEOs varies significantly across countries. At the country level, the share of leaders is positively associated with the log of countries' GDP per capita, as shown in Figure 2. In this figure we weight observations with the number of firms present in each country, but the relationship is similar without weights. Figure 2 also shows that the leader shares—conditional on income—are similar in the Survey data (blue circles and purple triangles) and the Diary data (red

diamonds) despite being measured differently.

Appendix Table G.1 shows the leader distribution by country income result while also controlling for a set of firm characteristics included in Bandiera et al. (2020) (log employment, firm age, listed status of the firm, family-owned status of the firm, and sector of the firm in groups of two-digit ISIC Rev 3.1).[7] There is still a significant positive association between country GDP per capita and leaders while controlling for these variables.

Figure 2 Correlation between log GDP per capita and leader share



Notes: This figure shows the correlation between a country's share of leaders and its 2019 GDP per capita (in constant 2017 international $). The linear fit estimate controls for the country's share of firms interviewed over the phone and uses the number of firms interviewed in each country as weights when estimating the linear fit. ISO country codes are linked to country names in Appendix Table A.1. The red diamond shapes represent Diary data countries in the training sample, purple triangles represent Survey data countries in the training sample, and blue dots represent Survey data countries in the testing sample.

3.2 Correlation of leader-type CEOs with firm productivity

In this subsection, we examine how CEO type correlates with firm performance. First, we do this in a specification that replicates the Bandiera et al. (2020) specification. We regress different measures of firm performance on the continuous CEO time use index, controlling for firm-level employment, capital, and material costs. To do this, we match firms in the Survey sample to Bureau van Dijk's Orbis data based on firm name and location, as the ES

[7]These variables are available in both datasets, and we use them for prediction in Section 4.1.

survey does not include capital and material costs.[8]

We estimate the following empirical model:

$$y_{ifts} = \alpha\hat{\theta}_i + \delta^E e_{ft} + \delta^K k_{ft} + \delta^M m_{ft} + \delta^{MNG} mng_{ft} + \psi_t + \nu_s + \epsilon_{ifts}, \quad (1)$$

where y_{ifts} is the log sales of firm f, led by CEO i, in period t and sector s; $\hat{\theta}_i$ is the time use index of CEO i; e_{ft}, k_{ft}, m_{ft} denote, respectively, the natural logarithms of the number of firm employees and, when available, capital and material costs; mng_{ft} denotes firms' management practices index; and ψ_t and ν_s are period and two-digit ISIC Rev 3.1 sector fixed effects (listed in Appendix E).

Table 4 shows that the CEO time use index (which increases with leader type) is positively and significantly associated with firm revenue, while controlling for employment to approximate labor productivity. This remains the case even when we include a measure of the quality of general management practices (column 4), which was also collected as part of the ES; both the CEO time use and general management practices are independently correlated with operating revenue. The CEO time index is thus distinct from other firm-wide management differences. Hence, we show that the CEO time use index based on four survey questions is correlated with firm performance in a large number of countries (corroborating results from fewer countries in Bandiera et al., 2020, Table 3).

Having established that the CEO time use index is correlated with performance for a subset of firms with additional information on capital, materials, and management (Table 4), we now explore the same relationship but expand the analysis to the broader sample (i.e., including firms for which we have data on sales and employment from the ES but no additional data from Orbis). We also examine this correlation using the coarser (binary) CEO-type (leader or manager) variable, which we use for the analysis of mismatches in subsequent sections of the paper. Specifically, we estimate

$$y_i = \beta_0 + \beta_1 * Leader_i + \beta_2 Leader_i \times Training_i + \zeta_c + \epsilon_i, \quad (2)$$

where y_i is the log sales, employment, or labor productivity of firm i; $Leader_i$ is equal to 1 if the CEO time use index of firm i is greater than 3, and 0 otherwise; $Training_i$ is an indicator for the observation belonging to the training sample rather than the testing sample; and ζ_c are country fixed effects.

The results in Table 5 indicate a positive correlation between leader-type CEOs and

[8]We can find an Orbis match for 4,576 manufacturing firms (out of 4,817 with the required survey variables), but only a subset of these have non-missing values for the required Orbis variables. As a result, only between 2,409 and 1,203 manufacturing firms are included in these regressions.

Table 4 Orbis firm outcomes and the CEO time use index (Survey sample)

Dependent variable →	Log (operating revenue)			
	(1)	(2)	(3)	(4)
Time use index	0.190***	0.145**	0.080***	0.072**
	(0.057)	(0.061)	(0.025)	(0.024)
Log (employment)	1.075***	1.030***	0.517***	0.513***
	(0.026)	(0.043)	(0.016)	(0.017)
Log (capital)		0.089***	0.044***	0.042***
		(0.017)	(0.011)	(0.011)
Log (materials costs)			0.436***	0.433***
			(0.030)	(0.030)
Management practices index				0.052**
				(0.019)
Country FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Observations	2,409	1,765	1,203	1,203
Adjusted R^2	0.733	0.738	0.885	0.886

Notes: This table presents OLS regressions to estimate the relation between the log transformation of firm-level operating revenue and the CEO time use index. The management practices index is based on the scoring of firms' management practices along five dimensions: 1) operations management, 2) performance monitoring, 3) target setting, 4) leadership initiative, and 5) talent management. The sample consists of all manufacturing firms in the Survey sample that can be matched to Bureau van Dijk's Orbis and for which the required variables do not have missing values. All regressions include country and sector fixed effects, an indicator for an interview held over the phone, fixed effects for the fiscal year the interview refers to and the day the interview was conducted, and a control for the duration of the interview. Standard errors clustered by sectors are reported in parentheses. ***, **, and * correspond to 1%, 5%, and 10% levels of statistical significance. *Sources*: EBRD-EIB-WB Enterprise Survey and Bureau van Dijk's Orbis.

employment, sales, and labor productivity in both the testing and training samples. Having a leader-type CEO is associated with 18% higher employment (column 1), 30% higher sales (column 3), and 12% higher sales per employee (labor productivity, column 5) in the testing sample, within country. The training sample coefficient only appears different ($p<0.1$) in terms of employment. Overall, this table shows that firms with leader-type CEOs generally perform better across a wide range of countries. This pattern may reflect sorting/matching of leader-type CEOs to more productive firms, an effect of leader-type CEOs on firms, or both.

Table 5 Firm characteristics and leader status

Dependent variable →	Log (employment)		Log (sales)		Log (labor productivity)	
	(1)	(2)	(3)	(4)	(5)	(6)
Leader	0.180***	0.161***	0.300***	0.264***	0.120***	0.103***
	(0.027)	(0.026)	(0.047)	(0.045)	(0.039)	(0.037)
Leader * Training Sample	0.169*	0.116	0.088	0.040	-0.081	-0.076
	(0.087)	(0.084)	(0.122)	(0.117)	(0.082)	(0.081)
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Sector FE	No	Yes	No	Yes	No	Yes
Listed firm	No	Yes	No	Yes	No	Yes
Family-owned firm	No	Yes	No	Yes	No	Yes
Firm age	No	Yes	No	Yes	No	Yes
Observations	5,212	5,212	5,212	5,212	5,212	5,212
Adjusted R^2	0.078	0.155	0.491	0.532	0.645	0.670

Notes: This table presents OLS regressions to estimate the relation between having a leader-type CEO and employment, sales, and sales per employee (labor productivity) in the pooled samples. Sales are in 2019 US$. Employment refers to the number of full-time, permanent employees. Leader is an indicator variable equal to 1 if the CEO time use index is above 3, and 0 otherwise. All regressions controls for whether the interview was conducted on the phone and the fiscal year it refers to. Table 2 provides summary statistics, and Table E.1 provides information on sectors. Robust standard errors are shown in parentheses. $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

4 Understanding CEO-firm matching: Methodology

4.1 Predicting which firms need leader-type CEOs

Under the assumption that talent allocation is more efficient in countries with higher income per capita, we can deduce "ideal" matching patterns by leveraging the correlation between

CEO type and firm characteristics for high-income countries in the training sample.[9] We use this to predict the "ideal" CEO type for firms in middle-income countries (the testing sample). The firm characteristics we can use for the prediction are five firm-level variables that are available in both samples. The first two variables relate to firm ownership, while the remaining three pertain to firm type: 1) whether the firm is listed on a stock exchange, 2) whether it is family-owned (defined as family-owned if a single family owns at least 25%), 3) industry (two-digit ISIC Rev 3.1 sector), 4) firm size (log of the number of employees), and 5) firm age.

We want to use these variables to create variation in the predicted leader-type CEO status without arbitrary assumptions about the included variables or their potential interactions. To achieve this, we allow the least absolute shrinkage and selection operator (LASSO) method to determine which of these variables and interactions are important predictors of whether the firm has a leader-type CEO in the training sample. We further allow LASSO to choose interactions of all variables except industry fixed effects, as the data cells would otherwise become quite small:[10]

$$Sector_i + Listed_i * FamilyOwned_i * Log(Employment_i) * FirmAge_i. \tag{3}$$

We predict leader-type CEO status using the selected variables, with coefficients penalized by LASSO, and a probit regression in the training sample.[11] The penalized prediction coefficients selected by LASSO are shown in Table 6.

The results indicate that while we allow LASSO great flexibility with the inclusion of a four-wise interaction, it arrives at a rather parsimonious specification, containing only one interaction (between firms' listed status and family ownership). This interaction can also be interpreted intuitively as showing that the family-owned status of a firm affects its CEO

[9]Our benchmark consists of high-income countries in the training sample. Structural estimates in Bandiera et al. (2020) suggest that 95% of CEOs in the Diary data's high-income countries are well matched. Nevertheless, matching frictions may exist in the CEO labor markets of affluent countries as well. In Appendix G.5, we investigate how one potential first-order friction—the tendency of family-owned firms to hire family CEOs (Bandiera, Lemos, Prat, and Sadun, 2018)—impacts our results. Our qualitative conclusions are unchanged when we separate the influence of this friction from the rest of the matching patterns.

[10]Appendix G.7 reports the number of observations in each possible interaction cell that could be selected by LASSO in both the training and testing samples. Including interactions between all sector codes and the other four variables would lead to cells with very few observations. In Appendix G.8, we arrive at qualitatively similar results even when allowing interactions with sector codes. However, we believe that our main specification is more reliable.

[11]In the linear model, with larger sample sizes, post-selection (unpenalized coefficients) may have better out-of-sample prediction performance than the penalized coefficients (Belloni and Chernozhukov, 2013). However, there is no theory for non-linear models. We select the coefficients with a lower mean square error (MSE)—the penalized coefficients. Nevertheless, we show in Appendix G.9 that using the post-selection coefficients leads to qualitatively similar results.

type needs only if the firm is not publicly listed. Potentially, this reflects that such non-listed family-owned firms have distinct funding needs, which are better addressed by manager-type CEOs. Indeed, Mullins and Schoar (2016) suggest that CEOs of family-owned firms are more likely to behave like "manager type" CEOs in that they supervise tasks more closely, have fewer managers reporting directly to them, value sector-specific knowledge more, and have closer relationships with banks than CEOs of non-family-owned firms.

Table 6 Predicting leader status with firm characteristics

Dependent variable →	Leader-type CEO status
Non-listed firm	-0.081
Non-listed firm * Family-owned firm	-0.361
Log (employment)	0.152
Manufacture of food products and beverages	-0.112
Manufacture of textiles	-0.330
Manufacture of rubber and plastics products	0.239
Manufacture of fabricated metal products	-0.004
Manufacture of office, accounting and computing machinery	0.297
Constant	-0.448
Observations	636

Notes: This table presents penalized coefficients from a LASSO probit estimating the relationship between leader-type CEO status and sector, and the interaction between listed status, family-owned status, log employment, and firm age in the training sample. Coefficients are selected by cross-validation. The table shows penalized coefficients from LASSO without standard errors as there is no consensus on how to define them (and they are not important for prediction). *Sources*: EBRD-EIB-WB Enterprise Survey.

Table 7 in turn shows the average marginal effects from regressing these selected variables (without interactions) on leader status in a probit regression in the training sample.[12] Larger firms and firms in some high-tech industries (computing machinery) are more likely to have a leader-type CEO, while non-listed and family-owned firms are less likely to have a leader-type CEO.

4.1.1 Applying the prediction to middle-income countries

Using the prediction with penalized coefficients (Table 6), we next predict which firms need leader-type CEOs in the testing sample (middle-income countries). We use a probit regression, meaning that our predicted ideal "leader-type CEO propensity" (interpreted as a firm-level characteristic) is a continuous variable bounded between 0 and 1. Regressing the

[12]In contrast to Table 6, we report non-penalized coefficients with standard errors in Table 7, as we aim to show how our predictors relate to leader-type CEO status.

Table 7 Relationship between actual leader-type CEO and LASSO-selected firm characteristics

Dependent variable →	Leader status
Non-listed firm	-0.127**
	(0.064)
Family-owned firm	-0.132***
	(0.035)
Log (employment)	0.066***
	(0.016)
Manufacturing of food products and beverages	-0.080
	(0.053)
Manufacture of textiles	-0.249*
	(0.149)
Manufacture of rubber and plastics products	0.158
	(0.101)
Manufacture of fabricated metal products	-0.043
	(0.050)
Manufacture of office, accounting and computing machinery	0.131**
	(0.060)
Observations	636
Pseudo R^2	0.081

Notes: This table presents average marginal effects from the probit regression to estimate the relation between firm characteristics and leader-type CEO status in the training sample, using the LASSO regression analysis method. Robust standard errors are shown in parentheses. ***, **, and * correspond to 1%, 5%, and 10% levels of statistical significance. *Sources*: EBRD-EIB-WB Enterprise Survey.

firm's actual leader-type CEO status on the predicted leader-type CEO propensity, we find a significant positive relationship between these two variables (Table 8) in the testing sample.[13] That is, there are commonalities in the underlying matching patterns between CEO types and firms in the high-income countries of the training sample and the middle-income countries of the testing sample. However, as Figure 3 shows, the accuracy of the prediction is not uniform across countries, as the figure indicates that the root mean square error (RMSE)

[13]The regressions that include the predicted leader-type CEO status have country fixed effects throughout. We prefer not to include sector fixed effects in our main specifications as some of the sector indicators were selected by LASSO for prediction. However, because this is not true for all sector categories, we have ascertained that controlling for sector and the other unselected predictor, CEO age, does not qualitatively change the results, as shown in Appendix G.10. Appendix G.11 further shows that using regional (NUTS1) fixed effects instead of country fixed effects does not change the results qualitatively. Moreover, all results incorporating predicted variables use bootstrapped standard errors with 1,000 replications without clustering. Appendix G.12 further shows that clustering bootstrapped standard errors at the locality level to capture potential interactions between firms in local markets also does not change the results qualitatively.

decreases with GDP per capita.[14]

Table 8 Relationship between actual leader-type CEO status and predicted leader propensity

Dependent variable →	Leader status
Leader propensity	0.247***
	(0.070)
Country FE	Yes
Observations	4,576
Adjusted R^2	0.061

Notes: This table presents estimates from an OLS regression to estimate the relation between predicted leader-type CEO propensity and actual leader-type CEO status. Leader-type CEO status is an indicator variable equal to 1 if the CEO time use index is greater than 3, and 0 otherwise. The regression controls for whether the interview was conducted on the phone and the fiscal year it refers to. Bootstrapped standard errors with 1,000 replications are shown in parentheses. ***, **, and * correspond to 1%, 5%, and 10% levels of statistical significance.

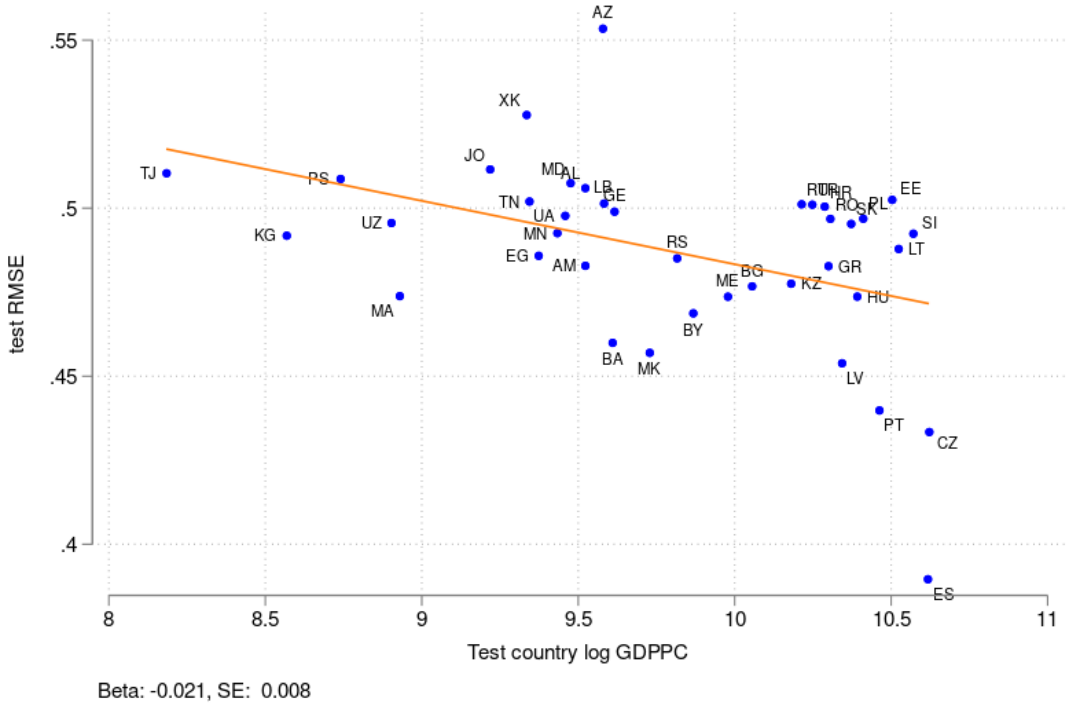
These patterns could be explained in two alternative ways. First, firms' needs may be similar across countries conditional on prediction covariates, but matching quality may vary with the level of development, which aligns with our assumption. Alternatively, firms in different countries may have different CEO needs conditional on prediction covariates. Table 8 provided some evidence that matching follows a pretty similar structure across countries. The next section provides more conclusive evidence of the mismatch hypothesis by showing that firms that deviate from the predicted CEO type have worse performance than those that conform to it.

To analyze the productivity differences between mismatched firms and those matched according to prediction, we first identify firms with correctly or incorrectly matched CEOs by mapping the continuous predicted values into a dummy variable equal to 1 if the predicted leader propensity is greater than or equal to 0.5.[15] The share of actual leader-type CEOs across firms in the testing sample is 54.0%, but the prediction assigns the ideal leader-type

[14]Appendix Figure G.5 shows the same prediction exercise with the training sample consisting only of US firms. We conduct this exercise to examine if the prediction performs similarly well for countries with similar income levels regardless of whether they belong to the training or the testing sample. The prediction errors of the UK, Germany, France, Cyprus, Malta, and Italy do not appear qualitatively different from that of higher-income testing sample countries. This is true despite the fact that the data in the UK, Germany, and France were collected differently than the data in the testing sample countries.

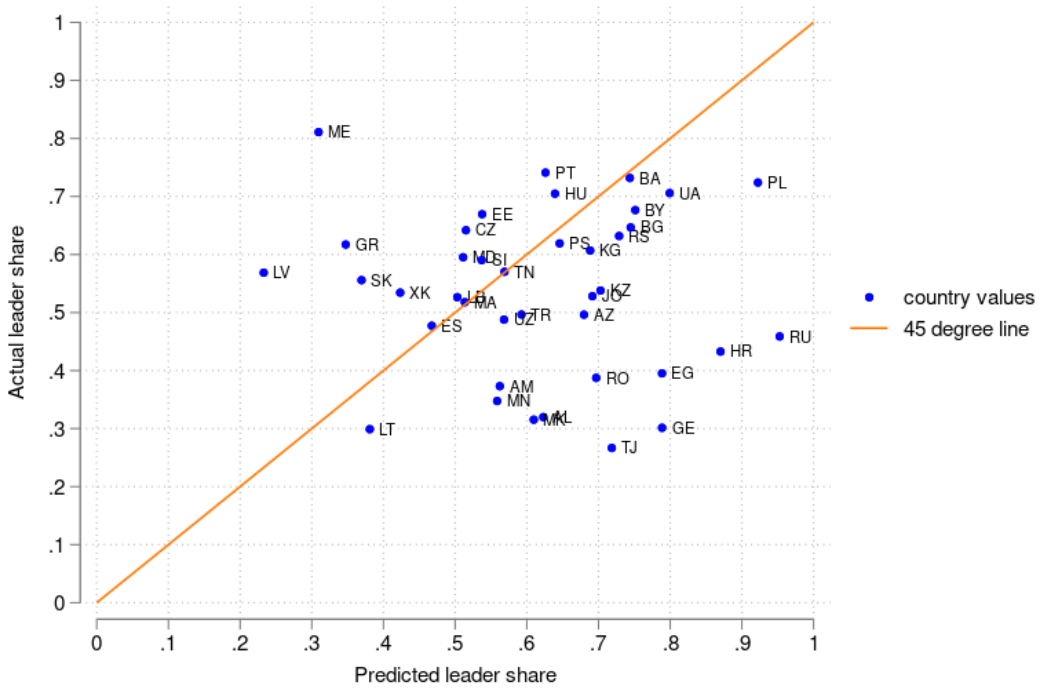
[15]As shown in Appendix G.14, the results are largely robust to using the share of leaders in the sample (0.55) instead of 0.5 as the threshold.

Figure 3 Out-of-sample prediction and log GDP per capita (training sample)



Notes: This figure shows the scatterplot of the log GDP per capita and root mean square error (RMSE) from predicting the type of CEO a firm needs using the estimates of Table 6 and the full training sample. The RMSE is calculated by regressing the actual leader status of firms in the country on the predicted leader-type CEO propensity measure. A list of countries in the samples can be found in Appendix A.

Figure 4 Actual and predicted leader-type CEO share across testing sample countries



Notes: This figure shows the actual and predicted leader-type CEO shares across countries in the testing sample.

CEO to 69.1% of the firms. This discrepancy is further illustrated in Figure 4, which shows that at the country level, most testing sample countries have lower actual leader shares than what they need based on the prediction.

