EIEF WORKING PAPER SERIES

EINAUDI INSTITUTE FOR ECONOMICS AND FINANCE

EIEF Working Paper 24/13

November 2024

Careers and Wages in Family Firms: Evidence from Matched Employer-Employee Data

By

Edoardo Di Porto
(University of Naples Federico II, CSEF and INPS)

Marco Pagano
(University of Naples Federico II, CSEF, ECGI, EIEF and CEPR)

Vincenzo Pezone
(Tilburg University)

Raffaele Saggio
(University of British Columbia and NBER)

Fabiano Schivardi
(Luiss University, EIEF and CEPR)

Careers and Wages in Family Firms: Evidence from Matched Employer-Employee Data*

Edoardo Di Porto[1], Marco Pagano[2], Vincenzo Pezone[3], Raffaele Saggio[4], and Fabiano Schivardi[5]

[1] *University of Naples Federico II, CSEF and INPS*
[2] *University of Naples Federico II, CSEF, ECGI, EIEF and CEPR*
[3] *Tilburg University*
[4] *University of British Columbia and NBER*
[5] *Luiss University, EIEF and CEPR*

November 2024

Abstract

We investigate compensation policies in family and non-family firms using a novel employer-employee matched dataset comprising nearly the universe of Italian incorporated firms and ownership information. Family firms pay significantly lower wages and offer slower and less rewarding careers. Differences in worker sorting account for half of the wage gap while productivity differences and compensating differentials explain little of the residual gap. The wage distribution in family firms is more compressed, with infrequent promotions. We rationalize this evidence with a model where family owners seek to maintain control, creating a "glass ceiling" that limits their employees' career progression.

Keywords: family firms, corporate control, wage, career, workers, human capital, productivity, management.

JEL Codes: D22, D23, D24, G32, G34, J24, J31, J32, J62, M12, M51, M52, M54.

*We are grateful to Jerome Adda, Manuel Adelino, Lorenzo Cappellari, Pietro Coretto, Andrew Ellul, Daniel Ferreira, Sara Formai, Giulia Giupponi, Isaac Hacamo, Marc Kaufmann, Yongseok Kim, Francis Kramarz, Salvatore Lattanzio, Virginia Minni, Daniel Paravisini, Vincenzo Scrutinio, Andrei Shleifer, and seminar participants at EIEF, Imperial College, LUISS University, Maastricht University, Mannheim University, the NBER Organizational Economics Spring Meeting, the Tilburg Finance Brown Bag Seminar, the UCSD Rady School of Management Brown Bag Finance Seminar, the University of Navarra, the XXXVIII AIEL Conference, the 6$^{\text{th}}$ Annual Conference of the Visitinps Scholars Program, the Lake District Workshop in Corporate Finance, the CSEF-RCFS Conference on Finance, Labor and Inequality, the EEA-ESEM Congress, the EFA Annual Meeting, the Organizational Economics Summer Symposium and the Dauphine Family Business Online Seminar. We also thank Valentina Rutigliano for providing excellent research assistance. Marco Pagano and Vincenzo Pezone acknowledge support by the VisitINPS Scholars program for providing access to data on workers' careers. This research project was supported by the ERC Advanced Grant on Industrial Structure and the European Productivity Growth Divergence (awarded to Fabiano Schivardi, ERC grant 835201), and by the PRIN grant on Information, Firms' Performance and Workers' Careers (awarded by the Italian Ministry of the University and Research to Marco Pagano). The findings and conclusions expressed are solely those of the authors and do not represent the views of INPS.

1 Introduction

Recent research shows that firm heterogeneity is important in determining wages (Card, 2022). Various factors may explain why firms pay different wages to the same type of worker. For example, in monopsony models, differences in firm productivity can affect wages (Syverson, 2011; Card et al., 2018). Compensating differentials too can rationalize pay differentials (Rosen, 1986; Sorkin, 2018). But in a frictional labor market wage disparities across firms can persist even after accounting for heterogeneity in productivity and amenities if, say, company owners and managers have different preferences on wage-setting and promotions practices (Bloom and Van Reenen, 2010; Acemoglu et al., 2022).

This paper studies the role of firm ownership in determining wages, focusing on the divide between family-owned and other firms. Family firms make up a significant share of businesses around the world (Aminadav and Papaioannou, 2020), but their impact on wages is still not clearly understood. Some studies find that family firm owners tend to value the private benefits of control (Bandiera et al., 2018) and family prestige (Burkart et al., 2003), which can result in lower wages for non-family employees (Bandiera et al., 2015; Lemos and Scur, 2019). Others note instead that family firms have higher profitability (Sraer and Thesmar, 2007; Anderson and Reeb, 2003) and stronger incentives to foster positive relationships with employees (Mueller and Philippon, 2011), which can result in more rent-sharing (Ellul et al., 2017). The existing empirical evidence on the impact of family firms on wages suffers from two key limitations. First, family firm status is often observed only for a subset of highly selected firms – typically, publicly listed companies. Second, the evidence is based on firm-level data, which does not permit to take into account the role of workers' sorting, i.e. the possibility that higher-skilled workers are more likely to work for, say, non-family firms.

We overcome these problems by constructing a novel dataset that combines matched Italian employer-employee data from social security records, firms' financial records, and ownership structure. We thus observe family firm status for about a million companies and track the career progression and job transitions of all their workers (around 20 million) over 14 years. We document that family firms (FFs) pay wages 16 log points lower than non-family firms (NFFs). The richness of our data permits us to decompose this gap along four main margins: worker selection, firm productivity, rent-sharing, and compensating differentials. Half of the family-firm wage discount found is due to worker sorting, as more skilled workers tend to work in NFFs. At the same time, FFs tend to be less productive but offer more rent-sharing. After accounting for differences in sorting, productivity, and rent-sharing, we observe a remaining FF wage discount of around 10 log points. Interestingly, none of this residual gap appears to be driven by compensating differentials. Instead, our evidence suggests the existence of a "glass ceiling" in FFs as promotions to top positions are less common and yield smaller salary increases than in NFFs.

To identify family firms in these administrative data, we proceed in two steps. First, based on the data for the near-universe of Italian limited liability companies, we reconstruct their ownership chains and identify their ultimate shareholders (Porta et al., 1998; Claessens et al., 2000; Faccio and Lang, 2002). Second, we identify family relationships between shareholders using last names and shared addresses: a family firm is defined as a company in which a single family holds the majority of shares. By this method, we classify over 900,000 firms in Italy.

Using the complete dataset of job transitions for Italian workers, we then estimate the wage policy of a given firm using the two-way fixed effects specification popularized by Abowd et al. (1999) (AKM). Firms' wage policies—captured by the AKM firm effects—are then averaged for FFs and for NFFs to determine the size of the pay gap between these two types of firms not explained by worker selection. We find that in Italy FFs pay around 16 log points less than NFFs. About 50% of the gap depends on FFs' tendency to employ less skilled workers, as proxied by their AKM worker fixed effect.

Firm heterogeneity thus explains the other half of the raw gap in pay between FFs and NFFs. To better understand the sources of this heterogeneity, we project the estimated firm effects on firm productivity—proxied by value added per worker—separately for FFs and NFFs, to assess the role of differences in productivity and rent-sharing. We then apply to this specification a Oaxaca decomposition to assess the role of rent-sharing and differences in productivity in explaining the gap in firm effects between family and non-family firms (Oaxaca, 1973; Card et al., 2015). The lower productivity of FFs explains 25% (or 2 log points) of the gap in firm effects, while rent-sharing does not contribute at all; in fact, we find more rent-sharing within family firms. All these findings are confirmed when we exclude the small number of workers who are also shareholders, thus ruling out the possibility that FF employees have lower salaries because they are paid partly via dividends.

All told, after accounting for differences in worker sorting, firm productivity, and rent-sharing there remains a 10-log-point difference in wages between FFs and NFFs. In a frictionless labor market, any remaining difference in pay between FFs and NFFs should be explained by FFs' providing more amenities (Rosen, 1974, 1986). However, using the revealed preference approach of Sorkin (2018) to rank firms based on voluntary moves by workers, we find that FFs actually offer *less* amenities.

If the FFs' wage discount does not depend on compensating differentials, what can explain this systematic gap in pay? We argue that it may flow from the presence of a "glass-ceiling" in family firms. A stark pattern that emerges from our data is that the gaps in pay between family and non-family employers are considerably larger for workers in the upper wage percentiles within the firm. In other words, the top earners in FFs make substantially less than in NFFs, while the wage gap is much smaller for median wages and for workers in the lower tail of the distribution. This suggests that FFs and NFFs differ substantially in career prospects.

Consistent with this hypothesis, FFs have a substantially smaller fraction of managers and middle managers and are less likely to promote their workers internally. Not only are promotions less frequent, but their returns to the worker are also lower: controlling for a host of firm and worker characteristics, the same type of promotion (from rank-and-file worker to middle manager, or from middle manager to manager) produces smaller earnings gains in family firms.

Our preferred explanation for this pattern is that family firms place greater value on "private benefits of control": as families do not want to relinquish control, they are reluctant to share technical and operational know-how with workers outside the family. We test one implication of this theory with an analysis of firms' performance following the death of the CEO. We hypothesize that the loss of the top manager should be highly disruptive in family firms, where decision-making is extremely concentrated. Conversely, in NFFs, with their greater delegation of responsibility, other employees should be prepared to step in and make strategic decisions.

As we predict, following the death of the CEO productivity drops more in FFs than in NFFs. Interestingly, we find that in FFs whose CEOs do not belong to the controlling family there is no effect of the death of the top manager, which dovetails with the idea that as in these firms family owners have already surrendered control, CEOs are more easily replaced.

The last part of the paper presents a simple theoretical framework explaining the differences in the wage and employment policies between FFs and NFFs. In the model, workers and managers can acquire firm-specific productivity-enhancing skills by exerting costly effort. As effort is not observable, firms offer performance-based pay and managerial promotions as incentives. Promoting skilled workers is profit-maximizing but decreases the private benefits of control. In this trade-off, FF owners place a greater weight on private benefits, in accordance with the literature. Simple as it is, the model rationalizes our empirical results. The basic mechanism is that FFs rely less on incentives in setting wages and promotion prospects. This results in lower effort, hence lower wages and lower promotion probability: as in the data, workers earn less and are less likely to become managers in FFs. Moreover, due to weaker managerial incentives, on average promotions are associated with smaller managerial bonuses in FFs. Finally, the greater effort exerted by employees in NFFs is more than compensated for by higher wages and better career prospects: as a result, NFFs deliver greater utility to their employees. We close the model by showing that NFFs receive more applications because of the greater utility they ultimately provide to their employees. As the fractions of FFs and NFFs are assumed fixed, employment is more sharply rationed in NFFs than in FFs. Hence, applicants to family firms are more likely to be employed, which in equilibrium compensates for their lower utility upon being hired. In other words, *applying* to either type of firm delivers the same expected utility, explaining why the two types of firm can coexist in equilibrium.

This paper contributes to three main strands of research. The first is the literature on risk-sharing in the wage and employment policies of FFs and NFFs (Sraer and Thesmar, 2007; Bassanini et al., 2013; Bach and Serrano-Velarde, 2015; Ellul et al., 2017). Here our contribution is threefold. To begin with, while most of this literature is based on a selected sample of firms (usually publicly listed ones), we have data for the majority of Italian limited liability companies and can study family firms' pay policies in the entire economy. Second, the availability of matched employer-employee data gives us the opportunity to isolate the role of worker sorting in explaining the wage gap between FFs and NFFs, whereas most previous estimates are based on firm-level data. Finally, micro-level data on voluntary job changes by Italian workers allows us to measure the overall impact of compensating differentials in explaining the lower wages at FFs applying the revealed preferences approach proposed by Sorkin (2018), which does not require focusing on any specific amenity.

A second strand of the literature relates family ownership with firm performance, and in particular with managerial practices. There is growing empirical evidence that family ownership is associated with *worse* managerial practices. Bloom and Van Reenen (2007), Bloom et al. (2012), Bandiera et al. (2018), and Lemos and Scur (2019) find that family firms tend to be poorly managed. Lippi and Schivardi (2014) and Bandiera et al.