5 Main results

5.1 CEO-firm matching and productivity

If a deviation from the predicted ideal CEO type captures mismatch, then this would be reflected in a firm's productivity. Therefore, we estimate the following regression at the firm level:

$$Y_i = \beta_0 + \beta_1 Mismatched_i + \zeta_c + \epsilon_i, \tag{4}$$

where Y_i is firm i's log labor productivity, $Mismatched_i$ is an indicator equal to 1 if the actual and the ideal CEO types are mismatched—that is, the firm is predicted to need a leader but has a manager, or vice versa—and equal to 0 if they match, and ζ_c are country

fixed effects.

When a firm deviates from the predicted ideal CEO type, it has significantly lower labor productivity (16% lower—see column 1 of Table 9). To test for horizontal differentiation, we further break the mismatch into "Needs Leader, Gets Manager" and "Needs Manager, Gets Leader." Column 2 of Table 9 shows that firms that are predicted to need a leader but obtain a manager-type CEO have 18% lower labor productivity than firms with correctly matched CEO type.

Table 9 Productivity and CEO type mismatch

Dependent variable →	Log (labor productivity)		
	(1)	(2)	(3)
Mismatched	-0.157***		
	(0.037)		
Needs Leader, Gets Manager		-0.177***	-0.197***
		(0.047)	(0.047)
Needs Manager, Gets Leader		-0.122**	-0.145***
		(0.050)	(0.054)
Needs Manager, Gets Manager			-0.076
			(0.064)
$\hat{\beta}_{LM} - \hat{\beta}_{ML} = 0$ (p-value)		0.324	0.344
Country FE	Yes	Yes	Yes
Observations	4,576	4,576	4,576
Adjusted R^2	0.342	0.342	0.342

Notes: This table presents OLS regressions to estimate the relation between the (mis)match of the CEO type the firm has and the CEO type it needs, on the one hand, and labor productivity, on the other hand. Mismatched is an indicator equal to 1 if the firm's ideal and actual CEO types are mismatched and 0 otherwise. Needs Leader, Gets Manager is an indicator equal to 1 if the firm's ideal CEO type is a leader and its actual CEO type is a manager, and 0 otherwise. LM refers to Needs Leader, Gets Manager, and ML refers to Needs Manager, Gets Leader. The regression controls for whether the interview was conducted on the phone and the fiscal year it refers to. Table 2 provides summary statistics. Bootstrapped standard errors with 1,000 replications are shown in parentheses. ***, **, and * correspond to 1%, 5%, and 10% levels of statistical significance.

Column 2 also shows that firms that are predicted to need a manager-type CEO but get a leader-type CEO have 12% lower labor productivity than firms with correctly matched CEO type. The estimated coefficients on the two mismatched categories are not statistically different from each other, as shown by the p-value in the row labeled "$\hat{\beta}_{LM} - \hat{\beta}_{ML} = 0$ (p-value)" in Table 9. The two negative coefficients, and the fact that they are not different,

suggests there is horizontal differentiation among CEOs: leader-type CEOs are better suited to some firms, while manager-type CEOs are better suited to others.

In column 3, we split the omitted category of correctly matched CEO types and firms into two categories, "Needs Manager, Gets Manager" and "Needs Leader, Gets Leader," and omit the latter. While correctly matched firms whose ideal CEO type is manager have slightly lower labor productivity than firms with a leader-type CEO and whose ideal CEO type is leader, this difference is not statistically significant. This result provides further evidence of horizontal differentiation. Moreover, it suggests that the overall benefit of leaders that appeard in the data when not considering firm needs (Table 3), stems from the fact that leaders are in low supply. That is, some firms with manager-type CEOs have lower productivity not because manager-type CEOs are inherently worse but because more mismatched firms need a leader-type CEO.[16]

5.1.1 Back-of-the-envelope calculation of total productivity loss

We use the estimates in Table 9 to calculate the total productivity loss from CEO-firm mismatch and from a low share of leaders in the sample. We calculate this by first deriving the counterfactual sales (S_i^C) for each firm under the hypothesis of correct allocation of CEO types to firms using the following formula:

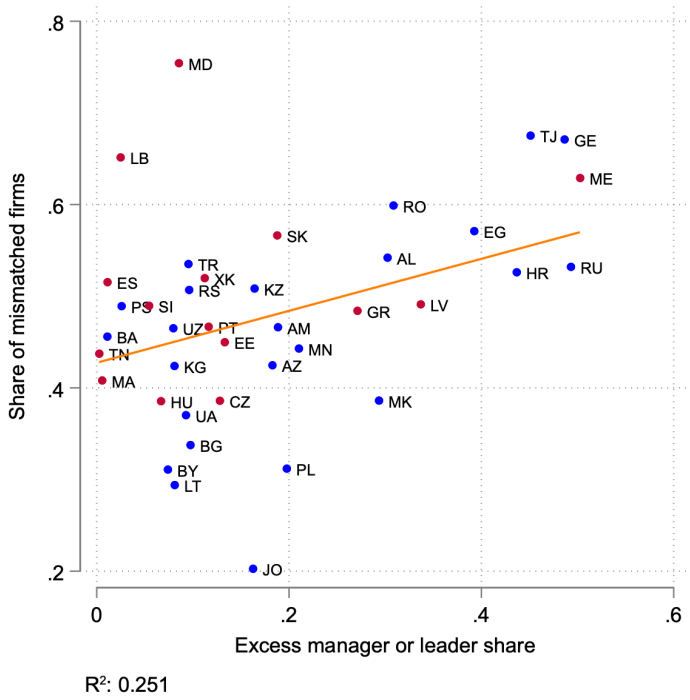
$$\begin{aligned} S_i^C &= S_i * [\mathbb{1}_{\text{Type}_i=\text{MM}} + \mathbb{1}_{\text{Type}_i=\text{LL}} \\ &\quad + \mathbb{1}_{\text{Type}_i=\text{LM}}/(1+\hat{\beta}_{LM}) + \mathbb{1}_{\text{Type}_i=\text{ML}} * (1+\hat{\beta}_{MM})/(1+\hat{\beta}_{ML})], \end{aligned} \tag{5}$$

where S_i represents the actual sales of firm i; $\mathbb{1}(.)$ is the indicator function; $Type_i = XY$ captures the firm's matching type, with X denoting the needed CEO type and Y the actual CEO type (M: manager and L: leader); and the $\hat{\beta}_{XY}$ coefficients come from Table 9. When the actual and the needed CEO type are the same ($X = Y$), the counterfactual sales are equal to the actual sales. In the case of mismatch ($X \neq Y$), the firm's sales are adjusted by the productivity impact of the mismatch. The derived counterfactual values are aggregated using the sampling weights to calculate the mean counterfactual sales in our sample. Comparing this counterfactual to the actual mean sales suggests that correcting the allocation would lead to a 9.1% increase in sales in the testing sample countries.[17]

[16] In this section, we use a binary variable to capture whether firms need leader- or manager-type CEOs. In Appendix G.15, we show that variation in the predicted $LeaderPropensity_i$ index can be used to also capture how much firms need their ideal CEO type. For this, we calculate the intensity of the mismatch for each firm as $|Leader_i - LeaderPropensity_i|$, which we then group into quartiles over the testing sample. We find that the more intensively mismatched firms are according to this measure, the lower their productivity. This is true for firm with both leader- and manager-type CEOs.

[17] Appendix Table G.39 shows the number of firms and total sales in each group.

Figure 5 Relationship between the share of mismatched firms and excess manager or leader share



Notes: The share of mismatched firms is the share of firms where the actual CEO type differs from the ideal CEO type. Excess manager or leader share is the absolute value of the difference between the share of firms with a manager-type CEO and the share of firms whose ideal CEO type is manager. Blue symbols indicate countries with excess manager shares, cranberry symbols indicate countries with excess leader shares, and the orange line indicates a linear fit.

In theory, this gain could be achieved by supplying all firms with their needed CEO type. However, the actual share of leader-type CEOs is lower than predicted in many countries. Could part of the productivity gap be closed by reallocating existing leader-type CEOs in a country, or does closing the gap entirely rely on increasing the supply of leader-type CEOs? Figure 5 shows that across countries, the share of mismatched firms and the excess manager or leader share (i.e., the difference between the share of firms with an abundant CEO type compared to the share of firms that need that CEO type) is not one-to-one (the estimated coefficient is 0.315, and the R^2 is 25.1%). This suggests that countries' productivity can be increased both by reallocation on the extensive margin (increasing the supply of the scarce CEO type) and the intensive margin (reallocating mismatched CEOs given the distribution of CEO types).

Since this distinction is policy-relevant, we examine the share of total productivity gains that could be achieved only by reallocating CEOs within countries without changing the supply of leaders and managers in a country. This scenario would apply to the short run,

while the earlier 9.1% gap could only be closed in the long run through an increase in the supply of leaders.

We study two short-run counterfactual reallocations. In the first, CEOs of the scarce type are randomly reallocated within each country to firms that are predicted to need them. This can be understood as a queuing allocation mechanism, where firms announce their needed CEO type. In the second, CEOs of the scarce CEO type are reallocated within each country to firms that need them most based on the ranking of the continuous leader-type CEO propensity index described in Section 4.1.[18]

To calculate the gains that could be achieved in these reallocations, we first identify firms by the type of CEOs they would receive in a particular short-run reallocation. Combined with data on the types of CEOs firms currently have and need, this distinction creates eight categories of firms.[19] Using these eight categories, we conduct counterfactual exercises similar to those in equation 5. However, instead of applying the coefficients to all the mismatched firms, we limit their application to those expected to change CEOs in the short term.

We then aggregate the firm-level counterfactual values using sampling weights to estimate the overall short-run effects. The result suggests that in the case of randomized reallocation, the short-run reallocation would lead to a 3.8% increase in productivity, less than half of the 9.1% long-run effect. When CEOs are matched with firms estimated to need them the most, short-run gains would amount to a 6.0% increase in productivity, more than half of the long-run effect. These findings indicate that substantial gains can be achieved in short-run reallocation—addressing misallocation alone, without changing the supply of leaders, accounts for around half of the total productivity loss.

5.2 Predictors of mismatch and low leader supply

We have shown that mismatches between actual and ideal CEO type are negatively associated with firm performance. In this section, we investigate the underlying factors associated with these mismatches and the low supply of leaders. First, we study to what extent mismatch is explained by country or industry, and in the following subsections we lay out other hypothesized explanatory variables and test them.

[18] More concretely, we define the firms that can get a leader in the short run, i.e. without increasing the share of leaders in the country, as $ShortRunLeader_i = \mathbb{1}[LeaderTypeC\hat{E}OPropensity_i > \phi_C]$, where i indexes a firm in country C, $\mathbb{1}(.)$ is the indicator function, $LeaderTypeC\hat{E}OPropensity_i$ is the estimated leader-type CEO propensity, and ϕ_C is a country-specific value taking the estimated leader-type CEO propensity value of the firm whose percentile rank is equal to the share of leaders in the country.

[19] Appendix Tables G.40 and G.41 show firm counts and total sales in each of these categories.

5.2.1 Is the relevant CEO market on the country level?

First, just as there are large barriers to mobility of other factors in production across regions wihtin countries (Gennaioli, LaPorta, de Silanes, and Shleifer, 2014), CEO mobility within a country may not be perfect. Moreover, firm-specific human capital and asymmetric information may further impact mobility of CEOs within countries (Cziraki and Jenter, 2022). For example, Sauvagnat and Schivardi (2024) use Italian matched employer-employee administrative data for all private firms with at least one employee and find that the relevant market for executives occurs at the sector-location (commuting zone) level. Yonker (2017) suggests that CEOs are often hired locally even by the largest US firms. In line with this, Panel A of Table 10 shows that country fixed effects explain only 2.1% of the variance in firm-level mismatch (deviation of actual from the ideal CEO type), while sector fixed effects explain even less (0.6%). The percentage of explained variance in firm-level mismatch increases to 10.4% when using country×sector-level fixed effects, yet it still leaves almost 90% of the variance unexplained. Similarly, while the explanatory power of country and sector is somewhat stronger for leader supply than for mismatches (Panel B of Table 10), a large portion of the variation is still left unexplained, justifying further analysis at the local market and firm level.

Table 10 Explained percentage of variance in firm-level mismatch and leader supply

	R^2 (%)	Adjusted R^2 (%)
Panel A: *Mismatched status*		
Country FE	2.1	1.3
Sector FE	0.6	0.4
Country×Sector FE	10.4	2.4
Panel B: *Leader-type CEO status*		
Country FE	6.5	5.8
Sector FE	0.5	0.3
Country×Sector FE	14.8	7.3

Notes: This table shows the percentage of variation in the firm-level deviation of actual CEO type from the ideal (predicted) CEO type (panel A) and leader-type CEO status (panel B) explained by country, sector, and country×sector fixed effects.

5.2.2 Hypothesized explanatory variables at the local and firm levels

Firm characteristics, the size of the local labor market, and proximity to business and/or management schools may play a role in predicting firm-level mismatch and the local supply of CEO types. In the following, we first describe some of these predictors and then conduct

a firm-level analysis of matching frictions.

Firm characteristics When making a hiring decision, a firm must be aware of the CEO type it needs to select the appropriate candidate. We hypothesize that the ideal and actual CEO types are more likely to match infirms that participate in international trade or are exposed to best practices through foreign ownership or the presence a board of directors or supervisory board. Additionally, credit-constrained firms may find it more difficult to attract the correct CEO type, as they may have difficulties offering a sufficiently high pay package.[20] On the other hand, in a sole proprietorship, the CEO is the sole proprietor, by definition.

Local labor market We hypothesize that firms located in the capital or main business city will have access to a larger pool of potential executives and greater exposure to international influences.

Proximity to business education Being a CEO requires a wide range of skills, including managing financial and operational resources; technical, administrative, problem-solving, and strategic thinking skills; and leadership and management skills (Sadun, Fuller, Hansen, and Neal, 2022), some of which are taught in business degree programs. Since business education likely focuses on teaching skills relevant to generalized, large firms, we expect that high-quality business education emphasizes leadership skills and could therefore increase the supply of leader-type CEOs. While we do not have information on business education of CEOs in the Survey data, we have detailed histories for CEOs in the Diary data. Using these data, we identify CEOs who attended MBA programs and find a significant positive relationship ($p < 0.05$) between MBA attendance and leader-type behavior—as measured by the continuous time use index or the leader CEO dummy—in this sample. This holds even after controlling for firms' countries.

To test the hypotheses that proximity to high-quality business education is associated with (i) a higher supply of leader-type CEOs and (ii) a lower mismatch between the actual CEO type a firm currently has and the type it needs, we collect data on HEIs and their characteristics (such as degrees offered, divisions, fields of study, year founded, public/private status) for the countries in our sample.[21] Figure 6 shows the number of schools offering business degrees in our testing (Panel A) and training (Panel B) samples split by ownership

[20] Credit-constrained firms are those that were either discouraged from applying for a loan or were rejected when they applied. See, for example, Beck, Degryse, De Haas, and Van Horen (2018).

[21] See Appendix F for details on these data and the construction of related measures.

(public versus private) over time (by the year the school was founded).[22] Two interesting patterns can be observed. In the training sample, the number of schools offering business degrees follows a linear trend, with no major differences between the number of public and private schools. In the testing sample, the number of public schools offering business degrees also follows a roughly linear trend in non-communist countries. In contrast, the trends in the number of private schools offering business degrees in all testing countries are anything but linear. The fast growth in the number of schools providing business education after the dissolution of Communism in 1990 is related to the liberalization of tertiary education (see Appendix F for more detailed background information). When there is such fast growth in schools, we expect that the average quality of newly formed schools will be lower than that of schools established before 1990 (henceforth "established schools").[23] While not much research is available on this, Domadenik, Drame, and Farčnik (2010) use a matched employer-employee micro dataset of 2007 business graduates in Slovenia and find that the probability of becoming employed after graduation was lower for those who graduated from newer private HEIs than from public HEIs (see Appendix F for more details).

Our hypothesis is that firms located near schools offering business degrees are more likely to find a CEO who matches their ideal CEO type, particularly as it is likely to increase the supply of leader-type CEOs which are often in short supply. Furthermore, based on the patterns in Figure 6, we expect that this effect will be driven by established schools.

5.2.3 Results

To test our hypotheses, we calculate the within-country distance from each firm to the nearest HEI of a particular type (such as business schools established before 1990). We focus on within-country distance for two main reasons. First, frictions along borders, such as language barriers or visa requirements, are likely to increase the cost of attending a particular HEI type abroad discontinuously, which suggests a focus on within-country education. Second, going beyond borders would require us to process education data from countries outside our sample, which would be challenging to execute carefully.[24]

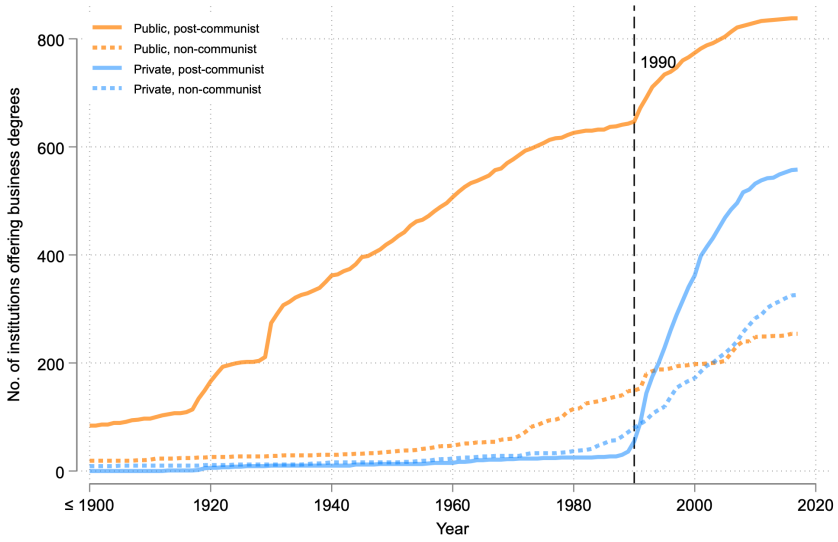
[22] We lack information on the year schools started offering business degrees.

[23] This is supported by the data QS World University Rankings. Only a fraction of HEIs in our sample are covered by rankings, but established schools are about three times more likely to be ranked in QS's world or regional rankings than new schools. This is true even when controlling for the schools' countries.

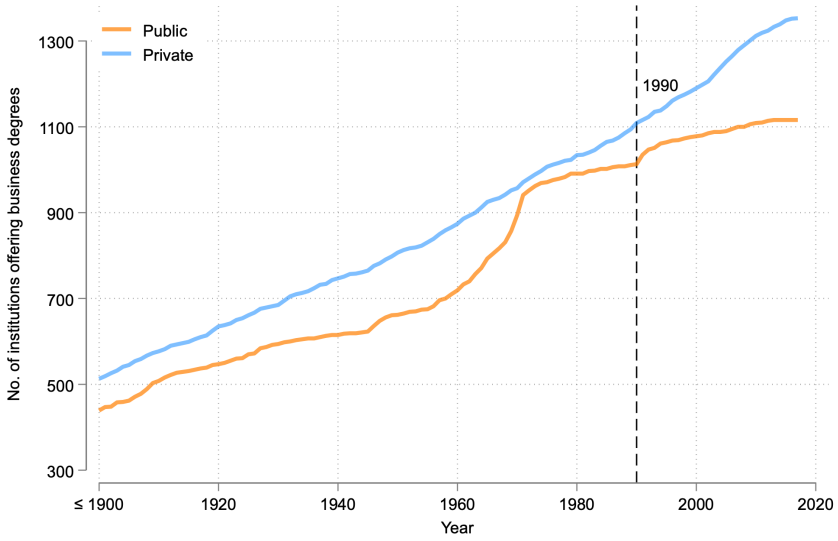
[24] For one country, Estonia (with only 49 firms), this strategy leads to missing minimum distance information from new schools offering business degrees, as this school type is unavailable in the country. For these firms, we impute distance information with a constant value, which is captured in the estimation by country fixed effects. We ensure that the results are unchanged if we exclude imputed observations from the estimation.

Figure 6 Number of schools offering business degrees over time

Panel A: Testing sample



Panel B: Training sample



Notes: Year refers to the year of establishment of the higher education institution to which the business school belongs, if the school is not a stand-alone entity. Appendix F provides more details on the data collection and definitions, and Appendix A presents a list of the countries in the testing and training samples. *Source*: Authors' calculations based on World Higher Education Database, Association to Advance Collegiate Schools of Business, Association of MBAs, European Quality Improvement System, and Kosovo Accreditation Agency.

We estimate the following firm-level regression:

$$M_i = \gamma_0 + \gamma_{\mathbf{E}}{}' \mathbf{E_i} + \gamma_{\mathbf{L}}{}' \mathbf{L_i} + \gamma_{\mathbf{F}}{}' \mathbf{X_i} + \zeta_c + \epsilon_i, \tag{6}$$

where M_i is a measure of mismatch for firm i, $\mathbf{E}$ are local education market characteristics (proximity to HEIs offering business degrees and other types of HEIs),[25] $\mathbf{L}$ are proxies for local labor market thickness (capital or main business city, size of the locality), $\mathbf{X}$ are firm-level characteristics (exporter/importer status, foreign ownership, sole proprietorship, board of directors/supervisory board presence, credit constraints), and ζ_c are country fixed effects. To study both misallocation and leader supply, we use two outcome variables: an indicator equal to 1 if the actual and ideal CEO types for firm i are mismatched, and an indicator equal to 1 if the firm has a leader-type CEO.[26]

Table 11 shows that distance to the nearest HEI offering business degrees has an overall weakly positive effect on mismatch, meaning that firms closer to such schools are less likely to experience a mismatch. Moreover, firms located farther from these schools are less likely to have a leader-type CEO. This aligns with our hypothesis that business education increases the supply of leader-type CEOs.[27]

Furthermore, Table 12, which has the same controls as Table 11, reveals that these effects are driven primarily by established universities offering business degrees, which are likely of higher quality. This provides further evidence that high-quality business education is indeed associated with a greater supply of leaders. We also control for proximity to other HEIs that do not offer business degrees.

The estimates in Table 11 confirm that being located in the capital or main business city affects leader supply. In particular, firms in these locations are more likely to have a leader-type CEO. According to Appendix Table G.42, this increase in leaders is also reflected in a significantly lower number of firms that need leaders but get managers, as well as a weakly larger number of firms that need managers but get leaders. In contrast, large cities that are not the main business city appear to have fewer leaders.[28] This suggests that there is something particular about capitals or main business cities which affects leader supply, possibly international influences, rather than the size of the city.

The results in Table 11 further indicate that several firm characteristics are correlated with mismatch and leader status. Firms that are exporters or importers, have at least 25% foreign ownership, or have a board of directors or a supervisory board are associated with a

[25] As shown in Appendix F, the distance measures are highly skewed, and therefore they are included in the logged form in our estimation. Appendix Table G.44 shows a version of the estimation where distance measure is included in levels.

[26] Appendix G.18 discusses a version of the estimation where only education market predictors are included.

[27] Appendix Table G.42 presents estimates with other measures of matching outcomes on the left-hand side as well: Needs Leader, Gets Manager; Needs Manager, Gets Leader; and Needs Leader. The results here indicate that the weakly positive association between mismatched status and the distance from schools offering business degrees is primarily driven by an increase in the number of firms that need leaders but get managers, an outcome significantly correlated with distance ($p < 0.05$).

[28] Most of the capital and main business cities in our sample have more than one million inhabitants.

Table 11 Firm-level mismatch predictors

Dependent variable →	Mismatch	Leader-type CEO
	(1)	(2)
Local education market		
Log minimum distance from HEIs	0.011*	-0.019***
offering business degrees	(0.006)	(0.006)
Log minimum distance from other	0.002	-0.001
HEIs	(0.006)	(0.006)
Local labor market thickness proxies		
Capital/main business city	-0.023	0.092***
	(0.029)	(0.029)
Population >1 million	0.068**	-0.079***
	(0.029)	(0.030)
250,000 < Population ⩽ 1 million	0.011	0.010
	(0.026)	(0.026)
50,000 < Population ⩽ 250,000	0.013	-0.006
	(0.023)	(0.023)
Firm-level characteristics		
Exporter or importer status	-0.039*	0.111***
	(0.021)	(0.019)
At least 25% foreign ownership	-0.059***	0.047**
	(0.021)	(0.021)
Has a board of directors or	-0.053***	0.044***
supervisory board	(0.017)	(0.016)
Credit constrained	0.005	-0.027
	(0.019)	(0.019)
Sole proprietorship	0.100***	-0.108***
	(0.029)	(0.028)
Country FE	Yes	Yes
Observations	4,576	4,576
Adjusted R^2	0.022	0.080

Notes: The regression controls for whether the interview was conducted on the phone and the fiscal year it refers to. For observations with missing education market predictors, we impute constant values within countries, which are captured by country fixed effects. For other categorical predictors, we include an additional (unreported) category if the value was originally missing (see Appendix D.2 for more details). Table 2 provides summary statistics. Bootstrapped standard errors with 1,000 replications are shown in parentheses. ***, **, and * correspond to 1%, 5%, and 10% levels of statistical significance.

Table 12 Firm-level mismatch predictors

Dependent variable →	Mismatch	Leader-type CEO
	(1)	(2)
Local education market		
Log minimum distance from established	0.017***	-0.029***
HEIs offering business degrees	(0.006)	(0.007)
Log minimum distance from new HEIs	-0.001	0.001
offering business degrees	(0.006)	(0.006)
Log minimum distance from other HEIs	0.001	0.000
	(0.006)	(0.006)
Firm-level characteristics	Yes	Yes
Local labor market thickness proxies	Yes	Yes
Country FE	Yes	Yes
$H_0 : \beta_{Established} = \beta_{New}$	0.069	0.002
Observations	4,576	4,576
Adjusted R^2	0.023	0.082

Notes: The regression controls for whether the interview was conducted on the phone and the fiscal year it refers to. For observations with missing education market predictors, we impute constant values within countries, which are captured by country fixed effects. For other categorical predictors, we include an additional (unreported) category if the value was originally missing (see Appendix D.2 for more details). Table 2 provides summary statistics. Bootstrapped standard errors with 1,000 replications are shown in parentheses. ***, **, and * correspond to 1%, 5%, and 10% levels of statistical significance.

lower likelihood of mismatch between the ideal and actual CEO type. These characteristics are also positively and significantly associated with having a leader CEO. Credit-constrained firms (based on ES survey responses) are not significantly more likely to be mismatched, suggesting pay packages may not be the constraining factor to finding the right CEO match. Finally, firms that are sole proprietorships are, as expected, associated with a higher likelihood of mismatch and a lower likelihood of having a leader-type CEO, as they are headed by their founder.

6 Conclusion

In this paper, we connect the literature on labor market frictions that lead to sub-optimal talent allocation with the literature on CEOs' impact on firm performance. We also connect to the literature on misallocation across firms as a driver of lower productivity in less-developed countries. Our two main contributions are, first, showing that it is possible to measure CEO leadership style at scale using four simple questions, and second, predicting the CEO type each firm needs and studying firm-CEO type matching across a sample of countries at varying stages of development.

Using novel data that uniquely combine information on CEO time use with data on firm performance across more than 40 countries, we find that firms and CEOs are matched differently across these nations. Furthermore, we find more similar patterns of matching in countries at similar stages of development. We also find that in middle-income countries, firm-CEO matching patterns resembling those in higher-income countries are correlated with higher firm productivity.

Connecting our results, we observe that smaller, unlisted firms or those active in certain sectors benefit from being matched with a manager-type CEO, who focuses on monitoring and implementation tasks to create systems that ensure precise and efficient plan execution. In particular, managers often meet suppliers and production workers or plant managers and have many 1:1 meetings but few meetings with other C-suite executives. Conversely, larger firms and those that are listed have higher productivity when matched with a leader-type CEO, who communicates with many stakeholders to foster organizational alignment. In particular, leaders have many meetings with several participants, both internal and external, and relatively few meetings with suppliers and production workers.

We find that firm-CEO mismatches are associated with a productivity loss of up to 20% for individual firms. Accounting for the total number of firms in the mismatched categories and their sales, we find that these mismatches, combined with a low leader share, are associated with a productivity loss of 9% across the entire sample. The overall benefit

of leaders that appears in the data when not considering firms' needs, stems from the fact that leaders are in low supply in many middle-income countries. That is, some firms with managers have lower productivity not because managers are inherently worse but because these firms need a leader type. We also analyze the influence of business education and find that established schools offering business degrees are associated with a higher supply of leaders in nearby firms.

To the best of our knowledge, this is the first paper to provide empirical evidence of mismatching between CEO behavior types and firms across a wide sample of countries, demonstrating that this stems from both a low leader supply and additional mismatches conditional on enough supply. The results suggest that particular attention should be paid to a CEO's leadership or management style, in relation to firm characteristics. This may be of interest to organizations that provide financing to firms in emerging economies. Policies that address the causes of mismatches, for instance focusing on high-quality business education, could have large potential to generate growth.

7 References

F. Abdelkhalek and R. Langsten (2019), "Track and sector in egyptian higher education: Who studies where and why?", *FIRE: Forum for International Research in Education* , *6*(2), 45–70.

O. Bandiera, S. Hansen, A. Prat, and R. Sadun (2020), "Ceo behavior and firm performance", *Journal of Political Economy*, *128*(4), 1325–1369.

O. Bandiera, A. Kotia, I. Lindenlaub, C. Moser, and A. Prat (2024), "Meritocracy across countries", Unpublished manuscript.

O. Bandiera, R. Lemos, A. Prat, and R. Sadun (2018), "Managing the family firm: Evidence from ceos at work", *The Review of Financial Studies*, *31*(5), 1605–1653.

T. Beck, H. Degryse, R. De Haas, and N. Van Horen (2018), "When Arm's Length is too Far: Relationship Banking over the Credit Cycle", *Journal of Financial Economics*, *127*(1), 174–196.

Belloni and Chernozhukov (2013), "Least squares after model selection in high-dimensional sparse models", *Bernoulli*, *19*(2), 521–547.

M. Bertrand and A. Schoar (2003), "Managing with style: The effect of managers on firm policies", *The Quarterly journal of economics*, *118*(4), 1169–1208.

D. Blei, A. Ng, and M. Jordan (2001), "Latent dirichlet allocation", *The Journal of Machine Learning Research*, *3*, 601–608.

F. Correia, A. Amaral, and A. Magalhães (2002), "Public and private higher education in portugal: unintended effects of deregulation", *European Journal of Education*, *37*(4), 457–472.

P. Cziraki and D. Jenter (2022), "The market for ceos", Working Paper 831/2022, ECGI.

N. Doghonadze (2009), "Business education in post soviet countries", *IBSU Scientific Journal*, *3*(2), 19–33.

P. Domadenik, M. Drame, and D. Farčnik (2010), "The school-to-work transitions of business graduates in slovenia", *Revija Za socijalnu politiku*, *17*(2), 299–324.

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

N. Gennaioli, R. LaPorta, F. L. de Silanes, and A. Shleifer (2013), "Human capital and regional development", *Quarterly Journal of Economics*, *128*(1), 105–164.

N. Gennaioli, R. LaPorta, F. L. de Silanes, and A. Shleifer (2014), "Growth in regions", *Journal of Economic Growth*, *19*(3), 259–309.

C. Goldin and L. F. Katz (1999), "The shaping of higher education: The formative years in the united states, 1980-1940", *Journal of Economic Perspectives*, *13*(1), 37–62.

M. Guclu (2020), "How turkish universities have evolved through constitutional changes", *Educational Research and Reviews*, *15*(3), 86–94.

D. D. Hendel and D. R. Lewis (2005), "Quality assurance of higher education in transition countries: Accreditation-accountability and assessment", *Tertiary Education and Management*, *11*(3), 239–258.

J. Hjort, H. Malmberg, and T. Schoellman (2024), "The missing middle managers: Labor costs, firm structure, and development", mimeo.

G. Hollinshead (2006), "Educating educators in a volatile climate- the challenge of modernising higher business schools in serbia and montenegro", *European Journal of Education*, *41*(1), 131–149.

C.-T. Hsieh and P. J. Klenow (2009), "Misallocation and manufacturing tfp in china and india", *The Quarterly Journal of Economics*, *124*(4), 1403–1448.

A. Kaplan (2014), "European management and european business schools: Insights from the history of business schools", *European Management Journal*, *32*(4), 529–534.

S. N. Kaplan, M. M. Klebanov, and M. Sorensen (2012), "Which ceo characteristics and abilities matter?", *Journal of Finance*, *67*(3), 973–1007.

J. P. Kotter (1990a), *A Force for Change: How Leadership Differs from Management*, New York: Free Press.

J. P. Kotter (1990b), "What Leaders Really Do", *Harvard Business Review*, *68*(3), 103–111.

E. Kraft and M. Vodopivec (2003), "The new kids on the block: The entry of private business schools in transition economies", *Education Economics*, *11*(3), 239–257.

A. Kuznetsov and H. Yakavenka (2005), "Barriers to the absorption of management knowledge in belarus", *Journal of Managerial Psychology*, *20*(7), 566–577.

R. LaPorta and A. Shleifer (2014), "Informality and development", *Journal of Economic Perspectives*, *28*(3), 109–126.

B. Leven (2010), "Poland's transition in business education", *American Journal of Business Education*, *3*(1), 53–60.

C. Lindsay and S. Sheather (2010), "Variable selection in linear reegression", *The Stata Journal*, *10*(4), 650–669.

Z. Mansouri (2023), "Moroccan university education: A history of a failing dualistic system", *The Journal of Quality in Education*, *13*(22), 125–141.

F. Marcos (2000), "Private and privatised higher educational institutions in jordan", *Mediterranean Journal of Educational Studies*, *5*(1), 95–117.

F. Marcos (2003), "Privatising higher education in spain", *European Business Organization Law Review*, *32*, 541–552.

A. Meziani (1999), "The system of higher education in morocco", *Mediterranean Journal of Educational Studies*, *4*(2), 215–219.

S. Mihai-Yiannaki (1987), "Managing change: The private university sector in cyprus, operations within the european context. a case study.", *Lex et Sciencia International Journal*, *XVIII*(1), 191–207.

A. Mitterle (2017), "In search of the private: on the specificities of private higher education in germany", in D. Cantini (Ed.), *Rethinking Private Higher Education*, 193–219, Leiden: Brill.

W. Mullins and A. Schoar (2016), "How do ceos see their roles? management philosophies and styles in family and non-family firms", *Journal of Financial Economics*, *119*(1), 24–43.

O. Palmer (2021), "Knowledge institutions in africa and their development 1960-2020: Tunisia", Unpublished manuscript.

V. M. Panaretos and Y. Zemel (2019), "Statistical aspects of wasserstein distances", *Annual Review of Statistics and its Application*, *6*(1), 405–431.

D. Restuccia and R. Rogerson (2008), "Policy distortions and aggregate productivity with heterogeneous establishments", *Review of Economic Dynamics*, *11*(4), 707–720.

M. Rosi, D. Tuček, V. Potočan, and M. Jurše (2018), "Market orientation of business schools: A development opportunity for the business model of university business schools in transition countries", *E+M Ekonomie a Management*, *21*(4), 175–194.

R. Sadun, J. B. Fuller, S. Hansen, and P. Neal (2022), "The C-Suite Skills That Matter Most", *Harvard Business Review*, *100*(4), 42–50.

J. Sauvagnat and F. Schivardi (2024), "Are executives in short supply? evidence from death events", *The Review of Economic Studies*, *91*(1), 519–559.

D. Scur, R. Sadun, J. Van Reenen, R. Lemos, and N. Bloom (2021), "The World Management Survey at 18: lessons and the way forward", *Oxford Review of Economic Policy*, *37*(2), 231–258.

G. K. Shaw and M. Blaug (1987), "The university of buckingham after ten years - a tentative evaluation", *Higher Education Quarterly*, *42*(1), 72–89.

A. Valero and J. Van Reenen (2019), "The economic impact of universities: Evidence from across the globe", *Economics of Education Review*, *68*, 53–67.

World Bank (2022), "Enterprise surveys: Sampling methodology", `https://www.enterprisesurveys.org/content/dam/enterprisesurveys/documents/methodology/Sampling_Note-Consolidated-2-16-22.pdf`.

S. E. Yonker (2017), "Geography and the market for ceos", *Management Science*, *63*(3), 609–630.

Appendices

Appendix Table of Contents

A Sample breakdown by country **3**

B The Enterprise Survey **4**

C Measuring CEO time use **5**

C.1 Selection of time use variables for the ES 5

C.2 Time use questions in the ES 6

C.3 Validating the ES CEO time use questions 8

C.4 Harmonization across the Diary and Survey data 10

D Sample restrictions and imputation **13**

D.1 Sample restrictions 13

D.2 Imputation of other variables 14

E Sector definition **15**

F Higher education institutions **16**

F.1 Data collection 16

F.2 Brief background on private and public education in the sample countries 17

G Robustness checks **20**

G.1 Leader-type frequency across countries with different income levels 20

G.2 Non-averaged performance measures in the Diary data 20

G.3 Alternative number of employees measure in the Survey data 22

G.4 Adding Brazil and India to the testing sample 24

G.5 Family-owned and family-managed firms 28

G.6 Including the multi-establishment indicator in the prediction 33

G.7 Prediction cells 37

G.8 Prediction with interacting all five prediction variables 38

G.9 Using the non-penalized coefficients for prediction 40

G.10 Sector fixed effects and firm age 43

G.11 Region fixed effects 45

G.12 Standard errors clustered at the locality level 50

G.13 RMSE and log GDP per capita relationship using only the United States as the training sample 51

G.14 Using the share of leaders as the Needs Leader threshold 52

G.15 Variation in the strength of the mismatch 53
G.16 Back-of-the-envelope calculation: Additional tables 55
G.17 Firm-level predictors with all outcome variables 57
G.18 Firm-level education predictors without other controls 57
G.19 Firm-level education predictors in levels 58

A Sample breakdown by country

Table A.1 Country summary statistics

Country	ISO 2 code	Observations	Diary data	Training sample	GDP p.c.
United States	US	118	1	1	62,459
Germany	DE	102	1	1	53,930
United Kingdom	GB	78	1	1	47,369
France	FR	97	1	1	46,018
Malta	MT	19	0	1	45,397
Italy	IT	203	0	1	42,746
Cyprus	CY	19	0	1	41,522
Czech Republic	CZ	132	0	0	40,981
Spain	ES	221	0	0	40,802
Slovenia	SI	65	0	0	38,947
Lithuania	LT	64	0	0	37,166
Estonia	EE	49	0	0	36,401
Portugal	PT	220	0	0	34,946
Poland	PL	134	0	0	33,185
Hungary	HU	150	0	0	32,554
Slovak Republic	SK	60	0	0	31,928
Latvia	LV	49	0	0	31,012
Romania	RO	182	0	0	29,875
Greece	GR	115	0	0	29,698
Croatia	HR	70	0	0	29,336
Türkıye	TR	393	0	0	28,197
Russia	RU	343	0	0	27,255
Kazakhstan	KZ	144	0	0	26,352
Bulgaria	BG	120	0	0	23,266
Montenegro	ME	12	0	0	21,534
Belarus	BY	134	0	0	19,279
Serbia	RS	47	0	0	18,307
North Macedonia	MK	57	0	0	16,773
Georgia	GE	36	0	0	14,989
Bosnia and Herzegovina	BA	43	0	0	14,897
Lebanon	LB	48	0	0	14,493
Azerbaijan	AZ	12	0	0	14,442
Armenia	AM	56	0	0	13,654
Albania	AL	85	0	0	13,653
Moldova	MD	58	0	0	13,027
Ukraine	UA	251	0	0	12,805
Mongolia	MN	41	0	0	12,486
Egypt	EG	571	0	0	11,763
Tunisia	TN	133	0	0	11,421
Kosovo	XK	19	0	0	11,318
Jordan	JO	36	0	0	10,071
Morocco	MA	134	0	0	7,547
Uzbekistan	UZ	199	0	0	7,348
West Bank and Gaza	PS	21	0	0	6,245
Kyrgyz Republic	KG	47	0	0	5,258
Tajikistan	TJ	25	0	0	3,581

Notes: GDP p.c. refers to the year 2019 GDP per capita and is expressed in 2017 billion USD terms.

B The Enterprise Survey

The Enterprise Survey (ES) is a firm-level survey of a representative sample of an economy's private sector. It covers a broad range of business environment topics, including access to finance, corruption, infrastructure, crime, and competition. The survey universe consists of the majority of manufacturing, retail and services sectors (wholesale, hotels, restaurants, transport, storage, communications, IT), and construction. This corresponds to firms classified with ISIC Rev 3.1 codes 15–37, 45, 50–52, 55, 60–64, and 72. In some larger countries, such as Türkıye and Ukraine, the survey design allows stratification by some of the sectors with the largest contribution to employment and value added.[1]

Only formal (registered) companies with five or more employees are eligible for an interview, and there are no restrictions on their age. Firms with 100% government/state ownership are ineligible to participate. Due to sensitive survey questions addressing business-government relations and bribery, private contractors (rather than a government agency) are hired to collect the data. Confidentiality of the survey respondents and the sensitive information they provide is necessary to ensure the greatest degree of survey participation, integrity, and confidence in the quality of the data.

The ES follows a robust methodology that has been in use since 2008. The survey data frame is ideally derived from the universe of eligible firms obtained from the country's statistical office or, if this is unavailable, from other sources. The sample frame contains data on firm size (number of employees or size class), business sector, and geographic region within a country; these are the strata used for stratified random sampling. Firm size classes are 5–19 (small), 20–99 (medium), and 100+ employees (large). The majority of firms in most countries are small and medium sized, and the ES oversamples large firms. The main sector breakdown is usually manufacturing, retail, and other services. For larger countries, specific manufacturing sub-sectors are selected as additional strata based on employment, value-added, and total number of establishments figures. Geographic regions within a country are selected based on which cities/regions collectively contain the majority of economic activity.

Once the simple random samples are selected within each size, sector, and region group, firms are contacted in a prescribed order over the phone using a screener questionnaire to determine their eligibility to participate in the survey. If they are eligible, the interviewer schedules a face-to-face interview.[2] Refusal to participate can occur at the screening stage or at the interview stage. The percentage of ineligible firms as well as the refusals at both the screener and interview stage vary by country.[3] The team working on the ES puts a high level of effort into making sure that all contractors follow the methodology closely, including by checking the detailed progress reports on

[1] Sample size is determined by the size of gross national income. See World Bank (2022) for more details.

[2] In Armenia and Spain, some of the interviews were conducted entirely over the phone. We include an indicator for these interviews as a control variable.

[3] Detailed information on this is available in the Report on sampling and implementation, available at `https://www.beeps-ebrd.com/data/2018-2020/`.

which firms were contacted on a weekly basis and identifying and correcting any issues as soon as possible.

In the Survey data, two variables are available for measuring employee count. The first one is the number of employees collected during the screener interview where no distinction was made between permanent versus temporary or full- versus part-time employees; it refers to the status at the time of the screener interview. The second is the number of permanent full-time employees in the establishment at the end of the completed fiscal year and is based on a response given in the survey. Both variables exhibit a strong (95%) correlation in log form. In the analysis presented in the main paper, we use the variable measuring the number of permanent full-time employees given that it was reported and refers to the fiscal year and establishment, consistent with other variables. In Appendix G.3, we nevertheless show that our results are robust to using the screener number of employees as well.

C Measuring CEO time use

In this section, we first describe how the time use variables were selected for inclusion in the ES and measured, and then provide details on how we combine the data for France, Germany, the United Kingdom, and the United States from Bandiera et al. (2020) (the Diary data) with the Survey data.

C.1 Selection of time use variables for the ES

In February 2017, an opportunity arose for the inclusion of five time use variables in the sixth wave of the ES. To select the set of survey questions that had a greater likelihood of capturing differences in CEO behavior—and, in particular, differences between leaders and managers discussed in Bandiera et al. (2020)—in the ES data, we leveraged the set of activities that a) were most predictive of the CEO leader index in the original Diary data and b) could be included in a standard ES instrument.

To identify the set of activities that were most predictive of the CEO leader index in the original Diary data, we adopted the following approach.

1. We computed the total counts of activities in which each CEO engaged during the sample week, separately for each of the 16 measures of CEO time use that were used to compute the leader measure by the LDA algorithm (see Bandiera et al. (2020) for details).

2. We randomly selected half of the sample, performed a regression of the leader variable against all the counts of activities computed in step 1 within this subsample, and selected the best five predictors within this set using the leaps-and-bounds algorithm (Lindsay and Sheather, 2010).

3. We regressed the leader variable against these five predictors in the rest of the sample (hold-out sample) and computed the adjusted R-square for the regression.