(2015) show that they hire managers of poorer quality. Mullins and Schoar (2016) present survey evidence that family firms are characterized by more hierarchical management. We show that these differences map systematically onto workers' wages and careers, resulting not only in lower wages, but also less propensity to hire and promote managers and lower within-firm salary inequality. The "glass ceiling" in FFs is a factor in their inferior performance, in line with previous work (Bennedsen et al., 2007; Pérez-González, 2006).[1] In our model, these differences can be rationalized by greater benefits of control in family firms (Burkart et al., 2003). However, this reluctance to delegate decision-making is costly privately, in that it leaves the firm vulnerable to disruptive events like the loss of the CEO, and costly socially, as it discourages employees' investment in human capital.

The third and final strand is the literature on the role of firm heterogeneity in wage inequality (Robinson, 1933; Card et al., 2013; Song et al., 2019; Bender et al., 2018; Card et al., 2018). We contribute here by showing the importance of ownership types in driving pay differences. Family firms may affect income distribution via four channels: the quality of their workers, their own productivity, their rent-sharing policies, and the amenities they provide. We show how the role of each channel can be quantified using estimates of AKM firm premia, data on value added per worker, and the ranking of firms' attractiveness as gauged by voluntary job moves. The fact that the wage gap between FFs and NFFs persists after accounting for these four channels is evidence of the latitude that firms have in setting wages (Card, 2022) and is consistent with a recent literature showing that standard profit-maximizing models are hard to reconcile with observed wage policies (Dube et al., 2018; Hjort et al., 2020; Hazell et al., 2022; Cullen et al., 2022).

The rest of the paper is organized as follows. Section 2 describes the data and Section 3 introduces the econometric framework. Section 4 estimates and decomposes the family firm wage discount, showing that it cannot be explained by compensating differentials. Section 5 analyzes the glass ceiling assumption and Section 6 introduces a simple model

[1] Note however that the evidence is not unequivocal in this respect, as Anderson and Reeb (2003) and Sraer and Thesmar (2007) find that family firms *outperform* non-family firms.

that rationalizes our findings. Finally, Section 7 concludes.

2 The data

This section details our data sources, describes the procedure for identifying family firms, and offers descriptive statistics for our key variables.

2.1 Data Sources

Our analysis relies on three main data sources. First, we obtain on-site access to the social security records of *all* Italian private sector employees through the *Visitinps* project. This matched employer-employee dataset has been made available to researchers selected through a refereed call for projects issued by the Italian National Institute for Social Security (INPS). The database has detailed information on earnings, type of contract (open-ended or fixed-term), part-time status, etc. It also presents a basic occupational classification, distinguishing between blue collar workers, white collar workers, middle management (*quadri*), and managers (*dirigenti*). The database also gives information on the reason for any job separation, which we use to identify voluntary moves (Di Addario et al., 2023).

Our second source is *Cerved*, a firm-level dataset with accounting information for all Italian nonfinancial limited liability companies.[2]

Information on ownership and management comes from our third data source, which we refer to as *Cerved-Infocamere*.[3] This is a digitized business registry, comprising three datasets. The first one, *Anagrafica*, covers the universe of Italian limited liability firms

[2]The primary data source is the business registry managed by Infocamere, an IT company owned by the Italian Chambers of Commerce. By law, all incorporated companies must file balance sheets and income statements with the business registry. Data providers, such as Cerved, can then acquire the data, process them, and sell them on the market.

[3]The primary data source is again Infocamere, to which all incorporate companies must file ownership and board composition. We purchased the data through Cerved, a private company that imported and made available to us the dataset in a user-friendly format.

for the period 2003–2019, with basic data such as date of creation, date of cessation (if any), and address. It numbers 2.47 million incorporated businesses. As we require all the firms to have basic accounting information, such as value added, we merge this dataset with the accounting information from *Cerved*, finding that 93% of the firms in the latter are also in *Anagrafica*.[4] The second file, *Soci*, has information on firms' owners, who could be either individuals or other firms. For each owner, the dataset includes name and tax ID, stake owned, and address. Since in Italy the tax ID is a known function of the individual's name, city of birth, date of birth, and gender, we can recover this information from the tax identifiers. The coverage of the ownership file is not complete, but it increases strongly over time – from 52% of the firms in the *Anagrafica* dataset for which we have accounting information in 2003 to 93% in 2019. The third file, *Esponenti*, lists the members of companies' boards of directors and their top executives, and contains individuals' tax ID, as well as their addresses and roles, as they appear in the companies' financial statements. Coverage is nearly complete.

Importantly, we can easily link these three datasets, as both individuals and companies are recorded with their respective tax ID in all three. In most of our analysis, the time span is 2005–2017.[5]

2.2 Identifying Family Firms

In most of our analysis, we define family firm status according to ownership. Our first step is to identify the ultimate owners of a firm. This distinction is relevant whenever a firm is at least partially owned by another firm, in which case we track the owning firm's shareholders, repeating the procedure along the entire control chain. Each ultimate shareholder is then assigned a share computed using the so-called "weakest link" principle:

[4] Conversely, we do not have accounting data for about half of the firms in *Anagrafica*, typically as they are cooperatives, public entities, financial or insurance companies.

[5] We choose 2005 as the starting year, as the coverage of the ownership database then becomes high enough, exceeding 80%. Conversely, we end the sample in 2017 because our administrative data extend up to 2020 and in the event-study evidence presented in Section 5 we use up to three years after the event year (CEO's or director's death or worker's promotion).

for example, if person A owns 10% of company B, which in turns owns 5% of company C, A is assumed to control 5% of company C.[6]

Neither *Cerved-Infocamere* nor *INPS* records family linkages; as a result, we need to make some assumptions in order to identify individuals who are likely to be relatives. First we assume that shareholders with the same last name belong to the same family.[7] Second, we assume that any two shareholders who differ in gender and surname but share the same address are life partners.

Throughout, we define a firm as family-owned if we can identify an individual or a group of family members who hold more than 50% of the shares. Given that the vast majority of our sample consists of small, non-listed firms, we chose a cutoff higher than those employed in most of the literature, which tends to study large, listed firms (Volpin, 2002; Sraer and Thesmar, 2007; Ellul et al., 2010; Ellul et al., 2017).

These assumptions entail two main risks of misclassification. First, two individuals sharing an address or last name may not actually be related. This type of error will tend to overstate the proportion of family firms. Second, when the controlling stake of a company is held by a firm for which we have no ownership information, we classify it as non-family. Yet the controlling firm may actually be controlled by a single individual or firm, for which we happen not to have ownership information (because, say, it is incorporated in a foreign country). This type of error will tend to underestimate the proportion of family firms.

Fortunately, for a subsample of firms we can cross-check the validity of our algorithm, and thus assess the magnitude of the possible misclassification: in 2018, *Cerved* developed an in-house classification of firm ownership for a sample of large and medium-sized firms, striving to accurately identify family relationships by consulting the real estate registries

[6]This is the standard approach followed in the governance literature (see for example La Porta et al., 1999, Faccio and Lang, 2002, Claessens et al., 2000, among others). See also Edwards et al. (2004) for a discussion.

[7]In Italy, women do not change their family names upon marriage. Hence, siblings always share the same family name, as well as children and their fathers (until 2022, children could not be given the mother's family name at birth). Hence, sharing a common last name is a highly accurate predictor of a family link.

and linking individuals who share real estate properties. Given the smaller sample, the *Cerved* classification can also verify online the ownership structure for the subset of firms that are incorporated abroad. The overlap of the *Cerved* sample with ours covers 68,367 firms (employing 3.95 million workers); the two classifications coincide in 93% of the observations. (The proportion remains identical if firms are weighted by workforce) This percentage is likely to represent a lower bound to the true accuracy of our classification algorithm, as this sample over-represents large firms, which tend to have more complex and diversified ownership structures, and where our procedure is more likely to deliver incorrect answers. Hence, the severity of our potential misclassification problem appears quite limited.

2.3 Sample Structure and Descriptive Statistics

As set out in the next section, our analysis is based on a two-way AKM fixed-effects decomposition of the logarithm of weekly wages. The resulting AKM firm effects are then regressed on value added per worker. As a result, our estimation sample includes the observations for the largest connected set of firms linked by worker mobility, as defined by Card et al. (2013), for which we have information on value added from CERVED.[8]

Table 1 presents descriptive statistics for the sample of firms included in our analysis, which comprises 900,000 unique firms (with 4.7 million firm-year observations), two thirds of which are family firms. As family firms are smaller on average (14 employees against 32), the distribution of workers between the two groups of firms is roughly even.

While our sample is much larger than those used in previous studies, the data patterns displayed in Table 1 are broadly in line with the literature. Family firms are substantially less productive, as measured by operating value added per worker (45,900 versus 63,300 euros per worker in nominal terms), consistent with Bandiera et al. (2018). However,

[8] Our firm-level data do not include financial and insurance firms, which have different accounting rules. The fraction of person-year observations where the worker's dominant firm is present in the CERVED data is around 60%. Within this sample, practically 99% of person-year observations belong to the connected set of firms. The dominant firm is defined as the employer that paid the worker the most in a given year.

family firms' profitability, as measured by return-on-assets, is slightly higher (4.42% versus 4.37%). Although this gap is fairly small, this finding is qualitatively in line with Anderson and Reeb (2003) and Sraer and Thesmar (2007). Leverage is also similar.

How can family firms match non-family firms' profitability despite their lower productivity? Part of the answer is simply lower wages. Average weekly wages are 64 euros lower in family firms, corresponding to a raw 14-log-point gap. The gap in the entry wage is of comparable magnitude, suggesting that the family firm discount emerges at the very start of a worker's career. But the gap tends to widen over time, as family firms also display slightly lower wage growth.

Interestingly, family firms are also marked by less within-firm wage inequality, whether gauged by the log-difference between the highest and the median wage or by that between the 90^{th} and 10^{th} percentiles. The gaps are 82 and 72 log-points, respectively, as opposed to 106 and 82 log-points for non-family firms. Part of this gap is due to organizational differences between the two types of firm: non-family firms have 2.5 times as many employees in managerial positions (1% versus 0.4%) and 4 times more in middle management positions (4% versus 1%).

Age, experience, and gender, instead, do not differ significantly between the two types of firm. Family firms have fewer workers with full-time or open-ended contracts, but here too the difference is fairly small.

3 Econometric Framework

This section presents our methodology. Section 3.1 introduces the econometric design used to identify differences in wage policies between family and non-family firms. Section 3.2 presents a method for decomposing the wage differences so as to take into account the role of worker- and firm-level heterogeneity in driving the wage gap between the two types of firm. Section 3.3 discusses the role of compensating differentials in driving differences in wages between family and non-family firms.

3.1 Wage Policies Across Family and Non-Family Firms

There are two fundamental challenges in interpreting the effect of family ownership on the wage structure. The first is that differences between the average wage paid by family and non-family firms could be driven by sorting. That is, more skilled workers may be systematically more likely to work in non-family firms. The second is that even after accounting for worker selection, differences in wages may be driven by differences in productivity and by how these map into wages, i.e., differences in rent-sharing elasticities (Bloom and Van Reenen, 2007; Sraer and Thesmar, 2007; Card et al., 2018).

We address both challenges by first studying how a given worker's wage changes following a job move. In particular, we estimate an AKM specification:

$$w_{it} = \alpha_i + \psi_{j(i,t)} + X'_{it}\beta + r_{it}, \tag{1}$$

where w_{it} is the log weekly wage of worker i in period t and the function $j(i,t)$ determines the identity of the worker i's (dominant) employer in period t.[9] The vector X_{it} includes a cubic in age and years of labor market experience, as well as year fixed effects. The worker fixed effect α_i represents the portable component of worker i's wage, i.e., the component of pay that the worker always receives irrespective of the employer's identity. Finally, the firm fixed effect ψ_j is the systematic pay premium (or discount) associated with working for firm j. The average wage effect (person-year weighted) of working for a family firm, net of worker selection, is therefore calculated as follows:

$$\Delta_{\psi,F} \equiv \mathbb{E}[\psi_{j(i,t)}|f(j(i,t)) = F] - \mathbb{E}[\psi_{j(i,t)}|f(j(i,t)) = NF], \tag{2}$$

[9] To handle situations where ownership changes (e.g., from family to non-family ownership), we estimate a separate firm effect depending on whether the firm is family or non-family owned.

where the function $f(\cdot) : (1, \ldots, J) \rightarrow (F, NF)$ determines whether j is a family firm (FF) or a non-family firm (NFF).