4. We repeated steps 2 and 3 1,000 times, each time randomly selecting different subsamples.

5. For each of the variables that were included in the set of best five predictors at least once across the 1,000 replications, we computed the average adjusted R-squared of the variable, averaging across the adjusted R-squared of the regressions in which the variable appeared among the set of regressors.

6. Finally, we selected five variables that had the highest adjusted R-squared, under the constraint that we wanted to include a maximum of three variables measuring who was involved in the activity, one variable describing other features of the activities (e.g., number of participants, planning horizon), and one variable measuring the duration of the meetings.

This exercise showed that the variables with the highest chance of predicting the CEO leader index were counts of activities involving 1) C-suite executives, board members, business unit managers, or managers from a parent company; 2) suppliers; 3) employees involved in production activities; 4) more than one participant; and 5) a duration greater than one hour. The response options for each question were also chosen based on the frequency with which they appeared in the Diary data.

Because the ES is conducted over a longer time period than the collection of Diary data took, ES questions refer to a *typical* week to increase comparability across firms. It was also decided that time use questions would only be asked to firms with at least 50 employees.

C.2 Time use questions in the ES

We measure CEO time use in the ES by asking how often the CEO engages in the five distinct activities described in the previous subsection in a typical week. Respondents can choose among five distinct choices, as shown below. The answers to each of these questions are converted into a scale ranging between 1 and 5, and then aggregated to build a single CEO index.

BMT.1 (Q1) In a typical week, how often does the top manager meet with one or more of the following: Chief Operating Officer (COO), Chief Administrative Officer (CAO), Chief Marketing Officer (CMO), Board members, Business Unit managers, or managers from a parent company?

	Original answer	Score
Never	1	1
Once a week	2	2
Between 2 and 4 times a week	3	3
Daily	4	4
More than once a day	5	5
Don't know (spontaneous)	-9	.

BMT.2 (Q2) In a typical week, how often does the top manager meet with suppliers?

	Original answer	Score
Never	1	5
Once a week	2	4
Between 2 and 4 times a week	3	3
Daily	4	2
More than once a day	5	1
Don't know (spontaneous)	-9	.

BMT.3 (Q3) In a typical week, how often does the top manager meet with employees involved in production activities (e.g., plant managers, front line production workers)?

	Original answer	Score
Never	1	5
Once a week	2	4
Between 2 and 4 times a week	3	3
Daily	4	2
More than once a day	5	1
Don't know (spontaneous)	-9	.

BMT.4 (Q4) In a typical week, how many meetings that involve the top manager include more than one other participant?

	Original answer	Score
Fewer than 5 meetings	1	1
Between 5 and 10 meetings	2	2
Between 11 and 15 meetings	3	3
Between 16 and 20 meetings	4	4
More than 20 meetings	5	5
Don't know (spontaneous)	-9	.

BMT.5 (Q5) In a typical week, how many meetings that involve the top manager last longer than one hour?

	Original answer	Score
Fewer than 5 meetings	1	1
Between 5 and 10 meetings	2	2
Between 11 and 15 meetings	3	3
Between 16 and 20 meetings	4	4
More than 20 meetings	5	5
Don't know (spontaneous)	-9	.

C.3 Validating the ES CEO time use questions

While the CEO time use data collected in the ES are directly inspired by measures collected in the Diary data, the two studies differ in two key methodological choices. First, the Diary data are collected by enumerators who measure actual quantities of CEO time spent in specific activities, while the Survey data classify time use in five broader options. Second, the Survey questions are based on the CEO recollecting the frequency of different types of activities over a typical week, whereas the Diary data was sampled and measured daily over an arbitrarily selected week by an enumerator.[4]

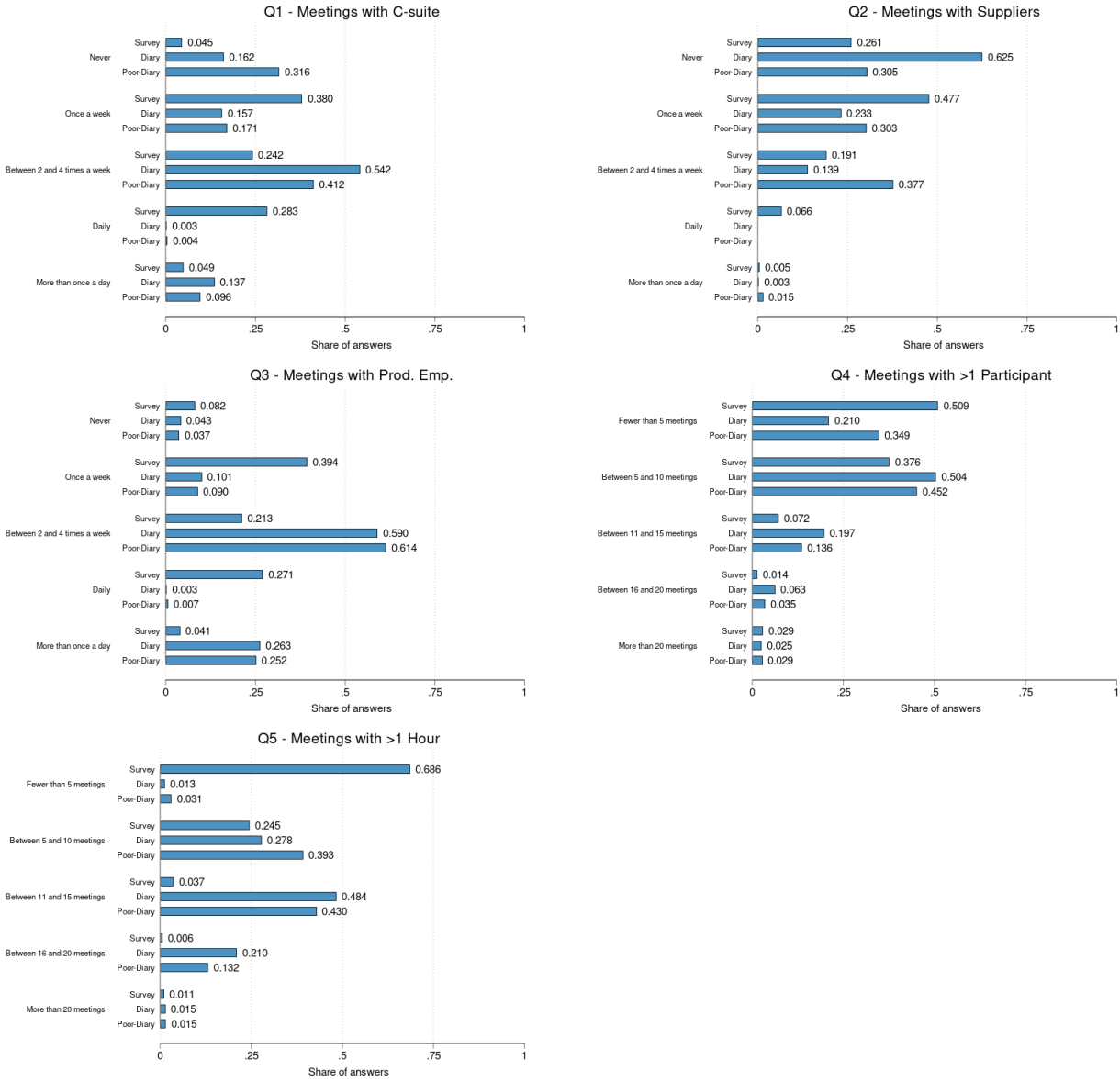
To investigate whether the questions included in the ES can capture meaningful variation in CEO time use, we compare the distributions of the individual time use variables across the two samples. The Diary data here include the two low-income countries—Brazil and India—to examine whether potential differences in the answers to the five CEO time use questions in the Diary and Survey data can be attributed to differences in income levels across countries. Figure C.1 presents the distribution of these answers. While the relative frequencies of answers do indeed differ between the two sources, the overall distribution of the answers for Q1 to Q4 are not too dissimilar across sources and heterogeneous across choices within the Survey data. Where we observe some changes across frequencies to individual responses—for example, in the case of Q3—the mass of the distribution seems to relocate between neighboring responses rather than distant answers. This suggests that this distribution changes due to the reallocation of edge cases across the discrete response options.

The only Survey question that shows a substantially different distribution from its Diary counterpart is Q5 (the frequency of meetings lasting over one hour). For this question, the option "Fewer than 5 meetings" was chosen by 69% of respondents, versus approximately 0.02% of respondents in the Diary data. Moreover, the mass of the distribution seems to reallocate between distant answers, suggesting that the change is not merely due to the reallocation of edge cases.

To quantify these differences across questions more concretely, we calculate, for each question, the Wasserstein distance (see, e.g., Panaretos and Zemel 2019) across Diary and Survey data distributions. In our context, this statistic measures the area between two cumulative distribution functions (CDFs) to express how dissimilar they are. Intuitively, if the two distributions are viewed as unit amounts of earth on their domain, the Wasserstein distance is the minimum "cost" of turning one pile into the other, where "cost" is assumed to equal the mass that has to be moved times the distance traveled during reallocation. This is a meaningful measure for us as it captures the differences in proximate versus distant reallocation of the mass highlighted above and could influence our results by skewing the share of leaders across the Diary and Survey data.

[4]See Appendix C.1 for details.

Figure C.1 Distribution of answers to CEO time use questions in the Diary and Survey data



Notes: The graph shows the distribution of responses across five response options for Survey, Diary, and lower-income Diary data for each time use question. The Survey data include all countries covered by the EBRD-EIB-WB Enterprise Survey; the Diary data include the US, the UK, France, and Germany; and lower-income Diary data include Brazil and India.
Source: EBRD-EIB-WB Enterprise Survey and Bandiera et al. (2020).

In our case, the Wasserstein distance index for a given question is calculated as

$$Distance^k = \frac{\sum_{x=1}^{5} |F^k_{Diary}(x) - F^k_{Survey}(x)|}{4},$$

where $F^k_j(x)$ is the share of firms with response $X \leqslant x$ for question k in sample j, x takes the cardinal values (1–5) assigned to the response options, and the sum is normalized with 4 to a 0–1 scale for interpretability.[5] With the normalization, the Wasserstein distance index ranges between 0 and 1. Appendix Table C.1 shows the results, and as expected, Q5 has the most dissimilar distribution.

Table C.1 Wasserstein distance between the distribution of responses across the Diary and Survey data

	Time use question				
	Q1	Q2	Q3	Q4	Q5
Wasserstein distance	0.12	0.07	0.16	0.10	0.36

Notes: The table reports the Wasserstein distance calculated for each time use question across Survey and Diary data. The Survey data include all countries covered by the EBRD-EIB-WB Enterprise Survey, while the Diary data include the US, the UK, France, Germany, Brazil, and India.
Source: EBRD-EIB-WB Enterprise Survey and Bandiera et al. (2020).

Figure C.2 plots the mean responses to individual questions for each country against the country's log GDP per capita instead of comparing the full distribution of responses across broad subsamples. The results suggest that for Q1–Q4, the mean of responses appear to change continuously across income levels—irrespective of whether a country belongs to the Diary or Survey data. In contrast, for Q5, the Diary data countries appear to have much larger mean responses than the Survey data countries, conditional on income.

These stark differences in the distribution of this variable lead us to believe that Q5 may have been interpreted differently in the Diary versus in the Survey data, and hence we exclude it from the computation of the aggregate CEO time use index.

C.4 Harmonization across the Diary and Survey data

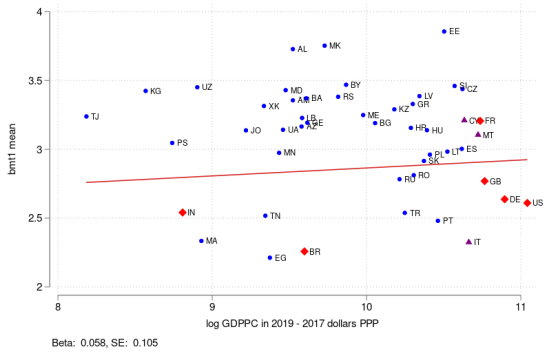
We take several steps to increase the comparability of the Diary and Survey data. First, because the Diary data only cover manufacturing firms, we also limit the Survey data to manufacturing firms. This reduces the sample from 8,899 to 6,011 firms. Second, to harmonize the control variables, we take the following steps:

1. We recode the Standard Industry Classification sector codes used in the Diary data into the ISIC Rev. 3.1 sector classification used in the Survey data.

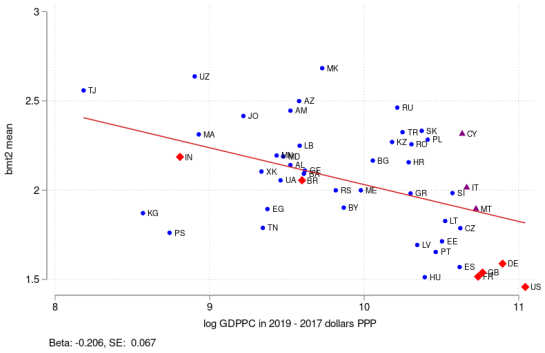
[5] We choose 4 as this is the maximum of the sum, which is attained when CDFs are (1,1,1,1,1) versus (0,0,0,0,1) for x values 1–5.

Figure C.2 Log GDP per capita and the mean of time use index components

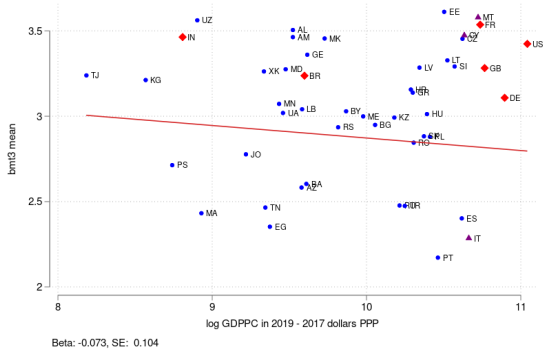
(a) Q1: Meetings with C-suite



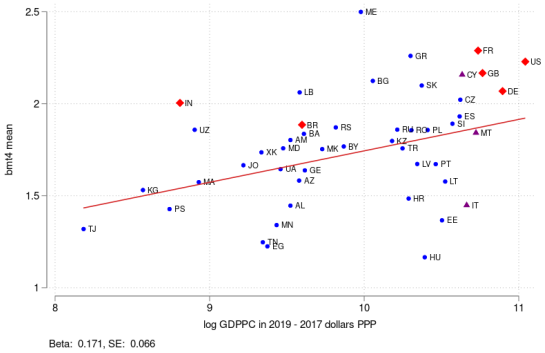
(b) Q2: Meetings with suppliers



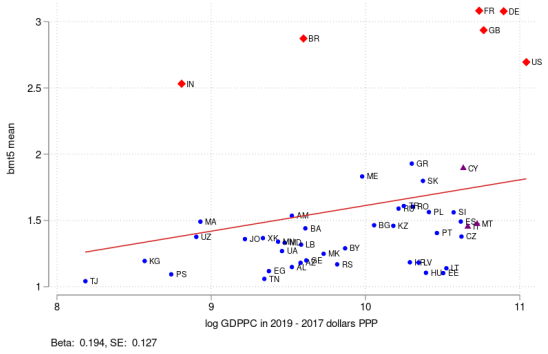
(c) Q3: Meetings with production employees



(d) Q4: Meetings with more than 1 participant



(e) Q5: Meetings longer than 1 hour



Notes:

Source: EBRD-EIB-WB Enterprise Survey and Bandiera et al. (2020).

2. Family-owned firms are identified in the Diary data using a 25% ownership threshold. For the Survey data, we use information on exact share of family ownership (question BMB.1 in the ES) to identify family-owned firms using the same threshold. The question asks, "What percentage of the firm is owned by the same family? (If more than one family, refer to the one with largest ownership)?"

3. Family-management is identified directly in the Diary data. For the Survey data, we use question BMB.2 in the ES. This question asks, "What percentage of the key management positions of this firm are occupied by members of [the family that owns the firm]?" To ensure that we only capture cases when it is indeed the CEO who is reported to be a family member—rather than other managers—we use 100 as the cutoff threshold in this question.

4. For the Diary data, we identify multi-establishment firms using the question: "how many production sites are in the firm?". For the Survey data, we use question A.7—"Multi-establishment firm" to identify multi-establishment firms. Establishment is defined as physical location where business is conducted. It must have its own control and its costs and revenues are accounted separately. This variable is recorded by the screener before the interview.

Table C.2 compares control variables unrelated to CEO time use. It shows that the steps detailed above lead to comparable distributions of most variables between the Survey and the Diary data, with a few exceptions. Firms in the Diary data are on average larger, older, and more likely to be multi-establishment firms than the Survey data firms. Firm age and the share of multi-establishment firms are significantly different at the 10% level across the two samples even conditional on firm size (the share of multi-establishment firms is discussed further in Section G.6). The same pattern emerges when we compare the Survey data with the Diary data for middle-income countries (India and Brazil), where the oversampling of large firms is even more pronounced. We only use Diary data for middle-income countries in Appendix G.4.

Table C.2 Distribution of variables across samples

Sample	Survey	High-income Diary	Middle-income Diary
Variable		Mean	
Listed firms (%)	13.8	20.0	53.7
Family-owned firms (%)	41.8	36.0	57.7
Family-managed firms (%)	17.8	22.1	56.3
Multi-establishment firms (%)	18.1	53.9	57.8
Log (employment)	5.0	5.7	5.9
Firm age	27.2	64.0	40.0
Number of firms	4,817	395	456

Notes: This table shows the mean values of variables available in the Survey, the high-income Diary, and the lower-income Diary data.

D Sample restrictions and imputation

D.1 Sample restrictions

The baseline samples used throughout the analysis include 4,817 observations from the Survey data and 395 firms from the Diary data. Tables D.1 and D.2 show the individual restrictions that lead to these sample sizes, starting from the 476 originally available firms in the Diary data and 28,778 firms in the Survey data, respectively. The tables also report additional (one-off) sample changes used in robustness checks, either due to changes in the predictor variables used in the robustness check or the inclusion of Brazil and India in the analysis.

Table D.1 Diary data sample restrictions and associated decrease in sample sizes

Sample change	Observation count change	Remaining sample size
Restrictions in high-income Diary data countries		
Unrestricted data	.	476
Employment data available	.	432
Sector data available (2-digit ISIC Rev 3.1)	17	415
Manufacturing firm (2-digit ISIC Rev 3.1 $\leqslant$ 36)	18	397
Q1–Q4 time use questions data available	0	397
Sales data available	0	397
Family-ownership data available	0	397
Firm age data available	2	395
Listed status data available	0	395
One-off sample changes in robustness checks		
Adding Brazil and India	456	851
Multi-establishment data available	11	384
Family-management data available	1	394

Table D.2 Survey data sample restrictions and associated decrease in sample sizes

Sample change	Observation count change	Remaining sample size
Restrictions in the main analysis		
Unrestricted data	.	28,778
Firm with at least 50 employees (as per screener size)	19,879	8,899
Manufacturing firm (2-digit ISIC Rev 3.1 $\leqslant$ 36)	2,888	6,011
Q1–Q4 time use questions data available	611	5,400
Sales data available	442	4,958
Family-ownership data available	112	4,846
Firm age data available	17	4,829
Coordinates data available	1	4,828
Listed status data available	0	4,828
Sampling weights available	0	4,828
Reported full-time employees data available	11	4,817
One-off sample changes in robustness checks		
Using number of employees (as per screener size)	11	4,828
Multi-establishment data available	0	4,817
Family-management data available	33	4,784

D.2 Imputation of other variables

Some variables, not used for the prediction, have missing values in our baseline sample, but we do not condition on their availability given that they are less central to our analysis. These variables belong to two main categories: indicator variables used as explanatory variables and continuous variables used as explanatory variables. For the former, we recode missing values as zero and add an indicator variable equal to 1 for the imputed observations in the regressions. The related variables, the number of missing observations, and their share in the total sample are reported in Table D.3.

The second group contains the "minimum distance to new business degree HEIs" variable from the set of local education market characteristics. Here, 49 observations in Estonia have no minimum distance information as there are no schools established after 1990 that offer business degrees. We impute these cases with a constant value, which (irrespective of the value used) is captured by the country fixed effect.

Table D.3 Indicators for missing observations

Variable	N	Sum	Mean
Missing *Exporter or importer status* indicator	4,576	95	0.021
Missing *At least 25% foreign ownership* indicator	4,576	26	0.006
Missing *Has a board of directors or supervisory board* indicator	4,576	4	0.001
Missing *Sole proprietorship* indicator	4,576	37	0.008

E Sector definition

The ES generally uses three stratification sectors (manufacturing, retail, and other services) but employs more for larger countries.[6] Importantly, it collects more detailed information on the main activity, product, or service, corresponding to four-digit ISIC Rev. 3.1 sectors. We use this information to define 23 two-digit ISIC Rev 3.1 sectors (manufacturing sectors covered by the ES) and then combine them into 11 sectors that we use in the analysis. Table E.1 contains the exact concordance between the two-digit ISIC Rev. 3.1 sectors and the 11 sectors we use.

Table E.1 Mapping between 2-digit ISIC Rev. 3.1 sectors and sectors used in the analysis

	2-digit ISIC 3.1 sectors covered by ES	Sectors used in the analysis
Manufacturing	Food (15) Tobacco (16)	15
	Textiles (17)	17
	Garments (18) Leather (19)	18
	Wood (20) Paper (21) Publishing, printing, and recorded media (22)	20
	Refined petroleum product (23) Chemicals (24)	24
	Plastics & rubber (25)	25
	Non-metallic mineral products (26)	26
	Basic metals (27) Fabricated metal products (28)	28
	Machinery and equipment (29)	29
	Office, accounting, and computing machinery (30) Electrical machinery (31) Radio, television, and communication equipment (32) Precision instruments (33) Motor vehicles, trailers, and semi-trailers (34) Other transport equipment (35)	30
	Furniture (36) Recycling (37)	36

Source: EBRD-EIB-WB Enterprise Survey and authors' calculations.

[6] Details can be found in the "Report on Sampling and Implementation" at `https://www.beeps-ebrd.com/data/2018-2020/`.