Identification: The key identifying assumption of equation (1) is that workers do not select their employer based on the unobserved component r_{it}, i.e., moves between employers occur under an "exogenous mobility" condition. Sorting based on the worker effects α_i, firm effects ψ_j, and observables X_{it} does not violate exogenous mobility. Endogenous mobility exists if workers select their employer according to an idiosyncratic productivity component of the job (i.e., a "match effect") or because of innovations r_{it} driven by employer learning (Gibbons et al., 2005) or changes to outside options (Postel-Vinay and Robin, 2002). However, evidence from a number of countries appears inconsistent with endogenous mobility.[10]

In the Italian context, using the same data used here, Casarico and Lattanzio (2024) check for the presence of endogenous mobility using the event-study tests pioneered by Card et al. (2013). These tests show flat pre-trends in wages in the years leading up to a job move—which is hard to reconcile with models of employer learning. Moreover, the wage gains from moving from a lower-paying to a high-paying employer are roughly symmetric to the wage losses from the inverse move. This is inconsistent with workers sorting to employers according to unobserved matched effects. Di Addario et al. (2023)—again using Italian data—consider an augmented AKM specification that incorporates insights from the sequential auction models of Postel-Vinay and Robin (2002) and estimate firm effects virtually identical to those obtained when fitting an AKM specification. In light of this evidence, we assume that endogenous mobility is not a first-order concern in the Italian data.

The importance of not aggregating moves: Note that in our estimates we do not adopt a "collapsed" version of equation (1), where firm effects are replaced by a dummy

[10] See Card et al. (2013) for Germany; Card et al. (2015) for Portugal; Song et al. (2019) for the US.

for whether the current employer is a family firm or not. As is noted by Card et al. (2024) in a study of industry-wage differentials, moves between family and non-family firms might occur within a selected sample of firms (say, movers tend to leave highly productive FFs for low-productivity NFFs). If so, estimates based on a specification with worker fixed effects and a family firm dummy might yield a biased estimate of the average wage effect of working for a family firm, even if exogenous mobility holds. By estimating equation (1)—and in particular by leveraging all the moves, including those within FFs or within NFFs—our estimate avoids this bias. What is more, as described below, with this approach we can gauge how much of the difference in pay between FFs and NFFs stems from differences in productivity.

3.2 Oaxaca Decomposition

If wages are determined by equation (1), then the average gap in log-wages between family and non-family firms can be decomposed as follows (for simplicity, we abstract from the impact of the vector of covariates X_{it}):

$$\begin{aligned}
\Delta w &\equiv E[w_{it}|f(j(i,t)) = F] - E[w_{it}|f(j(i,t)) = NF] \\
&= \underbrace{E[\alpha_i|f(j(i,t)) = F] - E[\alpha_i|f(j(i,t)) = NF]}_{\Delta_{\alpha,F}\equiv\text{selection component}} \\
&+ \underbrace{E[\psi_{j(i,t)}|f(j(i,t)) = F] - E[\psi_{j(i,t)}|f(j(i,t)) = NF]}_{\Delta_{\psi,F}\equiv\text{firm component}}.
\end{aligned} \tag{3}$$

The first term is the "selection component," which measures how much of the wage gap is due to differences in the unobserved abilities of the workers hired by the two types of firm. The second term is the "firm component," which captures differences in wage policies between family and non-family firms.

The availability of matched employer-employee data together with financial records and ownership permits us to identify the role of productivity and rent-sharing in explaining

$\Delta_{\psi,F}$. Consider a linear projection of the firm effects, ψ_j, on productivity, estimated separately for family and non-family firms:

$$\psi_j = \theta_{f(j)} + \pi_{f(j)} P_j + \nu_j, \tag{4}$$

where P_j measures the latent productivity of firm j and the coefficient $\pi_{f(j)}$ measures the pass-through of productivity to wages, after accounting for worker sorting (Guiso et al., 2005; Card et al., 2018). Importantly, we allow this coefficient to differ between family and non-family firms, in the light of recent evidence that the impact of sales shocks on wages differs between the two types of firm (e.g., Ellul et al., 2017). When taking expression (4) to the data, we proxy P_j with the logarithm of value added per worker of firm j. Finally, the term v_j is a random effect, normalized to have mean zero, that captures heterogeneity in wage policies orthogonal to differences in productivity.

Based on equation (4), the firm component $\Delta_{\psi,F}$ can be decomposed via a Oaxaca-Blinder (Oaxaca, 1973; Blinder, 1973) decomposition:

$$\begin{aligned}
\Delta_{\psi,F} &\equiv E[\psi_{j(i,t)}|f(j(i,t)) = F] - E[\psi_{j(i,t)}|f(j(i,t)) = NF] \\
&= \underbrace{\pi_{NF}\{E[P_j|f(j) = F] - E[P_j|f(j) = NF]\}}_{\text{productivity component}} \\
&+ \underbrace{(\pi_F - \pi_{NF})E[P_j|f(j) = F]}_{\text{bargaining component}} + \underbrace{\theta_F - \theta_{NF}.}_{\text{systematic component}}
\end{aligned} \tag{5}$$

The first term captures how much of the difference in firm effects between FFs and NFFs is driven by differences in productivity. The second component captures how much is driven by differences in the pass-through coefficient.[11] The last term captures systematic wage differences unexplained by differences in worker sorting, productivity, and rent-sharing elasticities.

[11] The productivity component is assessed using the pass-through coefficient of non-family firms. The bargaining component is assessed based on the average productivity of family firms. Results are similar with an alternative decomposition that calculates the sorting component based on the pass-through of family firms and uses the average productivity of non-family firms to weigh the bargaining component.

3.3 Compensating Differentials

Systematic pay differences between firms can occur if firms offer different amenities to workers, so that wage differences reflect compensating differentials for these amenities (Rosen, 1986). Assessing the role of compensating differentials is challenging for two reasons. First, one often lacks comprehensive data on amenities. Second, even when such data are available (e.g., Lavetti and Schmutte, 2016), their role is often assessed via a hedonic approach that typically amounts to a cross-sectional regression of pay on non-pay characteristics. This assumes that labor markets are perfectly competitive and so neglects the possibility that labor market frictions may generate dispersion in utilities across workers employed by different employers. In a frictional labor market, family firms may offer lower pay while also providing *lower* amenities to workers (Mortensen, 2003).[12]

We address these issues using the revealed preferences approach of Sorkin (2018), which consists in applying Google's Pagerank algorithm to voluntary job transitions in order to calculate the common value of working for a particular firm j. If family firms pay systematically lower wages because they offer better amenities, the average value of working for a FF should not be systematically different from that of working for a NFF.

The common value of working for firm j is calculated as follows. Assume that the utility obtained by individual i from being employed at firm j is

$$U_{ij} = v_j + e_{ij}, \tag{6}$$

where e_{ij} is extreme value Type-1 distributed. Sorkin (2018) shows that the utility offered by firm j net of the idiosyncratic preference shock, i.e., v_j, can be identified from the following recursive equation:

$$\exp(v_j) = \sum_{\ell \in \mathcal{B}_j} \omega_{j,\ell} \exp(v_\ell), \qquad \text{for } j = 1, \ldots, J, \tag{7}$$

[12] Henceforth, in explaining the role of compensating differential, we assume that family-firms pay systematically lower wages, consistent with the evidence to be presented in Section 4.

where $\omega_{j,\ell}$ is the number of workers who move voluntarily from employer ℓ to employer j, scaled by the number of workers who leave employer j voluntarily. Intuitively, "good" employers are unlikely to have workers leave and likely to attract workers from other good employers (just as a good web page, according to Google, is one that is linked to other good web pages). To distinguish between voluntary and involuntary job transitions, we leverage the fact that the INPS data specify the actual reason for each separation. A voluntary separation is one where the worker, according to INPS, has resigned from their previous job (as in Di Addario et al., 2023).

To assess the role of compensating differentials, we consider person-year-weighted linear projections of the PageRank values v_j on an indicator of whether firm j is a family firm:

$$v_j = \beta_0 + \beta_1 \mathbf{1}\{f(j) = F\} + Z_j'\gamma + \chi_j, \tag{8}$$

where the vector Z_j includes province and sector fixed effects. If differences in wage policies between family and non-family firms are driven by compensating differentials, then we should expect $\beta_1 = 0$. We also consider a version where Z_j includes the firm-wage effects ψ_j, so as to assess whether family firms offer better or worse amenities compared to non-family firms with similar wages, and therefore different utility to their employees.

4 The Family Firm Wage Discount

In this section we set out the main results. Section 4.1 shows the results from the decomposition of the raw gap in wages between family and non-family firms presented in equation (5). Section 4.2 assesses whether these gaps are driven by differences in the amenities offered by the two types of firm.

4.1 Estimates of the Family Firm Wage Discount

Table 2 reports the results from the decomposition of equation (5), obtained after fitting the two-way AKM specification in equation (1) to the largest connected set of firms in our data (Card et al., 2013). As shown in Table 1, the raw gap in wages is around 14 log points, in line with the results for France given by Sraer and Thesmar (2007) and Bassanini et al. (2013) and the international evidence of Ellul et al. (2017). Netting out the effects of the covariates X_{it} from equation (1), the gap increases to 16 log points.

Half of this covariate-adjusted wage gap is explained by the fact that workers with a higher portable component of wages are systematically more likely to sort into non-family firms. This underscores the importance of controlling for workforce composition in estimating the family firm wage discount and accordingly highlights the advantage of accessing individual worker-level data. The remaining half of the wage gap is explained by systematic differences in wage policies between family and non-family firms.

In models where firms have latitude in setting wages, inter-firm productivity differences map into differences in wages (Card et al., 2018). As family firms tend to be less productive (see Table 1), it is natural to ask how much of the average gap in firm-wage effects is driven by differences in productivity.

Figure 1 displays a binscatter plot where for each centile of value added per worker (calculated for both family and non-family firms) we overlay on the $x-$axis the corresponding average logarithm of value added per worker and on the $y-$axis the average firm effect for family and non-family firms. While the rate at which increases in productivity pass on to wages is broadly linear, and with a slope not too dissimilar between the two types of firm, it appears that non-family firm wage effects are invariably greater than the family firm effects, and by an almost constant amount, except that the gap tends to widen in the upper tail of the productivity distribution.

To isolate the role of gaps in productivity in driving the differences in average firm effects, we fit the linear specification of equation (4) to the relationship depicted in Figure

1 and then apply the decomposition of equation (5): the resulting estimates are shown in Table 2. There is a 17 log-point difference in average productivity between non-family and family firms. The rent-sharing elasticity for family firms—i.e., the pass-through coefficient π_F from equation (4)—is around 0.14, close to the rent-sharing coefficient found by Card et al. (2015) for Portugal and by Lamadon et al. (2019) for the US. For non-family firms, the rent-sharing elasticity is a bit smaller (0.13), and the difference from family firms is statistically significant.

The Oaxaca decomposition from equation (5) shows that the bargaining component reduces the wage gap by around 4 log points, given family firms' higher rent-sharing coefficient. Differences in productivity account for about 2 log points and explain around 25% of the average difference in firm effects. In line with the visual evidence of Figure 1, there is a substantial component unexplained by differences in productivity or bargaining power that factors into the pay difference. In a counterfactual in which we move each FF worker to a NFF with the same productivity and with the same rent-sharing coefficient as family firms, wages would increase on average by around 10 log points according to equation (4).

One important concern is that FF shareholders might get part of the compensation in the form of dividends, which are not captured by our wage measure. This would imply that we are underestimating the true compensations of FF. To check for this, in Appendix Table A1 we repeat the analysis of the FF wage discount excluding workers who are also shareholders. The results are virtually identical to those with the full sample (see Table 2), indicating that dividend policies are not a significant factor.

Heterogeneity by Firm Size, Location and Sector Figure 2 displays the systematic difference in wages between family and non-family firms—$\theta_F - \theta_{NF}$ from equation 5—fitting the Oaxaca decomposition described in Section 3.2 separately for three different firm workforce size buckets: 1–49; 50–249; ≥ 250. The family-firm discount is significantly smaller for small firms and quite substantial among the larger ones. As we argue in Section

5, this is consistent with the thesis that family firms do not offer the highest-paid positions, such as those of managers, or else that they pay these positions much less than comparably sized non-family firms. Given their simpler and less formalized organization (Bloom and Van Reenen, 2010), with fewer managers, small firms display a significantly narrower wage gap between family and non-family firms.

A similar pattern is observed in Figure 3 which plots $\theta_F - \theta_{NF}$ after fitting the Oaxaca decomposition described in Section 3.2 separately in each Italian province. The wage discount is significantly smaller in the South than in the North. Firms in the South are notoriously smaller and less productive than in the North (Boeri et al., 2021). As a result, consistent with Figure 1, the discount is much less pronounced in the South. The presence of highly productive firms—which can potentially offer high-paying jobs—thus appears to be a key factor in the geographical heterogeneity displayed in the Figure.[13]

Lastly, Figure 4 reports the systematic difference in pay between family and non-family firms across production sectors, again obtained by fitting our baseline Oaxaca decomposition separately within each 2-digit industry code. For the vast majority of sectors, we estimate an economically significant family firm wage discount. Only for a handful of sectors does the discount turn positive, suggesting that sectoral heterogeneity as such (unrelated to differences in productivity) plays a modest role at best in determining the wage discount estimated in Table 2.

4.2 The Role of Compensating Differentials

The systematic wage gap between FFs and NFFs might be offset by greater amenities offered by the former. Previous work notes that family firms might offer greater job stability, as family ownership decreases employers' incentive to break the implicit contract,

[13] In Appendix-Figure A1 we further address the possibility that the family firm wage discount may reflect geographical variation, say an "urban wage premium." We regress the firm wage fixed effect on a combination of city of birth, city of residence, and city of workplace location fixed effects. Even the most saturated specification, which includes separate dummies for every possible combination of city of birth, residence, and work, cannot explain more than one log-point of the wage difference between family and non-family firms.

shielding employees from negative shocks (Shleifer and Summers, 1988; Sraer and Thesmar, 2007; Ellul et al., 2017). The existence of a job security premium among family firms may explain why they tend to pay less. Other scholars argue that family firms invest more in the relationship with their employees (Mueller and Philippon, 2011; Kang and Kim, 2020), which may also compensate for lower pay.

To assess the overall role of compensating differentials in driving the family firm wage discount, we adopt the revealed preference approach of Sorkin (2018) described in Section 3.3. That is, we estimate the ranking of employers based on voluntary job moves by Italian workers. In deriving this ranking, we do *not* assume that utility must be equalized across jobs, as in the typical hedonic-style approach. If compensating differentials account for the discount, then we should find that, on average, the utility of working for a FF is similar to that of working for a NFF.

Figure 5 displays the common valuation of firms obtained from the recursive formulation of equation (7) by overlaying the Pagerank of family and non-family firms, for each centile of value added per worker. The utility pattern that emerges from Figure 5 is very similar to that in Figure 1. More productive firms tend to be more desirable, and NFFs tend to be more desirable than FFs, even when their underlying productivity is similar. Based on the voluntary job transitions of Italian workers, it would appear that family firms have a lower Pagerank; that is, they offer *lower* utility from employment than non-family firms.

This is further confirmed by Table 3, which reports estimates from equation (8). Moving a worker from a family to a non-family firm would increase average utility by a margin equal to 48% of the overall standard deviation of the Pagerank index in our data. This very large utility gap is explained in part by the fact that family firms pay lower wages, as measured by the firm effect (Column 2), and tend to be found in sectors and areas where utility from work is lower (Columns 3 and 4). However, as is shown in Column 5, even after controlling flexibly for province-by-sector indicators as well as the firm effect, it still holds that family firms are lower-ranked. It is accordingly hard to reconcile this evidence with the compensating differential thesis that family firms pay lower wages solely because they offer greater amenities.

We perform a few robustness checks on this result. First, we re-estimate equation (8) using a version of the Pagerank index that accounts for the possibility that a portion of the voluntary job moves are driven by FFs differing in offer intensity as well as firm size, along the lines suggested by Sorkin (2018). As shown by Table A4 in the Appendix, when using this augmented Pagerank index our qualitative conclusions remain unchanged.

Second, we complement the "top-down" approach of Sorkin (2018) with one where we estimate the provision of certain amenities in our data and explore whether they differ between the two types of firm. Thanks to the INPS data, we can measure the number of fully paid sick days (n_{it}) a worker takes in a given year. We identify a firm-level component of this measure by estimating equation (1), using $\log n_{it}$ as the outcome, a specification akin to that used by Lachowska et al. (2022) to isolate the importance of firms in unemployment insurance take-ups or by Bana et al. (2020) to isolate the importance of firms in maternity leave take-up. Appendix-Figure A2 shows the distribution of this firm-level effect for centiles of value added per worker. The evidence gives no indication that FFs are systematically associated with more sick leave.

All in all, the evidence provided here is at odds with the thesis that it is thanks to better amenities that family firms can pay lower wages. Ranking employers according to voluntary transitions by workers, we find that family firms offer lower utility, even after controlling for differences in pay. That is, family firms not only pay lower wages but also offer lesser amenities. In other words, they are systematically worse employers, amplifying inequalities not only in pay but also in total utility. This suggests that these firms are able to recruit and retain workers because frictions (such as asymmetric information, search frictions, queueing, horizontal differentiation) prevent workers from leaving them for employers that offer higher utility (Mortensen, 2003). The following sections will present a theoretical rationale, and empirical evidence, for an economic mechanism that can explain why family firms can operate despite offering lower pay and utility.

5 The Glass Ceiling in Family Firms

Having established that large, consistent differences in pay between FFs and NFFs persist even accounting for differences in worker sorting, productivity, rent-sharing, and compensating differentials, we now explore other potential contributing factors.

Section 5.1 presents three sets of evidence that, taken together, suggest that workers in family firms are subject to a "glass ceiling." First, the wage gap between FFs and NFFs is mainly observed in the right tail of the wage distribution: the best-paid employees in NFFs are paid consistently more than in FFs. Second, family firms feature significantly fewer promotions. And third, their promotions are associated with smaller salary increases. As a result, the wage distribution of family firms is more compressed, since they are systematically less likely to promote workers to managerial—and thus higher paying—positions.

But why are owners of family firms less likely to promote? We argue that this is because family owners attach a high value to the private benefits of control (Bandiera et al., 2018), entailing reluctance to share technical and operational know-how with non-relatives. As we argue below, a testable implication of this hypothesis is that the loss of the top manager should be more disruptive in family firms, where decision-making is extremely concentrated. Section 5.2 confirms that this is indeed the case, as CEO deaths are more harmful, in terms of productivity losses, in family firms (especially family-managed firms) than in non-family ones.

5.1 Top Wages and Promotions in Family vs. Non-Family Firms

Figure 6 plots the within-firm log-wage distribution by centile of log-value added per worker. It shows the 5^{th}, 10^{th}, 25^{th}, 75^{th}, 90^{th}, and 95^{th} percentiles of the within-firm distribution of the logarithm of wages (Panels A, B, C, D, E, and F,respectively). The percentiles for FFs and NFFs are fairly close for the 5^{th} and 10^{th} percentiles, but a perceptible gap begins to emerge at the 25^{th} percentile, growing more evident, and economically large, for the 75^{th}, 90^{th}, and 95^{th} percentiles, especially for high values of log-value added per worker, i.e., for the most productive firms. While this evidence is merely descriptive, as it does not account for sorting, it still shows that the gap in average earnings depends mainly on differences in compensation at the high end of the salary distribution.

To further investigate this point, we now look at how FFs and NFFs differ in promotion practices. We consider promotions to the level of middle managers (*quadri*), and from middle to top managers (*dirigenti*). We consider all employees who are with the same firm for two consecutive years.

For each firm and year, Panels A and B of Figure 7 display the fraction of workers promoted, either to middle or to top posts, by percentile of value added per worker. Panel A shows that for about three fourths of the labor productivity distribution, promotions to middle management are relatively rare and fairly similar between FFs and NFFs. For more productive firms, however, they become more frequent. These firms also show a sizable gap between NFFs and FFs: in the top decile of the productivity distribution, the former's propensity to promote is two to three times higher than the latter's. For example, in the highest productivity percentile, the fraction of employees promoted every year is close to 0.6% in non-family firms, against 0.2% in family firms.

The results shown in Panel B, which plots the average promotion rate from middle to top management, are qualitatively similar. Unsurprisingly, the average promotion rates are lower, but the difference is even greater: in the top percentile, the average rates for FFs and NFFs are 0.02% and 0.1%, respectively.

While NFFs promote more to managerial positions, this could be due to higher turnover of employees in these positions, due to dismissals or resignations. In this case, interpreting the higher promotion rate as an indication of better career opportunities would be questionable. Panels C and D of Figure 7 plot the fractions of middle and top managers in the workforce, which mirror those of promotions: at higher levels of productivity, NFFs have a larger share of employees in managerial positions. Hence, FFs appear to provide significantly lower career opportunities.

Returns to Promotion: In principle, family firms might use promotion more sparingly but reward those promoted more generously. From the employees' perspective, the larger raise could compensate for the lower chance of promotion. Hence, in our third step we compare the returns to promotion of workers in FFs and NFFs, using a matching approach. Intuitively, we select workers who have similar characteristics, including type of promotion, but differ in type of employer. Every worker promoted in a family firm is paired with a "twin" in a non-family firm, matching on the log of firm size, the log of value added per worker, two lags of the log of weekly wage, and age, as well as (exactly) promotion type (to middle manager or top manager), industry (using the NACE 2-digit classification), contract type, full-time status, and gender.[14] Imposing a two-year tenure requirement and following workers over a $(-3, +3)$-year window, we are left with 22,002 promotions to middle management and 938 from middle to top manager. We then estimate the following model:

$$Y_{it} = \sum_{k=-3}^{3} \alpha_k D_{it}^k + \sum_{k=-3}^{3} \beta_k D_{it}^k \times FF_i + \gamma_i + \delta_t + \varepsilon_{it}. \tag{9}$$

Here $D_{it}^k \equiv \mathbb{1}(t = t_i^* + k)$, where t_i^* is the event year for worker i and Y_{it} is an outcome of interest. FF is a dummy equal to 1 if the worker has been promoted in a family firm, γ and δ are vectors of worker and year fixed effects, respectively. We are interested in the

[14] We implement a caliper matching algorithm without replacement. For the logarithm of firm size, the logarithm of value added per worker, two lags of the logarithm of weekly wage, and age (the variables for which the match is approximate) the caliper widths are 1, 0.5, 0.1, 0.1, and 5, respectively. All matching variables are measured in the year prior to the promotion.

estimated β_k coefficients, which track the differential evolution of the dependent variable before and after the promotion for a worker who has been promoted within a family firm relative to a worker who instead has been promoted in a non-family firm. As β_{-1} is normalized to 0, our baseline year is the year prior to the event.

The results, plotted in Figure 8 indicate a 1.2% smaller salary raise in family firms three years after the event. The effect on "stayers," i.e., workers who remain with their t_i^* employer and are thus fully exposed to the promoting firm's compensation policies, are very similar. As to earnings, workers promoted in family firms earn €467 less than their matched counterparts three years after the event, but this effect is not statistically significant.

Panel B, on promotions from middle to top management, shows qualitatively similar results, albeit with larger magnitudes. Workers promoted in a family firm earn 4.98% less three years after the promotion than those in non-family firms and the difference increases to 5.96% for stayers. The difference in the earnings trajectory is economically large as well: three years after event, we find a difference of €5,667 relative to the baseline year.[15]

5.2 Evidence from CEO Deaths

A possible inference from the evidence presented above is that controlling families seek to reserve most if not all managerial positions for relatives and thus keep most of the firm's technical and organizational know-how within the family circle. In other words, they are reluctant to delegate key decisions to non-relatives.

This hypothesis can explain why family firms tend to promote fewer workers to managerial positions, and also to reward them less upon promotion. It is also consistent with the family wage discount being especially pronounced for top earners: this likely reflects

[15] To get a better sense of the magnitude of this effect, we also scale realized earnings by $t^* - 1$ earnings. In unreported results, we find that promotions to managerial roles produce significantly lower returns for FF employees, with coefficients equal to -0.041, -0.064, and -0.073 one, two, and three years after the promotion. We again find negative but insignificant differences in the returns to promotion to middle manager.

family firms' reluctance to compete keenly to attract or retain non-family employees in major decision-making roles. An implication of this hypothesis is that losing the top decision maker should have a much more disruptive impact on FFs than NFFs, as the latter presumably have a larger pool of employees able to step in and effectively replace the deceased CEO.