F Higher education institutions

F.1 Data collection

Data on higher education institutions (HEIs) and their characteristics (such as degrees offered, divisions, fields of study, year founded, public/private status) were scraped from the International Association of Universities' World Higher Education Database (WHED) and augmented with business school accreditation data scraped from the membership websites of three main agencies: the Association to Advance Collegiate Schools of Business (AACSB), the Association of MBAs (AMBA), and the European Quality Improvement System (EQUIS). Data on HEIs in Kosovo were obtained from the Kosovo Accreditation Agency (`https://akreditimi.rks-gov.net/`). Where the data on some or all characteristics were missing in these databases, we searched for them on the websites of individual HEIs. We retrieved the GPS coordinates of all HEIs through Google Maps API.

We used a five-step procedure to identify schools offering business degrees. First, we marked schools that reported to have business-related degrees (banking, business administration, business and commerce, business and finance, finance, international business, management, and business computing). Second, we marked schools that reported to have business or management divisions. Third, any HEI whose name includes specific word combinations such as "School of Business," "Business School," "School of Management," and "Management School" was added to the list of business schools. Fourth, HEIs found in the WHED were matched to HEIs in the AACSB, AMBA, or EQUIS databases, with any HEIs with an existing accreditation membership in at least one of these agencies labeled as business schools. In addition, HEIs found in at least one of these databases but not in the WHED were also added to the list of HEIs and business schools. Fifth, the list of business schools identified as described above was verified by EBRD regional economists and analysts to ensure there were no major omissions. Table F.1 shows how many schools are added as a result of these steps.

Table F.1 Capturing schools offering business degrees (testing sample)

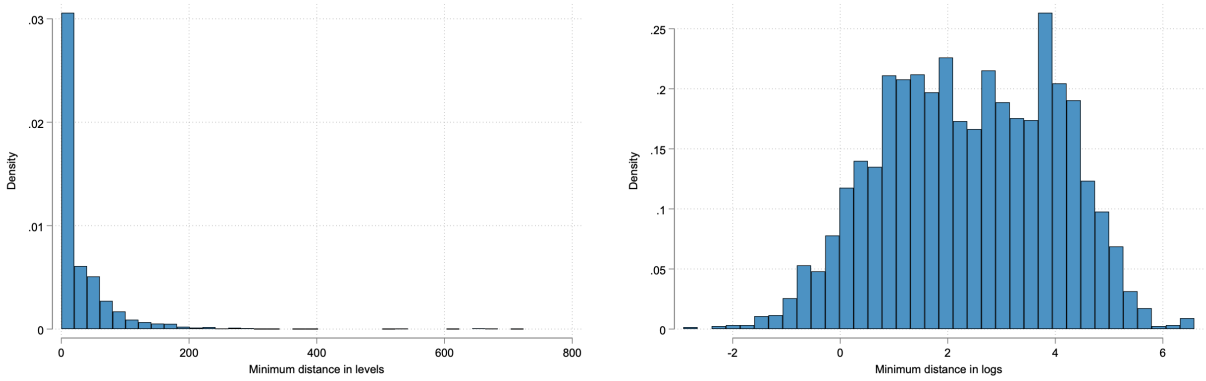
Type	Total number	New addition
Business degree (step 1)	1,854	1,854
Business division (step 2)	508	105
Business school (steps 3-5)	392	21
Total schools offering business degrees	.	1,980
Total higher education institutions	.	3,051

We further distinguish HEIs based on their year of founding to create established and new school categories. Using the identified categories, we calculate the distance between firms and institutions within countries. We focus on within-country distance for two main reasons. First, frictions along

borders, such as language barriers or visa requirements, are likely to increase the cost of attending a particular HEI type abroad discontinuously, making it impractical to disregard national borders. Second, going beyond borders would require us to process education data from countries outside our sample, which would be challenging to execute carefully. For 49 firms in Estonia, this strategy leads to missing minimum distance information from new schools offering business degrees, as this school type is unavailable in the country. For these firms, we impute distance information with a constant value, which is captured in the estimation by country fixed effects. We ensure that the results are unchanged if we exclude imputed observations from the estimation.

Figure F.1 plots the distribution of the measure of minimum distance from schools offering business degrees, in both levels and log form. The skewed nature of this variable motivates us to include these variables in the main analysis in the logged form. Appendix G.19 discusses the results without the log transformation.

Figure F.1 Distribution of the minimum distance to schools offering business degrees (testing sample)



Notes: The figures show the distribution of the "minimum distance to schools offering business degrees" variable in the testing sample. The left panel shows the variable in levels, while the right panel shows the distribution in logs.

F.2 Brief background on private and public education in the sample countries

This subsection provides background information on the observed patterns depicted in Figure 6 for private and public HEIs offering business degrees.

Among the training sample countries, private HEIs were initially more common than public HEIs in the US (Goldin and Katz, 1999). In contrast, they are a minority in France, Germany, Italy, and the UK, and were established in 1902 in Italy, in the early 1900s in Germany (Mitterle, 2017), and 1973 in the UK (Shaw and Blaug, 1987). In Cyprus, the first HEI, established in 1961, was private (Mihai-Yiannaki, 1987). Additionally, Kaplan (2014) notes that the first business

schools were established in France (in 1819), Belgium (1852), the US (1881), and Germany (1898). After World War II, business schools in the UK and Spain adopted the US-style management more readily than those in France and Germany (see Kaplan, 2014).[7]

In non-communist countries (Egypt, Greece, Jordan, Lebanon, Morocco, Portugal, Spain, Tunisia, Türkıye, and the West Bank and Gaza), the observed trends reflect the fact that business education started becoming more widespread in the late 1950s/early 1960s, coinciding with programs like the Ford Foundation promoting US-style management across Europe. Private business education in non-communist countries followed in much greater numbers from the 1990s, as educational policies shifted toward a more liberal model in response to increased demand for higher education and its rising costs.

Some examples follow. In Türkıye, private universities did not exist until 1982, when the constitution permitted the establishment of foundation universities (Guclu, 2020). Spain recognized the legal status of private universities in 1983 with the University Reform Act (Marcos, 2003); Portugal also embraced the idea of private higher education in the mid-1980s (Correia, Amaral, and Magalhães, 2002). Morocco allowed private HEIs to be established from 1986 onward (Meziani, 1999; Mansouri, 2023), while Jordan's first private universities were established in 1989, in response to the demands of the expatriate community (Marcos, 2000). In Egypt, a 1992 presidential decree opened the path for setting up private universities, though the first private university was established in 1919 (American University in Cairo) (Abdelkhalek and Langsten, 2019). In Tunisia, investment in private HEIs picked up in 2000, after the creation of a strict legal framework for regulating the private higher education sector (Palmer, 2021).

In post-communist countries, which constitute the majority of our testing sample, private business education was virtually nonexistent until the end of Communism in the very end of the 1980s and early 1990s.[8] Universities were state run (Kraft and Vodopivec, 2003; Leven, 2010), and there were no private HEIs. This changed in the late 1980s, when business education became possible and prestigious in the period of Perestroika (Doghonadze, 2009). There was a significant increase in demand for law, business, and economics graduates, which could not be met by public HEIs alone. Most countries thus allowed private HEIs to be established in the 1990s, though sometimes with significant legal restrictions.[9] Between 1990 and 2000, the share of business education students in total tertiary enrollment increased substantially in Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Romania, and Slovenia, along with the number of private providers (Kraft and Vodopivec, 2003).

The rapid expansion of private HEIs and enrollment was accompanied by concerns about dis-

[7]INSEAD, founded in 1958, was the first business school to offer an MBA degree in Europe. In Spain, Escuela de Organización Industrial was established in 1955, followed by ESADE and IESE in 1958, while the London Business School was established in 1964 (see Kaplan, 2014).

[8]In fact, Doghonadze (2009, p.20) notes that the "Soviet system of education did not include any business education, as the very word 'business' sounded inappropriate."

[9]For example, Kraft and Vodopivec (2003) note that in Lithuania, private HEIs were not allowed to offer undergraduate education, while in Slovenia, they were not allowed to offer university-level education.

tinguishing between "degree mills" and legitimate HEIs, and quality more broadly (Hendel and Lewis, 2005). Efforts to improve management education in transition countries have been led by international organizations and Western governments (such as the World Bank, USAID, the European Commission, the British Council, and the German Technical Cooperation (GTZ, now the German Development Cooperation). However, despite good intentions, the knowledge transfer was not as successful as envisaged, a major reason being that the context on the ground differed from that in the West (Kuznetsov and Yakavenka, 2005; Hollinshead, 2006).

Rosi, Tuček, Potočan, and Jurše (2018) note that even business schools belonging to public HEIs have been facing rigid institutional, administrative, and academic regulations that hindered their ability to compete in the increasingly globalized business education market. Nonetheless, Domadenik et al. (2010), use a matched employer-employee micro dataset of 2007 business graduates in Slovenia and find that the probability of becoming employed after graduation was lower for those who graduated from newer private HEIs than from public HEIs. This could be either due to the lower quality of academic programs offered by the former or to the lower ability of their students compared with those from the latter.

G Robustness checks

G.1 Leader-type frequency across countries with different income levels

Table G.1 Distribution of leader-type CEOs across countries (average marginal effects)

Dependent variable →	Leader status	
	(1)	(2)
Log (GDP per capita)	0.120***	0.103***
	(0.011)	(0.012)
Log (employment)	No	Yes
Listed firm	No	Yes
Family-owned firm	No	Yes
Sector FE	No	Yes
Observations (Testing sample)	4,576	4,576
Observations (Training sample)	636	636
Observations	5,212	5,212
Pseudo R^2	0.015	0.028

Notes: This table presents average marginal effects from probit regressions to estimate the relation between having a leader-type CEO and log GDP per capita in 2019 (in PPP constant 2017 international $). Table 2 provides summary statistics, and Table E.1 provides information on sectors. Robust standard errors are shown in parentheses. $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

G.2 Non-averaged performance measures in the Diary data

The analysis presented in the main paper averages log sales and log employment information from Orbis for several years around the collection of time use data for Diary data firms. This subsection shows that the results are robust to using only the data point closest to the year of the time use data collection. For 98% of firms, this is at most one year away.

Table G.2 presents the equivalent of Table 6 to show that the variables selected by LASSO are unchanged by the inclusion of an extra sector dummy. Table G.3 reports the robust significant association between the predicted leader-type CEO propensity and actual leader status. In Table G.4, we test the robustness of the results in Table 9. The horizontal differentiation finding still holds, as indicated by the negative significant coefficients of the Needs Manager, Gets Leader and Needs Leader, Gets Manager variables, the insignificant coefficient of the Needs Manager, Get Manager variable, and the non-significant p-value in the "$\hat{\beta}_{LM} - \hat{\beta}_{ML} = 0$ (p-value)" row in Table G.4.

Table G.2 Predicting leader status with firm characteristics (using only non-averaged performance measure in Diary data)

Dependent variable →	Leader-type CEO status
Non-listed firm	-0.106
Non-listed firm * Family-owned firm	-0.367
Log (employment)	0.137
Manufacture of food products and beverages	-0.120
Manufacture of textiles	-0.363
Manufacture of rubber and plastics products	0.254
Manufacture of fabricated metal products	-0.014
Manufacture of machinery and equipment	0.006
Manufacture of office, accounting and computing machinery	0.313
Constant	-0.335
Observations	636

Notes: This table presents penalized coefficients from a LASSO probit estimating the relationship between leader-type CEO status and sector, and the interaction between listed status, family-owned status, log employment, and firm age in the training sample. Coefficients are selected by cross-validation. *Sources*: EBRD-EIB-WB Enterprise Survey.

Table G.3 Leader status and predicted leader propensity (using only non-averaged performance measure in Diary data)

Dependent variable →	Leader status
Leader propensity	0.227***
	(0.070)
Country FE	Yes
Observations	4,576
Adjusted R^2	0.060

Notes: This table presents estimates from an OLS regression to estimate the relation between the predicted leader-type CEO propensity and the actual leader-type CEO status in the testing sample. Leader-type CEO status is an indicator variable equal to 1 if the CEO time use index is greater than 3, and 0 otherwise. Leader-type CEO propensity in the testing sample is predicted using coefficients estimated by regressing the actual leader-type CEO status of training sample firms on firm characteristics. The regression controls for whether the interview was conducted on the phone and the fiscal year it refers to. Table 2 provides summary statistics, and Table E.1 provides information on sectors. Bootstrapped standard errors with 1,000 replications are shown in parentheses. ***, **, and * correspond to 1%, 5%, and 10% levels of statistical significance.

Table G.4 Productivity and leader mismatch (using only non-averaged performance measure in Diary data)

Dependent variable →	Log (labor productivity)		
	(1)	(2)	(3)
Mismatched	-0.154***		
	(0.037)		
Needs Leader, Gets Manager		-0.180***	-0.195***
		(0.047)	(0.047)
Needs Manager, Gets Leader		-0.106**	-0.124**
		(0.050)	(0.054)
Needs Manager, Gets Manager			-0.059
			(0.065)
$\hat{\beta}_{LM} - \hat{\beta}_{ML} = 0$ (p-value)		0.172	0.205
Country FE	Yes	Yes	Yes
Observations	4,576	4,576	4,576
Adjusted R^2	0.341	0.342	0.341

Notes: This table presents OLS regressions to estimate the relationship between the (mis)match of the CEO type the firm has and the CEO type it needs, on the one hand, and labor productivity, on the other hand, in the testing sample. Mismatched is an indicator equal to 1 if the firm's ideal and actual CEO types are mismatched and 0 otherwise. Needs Leader, Gets Manager is an indicator equal to 1 if the firm's ideal CEO type is leader and its actual CEO type is manager and 0 otherwise. LM refers to Needs Leader, Gets Manager, and ML refers to Needs Manager, Gets Leader. The regression controls for whether the interview was conducted on the phone and the fiscal year it refers to. Table 2 provides summary statistics. Bootstrapped standard errors with 1,000 replications are shown in parentheses. ***, **, and * correspond to 1%, 5%, and 10% levels of statistical significance.

G.3 Alternative number of employees measure in the Survey data

The analysis presented in the main paper uses the number of permanent, full-time employees to measure employee numbers in the Survey data. This subsection shows that the results are robust to using the number of employees collected during the screener interview.

Table G.5 presents the equivalent of Table 6 to show that despite the change affecting the employee numbers of firms in Italy, Cyprus, and Malta in the training sample, the variables selected by LASSO are largely unchanged modulo a sector dummy. Table G.6 reports the robust significant association between the predicted leader-type CEO propensity and actual leader status. In Table G.7, we test the robustness of the results in Table 9. The horizontal differentiation finding still holds, as indicated by the negative significant coefficients of the Needs Manager, Gets Leader and Needs Leader, Gets Manager variables, the insignificant coefficient of the Needs Manager, Get Manager variable, and the non-significant p-value in the "$\hat{\beta}_{LM} - \hat{\beta}_{ML} = 0$ (p-value)" row in Table

G.7.

Table G.5 Predicting leader status with firm characteristics (using screener number of employees for the Survey data)

Dependent variable →	Leader-type CEO status
Non-listed firm	-0.066
Non-listed firm * Family-owned firm	-0.358
Log (employment)	0.156
Manufacture of food products and beverages	-0.109
Manufacture of textiles	-0.292
Manufacture of rubber and plastics products	0.218
Manufacture of office, accounting and computing machinery	0.285
Constant	-0.524
Observations	636

Notes: This table presents penalized coefficients from a LASSO probit estimating the relationship between leader-type CEO status and sector, and the interaction between listed status, family-owned status, log employment, and firm age in the training sample. Coefficients are selected by cross-validation. *Sources*: EBRD-EIB-WB Enterprise Survey.

Table G.6 Leader status and predicted leader propensity (using screener number of employees for the Survey data)

Dependent variable →	Leader status
Leader propensity	0.263***
	(0.070)
Country FE	Yes
Observations	4,587
Adjusted R^2	0.061

Notes: This table presents estimates from an OLS regression to estimate the relation between the predicted leader-type CEO propensity and the actual leader-type CEO status in the testing sample. Leader-type CEO status is an indicator variable equal to 1 if the CEO time use index is greater than 3, and 0 otherwise. Leader-type CEO propensity in the testing sample is predicted using coefficients estimated by regressing the actual leader-type CEO status of training sample firms on firm characteristics. The regression controls for whether the interview was conducted on the phone and the fiscal year it refers to. Table 2 provides summary statistics, and Table E.1 provides information on sectors. Bootstrapped standard errors with 1,000 replications are shown in parentheses. ***, **, and * correspond to 1%, 5%, and 10% levels of statistical significance.

Table G.7 Productivity and leader mismatch (using screener number of employees for the Survey data)

Dependent variable →	Log (labor productivity)		
	(1)	(2)	(3)
Mismatched	-0.157***		
	(0.039)		
Needs Leader, Gets Manager		-0.185***	-0.199***
		(0.049)	(0.051)
Needs Manager, Gets Leader		-0.103**	-0.118**
		(0.049)	(0.054)
Needs Manager, Gets Manager			-0.052
			(0.067)
$\hat{\beta}_{LM} - \hat{\beta}_{ML} = 0$ (p-value)		0.141	0.179
Country FE	Yes	Yes	Yes
Observations	4,587	4,587	4,587
Adjusted R^2	0.332	0.332	0.332

Notes: This table presents OLS regressions to estimate the relationship between the (mis)match of the CEO type the firm has and the CEO type it needs, on the one hand, and labor productivity, on the other hand, in the testing sample. Mismatched is an indicator equal to 1 if the firm's ideal and actual CEO types are mismatched and 0 otherwise. Needs Leader, Gets Manager is an indicator equal to 1 if the firm's ideal CEO type is leader and its actual CEO type is manager and 0 otherwise. LM refers to Needs Leader, Gets Manager, and ML refers to Needs Manager, Gets Leader. The regression controls for whether the interview was conducted on the phone and the fiscal year it refers to. Table 2 provides summary statistics. Bootstrapped standard errors with 1,000 replications are shown in parentheses. ***, **, and * correspond to 1%, 5%, and 10% levels of statistical significance.

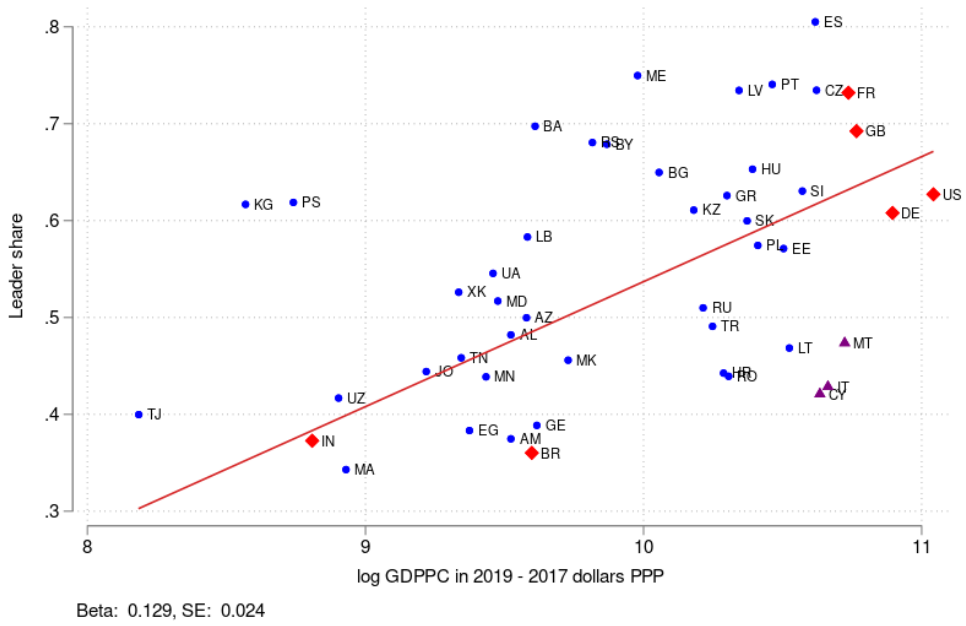
G.4 Adding Brazil and India to the testing sample

In this subsection, we add Brazil and India to the testing sample countries, which adds an additional 236 firms from Brazil and 220 firms from India to the analysis. As these countries were not included in the ES, their data do not cover several variables used in the analysis, which restricts the results we can reproduce. As shown below, however, our conclusions are robust to including data from Brazil and India whenever possible.

We first find that the actual leader share in Brazil and India is similar to what we would expect based on our linear prediction for countries at their income level (Figure G.1). Second, the positive unconditional correlation between leader status and firm-level outcomes is unchanged, with the Leader × Training sample interaction now not being a significant predictor even for the number of employees (Table G.8). Third, the country-level RMSE still decreases in the income level of the testing sample countries, and India and Brazil are not far away from where the linear

fit would predict (Figure G.2). Fourth, the relationship between actual leader status and predicted leader-type CEO propensity remains significant when Brazilian and Indian firms are included in the testing sample (Table G.9). Last, the horizontal differentiation finding holds when Brazil and India are included in the sample, as indicated by the negative significant coefficients of the Needs Manager, Gets Leader and Needs Leader, Gets Manager variables, the insignificant coefficient of the Needs Manager, Get Manager variable, and the non-significant p-value in the "$\hat{\beta}_{LM} - \hat{\beta}_{ML} = 0$ (p-value)" row in Table G.10.

Figure G.1 Correlation between log GDP per capita and leader share (Brazil and India added to the testing sample)



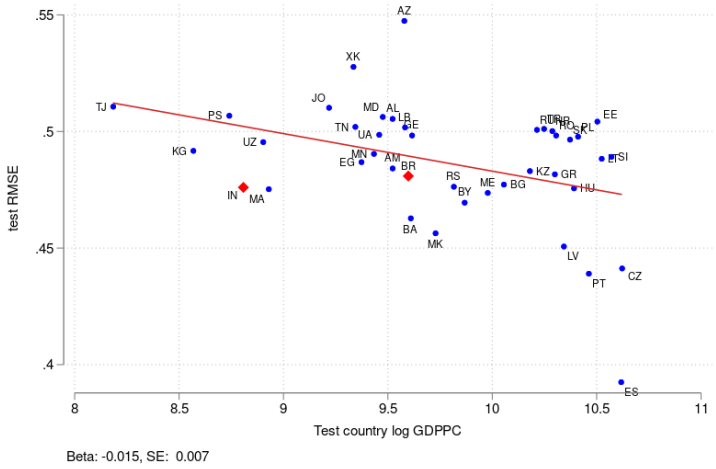
Notes: This figure shows the correlation between a country's share of leaders and its 2019 GDP per capita (in constant 2017 international $) when Brazil and India are added to the testing sample. When estimating the linear fit, observations are not weighted by the number of firms present in each country. This regression also controls for the share of firms interviewed over the phone in the country. A translation of the country ISO codes to country names is available in Appendix Table A.1. The red diamond shapes represent countries from the Diary data, purple triangles represent countries from the Survey data included in the training sample, and blue dots represent Survey data countries in the testing sample.