To test the implication of this hypothesis, we use the *Esponenti* file in our data, which lists the top executives and board members of the companies in our sample. For each firm, we identify the top decision maker, namely the chief executive officer (*amministratore delegato*) or, if absent, the president (*presidente*). For brevity, in what follows both are referred to as CEOs. INPS records the year of death of all the individuals who, at any point in time, paid social security contributions. We can match these records with data on CEO deaths, so as to build a panel of firms affected by a CEO decease event.[16]

In the sample of firms subject to a CEO death event, we then distinguish between FFs and NFFs as treatment and control firms, respectively. For every treated firm, we select a control firm in the same industry (NACE 2-digit classification), further matching on the logarithm of size and the logarithm of value added per worker, measured at the end of the year prior to the event.[17] This matching procedure leaves us with a sample of 1,268 events, half of them involving the death of a family CEO. Then, in order to investigate whether the loss of the firm's CEO affects productivity differently in family non-family firms, we estimate the following regression:

$$\log(Productivity)_{jt} = \sum_{k=-3}^{3} \alpha_k D_{jt}^{k} + \sum_{k=-3}^{3} \beta_k D_{jt}^{k} \times FF_j + \gamma_j + \delta_t + \varepsilon_{jt}, \tag{10}$$

where the j and t subscripts identify firms and years and γ and δ are firm and year fixed

[16] See Sauvagnat and Schivardi (2024) and Smith et al. (2019) for general analyses of the effects on firm performance of CEOs' and owners' deaths, respectively. Bennedsen et al. (2020) examine the effects of CEO hospitalization events on firm performance and find that the effects are worse in family firms, consistent with our evidence. Aside from a focus on a different type of event, in our analysis we additionally emphasize the distinction between family-managed and non-family-managed firms.

[17] As before, we implement a caliper matching algorithm without replacement. We use the caliper width 0.2 for both the logarithm of size and the logarithm of value added per worker.

effects, respectively. As before, $D^k \equiv \mathbb{1}(t = t_j^* + k)$, where t_j^* is the CEO death year for firm j. As usual, productivity is measured as the logarithm of value added per worker. The coefficients of interest are the β_ks, with β_{-1} being normalized to zero. Standard errors are clustered at firm level.

Figure 9, Panel A.i., shows that after a CEO death productivity drops more in family than non-family firms. (Appendix Table A3 reports the coefficients and standard errors in tabular form) While the effect is initially limited, two and three years after the event FFs experience productivity losses 7.6 and 5.7 percent greater than their non-family counterparts, although coefficient β_3 is not very precisely estimated.

In our ownership-based classification, we have so far abstracted from the fact that some FFs are run by managers not belonging to the controlling family. In this setting, however, this distinction is potentially relevant. In non-family-managed firms the decision to delegate power has already been taken by the controlling family. Nor is there any obvious reason why non-family CEOs should expect to pass managerial roles on to relatives or heirs, which reduces their incentive to guard technical knowledge. Indeed, Mullins and Schoar (2016) provide survey evidence that professional CEOs in FFs behave very similarly to the managers of NFFs in propensity to delegate. Hence, one should expect these negative effects to be particularly pronounced in family-managed firms. This is confirmed by the evidence in Panel A.ii, which shows larger and more precisely estimated negative effects of CEO deaths than Panel A.i. Interestingly, Panel A.iii reveals no detectable differential effect of CEO deaths in non-family-managed firms.

Trauma or Losses in Know-How? The drop in value added per worker documented in Figure 9 might reflect the "emotional" impact of loss of the owner rather than losses in know-how. Panel B presents a "placebo test" based on the same design as equation (10) but where the relevant event is the death of a director (*amministratore*), rather than the CEO. The death of a family member should be an emotionally disruptive event independent of their role; however, we would expect a substantial loss in technical and operational know-how only in the case of loss of a family member with a top executive role. This is precisely what Panel B shows: there is no evidence of a differential productivity response for director deaths, whether we focus on all firms (Panel B.i), on family-managed firms (Panel B.ii), or on non-family-managed firms (Panel B.iii).

6 A Model of Career Paths in Family Firms

The evidence presented so far can be broadly summarized as follows. First, careers are slower in FFs, given the significantly smaller frequency of promotions. Second, FFs pay lower average compensation at each stage of the career, even after accounting for sorting of workers into firms with different characteristics. Third, workers appear to prefer NFFs, suggesting that FFs do not offer greater amenities to make up for their other disadvantages. In what follows, we provide a simple theoretical framework for these empirical regularities.

In the model, owners set promotion policies, balancing the efficiency gains from promoting skilled employees against the personal advantages of retaining control over the firm. Specifically, the owners' relative preference for retaining power is assumed to be greater in FFs than in NFFs, consistent with the evidence in Lippi and Schivardi (2014). Consequently, they are less likely to promote their employees. Workers exert unobservable effort to accumulate firm-specific human capital. Only those who are successful in this are considered for promotion to managerial positions. The poorer career prospects offered by FFs yield weaker incentives to invest in firm-specific human capital, so that employees accumulate less (non-transferable) skills and thus earn commensurately lower wages than in a counterfactual where they are instead hired by a NFF. Given the better earnings prospects of NFFs, in equilibrium they receive more job applications in proportion to vacancies, increasing the probability of applicants being rejected and remaining unemployed. This higher probability of unemployment offsets the greater expected utility of working for a non-family firm: in equilibrium, applying for a job to FFs and NFFs delivers the same expected utility.

6.1 Setup of the Model

The economy is populated by identical risk-neutral workers and heterogeneous firms. There are N_{NF} non-family firms and N_F family firms. Each firm hires a mass 1 of workers, so that total employment is $N = N_{NF} + N_F$. The labor force exceeds total employment, and the utility of unemployed workers is normalized to zero. Firm j's employees produce y_j if they acquire skills on the job; this happens with probability e_w, which coincides with their effort, which is costly. Effort is unobservable and its cost is $c_w e_w^2/2$, but once they have acquired skills can be observed. Each firm draws its productivity parameter y_j from a probability distribution $F(y)$ with strictly positive support, assumed to be the same for family and non-family firms. Unskilled workers produce a low output, which for simplicity we normalize to 0.

Each firm j chooses to promote a fraction ϕ_j of skilled workers to managerial positions. The employees promoted produce additional output δy_j only if they become skilled managers; this occurs with probability e_m, which coincides with their additional effort at a cost $c_m e_m^2/2$.[18] Firm j chooses its promotion rate ϕ_j by trading off its expected profits against the loss of private benefits $\beta_j \phi_j^2/2$ associated with promotion rate ϕ_j, where the parameter β_j captures the firm's "taste for control over promotions" and is assumed to be greater in FFs than NFFs ($\beta_F > \beta_{NF}$), and for simplicity constant within each of the two groups. The quadratic loss function captures the idea that incremental control losses generate increasing disutility for controlling shareholders.

The timeline of the model, illustrated in Figure 10, comprises four stages:

- **Matching and hiring**. At $t = 0$, each worker decides whether to apply to a randomly chosen FF or NFF, taking into account the expected lifetime income each

[18] To keep our notation simple, we do not index by j the effort chosen by the employees and the managers of firm j.

offers. Firms hire by picking workers randomly from their applicant pool, with unsuccessful applicants remaining unemployed. Upon hiring, each firm j offers the worker a (possibly state-contingent) wage contract.

- **On-the-job learning.** At $t = 1$, the employees of firm j choose effort e_w, anticipating the promotion probability ϕ_j of their employer j and the bonus contract that it will offer at stage $t = 2$.
- **Promotion.** At $t = 2$, each firm j promotes a fraction ϕ_j of skilled workers to managerial positions and offers them a (possibly state-contingent) bonus contract. Managers can generate additional output δy_j only if they acquire managerial skills *on top of* worker skills. They do so by exerting unobservable effort e_m.
- **Production and compensation.** At $t = 3$, skilled workers and skilled managers produce output y_j and $(1 + \delta)y_j$, respectively, and firms pay the agreed wages and bonuses.

At $t = 0$ firms are assumed to be unable to commit to a specific promotion probability or a specified bonus contract at $t = 2$.

6.2 Main Predictions

The model is solved by backward induction: at each stage, workers and firms act optimally, following their time-consistent policies. In what follows, we present the main results in equilibrium. Appendix A.2 discusses the choice problems of workers and firms at each stage and gives the proof of the results.

Prediction 1. *Family firms pay lower average wages than non-family firms, and the gap increases with productivity. The average gap results both from differences in the compensation of high-pay workers and in the fraction of these workers in the wage distribution of each firm.*

Under the optimal contract, firm j's employees are paid a contingent wage. If they manage to acquire productive skills, they earn the efficiency wage:

$$w_j^* = \frac{1}{2}\left[y_j + \frac{\phi_j^*}{2c_m}\left(\frac{\delta y_j}{2}\right)^2\right] = \frac{1}{2}\left[y_j + \frac{1}{2\beta_j c_m^2}\left(\frac{\delta y_j}{2}\right)^4\right]. \tag{11}$$

Otherwise, they receive their reservation wage, which we normalize to zero with no loss of generality. The efficiency wage (11) is not only increasing in the firm's productivity parameter (y_j) but also in the probability of promotion (ϕ^*) and in workers' incremental future productivity as managers (δy_j), since all of these increase the surplus that skilled workers are able to generate. Substituting for the optimal promotion probability, we see that NFFs, which feature a lower β_j, pay higher salaries to their best-paid employees: in setting them, they take into account that their skill acquisition generates a larger expected surplus, thanks to their greater probability of promotion (as we show next).

Moreover, NFF employees also have a greater probability of acquiring productive skills than FF employees, and thus of earning the efficiency wage (11). This is because in equilibrium this probability coincides with a worker's optimal effort level, which is increasing in the same parameters that drive the efficiency wage and therefore also in the probability of promotion:

$$e_w^*(w_j^*) = \frac{1}{2c_w}\left[y_j + \frac{3\phi_j^*}{2c_m}\left(\frac{\delta y_j}{2}\right)^2\right] = \frac{1}{2c_w}\left[y_j + \frac{3}{2\beta_j c_m^2}\left(\frac{\delta y_j}{2}\right)^4\right]. \tag{12}$$

Hence, the difference in the average wage between NFF and FF employees stems both from the difference in the efficiency wages of their respective best-paid workers and also from the the difference in their chances of achieving high-pay positions within the firm's wage distribution. Moreover, on both accounts the wage gap widens for the more productive firms, as both the high-pay wage (11) and the probability of achieving it (12) are increasing in the firm productivity parameter y_j.

These predictions are consistent with the results in Table 2 and Figure 1, which show that, even after controlling for workers' sorting, family firms pay lower wages, and more so for high productivity levels. They are also consistent with Figure 6, which documents that in the low percentiles of the within-firm wage distribution there is no wage gap. In the model, these are the workers who do not develop skills, and therefore earn only the reservation wage in both FF and NFFs.

Prediction 2. *Family firms promote fewer workers than non-family firms, and the gap in promotion rates increases with firm productivity.*

The optimal promotion probability chosen by firm j is

$$\phi_j^* = \min\left[\frac{1}{\beta_j c_m}\left(\frac{\delta y_j}{2}\right)^2, 1\right], \tag{13}$$

so that FFs, which show greater taste for private benefits β_j, promote fewer employees to managerial positions. Given that the cross-derivative of ϕ_j^* with respect to β_j and y_j is negative, the gap in promotion rates between NFFs and FFs is greater for more productive firms. This is in line with the evidence of Figure 7, which characterizes promotion policies for different classes of value added per worker, and shows that family firms have lower promotion rates.[19]

Prediction 3. *Promotions in family firms produce a smaller increase in compensation.*

This result holds under a mild regularity assumption on the moments of the distribution of y_j (stated formally in the Appendix), and is satisfied by most common distributions with positive support, namely, uniform, beta, gamma and log-normal.[20] First, notice that

[19]Note however that the figure plots the fraction of promotions against the observed value added per worker rather than the unobserved y_j. But, since (i) from Figure 1 we know that, for given value added per worker, family firms promote fewer workers and (ii) promotions are increasing in y_j, for a family firm to have the same abscissa (same value added per worker) as a non-family firm, it must have a higher value of y_j (that is, a higher value y_j must compensate the negative effect on promotions due to the higher β_j). Hence, if we were to plot promotion rates against y_j, the data points corresponding to family firms would be shifted to the right relative to that shown in Figure 7, making the difference between promotion rates in family and non-family firms even larger.