Table G.8 Firm characteristics and leader status (Brazil and India added to the testing sample)

Dependent variable →	Log (employment)		Log (sales)		Log (labor productivity)	
	(1)	(2)	(3)	(4)	(5)	(6)
Leader	0.206***	0.182***	0.337***	0.295***	0.131***	0.113***
	(0.027)	(0.026)	(0.045)	(0.043)	(0.036)	(0.035)
Leader * Training sample	0.143	0.104	0.051	0.018	-0.092	-0.086
	(0.087)	(0.084)	(0.121)	(0.117)	(0.081)	(0.080)
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Sector FE	No	Yes	No	Yes	No	Yes
Listed firm	No	Yes	No	Yes	No	Yes
Family-owned firm	No	Yes	No	Yes	No	Yes
Firm age	No	Yes	No	Yes	No	Yes
Observations	5,668	5,668	5,668	5,668	5,668	5,668
Adjusted R^2	0.128	0.194	0.620	0.649	0.770	0.785

Notes: This table presents OLS regressions to estimate the relation between having a leader-type CEO and employment, sales, and sales per employee (labor productivity) in the pooled samples, with Brazil and India added to the testing sample. Sales are in 2019 US$. Employment refers to the number of full-time, permanent employees. Leader is an indicator variable equal to 1 if the CEO time use index is above 3, and 0 otherwise. All regressions control for whether the interview was conducted on the phone and the fiscal year it refers to. Table 2 provides summary statistics, and Table E.1 provides information on sectors. Robust standard errors are shown in parentheses. $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure G.2 Out-of-sample prediction and log GDP per capita (Brazil and India added to the testing sample)



Notes: This figure shows the scatterplot of the log GDP per capita and root mean square error (RMSE) from predicting the type of CEO a firm needs using the estimates of Table 6 and the full training sample. The RMSE is calculated by regressing the actual leader status of firms in the country on the predicted leader-type CEO propensity measure. A list of countries in the training and testing samples, in addition to Brazil and India, can be found in Appendix A. The red diamond shapes represent Brazil and India from the Diary data, and blue dots represent Survey data countries in the testing sample. *Source*: Authors' calculations based on the Diary and Survey data.

Table G.9 Leader status and predicted leader propensity (Brazil and India added to the testing sample)

Dependent variable →	Leader status
Leader propensity	0.278***
	(0.066)
Country FE	Yes
Observations	5,032
Adjusted R^2	0.065

Notes: This table presents estimates from an OLS regression to estimate the relation between the predicted leader-type CEO propensity and the actual leader-type CEO status in the testing sample, including Brazil and India. Leader-type CEO status is an indicator variable equal to 1 if the CEO time use index is greater than 3, and 0 otherwise. Leader-type CEO propensity in the testing sample is predicted using coefficients estimated by regressing the actual leader-type CEO status of training sample firms on firm characteristics. The regression controls for whether the interview was conducted on the phone and the fiscal year it refers to. Table 2 provides summary statistics, and Table E.1 provides information on sectors. Bootstrapped standard errors with 1,000 replications are shown in parentheses. ***, **, and * correspond to 1%, 5%, and 10% levels of statistical significance.

Table G.10 Productivity and leader mismatch (Brazil and India added to the testing sample)

Dependent variable →	Log (labor productivity)		
	(1)	(2)	(3)
Mismatched	-0.158***		
	(0.035)		
Needs Leader, Gets Manager		-0.181***	-0.200***
		(0.042)	(0.044)
Needs Manager, Gets Leader		-0.112**	-0.135***
		(0.048)	(0.051)
Needs Manager, Gets Manager			-0.080
			(0.058)
$\hat{\beta}_{LM} - \hat{\beta}_{ML} = 0$ (p-value)		0.225	0.250
Country FE	Yes	Yes	Yes
Observations	5,032	5,032	5,032
Adjusted R^2	0.706	0.706	0.706

Notes: This table presents OLS regressions to estimate the relationship between the (mis)match of the CEO type the firm has and the CEO type it needs, on the one hand, and labor productivity, on the other hand, with Brazil and India added to the testing sample. Mismatched is an indicator equal to 1 if the firm's ideal and actual CEO types are mismatched and 0 otherwise. Needs Leader, Gets Manager is an indicator equal to 1 if the firm's ideal CEO type is leader and its actual CEO type is manager and 0 otherwise. LM refers to Needs Leader, Gets Manager, and ML refers to Needs Manager, Gets Leader. The regression controls for whether the interview was conducted on the phone and the fiscal year it refers to. Table 2 provides summary statistics. Bootstrapped standard errors with 1,000 replications are shown in parentheses. ***, **, and * correspond to 1%, 5%, and 10% levels of statistical significance.

G.5 Family-owned and family-managed firms

In high-income countries, family-owned firms tend to hire family CEOs (Bandiera et al., 2018). In this subsection we investigate how this CEO labor market friction affects our results. Tables 6 and 7 in Section 4.1 show that family-owned firms are more likely to have manager-type CEOs. This aligns with Mullins and Schoar (2016), who find that CEOs of family-owned firms tend to delegate less and value sector-specific knowledge more than CEOs of non-family-owned firms, even conditional on other characteristics like firm size, sector, and listed status. However, the estimated overall association between CEO time use type and family ownership may be influenced by both the genuinely differing needs of family-owned firms and the tendency of these firms to pick family CEOs from a constrained talent pool where the abundant manager-type CEO is more likely to be common. If the latter has a substantial influence on our estimate, it would make the estimate less

informative about firms' CEO type needs.

Family-owned firms make up 49.5% of our training sample, but just over half of them (52.3%) are managed by a family CEO. Given that the share of family-managed firms is not negligible, we attempt to isolate the influence of them on our prediction. To this end, we include an indicator for family-managed firms in the prediction equation in addition to family-owned firms:[10]

$$Sector_i + Listed_i * FamilyOwned_i * Log(Emp_i) * FirmAge_i * FamilyManaged_i.$$

While the $FamilyManaged$ coefficient aims to capture the influence of CEO labor market frictions, the $FamilyOwned$ coefficient aims to capture the genuinely differing needs of family-owned firms. The penalized coefficients estimated by LASSO are reported in Table G.11. Family management is taken to predict CEO type only if the firm is not listed and is family-owned. The interaction between non-listed and family-owned firms was selected in our baseline specification in Table 6, and thus interacting it with family-managed firms essentially divides it into family-managed and non-family-managed categories. The coefficient of this interaction is negative, which aligns with the expectation that family-managed firms often struggle to recruit scarce leader-type CEOs.

Table G.11 Predicting leader status with firm characteristics (including family-managed firm status)

Dependent variable →	Leader status
Manufacturing of food products and beverages	-0.093
Manufacture of textiles	-0.229
Manufacture of rubber and plastics products	0.203
Manufacture of office, accounting and computing machinery	0.305
Non-listed firm	-0.067
Non-listed firm * Family-owned firm	-0.232
Log (employment)	0.145
Non-listed firm * Family-owned firm * Family-managed firm	-0.231
Constant	-0.430
Observations	635

Notes: This table presents penalized coefficients from a LASSO probit estimating the relation between leader-type CEO status, sector, and the interaction between listed status, family-owned status, log employment, firm age, and family-managed firm status in the training sample. *Sources*: Training sample.

Comparing the average marginal effects in Table G.12 shows that family ownership has a larger average marginal effect on a firm having a manager-type CEO than having a family-managed CEO (–0.111 versus –0.055), and the coefficient is similar to when family management is not included

[10]We prefer not to include this variable in our baseline specification as it can be seen as an outcome of the matching process rather than a fundamental firm characteristic affecting firm needs.

(–0.132; see Table 7). This suggests that this original negative estimate captures, in large part, firms' needs independent of the family-management constraint.[11]

Table G.12 Relationship between actual leader-type CEO and LASSO-selected firm characteristics (family-managed firm status)

Dependent variable →	Leader-type CEO status
Manufacturing of food products and beverages	-0.066
	(0.051)
Manufacture of textiles	-0.219
	(0.148)
Manufacture of rubber and plastics products	0.175*
	(0.100)
Manufacture of office, accounting and computing machinery	0.153***
	(0.059)
Listed firm	0.125**
	(0.063)
Family-owned firm	-0.111***
	(0.037)
Log (employment)	0.066***
	(0.016)
Family-managed firm	-0.055**
	(0.026)
Observations	635
Pseudo R^2	0.086

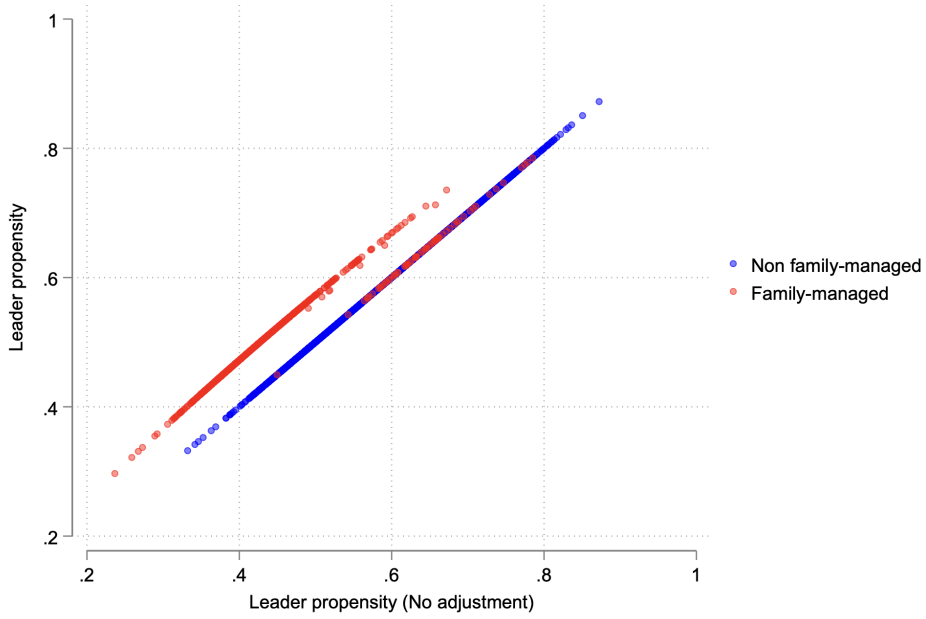
Notes: This table presents average marginal effects from the probit regression to estimate the relation between firm characteristics and leader-type CEO status in the training sample, using the LASSO regression analysis method. Robust standard errors are shown in parentheses. ***, **, and * correspond to 1%, 5%, and 10% levels of statistical significance. *Sources*: EBRD-EIB-WB Enterprise Survey.

Given these estimates, we can attempt to disregard the frictions associated with family management by taking non-family-managed firms as our normative basis for the prediction and thus not using the extra "manager-type preference" component captured by the coefficient of the family-managed firm variable. This essentially means adjusting the predicted leader-type CEO propensity of listed family-owned, family-managed firms upward (disregarding the negative component predicted in this interaction), illustrated in Figure G.3. As a result of this change, the categorical Needs Leader status of family-managed firms changes, as shown in Table G.13. Therefore, about one-third of family-managed firms originally classified as needing a manager are now reclassified

[11] Using these estimates for prediction—without any additional changes—would lead to similar results to those in the analysis presented in the main paper, with the MM coefficient now significant at the 10% level. The magnitude of this coefficient is -0.116 against LM and ML coefficients of -0.192 and -0.185, respectively. The results are available upon request.

as needing a leader. Making these adjustments leads to no qualitative change in the results. Table G.14 shows the relationship between leader-type CEO propensity and actual leader-type CEO status.

Figure G.3 Adjustment of predicted leader-type CEO propensity for listed family-owned family-managed firms



Notes: This figure illustrates how the leader-type CEO propensity of family-managed firms changes as a result of disregarding the triple interaction's coefficient in Table G.11. The horizontal axis measures the estimated leader-type CEO propensities with the coefficient, while the vertical axis measures the estimated propensities without it. The leader-type CEO propensity of firms that are 1) not family-managed or 2) are family-managed but not listed or family-owned remains unchanged.

Table G.15 indicates that our finding of horizontal differentiation between the leader and manager CEO types in Table 9 still holds. The coefficients on the Needs Manager, Gets Leader and Needs Leader, Gets Manager variables remain negative and significant, while the coefficient on the Needs Manager, Get Manager variable and the p-value in the "$\hat{\beta}_{LM} - \hat{\beta}_{ML} = 0$ (p-value)" row remain insignificant. Other results are available upon request.

Taken together, the results in this subsection confirm that the presence of family-managed firms does not affect our identification strategy.

Table G.13 Changes in family-managed firms' needs as a result of the family-management adjustment

	Needs Leader (adjusted)	Needs Manager (adjusted)
Needs Leader	128	0
Needs Manager	200	426

Notes: This table shows how the leader-type CEO status of family-managed firms changes if the triple interaction's coefficient in Table G.11 is disregarded. The rows indicate the estimated leader-type CEO status with the coefficient, while the columns indicate the estimated leader-type CEO status without it. The status can remain unchanged either because a firm is 1) family-managed but not listed or family-owned remains unchanged or 2) the estimated leader-type CEO propensity does not change sufficiently to alter its binary leader-type CEO status.

Table G.14 Leader status and predicted leader propensity (disregarding the family-managed firm status)

Dependent variable →	Leader-type CEO status
Leader propensity	0.380***
	(0.087)
Country FE	Yes
Observations	4,543
Adjusted R^2	0.062

Notes: This table presents estimates from an OLS regression to estimate the relation between predicted leader-type CEO propensity and actual leader-type CEO status. Leader-type CEO status is an indicator variable equal to 1 if the CEO time use index is greater than 3, and 0 otherwise. Leader-type CEO propensity in the testing sample is predicted using coefficients estimated by regressing the actual leader-type CEO status of training sample firms on firm characteristics. The regression controls for whether the interview was conducted on the phone and the fiscal year it refers to. Table 2 provides summary statistics, and Table E.1 provides information on sectors. Bootstrapped standard errors with 1,000 replications are shown in parentheses. ***, **, and * correspond to 1%, 5%, and 10% levels of statistical significance.

Table G.15 Productivity and leader mismatch (disregarding the family-managed firm status)

Dependent variable →	Log (labor productivity)		
	(1)	(2)	(3)
Mismatched	-0.154***		
	(0.039)		
Needs Leader, Gets Manager		-0.162***	-0.177***
		(0.043)	(0.046)
Needs Manager, Gets Leader		-0.131**	-0.148**
		(0.057)	(0.060)
Needs Manager, Gets Manager			-0.079
			(0.068)
$\hat{\beta}_{LM} - \hat{\beta}_{ML} = 0$ (p-value)		0.618	0.638
Country FE	Yes	Yes	Yes
Observations	4,543	4,543	4,543
Adjusted R^2	0.348	0.348	0.348

Notes: This table presents OLS regressions to estimate the relation between the (mis)match of the CEO type the firm has and the CEO type it needs, on the one hand, and labor productivity, on the other hand. Mismatched is an indicator equal to 1 if the firm's ideal and actual CEO types are mismatched and 0 otherwise. Needs Leader, Gets Manager is an indicator equal to 1 if the firm's ideal CEO type is leader and its actual CEO type is manager and 0 otherwise. LM refers to Needs Leader, Gets Manager, and ML refers to Needs Manager, Gets Leader. The regression controls for whether the interview was conducted on the phone and the fiscal year it refers to. Table 2 provides summary statistics. Bootstrapped standard errors with 1,000 replications are shown in parentheses. ***, **, and * correspond to 1%, 5%, and 10% levels of statistical significance.

G.6 Including the multi-establishment indicator in the prediction

In addition to the five variables listed in specification (3), an indicator for multi-establishment firms can be constructed using both the Diary and Survey data, and it could be regarded as a fundamental firm characteristic. This subsection discusses the reasons for not including this variable and shows that the results are nevertheless robust to its inclusion.

As discussed in Appendix C, the multi-establishment firm variable is based on the following definitions: "how many production sites are in the firm?" for Diary data firms, and "Multi-establishment firm" for Survey data firms (recorded during the screener interview). Unlike the Diary data, the Survey data establishment definition requires that its costs and revenue are accounted separately and that it has its own control. The definitions in the Diary and Survey data thus do not match. Table G.16 shows that the distribution of the variable in Diary and Survey data is quite dissimilar.

Table G.16 Multi-establishment firms by sample

	Multi-establishment		Not multi-establishment	
Data	Number	%	Number	%
Survey	871	18.08	3,946	81.92
Diary	207	53.91	177	46.09
Lower-income Diary	257	57.75	188	42.25
Total	1,335	23.62	4,311	76.35

Notes: This table presents the number of multi-establishment firms and their share in all observations of Survey, Diary, and lower-income Diary data.

We also note in Appendix C that this is the only predictor variable where the distribution across samples appears dissimilar even conditional on employment. As lower-income Diary data countries have a similar share of multi-establishment firms as the high-income Diary data countries, we suspect that the distributions are influenced by differences in the definition across the Diary and Survey data.

We nevertheless incorporate the multi-establishment variable into the prediction as follows:

$$Sector_i + Listed_i * FamilyOwned_i * Log(emp_i) * FirmAge_i * MultiEst \tag{G.1}$$

The number of observations that this specification uses in the prediction cells created by the interaction of categorical variables is shown in Table G.17. The cells in the interactions contain very few firms for the training sample, which makes it more likely that the prediction overfits on this sample (despite using LASSO for regularization).

Table G.17 Prediction cells in the training sample (sector and 5-wise interaction)

		Family-owned	Not Family-owned
Multi-establishment	Listed	8	41
	Not listed	135	109
Not multi-establishment	Listed	7	20
	Not listed	161	142

Notes: This table tabulates the number of observations in cells created by the interaction of family-owned, listed, and multi-establishment firm variables in specification G.1 for the training sample.

This aligns with what we observe when applying the specification to the LASSO regression. As Table G.18 indicates, while most LASSO selections remain intuitive, the estimation also selects an interaction cell, "Listed firm * Family-owned firm * Non-multi-establishment firm," which has only seven observations and no variation in the left-hand-side variable.

Along with the implausible distribution of the multi-establishment variable, the increased likelihood of overfitting and the consequently non-intuitive interaction selection motivate us to exclude

this variable from the prediction. Moreover, the multi-establishment variable also has the strongest correlation with the log employment variable (0.34) out of all predictors (see Table G.19).

Table G.18 Predicting leader status with firm characteristics (sector and 5-wise interaction)

Dependent variable →	Leader status
Non-listed firm * Family-owned firm	-0.340
Log (employment)	0.136
Listed firm * Non-family-owned firm * Multi-establishment firm	0.126
Listed firm * Family-owned firm * Non-multi-establishment firm	0.414
Manufacturing of food products and beverages	-0.087
Manufacture of textiles	-0.268
Manufacture of rubber and plastics products	0.284
Manufacture of office, accounting and computing machinery	0.285
Constant	-0.466
Observations	616

Notes: This table presents penalized coefficients from a LASSO probit estimating the relation between leader-type CEO status, sector, and the five-wise interaction of 1) listed status, 2) family-owned status, 3) log employment, 4) firm age, and 5) multi-establishment firm status in the training sample. *Sources*: Diary data.

Table G.19 Correlation between predictors in the training sample

	Multi-establishment firm	Listed firm	Family-owned firm	Log (employment)	Firm age
Multi-establishment firm	1.00				
Listed firm	0.13	1.00			
Family-owned firm	-0.02	-0.23	1.00		
Log (employment)	0.35	0.32	-0.07	1.00	
Firm age	0.15	0.11	0.03	0.23	1.00

Notes: This table presents the correlation coefficients between the predictor variables in the training sample. *Sources*: Training sample.

We also conduct an exercise where, starting from specification (G.1), we rerun the variable selection with each predictor discarded individually. Table G.20 shows the number of times any individual predictor has been selected either alone or in an interaction by LASSO. This also motivates the exclusion of the multi-establishment variable from the set of categorical variables.

Nevertheless, the results using specification (G.1) or with the exclusion of one individual predictor (available upon request) are largely unchanged from our main results. We show selected

Table G.20 Number of times individual predictors have been selected by LASSO (sector and 5-wise interaction)

Prediction variable name	Number of times the variable is selected
Log (employment)	5
Sector dummies	5
Listed firm	5
Family-owned firm	5
Multi-establishment	3
Firm age	0

Notes: This table shows the number of times a variable was selected by LASSO, either alone or in an interaction.

results with specification (G.1) below. Table G.21 shows the relationship between leader-type CEO propensity and actual leader-type CEO status, and the estimated coefficient is very similar to the one in Table 8. The results in Table G.22 are also similar to those in Table 9. The finding of horizontal differentiation between the leader and manager CEO types is robust, as indicated by the negative and significant coefficients on the Needs Manager, Gets Leader and Needs Leader, Gets Manager variables, the insignificant coefficient of the Needs Manager, Get Manager variable, and the non-significant p-value in the "$\hat{\beta}_{LM} - \hat{\beta}_{ML} = 0$ (p-value)" row.

Table G.21 Leader status and predicted leader propensity (sector and 5-wise interaction)

Dependent variable →	Leader status
Leader propensity	0.233***
	(0.075)
Country FE	Yes
Observations	4,576
Adjusted R^2	0.060
Number of successful replications	1,000

Notes: This table presents estimates from an OLS regression to estimate the relation between predicted leader-type CEO propensity and actual leader-type CEO status in the sample using sector and five-wise interaction. Leader-type CEO status is an indicator variable equal to 1 if the CEO time use index is greater than 3, and 0 otherwise. Leader-type CEO propensity in the testing sample is predicted using coefficients estimated by regressing the actual leader-type CEO status of training sample firms on firm characteristics. The regression controls for whether the interview was conducted on the phone and the fiscal year it refers to. Table 2 provides summary statistics, and Table E.1 provides information on sectors. Bootstrapped standard errors with 1,000 replications are shown in parentheses. ***, **, and * correspond to 1%, 5%, and 10% levels of statistical significance.