[20]We thank Pietro Coretto for identifying and formally stating this regularity assumption.

the expected bonus b for a worker who is promoted, conditional on y_j, is equal for FFs and NFFs, and given by:

$$\mathbb{E}\left[b|y_j\right] = x^* e_m^* = \frac{\delta y_j}{2}\frac{\delta y_j}{2c_m} = \frac{1}{c_m}\left(\frac{\delta y_j}{2}\right)^2 \tag{14}$$

However, *unconditionally*, the expected bonus paid by NFFs is higher. This result, formally proven in Appendix A.2, has a simple intuition. Recall from Prediction 1 that NFFs are more likely than family firms to promote, and that the difference increases with productivity y_j. Hence, workers promoted in NFFs are going to be disproportionately concentrated in high-productivity firms. As the expected bonus b is increasing in firms' productivity, averaging b over the productivity distribution will result in a higher expected value for NFFs. This prediction is consistent with the event study evidence in Figure 8, which shows that the return to promotion (i.e., the bonus paid to promoted workers) is significantly lower in family firms.

Prediction 4. *Family firm employees have lower expected utility than non-family firm employees, but job applicants to the former have a greater probability of being hired.*

The equilibrium expected utility from working in firm j, computed as of $t = 1$, is increasing in the probability of skilled workers' promotions, ϕ_j^*:

$$\mathbb{E}(U_{1j}^*) = \frac{1}{8c_w}\left[y_j + \frac{3}{2}\frac{\phi_j^*}{c_m}\left(\frac{\delta y_j}{2}\right)^2\right]^2 = \frac{1}{8c_w}\left[y_j + \frac{3}{2\beta_j c_m^2}\left(\frac{\delta y_j}{2}\right)^4\right]^2. \tag{15}$$

Recalling that y_j is identically distributed across FFs and NFFs, the average $\mathbb{E}(U_{1j}^*)$, computed over the distribution of y_j, is higher for NFFs: by offering more rewarding careers, they foster more human capital accumulation and offer higher lifetime expected utility. Formally, $\mathbb{E}(U_{NF}^*) > \mathbb{E}(U_F^*)$, where $\mathbb{E}(U_{NF}^*) \equiv \mathbb{E}_{j\in NF}[\mathbb{E}(U_{1j}^*)]$ and $\mathbb{E}(U_F^*) \equiv \mathbb{E}_{j\in F}[\mathbb{E}(U_{1j}^*)]$,

NF and F respectively being the sets of non-family and family firms.[21] This prediction is consistent with Figure 5, based on the methodology proposed by Sorkin (2018), which shows that workers attach less expected utility to being employed in a FF.

The model helps to see that this fact can be reconciled with labor market equilibrium: even though FF employees enjoy lower average overall utility, at the job application stage, the expected utility of the two groups can be equal, as applicants to NFFs run a greater risk of unemployment, given that these jobs are more sought-after and thus have more applicants. The reason why NFFs do not cut their wages is that they are optimally set to elicit effort from employees.

Finally, note that our comparative statics are in relation to the intrinsic productivity parameter y_i, which is not observed in the data. Conversely, our productivity proxy is value added per worker, which, as our model shows, is itself a function of family firm status. In a simple simulation exercise, discussed in Appendix A.1, we show that our model yields relationships between wages, promotion rates, productivity, and family firm status that are qualitatively similar to those observed in the data.

7 Conclusion

This paper inquires into the differences in compensation and promotion policies between family and non-family firms, and what they imply for firms' performance. Based on a much larger sample of Italian firms than previous literature, we confirm the finding of a significant salary gap, family firms paying substantially lower wages. Our comprehensive matched employer-employee data permit exploration of the possible reasons. We find that about half of the gap, corresponding to 8 log-points, is due to workers' time-invariant characteristics, but the rest cannot be explained by sorting. Nor does accounting for

[21] Since our evidence indicates that FFs feature lower productivity than NFFs (see Table 1), it is worth pointing out that this result generalizes to the case where NFFs are more productive, in the sense that their productivity first-order stochastically dominates that of FFs, rather than being identical as assumed here. This is because expression (15) is increasing in y_j, so that the inequality $\mathbb{E}(U^*_{NF}) > \mathbb{E}(U^*_F)$ would hold *a fortiori* if NFFs were more productive than FFs.

differences in rent-sharing, productivity, industry, and location entirely eliminate the gap.

Moreover, the difference is greater in high-productivity firms and is driven mainly by differences in compensation for relatively high-paid employees, i.e., those in the top percentiles in each firm's earnings distribution. In addition, controlling for observable worker and firm characteristics, family firms tend to have fewer top and middle management positions, to promote less, and to offer lower returns to promotion. These findings suggest the presence of a "glass ceiling" in family firms: the controlling family may be reluctant to share organizational and technical know-how with non-relatives. To test this hypothesis, we conduct event studies around CEO deaths, showing that productivity is undercut more severely following the death of the CEO in family firms, and especially in family-managed ones, after matching for a host of observable characteristics. This corroborates our glass ceiling hypothesis: if organizational and technical know-how is not shared with other employees in family firms, then, with the loss of the top decision maker, part of it is irremediably lost.

We also investigate an alternative "compensating differential" hypothesis, namely the possibility that the lower wage paid by family firms may be offset by superior non-monetary amenities, such as a less confrontational work environment (Mueller and Philippon, 2011) or greater job security (Ellul et al., 2017). However, measuring the systematic utility associated with family and non-family firms following Sorkin (2018)'s revealed-preference approach, we still find that family firms appear to be inferior employers.

We complement this analysis with a simple theoretical model. In our setting, the only difference between family and non-family firms is that that the former's shareholders suffer a "disutility" from promoting workers, consistent with the presence of greater private benefits of control. Simple though it is, our model successfully rationalizes most of our main empirical findings: that family firms pay lower wages, have fewer managers, give smaller salary raises upon promotion, and provide lower overall expected utility for their workers. The model also suggests that despite their lower wages and slower careers, family firms may still be able to compete with family firms for job applicants, thanks to their greater likelihood of being hired.

References

Abowd, John M., Francis Kramarz, and David N. Margolis, 1999, "High Wage Workers and High Wage Firms", *Econometrica* 67, 251–333.

Acemoglu, Daron, Alex He, and Daniel le Maire, 2022, "Eclipse of Rent-Sharing: The Effects of Managers' Business Education on Wages and the Labor Share in the US and Denmark", Working Paper 29874, National Bureau of Economic Research.

Aminadav, Gur, and Elias Papaioannou, 2020, "Corporate Control around the World", *The Journal of Finance* 75, 1191–1246.

Anderson, Ronald C., and David M. Reeb, 2003, "Founding-Family Ownership and Firm Performance: Evidence from the S&P 500", *Journal of Finance* 58, 1301–1328.

Bach, Laurent, and Nicolas Serrano-Velarde, 2015, "CEO identity and labor contracts: Evidence from CEO transitions", *Journal of Corporate Finance* 33, 227–242.

Bana, Sarah H., Kelly Bedard, and Maya Rossin-Slater, 2020, "The Impacts of Paid Family Leave Benefits: Regression Kink Evidence from California Administrative Data", *Journal of Policy Analysis and Management* 39, 888–929.

Bandiera, Oriana, Luigi Guiso, Andrea Prat, and Raffaella Sadun, 2015, "Matching Firms, Managers, and Incentives", *Journal of Labor Economics* 33, 623–681.

Bandiera, Oriana, Renata Lemos, Andrea Prat, and Raffaella Sadun, 2018, "Managing the Family Firm: Evidence from CEOs at Work", *The Review of Financial Studies* 31, 1605–1653.

Bassanini, Andrea, Thomas Breda, Eve Caroli, and Antoine Rebérioux, 2013, "Working in Family Firms: Paid Less but More Secure? Evidence from French Matched Employer-Employee Data", *ILR Review* 66, 433–466.

Bender, Stefan, Nicholas Bloom, David Card, John Van Reenen, and Stefanie Wolter, 2018, "Management Practices, Workforce Selection, and Productivity", *Journal of Labor Economics* 36, S371–S409.

Bennedsen, Morten, Kasper Meisner Nielsen, Francisco Perez-Gonzalez, and Daniel Wolfenzon, 2007, "Inside the Family Firm: The Role of Families in Succession Decisions and Performance", *The Quarterly Journal of Economics* 122, 647–691.

Bennedsen, Morten, Francisco Pérez González, and Daniel Wolfenzon, 2020, "Do CEOs Matter? Evidence from Hospitalization Events", *The Journal of Finance* 75, 1877–1911.

Blinder, Alan S., 1973, "Wage Discrimination: Reduced Form and Structural Estimates", *The Journal of Human Resources* 8, 436–455.

Bloom, Nicholas, Christos Genakos, Raffaella Sadun, and John Van Reenen, 2012, "Management Practices Across Firms and Countries", *Academy of Management Perspectives* 26, 12–33.

Bloom, Nicholas, and John Van Reenen, 2007, "Measuring and Explaining Management Practices Across Firms and Countries", *The Quarterly Journal of Economics* 122, 1351–1408.

Bloom, Nicholas, and John Van Reenen, 2010, "Why Do Management Practices Differ across Firms and Countries?", *Journal of Economic Perspectives* 24, 203–24.

Boeri, Tito, Andrea Ichino, Enrico Moretti, and Johanna Posch, 2021, "Wage Equalization and Regional Misallocation: Evidence from Italian and German Provinces", *Journal of the European Economic Association* 19, 3249–3292.

Burkart, Mike, Fausto Panunzi, and Andrei Shleifer, 2003, "Family Firms", *The Journal of Finance* 58, 2167–2201.

Card, David, 2022, "Who Set *Your* Wage?", *American Economic Review* 112, 1075–90.

Card, David, Ana Rute Cardoso, Joerg Heining, and Patrick Kline, 2018, "Firms and Labor Market Inequality: Evidence and Some Theory", *Journal of Labor Economics* 36, S13–S70.

Card, David, Ana Rute Cardoso, and Patrick Kline, 2015, "Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women", *Quarterly Journal of Economics* 131, 633–686.

Card, David, Jörg Heining, and Patrick Kline, 2013, "Workplace Heterogeneity and the Rise of West German Wage Inequality", *Quarterly Journal of Economics* 128, 967–1015.

Card, David, Jesse Rothstein, and Moises Yi, 2024, "Industry Wage Differentials: A Firm-Based Approach", *Journal of Labor Economics* 42, S11–S59.

Casarico, Alessandra, and Salvatore Lattanzio, 2024, "What Firms Do: Gender Inequality in Linked Employer-Employee Data", *Journal of Labor Economics* 42, 325–355.

Claessens, Stijn, Simeon Djankov, and Larry H. P. Lang, 2000, "The separation of ownership and control in East Asian Corporations", *Journal of Financial Economics* 58, 81–112.

Cullen, Zoe B, Shengwu Li, and Ricardo Perez-Truglia, 2022, "What's my employee worth? the effects of salary benchmarking", Working Paper 30570, National Bureau of Economic Research.

Di Addario, Sabrina, Patrick Kline, Raffaele Saggio, and Mikkel Sølvsten, 2023, "It ain't where you're from, it's where you're at: Hiring origins, firm heterogeneity, and wages", *Journal of Econometrics* 233, 340–374.

Dube, Arindrajit, Alan Manning, and Suresh Naidu, 2018, "Monopsony and Employer Mis-optimization Explain Why Wages Bunch at Round Numbers", Working Paper 24991, National Bureau of Economic Research.

Edwards, Jeremy, Alfons J. Weichenrieder, and Alfons Weichenrieder, 2004, "How Weak is the Weakest-Link Principle? On the Measurement of Firm Owners' Control Rights", CESifo Working Paper Series 1255.

Ellul, Andrew, Marco Pagano, and Fausto Panunzi, 2010, "Inheritance Law and Investment in Family Firms", *American Economic Review* 100, 2414–50.

Ellul, Andrew, Marco Pagano, and Fabiano Schivardi, 2017, "Employment and Wage Insurance within Firms: Worldwide Evidence", *The Review of Financial Studies* 31, 1298–1340.

Faccio, Mara, and Larry H.P Lang, 2002, "The ultimate ownership of Western European corporations", *Journal of Financial Economics* 65, 365–395.

Gibbons, Robert, Lawrence F Katz, Thomas Lemieux, and Daniel Parent, 2005, "Comparative Advantage, Learning, and Sectoral Wage Determination", *Journal of Labor Economics* 23, 681–724.

Guiso, Luigi, Luigi Pistaferri, and Fabiano Schivardi, 2005, "Insurance within the Firm", *Journal of Political Economy* 113, 1054–1087.

Hazell, Jonathon, Christina Patterson, Heather Sarsons, and Bledi Taska, 2022, "National Wage Setting", Working Paper 30623, National Bureau of Economic Research.

Hjort, Jonas, Xuan Li, and Heather Sarsons, 2020, "Across-Country Wage Compression in Multinationals", Working Paper 26788, National Bureau of Economic Research.

Kang, Jun-Koo, and Jungmin Kim, 2020, "Do Family Firms Invest More than Nonfamily Firms in Employee-Friendly Policies?", *Management Science* 66, 1300–1324.

La Porta, Rafael, Florencio Lopez-De-Silanes, and Andrei Shleifer, 1999, "Corporate Ownership Around the World", *The Journal of Finance* 54, 471–517.

Lachowska, Marta, Isaac Sorkin, and Stephen A Woodbury, 2022, "Firms and Unemployment Insurance Take-up", Working Paper 30266, National Bureau of Economic Research.

Lamadon, Thibaut, Magne Mogstad, and Bradley Setzler, 2019, “Imperfect Competition, Compensating Differentials and Rent Sharing in the US Labor Market, institution = National Bureau of Economic Research”, Working Paper 25954.

Lavetti, Kurt, and Ian M Schmutte, 2016, “Estimating Compensating Wage Differentials with Endogenous Job Mobility” .

Lemos, Renata, and Daniela Scur, 2019, “The Ties that Bind: Implicit Contracts and Management Practices in Family-Run Firms”, *Centre for Economic Policy Research, Discussion Paper DP13794* .

Lippi, Francesco, and Fabiano Schivardi, 2014, “Corporate control and executive selection”, *Quantitative Economics* 5, 417–456.

Mortensen, Dale, 2003, *Wage Dispersion: Why Are Similar Workers Paid Differently?* (MIT press).

Mueller, Holger M., and Thomas Philippon, 2011, “Family Firms and Labor Relations”, *American Economic Journal: Macroeconomics* 3, 218–45.

Mullins, William, and Antoinette Schoar, 2016, “How do CEOs see their roles? Management philosophies and styles in family and non-family firms”, *Journal of Financial Economics* 119, 24–43.

Oaxaca, Ronald, 1973, “Male-Female Wage Differentials in Urban Labor Markets”, *International Economic Review* 14, 693–709.

Page, Lawrence, Sergey Brin, Rajeev Motwani, and Terry Winograd, 1998, “The PageRank Citation Ranking: Bringing Order to the Web”, Working paper, Stanford University.

Pérez-González, Francisco, 2006, “Inherited Control and Firm Performance”, *American Economic Review* 96, 1559–1588.

Porta, Rafael La, Florencio Lopez de Silanes, Andrei Shleifer, and Robert W. Vishny, 1998, “Law and Finance”, *Journal of Political Economy* 106, 1113–1155.

Postel-Vinay, Fabien, and Jean-Marc Robin, 2002, “Equilibrium Wage Dispersion with Worker and Employer Heterogeneity”, *Econometrica* 70, 2295–2350.

Robinson, Joan, 1933, *The Economics of Imperfect Competition* (Macmillan Company, New York).

Rosen, Sherwin, 1974, “Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition”, *Journal of Political Economy* 82, 34–55.

Rosen, Sherwin, 1986, "The theory of equalizing differences", in Orley C. Ashenfelter, and Richard Layard, eds., *Handbook of Labor Economics*, volume 1 of *Handbook of Labor Economics*, chapter 12, 641–692 (Elsevier).

Sauvagnat, Julien, and Fabiano Schivardi, 2024, "Are Executives in Short Supply? Evidence from Death Events", *Review of Economic Studies* 91, 519–559.

Shleifer, Andrei, and Lawrence H. Summers, 1988, "Breach of Trust in Hostile Takeovers", in Alan J. Auerbach, ed., *Corporate Takeovers: Causes and Consequences*, 33–56 (University of Chicago Press, Chicago).

Smith, Matthew, Danny Yagan, Owen Zidar, and Eric Zwick, 2019, "Capitalists in the Twenty-First Century", *The Quarterly Journal of Economics* 134, 1675–1745.

Song, Jae, David J. Price, Fatih Guvenen, Nicholas Bloom, and Till von Wachter, 2019, "Firming Up Inequality", *Quarterly Journal of Economics* 134, 1–50.

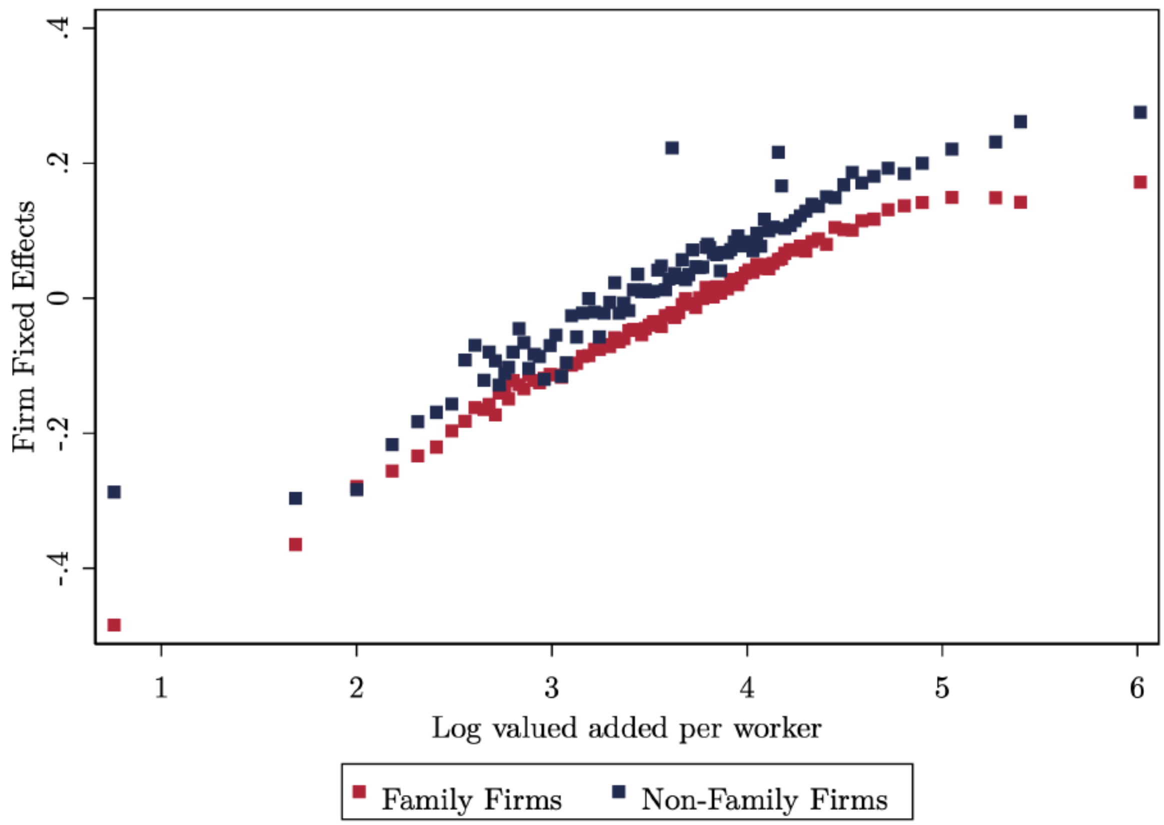
Sorkin, Isaac, 2018, "Ranking Firms Using Revealed Preference", *The Quarterly Journal of Economics* 133, 1331–1393.

Sraer, David, and David Thesmar, 2007, "Performance and Behavior of Family Firms: Evidence from the French Stock Market", *Journal of the European Economic Association* 5, 709–751.

Syverson, Chad, 2011, "What determines productivity?", *Journal of Economic literature* 49, 326–365.

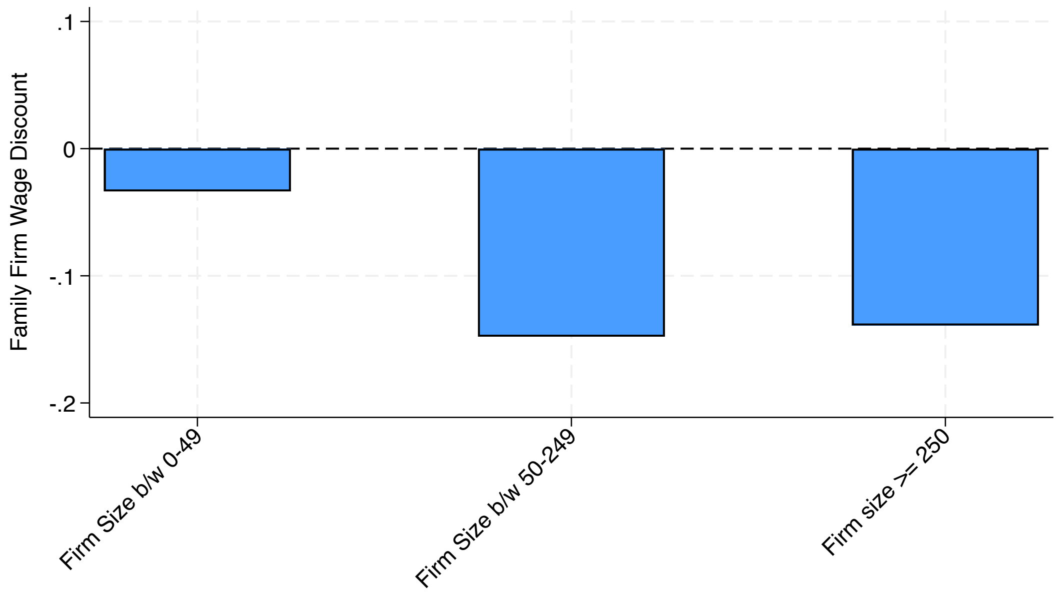
Volpin, Paolo F., 2002, "Governance with poor investor protection: Evidence from top executive turnover in Italy", *Journal of Financial Economics* 64, 61–90.

Figure 1: Firm Wage Premiums of Family and Non-Family Firms



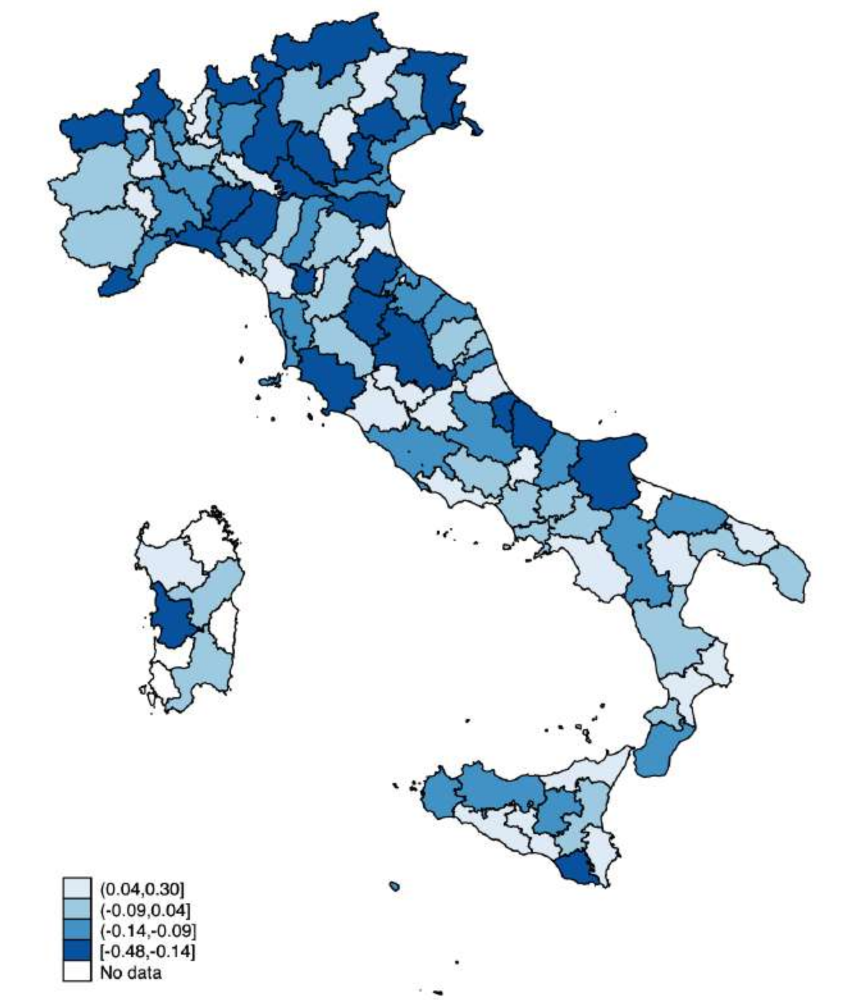
Notes: Each point represents the average of the estimated firm-effect from the AKM model of equation (1) for family and non-family firms belonging to a specific percentile of the overall distribution of log value added per worker. A firm's log value added per worker is calculated by averaging this measure across years. Family firms are defined as those in which an individual or a group of family members own more than 50% of the shares.