Table G.22 Productivity and leader mismatch (sector and 5-wise interaction)

Dependent variable →	Log (labor productivity)		
	(1)	(2)	(3)
Mismatched	-0.162***		
	(0.038)		
Needs Leader, Gets Manager		-0.177***	-0.197***
		(0.045)	(0.047)
Needs Manager, Gets Leader		-0.134***	-0.159***
		(0.049)	(0.053)
Needs Manager, Gets Manager			-0.082
			(0.066)
$\hat{\beta}_{LM} - \hat{\beta}_{ML} = 0$ (p-value)		0.457	0.499
Country FE	Yes	Yes	Yes
Observations	4,576	4,576	4,576
Adjusted R^2	0.342	0.342	0.342
Number of successful replications	998	1,000	1,000

Notes: This table presents OLS regressions to estimate the relation between the (mis)match of the CEO type the firm has and the CEO type it needs, on the one hand, and labor productivity, on the other hand, in the sample using sector and five-wise interaction. Mismatched is an indicator equal to 1 if the firm's ideal and actual CEO types are mismatched and 0 otherwise. Needs Leader, Gets Manager is an indicator equal to 1 if the firm's ideal CEO type is leader and its actual CEO type is manager and 0 otherwise. LM refers to Needs Leader, Gets Manager, and ML refers to Needs Manager, Gets Leader. The regression controls for whether the interview was conducted on the phone and the fiscal year it refers to. Table 2 provides summary statistics. Bootstrapped standard errors with 1,000 replications are shown in parentheses. ***, **, and * correspond to 1%, 5%, and 10% levels of statistical significance.

G.7 Prediction cells

The following two tables tabulate the number of observations within each cell formed by the categorical variables in specification (3) for the training and testing samples. Table G.23 tabulates observations within cells created by the interaction of the family-owned and listed variables. Table G.24 tabulates observations within the separately added sector fixed effects.

Table G.23 Prediction cells in specification (3) (interactions)

		Family-owned	Not Family-owned
Training sample	Listed	16	63
	Not Listed	300	257
Testing sample	Listed	219	431
	Not Listed	1,622	2,287

Notes: This table, along with Table G.24, tabulates the number of observations in cells created by the categorical variables in specification (3) for the training and testing samples. It presents the cells created by the interaction of family-owned and listed variables.

Table G.24 Prediction cells in specification (3) (sector FE)

	Training sample	Testing sample
Manufacturing of food products and beverages	105	1,032
Manufacture of textiles	11	298
Manufacture of wearing apparel	13	604
Manufacture of wood and of products of wood	77	242
Manufacture of chemicals and chemical products	40	232
Manufacture of rubber and plastics products	26	247
Manufacture of other non-metallic mineral products	22	366
Manufacture of fabricated metal products	119	583
Manufacture of machinery and equipment	102	447
Manufacture of office, accounting and computing machinery	92	361
Manufacture of furniture; manufacturing n.e.c.	29	164

Notes: This table, along with Table G.23, tabulates the number of observations in cells created by the categorical variables in specification 3 for the training and testing samples. It presents the cells created by sector fixed effects.

G.8 Prediction with interacting all five prediction variables

This subsection checks the robustness of our findings when we modify the LASSO specification from specification (3) to the following specification:

$$Sector_i * Listed_i * FamilyOwned_i * Log(Employment_i) * FirmAge_i.$$

We prefer not to use this five-wise interaction in our main specification due to the low number of observations per cell, which may make LASSO prone to overfitting. Table G.25 shows the LASSO estimates of the penalized coefficients for prediction. Some of these interactions contain very few observations and are difficult to interpret. Nevertheless, comparing the actual leader status of firms and predicted leader-type CEO propensity, we continue to observe a positive and significant relationship (Table G.26).

Table G.25 Predicting leader status with firm characteristics (5-wise interaction)

Dependent variable →	Leader status
Non-listed firm	-0.025
Manufacture of textiles * Non-listed firm	-0.008
Manufacture of wearing apparel * Listed firm	-0.281
Manufacture of food products and beverages * Family-owned firm	-0.008
Manufacture of rubber and plastics products * Family-owned firm	0.167
Non-listed firm * Family-owned firm	-0.304
Manufacture of fabric * Non-listed firm * Family-owned firm	-0.000
Log (employment)	0.125
Manufacture of office, accounting and computing machinery * Log (employment)	0.014
Manufacture of textiles * Non-listed firm * Family-owned firm * Log (employment)	-0.019
Manufacture of office, accounting and computing machinery * Firm age	0.074
Manufacture of food products and beverages * Family-owned firm * Firm age	-0.048
Manufacture of textiles * Non-listed firm * Non-family-owned firm * Firm age	-0.002
Manufacture of fabric * Listed firm * Log (employment) * Firm age	-0..005
Manufacture of textiles * Family-owned firm * Log (employment) * Firm age	-0..005
Constant	-0.369
Observations	636

Notes: This table presents penalized coefficients from a LASSO probit estimating the relation between leader-type CEO status and the five-wise interaction of sector, listed status, family-owned status, log employment, and firm age in the training sample. *Sources*: Training sample.

Table G.26 Leader status and predicted leader propensity (5-wise interaction)

Dependent variable →	Leader status
Leader propensity	0.223**
	(0.092)
Country FE	Yes
Observations	4,576
Adjusted R^2	0.059
Number of successful replications	1,000

Notes: This table presents estimates from an OLS regression to estimate the relation between predicted leader-type CEO propensity and actual leader-type CEO status in the sample using five-wise interaction. Leader-type CEO status is an indicator variable equal to 1 if the CEO time use index is greater than 3, and 0 otherwise. Leader-type CEO propensity in the testing sample is predicted using coefficients estimated by regressing the actual leader-type CEO status of training sample firms on firm characteristics. The regression controls for whether the interview was conducted on the phone and the fiscal year it refers to. Table 2 provides summary statistics, and Table E.1 provides information on sectors. Bootstrapped standard errors with 1,000 replications are shown in parentheses. ***, **, and * correspond to 1%, 5%, and 10% levels of statistical significance.

Table G.27 indicates that our finding of horizontal differentiation between the leader and manager CEO types in Table 9 still holds. The coefficients on the Needs Manager, Gets Leader and Needs Leader, Gets Manager variables remain negative and significant, while the coefficient of the Needs Manager, Get Manager variable and the p-value in the "$\hat{\beta}_{LM} - \hat{\beta}_{ML} = 0$ (p-value)" row remain insignificant. Other results are available upon request.

Table G.27 Productivity and leader mismatch (5-wise interaction)

Dependent variable →	Log (labor productivity)		
	(1)	(2)	(3)
Mismatched	-0.161***		
	(0.038)		
Needs Leader, Gets Manager		-0.171***	-0.186***
		(0.046)	(0.045)
Needs Manager, Gets Leader		-0.138***	-0.156***
		(0.054)	(0.055)
Needs Manager, Gets Manager			-0.070
			(0.067)
$\hat{\beta}_{LM} - \hat{\beta}_{ML} = 0$ (p-value)		0.577	0.613
Country FE	Yes	Yes	Yes
Observations	4,576	4,576	4,576
Adjusted R^2	0.342	0.342	0.342

Notes: This table presents OLS regressions to estimate the relation between the (mis)match of the CEO type the firm has and the CEO type it needs, on the one hand, and labor productivity, on the other hand, in the sample using five-wise interaction. Mismatched is an indicator equal to 1 if the firm's ideal and actual CEO types are mismatched and 0 otherwise. Needs Leader, Gets Manager is an indicator equal to 1 if the firm's ideal CEO type is leader and its actual CEO type is manager and 0 otherwise. LM refers to Needs Leader, Gets Manager, and ML refers to Needs Manager, Gets Leader. The regression controls for whether the interview was conducted on the phone and the fiscal year it refers to. Table 2 provides summary statistics. Bootstrapped standard errors with 1,000 replications are shown in parentheses. ***, **, and * correspond to 1%, 5%, and 10% levels of statistical significance.

G.9 Using the non-penalized coefficients for prediction

Our baseline analysis uses the penalized coefficients presented in Table 6 for prediction. This subsection shows that using the non-penalized/postselection coefficients—the estimates we obtain from regressing the variables selected by LASSO on the leader-type CEO status outcome variable in the training sample in a probit regression—leads to qualitatively similar results. Table G.28 presents the non-penalized coefficients.

Using these coefficients changes the results in Table 8 to those in Table G.29 and the results

in Table 9 to those in Table G.30. Our finding of horizontal differentiation between the leader and manager CEO types is robust to using non-penalized coefficients. The coefficients on the Needs Manager, Gets Leader and Needs Leader, Gets Manager variables remain negative and significant, while the coefficient of the Needs Manager, Get Manager variable remains insignificant. The p-value in the "$\hat{\beta}_{LM} - \hat{\beta}_{ML} = 0$ (p-value)" row indicates a significant difference ($p = 0.05$) between Needs Manager, Gets Leader and Needs Leader, Gets Manager firms, with the former associated with more negative productivity. Other results are available upon request.

Table G.28 Relationship between actual leader-type CEO and LASSO-selected firm characteristics (non-penalized LASSO coefficients)

Dependent variable →	Leader-type CEO status
Manufacture of food products and beverages	-0.224
	(0.150)
Manufacture of textiles	-0.696*
	(0.418)
Manufacture of rubber and plastics products	0.443
	(0.284)
Manufacture of office, accounting and computing machinery	0.366**
	(0.169)
Non-listed firm	-0.143
	(0.188)
Non-listed firm * Family-owned firm	-0.407***
	(0.113)
Log (employment)	0.185***
	(0.047)
Constant	-0.519
	(0.340)
Observations	636
Pseudo R^2	0.081

Notes: This table presents average marginal effects from the probit regression to estimate the relation between the firm characteristics and leader-type CEO status in the training sample, using the LASSO regression analysis method with non-penalized LASSO coefficients. Robust standard errors are shown in parentheses. ***, **, and * correspond to 1%, 5%, and 10% levels of statistical significance. *Sources*: EBRD-EIB-WB Enterprise Survey.

Table G.29 Leader status and predicted leader propensity (non-penalized LASSO coefficients)

Dependent variable →	Leader status
Leader propensity	0.186***
	(0.053)
Country FE	Yes
Observations	4,576
Adjusted R^2	0.061
Number of successful replications	1,000

Notes: This table presents estimates from an OLS regression to estimate the relation between predicted leader-type CEO propensity and actual leader-type CEO status. Leader-type CEO status is an indicator variable equal to 1 if the CEO time use index is greater than 3, and 0 otherwise. Leader-type CEO propensity in the testing sample is predicted using coefficients estimated by regressing the actual leader-type CEO status of training sample firms on firm characteristics. The regression controls for whether the interview was conducted on the phone and the fiscal year it refers to. Table 2 provides summary statistics, and Table E.1 provides information on sectors. Bootstrapped standard errors with 1,000 replications are shown in parentheses. ***, **, and * correspond to 1%, 5%, and 10% levels of statistical significance.

Table G.30 Productivity and leader mismatch (non-penalized LASSO coefficients)

Dependent variable →	Log (labor productivity)		
	(1)	(2)	(3)
Mismatched	-0.151***		
	(0.037)		
Needs Leader, Gets Manager		-0.195***	-0.210***
		(0.048)	(0.048)
Needs Manager, Gets Leader		-0.086*	-0.104**
		(0.048)	(0.052)
Needs Manager, Gets Manager			-0.051
			(0.060)
$\hat{\beta}_{LM} - \hat{\beta}_{ML} = 0$ (p-value)		0.051	0.050
Country FE	Yes	Yes	Yes
Observations	4,576	4,576	4,576
Adjusted R^2	0.341	0.342	0.342

Notes: This table presents OLS regressions to estimate the relation between the (mis)match of the CEO type the firm has and the CEO type it needs, on the one hand, and labor productivity, on the other hand. Mismatched is an indicator equal to 1 if the firm's ideal and actual CEO types are mismatched and 0 otherwise. Needs Leader, Gets Manager is an indicator equal to 1 if the firm's ideal CEO type is leader and its actual CEO type is manager and 0 otherwise. LM refers to Needs Leader, Gets Manager, and ML refers to Needs Manager, Gets Leader. The regression controls for whether the interview was conducted on the phone and the fiscal year it refers to. Table 2 provides summary statistics. Bootstrapped standard errors with 1,000 replications are shown in parentheses. ***, **, and * correspond to 1%, 5%, and 10% levels of statistical significance.

G.10 Sector fixed effects and firm age

Our baseline results in the main paper do not include sector fixed effects and firm age. This subsection shows that our findings are robust to including these additional variables.

Table G.31 shows the significant positive relationship between actual leader status and predicted leader-type CEO propensity. Table G.32 confirms that our finding of horizontal differentiation between the leader and manager CEO types in Table 9 still holds. The coefficients on the Needs Manager, Gets Leader and Needs Leader, Gets Manager variables remain negative and significant, while the coefficient of the Needs Manager, Get Manager variable and the p-value in the "$\hat{\beta}_{LM} - \hat{\beta}_{ML} = 0$ (p-value)" row remain insignificant. Other results are available upon request.

Table G.31 Leader status and predicted leader propensity (additional controls)

Dependent variable →	Leader status
Leader propensity	0.206**
	(0.084)
Country FE	Yes
Sector FE	Yes
Firm age	Yes
Observations	4,576
Adjusted R^2	0.061

Notes: This table presents estimates from an OLS regression to estimate the relation between predicted leader-type CEO propensity and actual leader-type CEO status. Leader-type CEO status is an indicator variable equal to 1 if the CEO time use index is greater than 3, and 0 otherwise. Leader-type CEO propensity in the testing sample is predicted using coefficients estimated by regressing the actual leader-type CEO status of training sample firms on firm characteristics. The regression controls for firm age, whether the interview was conducted on the phone and the fiscal year it refers to, as well as sector fixed effects. Table 2 provides summary statistics, and Table E.1 provides information on sectors. Bootstrapped standard errors with 1,000 replications are shown in parentheses. ***, **, and * correspond to 1%, 5%, and 10% levels of statistical significance.

Table G.32 Productivity and leader mismatch (additional controls)

Dependent variable →	Log (labor productivity)		
	(1)	(2)	(3)
Mismatched	-0.147***		
	(0.036)		
Needs Leader, Gets Manager		-0.174***	-0.183***
		(0.044)	(0.046)
Needs Manager, Gets Leader		-0.097*	-0.110**
		(0.049)	(0.056)
Needs Manager, Gets Manager			-0.038
			(0.064)
$\hat{\beta}_{LM} - \hat{\beta}_{ML} = 0$ (p-value)		0.153	0.183
Country FE	Yes	Yes	Yes
Sector FE	Yes	Yes	Yes
Firm age	Yes	Yes	Yes
Observations	4,576	4,576	4,576
Adjusted R^2	0.387	0.387	0.387

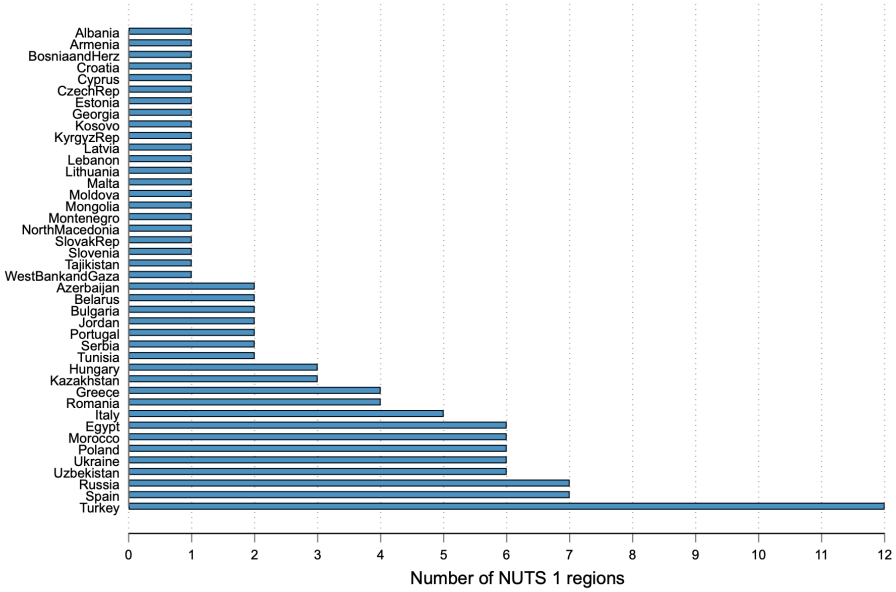
Notes: This table presents OLS regressions to estimate the relation between the (mis)match of the CEO type the firm has and the CEO type it needs, on the one hand, and labor productivity, on the other hand. Mismatched is an indicator equal to 1 if the firm's ideal and actual CEO types are mismatched and 0 otherwise. Needs Leader, Gets Manager is an indicator equal to 1 if the firm's ideal CEO type is leader and its actual CEO type is manager and 0 otherwise. LM refers to Needs Leader, Gets Manager, and ML refers to Needs Manager, Gets Leader. The regression controls for firm age, whether the interview was conducted on the phone and the fiscal year it refers to, and sector fixed effects. Table 2 provides summary statistics. Bootstrapped standard errors with 1,000 replications are shown in parentheses. ***, **, and * correspond to 1%, 5%, and 10% levels of statistical significance.

G.11 Region fixed effects

Our baseline results in the main paper use country fixed effects. In this subsection, we confirm that our results are unaffected by using more refined NUTS1 region fixed effects instead. For non-EU member or candidate countries, we follow the NUTS regulation on minimum and maximum population thresholds for the size of the NUTS regions.[12] Specifically, NUTS1 regions have a minimum of three and maximum of seven million inhabitants. There are two exceptions to this rule: a) the West region in Ukraine has more than seven million inhabitants but cannot be split as this was the ES stratification region, and b) the West region in Azerbaijan has fewer than three million inhabitants, but joining it with another region in Azerbaijan is impractical due to its location and the underlying ES stratification regions. Figure G.4 shows the number of NUTS1 regions in each of the testing sample countries.

[12]See `https://ec.europa.eu/eurostat/web/nuts/principles-and-characteristics`.

Figure G.4 Number of NUTS1 regions by country



Notes: This figure shows the number of NUTS1 regions by country. Regions are defined at the NUTS1 level for EU members and candidate countries and at the equivalent level for non-EU member or candidate countries. The latter were defined following the NUTS regulation on minimum and maximum population thresholds for the size of the NUTS regions (see `https://ec.europa.eu/eurostat/web/nuts/principles-and-characteristics`) where possible. Further details are available upon request.

Table G.33 shows the significant relationship between actual leader status and predicted leader-type CEO propensity. Table G.34 confirms that our finding of horizontal differentiation between the leader and manager CEO types in Table 9 still holds. The coefficients on the Needs Manager, Gets Leader and Needs Leader, Gets Manager variables remain negative and significant, while the coefficient of the Needs Manager, Get Manager variable and the p-value in the "$\hat{\beta}_{LM} - \hat{\beta}_{ML} = 0$ (p-value)" row remain insignificant.

Table G.35 shows that using NUTS1 region fixed effects rather than country fixed effects slightly weakens some of the results in Table 11. This is not surprising given the strong regional clustering of some these variables. For example, for the "Log minimum distance from schools offering business degrees" variable—which nevertheless is still significantly associated with leader-type CEOs at the 10% level—NUTS1 region fixed effects explain 54.6% of the variation in the variable in contrast to 30% with country fixed effects. Other results are available upon request.

Table G.33 Leader status and predicted leader propensity (NUTS1 region fixed effects)

Dependent variable →	Leader status
Leader propensity	0.299***
	(0.070)
NUTS1 FE	Yes
Observations	4,576
Adjusted R^2	0.094

Notes: This table presents estimates from an OLS regression to estimate the relation between predicted leader-type CEO propensity and actual leader-type CEO status. Leader-type CEO status is an indicator variable equal to 1 if the CEO time use index is greater than 3, and 0 otherwise. Leader-type CEO propensity in the testing sample is predicted using coefficients estimated by regressing the actual leader-type CEO status of training sample firms on firm characteristics. The regression controls for whether the interview was conducted on the phone and the fiscal year it refers to. Table 2 provides summary statistics, and Table E.1 provides information on sectors. Bootstrapped standard errors with 1,000 replications are shown in parentheses. ***, **, and * correspond to 1%, 5%, and 10% levels of statistical significance.

Table G.34 Productivity and leader mismatch (NUTS1 region fixed effects)

Dependent variable →	Log (labor productivity)		
	(1)	(2)	(3)
Mismatched	-0.141***		
	(0.036)		
Needs Leader, Gets Manager		-0.160***	-0.184***
		(0.045)	(0.045)
Needs Manager, Gets Leader		-0.108**	-0.135**
		(0.048)	(0.053)
Needs Manager, Gets Manager			-0.090
			(0.062)
$\hat{\beta}_{LM} - \hat{\beta}_{ML} = 0$ (p-value)		0.365	0.386
NUTS1 FE	Yes	Yes	Yes
Observations	4,576	4,576	4,576
Adjusted R^2	0.366	0.366	0.366

Notes: This table presents OLS regressions to estimate the relation between the (mis)match of the CEO type the firm has and the CEO type it needs, on the one hand, and labor productivity, on the other hand. Mismatched is an indicator equal to 1 if the firm's ideal and actual CEO types are mismatched and 0 otherwise. Needs Leader, Gets Manager is an indicator equal to 1 if the firm's ideal CEO type is leader and its actual CEO type is manager and 0 otherwise. LM refers to Needs Leader, Gets Manager, and ML refers to Needs Manager, Gets Leader. The regression controls for whether the interview was conducted on the phone and the fiscal year it refers to. Table 2 provides summary statistics. Bootstrapped standard errors with 1,000 replications are shown in parentheses. ***, **, and * correspond to 1%, 5%, and 10% levels of statistical significance.

Table G.35 Firm-level mismatch predictors

Dependent variable →	Mismatch	Leader-type CEO
	(1)	(2)
Local education market		
Log minimum distance from HEIs	0.006	-0.012*
offering business degrees	(0.007)	(0.007)
Log minimum distance from other	0.002	-0.003
HEIs	(0.007)	(0.007)
Local labor market thickness proxies		
Capital/main business city	-0.030	0.062*
	(0.036)	(0.036)
Population >1 million	0.043	0.004
	(0.037)	(0.037)
250,000 < Population ⩽ 1 million	-0.002	0.021
	(0.029)	(0.029)
50,000 < Population ⩽ 250,000	0.004	0.007
	(0.024)	(0.023)
Firm-level characteristics		
Exporter or importer status	-0.029	0.093***
	(0.021)	(0.020)
At least 25% foreign ownership	-0.062***	0.045**
	(0.021)	(0.020)
Has a board of directors or	-0.062***	0.051***
supervisory board	(0.017)	(0.017)
Credit constrained	0.010	-0.021
	(0.020)	(0.019)
Sole proprietorship	0.091***	-0.106***
	(0.030)	(0.030)
NUTS1 FE	Yes	Yes
Observations	4,576	4,576
Adjusted R^2	0.022	0.080

Notes: The regression controls for whether the interview was conducted on the phone and the fiscal year it refers to. For observations with missing education market predictors, we impute constant values within countries, which are captured by country fixed effects. For other categorical predictors, we include an additional (unreported) category if the value was originally missing (see Appendix D.2 for more details). Table 2 provides summary statistics. Bootstrapped standard errors with 1,000 replications are shown in parentheses. ***, **, and * correspond to 1%, 5%, and 10% levels of statistical significance.

G.12 Standard errors clustered at the locality level

Our baseline results use bootstrapped standard errors with 1,000 replications, without clustering. This subsection shows that our findings are robust to clustering standard errors at the locality level (the city, town, or village where the firm is located), which may capture interactions across firms within local markets. In our data, the locality clusters are fairly small: the 4,576 firms in our training sample come from 2,034 localities. Table G.36 shows that the relationship between actual leader-type CEO status and predicted leader-type CEO propensity remains positive and significant when we cluster standard errors by locality.

Table G.36 Leader status and predicted leader propensity (clustered standard errors)

Dependent variable →	Leader status
Leader propensity	0.247***
	(0.075)
Country FE	Yes
Observations	4,576
Adjusted R^2	0.061

Notes: This table presents estimates from an OLS regression to estimate the relation between predicted leader-type CEO propensity and actual leader-type CEO status. Leader-type CEO status is an indicator variable equal to 1 if the CEO time use index is greater than 3, and 0 otherwise. Leader-type CEO propensity in the testing sample is predicted using coefficients estimated by regressing the actual leader-type CEO status of training sample firms on firm characteristics. The regression controls for whether the interview was conducted on the phone and the fiscal year it refers to. Table 2 provides summary statistics, and Table E.1 provides information on sectors. Standard errors clustered by locality are shown in parentheses. ***, **, and * correspond to 1%, 5%, and 10% levels of statistical significance.

Similarly, Table G.37 confirms our finding of horizontal differentiation between the leader and manager CEO types, as in Table 9. The coefficients on the Needs Manager, Gets Leader and Needs Leader, Gets Manager variables remain negative and significant, while the coefficient of the Needs Manager, Get Manager variable and the p-value in the "$\hat{\beta}_{LM} - \hat{\beta}_{ML} = 0$ (p-value)" row remain insignificant.

Table G.37 Productivity and leader mismatch (clustered standard errors)

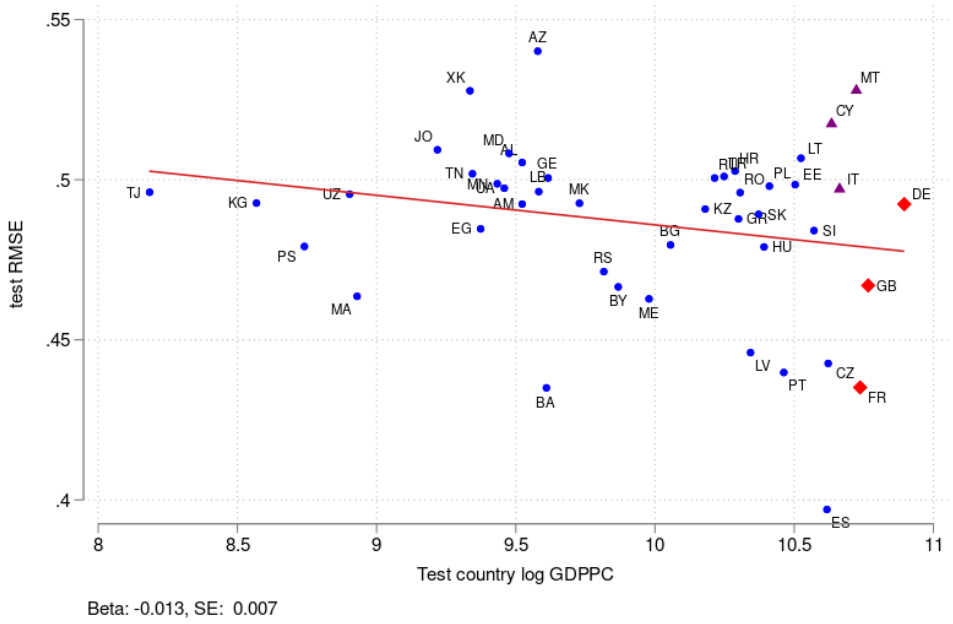
Dependent variable →	Log (labor productivity)		
	(1)	(2)	(3)
Mismatched	-0.157***		
	(0.040)		
Needs Leader, Gets Manager		-0.177***	-0.197***
		(0.051)	(0.052)
Needs Manager, Gets Leader		-0.122**	-0.145***
		(0.051)	(0.054)
Needs Manager, Gets Manager			-0.076
			(0.067)
$\hat{\beta}_{LM} - \hat{\beta}_{ML} = 0$ (p-value)		0.356	0.387
Country FE	Yes	Yes	Yes
Observations	4,576	4,576	4,576
Adjusted R^2	0.342	0.342	0.342

Notes: This table presents OLS regressions to estimate the relation between the (mis)match of the CEO type the firm has and the CEO type it needs, on the one hand, and labor productivity, on the other hand. Mismatched is an indicator equal to 1 if the firm's ideal and actual CEO types are mismatched and 0 otherwise. Needs Leader, Gets Manager is an indicator equal to 1 if the firm's ideal CEO type is leader and its actual CEO type is manager and 0 otherwise. LM refers to Needs Leader, Gets Manager, and ML refers to Needs Manager, Gets Leader. The regression controls for whether the interview was conducted on the phone and the fiscal year it refers to. Table 2 provides summary statistics. Standard errors clustered by locality are shown in parentheses. ***, **, and * correspond to 1%, 5%, and 10% levels of statistical significance.

G.13 RMSE and log GDP per capita relationship using only the United States as the training sample

Figure G.5 conducts the same prediction exercise as Figure 3, using only US firms as the training sample. That is, firms in the US sample are used to reestimate the coefficients of variables selected by LASSO in Table 6. We choose not to rerun the selection of variables as that may be too demanding for this smaller training sample and would make these results more difficult to compare with the prediction, on which we base our results.

Figure G.5 Out-of-sample prediction and log GDP per capita (US training sample)



Notes: This figure uses firms in the US sample to reestimate the coefficients of variables selected by LASSO in Table 6. These estimates are then used to predict the leader-type CEO propensity of firms in other countries. We derive the root mean squared error values by regressing these estimates on firms' leader-type CEO status in individual countries. GDP per capita is measured in 2019 (in constant 2017 international $). *Source*: Authors' calculations based on the Diary and Survey data.

G.14 Using the share of leaders as the Needs Leader threshold

Our baseline results use 0.5 as the threshold for the discretization of our continuous predicted leader-type CEO propensity variable (see Section 4.1). This subsection uses the share of leaders in the combined Diary and Survey data (0.55) instead.

Table G.38 shows that this threshold change results in slightly different conclusions than Table 9. The coefficients on the Needs Manager, Gets Leader and Needs Leader, Gets Manager variables are still negative and significant, and the coefficient on the Needs Manager, Gets Manager variable is not significant. However, the p-value in the "$\hat{\beta}_{LM} - \hat{\beta}_{ML} = 0$ (p-value)" row is now significant, at a 5% significance level. This implies that Needs Leader, Gets Manager firms are more negatively affected than Needs Manager, Gets Leader firms. Other results are unaffected and available upon request.

Table G.38 Productivity and leader mismatch (using the share of leaders as the Needs Leader threshold)

Dependent variable →	Log (labor productivity)		
	(1)	(2)	(3)
Mismatched	-0.153***		
	(0.036)		
Needs Leader, Gets Manager		-0.208***	-0.234***
		(0.050)	(0.052)
Needs Manager, Gets Leader		-0.089**	-0.117**
		(0.045)	(0.051)
Needs Manager, Gets Manager			-0.070
			(0.057)
$\hat{\beta}_{LM} - \hat{\beta}_{ML} = 0$ (p-value)		0.033	0.030
Country FE	Yes	Yes	Yes
Observations	4,576	4,576	4,576
Adjusted R^2	0.341	0.342	0.342

Notes: This table presents OLS regressions to estimate the relation between the (mis)match of the CEO type the firm has and the CEO type it needs, on the one hand, and labor productivity, on the other hand. Mismatched is an indicator equal to 1 if the firm's ideal and actual CEO types are mismatched and 0 otherwise. Needs Leader, Gets Manager is an indicator equal to 1 if the firm's ideal CEO type is leader and its actual CEO type is manager and 0 otherwise. LM refers to Needs Leader, Gets Manager, and ML refers to Needs Manager, Gets Leader. The regression controls for whether the interview was conducted on the phone and the fiscal year it refers to. Table 2 provides summary statistics. Bootstrapped standard errors with 1,000 replications are shown in parentheses. ***, **, and * correspond to 1%, 5%, and 10% levels of statistical significance.

G.15 Variation in the strength of the mismatch

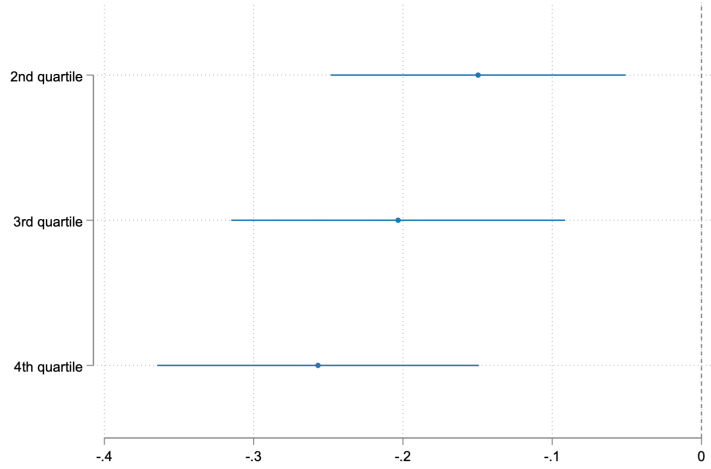
In the main paper, we use a binary variable to capture whether firms need leader- or manager-type CEOs. In this subsection, we show that variation in the predicted $LeaderPropensity_i$ index can also be used to capture how much firms need their ideal CEO type. For this, we calculate the intensity of the mismatch for each firm as $MismatchIntensity = |Leader_i - LeaderPropensity_i|$, which we then group into quartiles over the testing sample. We then use these categories to estimate

$$Y_i = \beta_0 + \sum_{k=2}^{4} \beta_1^k MismatchIntensityQ_i^k + \zeta_c + \epsilon_i,$$

where Y_i is firm i's log labor productivity, $MismatchIntensityQ_i^k$ is an indicator equal to 1 if the firm's $MismatchIntensity_i$ falls into the kth quartile, and ζ_c are country fixed effects. Figure G.6 presents the mismatch coefficients estimated in this regression, which indicate that the more

intensively firms are mismatched, the lower their estimated productivity.

Figure G.6 Estimated productivity-mismatch coefficients by intensity of mismatch



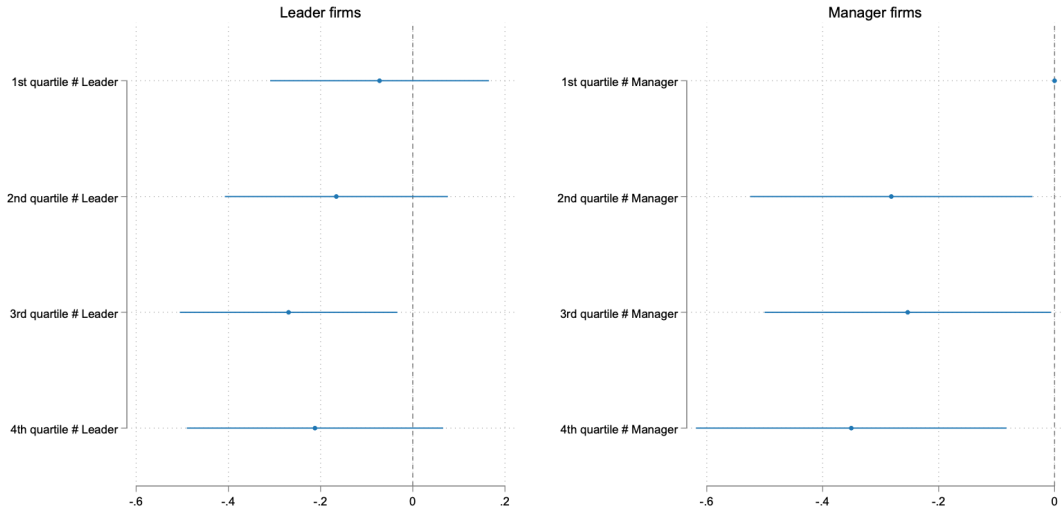
Notes: The figure presents estimated productivity-mismatch coefficients and 95% confidence intervals by intensity of mismatch quartiles.

If we instead differentiate between firms that currently have manager- or leader-type CEOs by running

$$\begin{aligned} Y_i &= \beta_0 + \sum_{k=2}^{4} \beta_1^{k,M} Manager_i \cdot MismatchIntensityQ_i^k + \\ &\quad + \sum_{k=1}^{4} \beta_1^{k,L} Leader_i \cdot MismatchIntensityQ_i^k + \zeta_c + \epsilon_i, \end{aligned}$$

we get the coefficients shown in Figure G.7. These two indicate a similar, increasingly negative association for more firms which are more mismatched.

Figure G.7 Estimated productivity-mismatch coefficients by intensity of mismatch



Notes: The figure presents estimated productivity-mismatch coefficients and 95% confidence intervals by intensity of mismatch quartiles.

G.16 Back-of-the-envelope calculation: Additional tables

Table G.39 Actual and ideal CEO types

	Ideal CEO type			
	Leader		Manager	
Actual CEO type	No. of firms	Sales	No. of firms	Sales
Leader	1,677	65,326	737	9,780
Manager	1,477	31,345	561	7,726

Notes: This table shows the distribution of firms in the testing sample by actual and ideal CEO type. In addition to the number of firms, total sales by type (in 2019 million USD) are also reported.

Table G.40 Productivity and leader mismatch (randomized reallocation)

Firm type	Number of firms	Sales (million US$)
Short-Run Leader, ML	137	3,342.04
Short-Run Manager, ML	600	6,438.62
Short-Run Leader, LM	966	21,045.60
Short-Run Manager, LM	511	10,328.10
Short-Run Leader, MM	54	858.71
Short-Run Manager, MM	507	6,868.09
Short-Run Manager, LL	421	15,708.80
Short-Run Leader, LL	1,256	49,617.60
Total	4,453	114,208.00

Notes: This table presents the number of firms and total sales for each group created in the randomized short-run reallocation exercise of Section 5.1.1. Short-Run Leader and Short-Run Manager indicate which type of CEOs the firms are assigned to in the hypothetical reallocation. LM refers to Needs Leader, Gets Manager firms; ML refers to Needs Manager, Gets Leader firms; LL refers to Needs Leader, Gets Leader firms; and MM refers to Needs Manager, Gets Manager firms.

Table G.41 Productivity and leader mismatch (need-based allocation)

Firm type	Number of firms	Sales (million US$)
Short-Run Leader, ML	134	2,831.84
Short-Run Manager, ML	603	6,948.81
Short-Run Leader, LM	966	19,814.5
Short-Run Manager, LM	511	11,531
Short-Run Leader, MM	57	942.89
Short-Run Manager, MM	504	6,783.91
Short-Run Leader, LL	1,257	53,259.2
Short-Run Manager, LL	420	12,067.2
Total	4,452	114,208.00

Notes: This table presents the number of firms and total sales for each group created in the need-based short-run reallocation exercise of Section 5.1.1. Short-Run Leader and Short-Run Manager indicate which type of CEOs the firms are assigned to in the hypothetical reallocation. LM refers to Needs Leader, Gets Manager firms; ML refers to Needs Manager, Gets Leader firms; LL refers to Needs Leader, Gets Leader firms; and MM refers to Needs Manager, Gets Manager firms.

G.17 Firm-level predictors with all outcome variables

Table G.42 Firm-level mismatch predictors

Dependent variable →	Mismatch	Needs Leader, Gets Manager	Needs Manager, Gets Leader	Leader-type CEO	Needs Leader Firm
	(1)	(2)	(3)	(4)	(5)
Local education market					
Minimum distance from schools offering business degrees	0.011*	0.014**	-0.003	-0.019***	-0.001
	(0.006)	(0.006)	(0.004)	(0.006)	(0.005)
Minimum distance from other HEIs	0.002	-0.001	0.003	-0.001	-0.004
	(0.006)	(0.006)	(0.004)	(0.006)	(0.005)
Local labor market thickness proxies					
Capital/main business city	-0.023	-0.060**	0.037*	0.092***	-0.005
	(0.029)	(0.028)	(0.020)	(0.029)	(0.025)
Population: >1 million	0.068**	0.081***	-0.013	-0.079***	0.015
	(0.029)	(0.029)	(0.018)	(0.030)	(0.025)
250,000 < Population ⩽ 1 million	0.011	0.026	-0.015	0.010	0.051**
	(0.026)	(0.024)	(0.018)	(0.026)	(0.023)
50,000 < Population ⩽ 250,000	0.013	0.016	-0.002	-0.006	0.013
	(0.023)	(0.020)	(0.016)	(0.023)	(0.019)
Firm-level characteristics					
Exporter or importer status	-0.039*	-0.058***	0.019	0.111***	0.034**
	(0.021)	(0.019)	(0.013)	(0.019)	(0.017)
At least 25% foreign ownership	-0.059***	0.027	-0.086***	0.047**	0.161***
	(0.021)	(0.019)	(0.014)	(0.021)	(0.017)
Has a board of directors or supervisory board	-0.053***	0.016	-0.069***	0.044***	0.128***
	(0.017)	(0.015)	(0.012)	(0.016)	(0.015)
Credit constrained	0.005	0.027	-0.022*	-0.027	0.022
	(0.019)	(0.018)	(0.012)	(0.019)	(0.017)
Sole proprietorship	0.100***	0.237***	-0.136***	-0.108***	0.264***
	(0.029)	(0.030)	(0.014)	(0.028)	(0.022)
Country FE	Yes	Yes	Yes	Yes	Yes
Observations	4,576	4,576	4,576	4,576	4,576
Adjusted R^2	0.022	0.098	0.109	0.08	0.182

Notes: This table presents OLS regressions to estimate the relation between, on the one hand, local education and labor market characteristics and firm-level characteristics and, on the other hand, the mismatch between the CEO type the firm needs and the CEO type it has, the direction of this mismatch, and whether it has a leader-type CEO. The regression controls for whether the interview was conducted on the phone and the fiscal year it refers to. For 49 observations with missing education market predictors, we impute constant values within countries, which are captured by country fixed effects. For other categorical predictors, we include an additional (unreported) category if the value was originally missing (see Appendix D.2 for more details). Table 2 provides summary statistics. Bootstrapped standard errors with 1,000 replications are shown in parentheses. ***, **, and * correspond to 1%, 5%, and 10% levels of statistical significance.

G.18 Firm-level education predictors without other controls

Table G.43 conducts the same estimation as Table 11, with distance variables included in levels. The results indicate that with this functional form, distance to schools offering business degrees is also significantly associated with firms being more likely to have leader-type CEOs.

Table G.43 Firm-level mismatch predictors

Dependent variable →	Mismatch	Leader-type CEO
	(1)	(2)
Local education market		
Minimum distance from schools offering business degrees	0.008 (0.006)	-0.020*** (0.006)
Minimum distance from other HEIs	0.000 (0.006)	-0.004 (0.006)
Firm-level characteristics	No	No
Labor market thickness proxies	No	No
Country FE	Yes	Yes
Observations	4,576	4,576
Adjusted R^2	0.014	0.063

Notes: This table presents OLS regressions to estimate the relation between, on the one hand, local education and labor market characteristics and firm-level characteristics and, on the other hand, the mismatch between the CEO type the firm needs and the CEO type it has, the direction of this mismatch, and whether it has a leader-type CEO. The regression controls for whether the interview was conducted on the phone and the fiscal year it refers to. For observations with missing education market predictors, we impute constant values within countries, which are captured by country fixed effects. For other categorical predictors, we include an additional (unreported) category if the value was originally missing (see Appendix D.2 for more details). Table 2 provides summary statistics. Bootstrapped standard errors with 1,000 replications are shown in parentheses. ***, **, and * correspond to 1%, 5%, and 10% levels of statistical significance.

G.19 Firm-level education predictors in levels

Table G.44 conducts the same estimation as Table 11, with distance variables included in levels. The results indicate that with this functional form, distance from schools offering business degrees is also significantly associated with firms having leader-type CEOs.

Table G.44 Firm-level mismatch predictors

Dependent variable →	Mismatch	Leader- type CEO
	(1)	(2)
Local education market		
Minimum distance from schools	0.219	-0.530***
offering business degrees (level)	(0.170)	(0.149)
Minimum distance from other HEIs	0.053	-0.127*
(level)	(0.071)	(0.069)
Firm-level characteristics	Yes	Yes
Labor market thickness proxies	Yes	Yes
Country FE	Yes	Yes
Observations	4,576	4,576
Adjusted R^2	0.022	0.081

Notes: This table presents OLS regressions to estimate the relation between, on the one hand, local education and labor market characteristics and firm-level characteristics and, on the other hand, the mismatch between the CEO type the firm needs and the CEO type it has, the direction of this mismatch, and whether it has a leader-type CEO. The regression controls for whether the interview was conducted on the phone and the fiscal year it refers to. For observations with missing education market predictors, we impute constant values within countries, which are captured by country fixed effects. For other categorical predictors, we include an additional (unreported) category if the value was originally missing (see Appendix D.2 for more details). Table 2 provides summary statistics. Bootstrapped standard errors with 1,000 replications are shown in parentheses. ***, **, and * correspond to 1%, 5%, and 10% levels of statistical significance.