Figure 2: Systematic Differences in Pay of Family Firms by Size



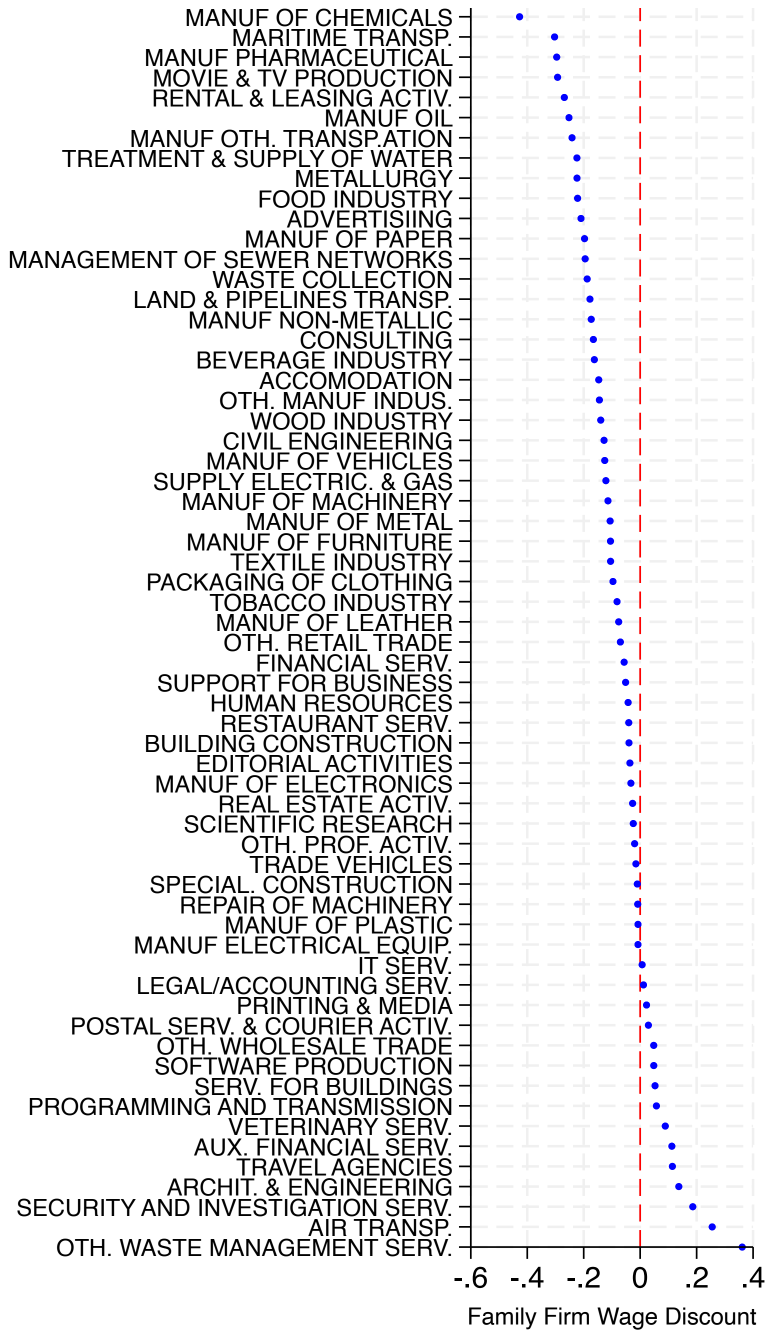
Notes: Each bar displays $\theta_F - \theta_{NF}$ from equation (5), i.e., the difference in average firm-effects between family and non-family firms after netting out differences in productivity and rent-sharing elasticities between these two type of firm, for different firm sizes. Each bar is computed after fitting equation (4) separately for the three firm-size buckets displayed on the $x-$axis.

Figure 3: Systematic Differences in Pay of Family Firms by Province



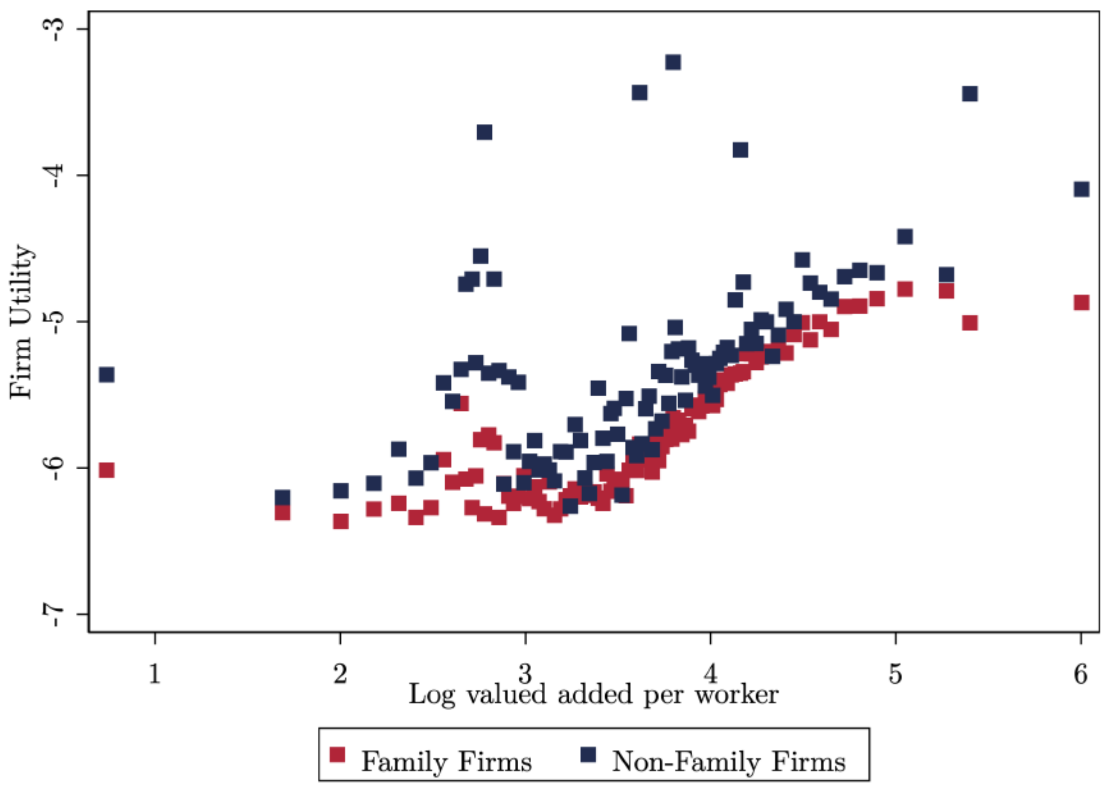
Notes: This figure reports $\theta_F - \theta_{NF}$ from equation (5), i.e., the difference in average firm effects between family and non-family firms after netting out differences in productivity and rent-sharing elasticities between the two types of firm, across Italian provinces. For each province, we report the resulting $\theta_F - \theta_{NF}$ from equation (5) after fitting equation (4) within the province.

Figure 4: Systematic Differences in Pay of Family Firms by Sector



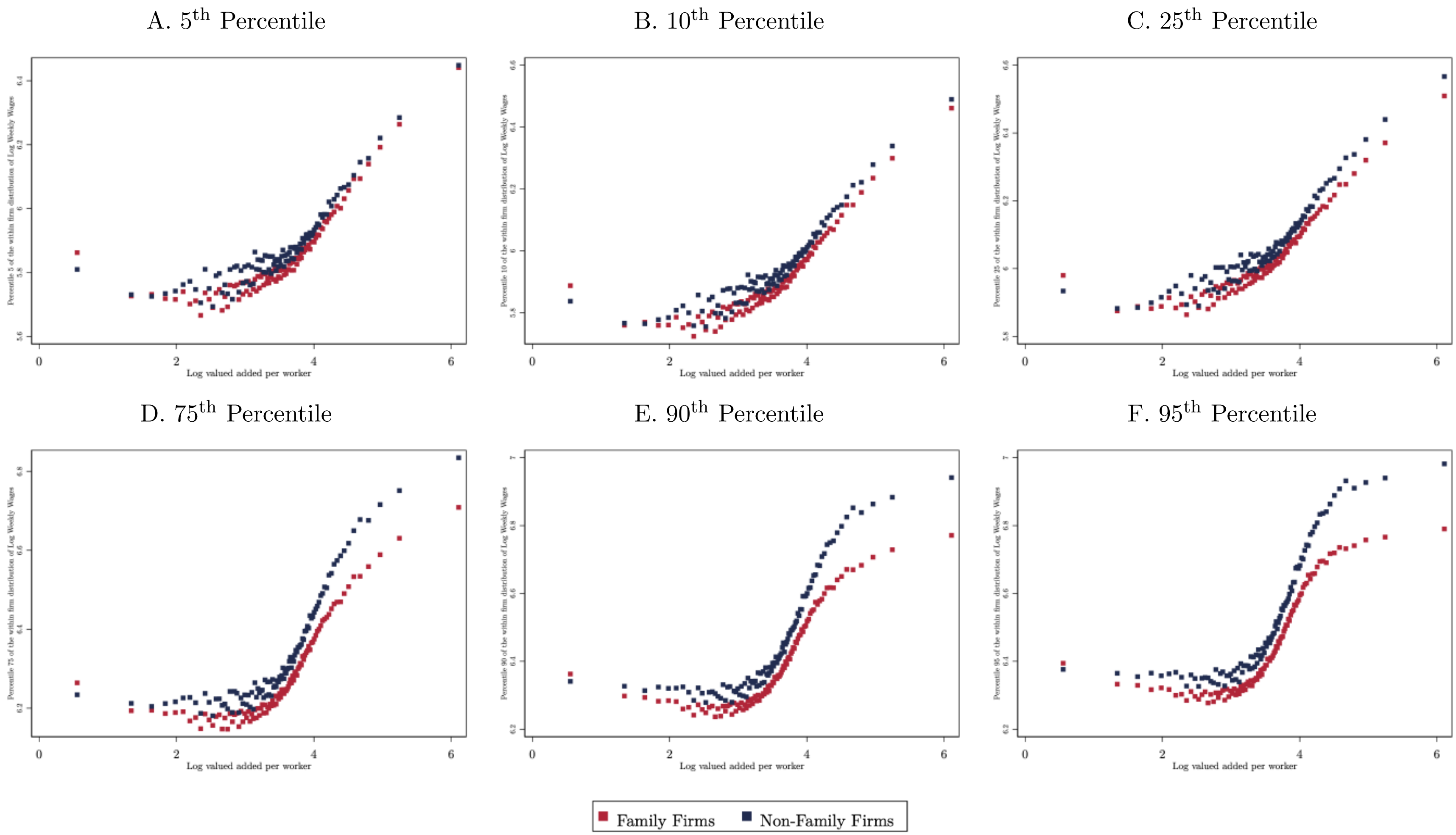
Notes: This figure reports $\theta_F - \theta_{NF}$ from equation (5), i.e., the difference in average firm effects between family and non-family firms after netting out productivity and rent-sharing elasticity differences between the two types of firm, across NACE 2-Digit sector codes. For each province, we report the resulting $\theta_F - \theta_{NF}$ from equation (5) after fitting equation (4) within the sector.

Figure 5: PageRank index of Family and Non-Family Firms



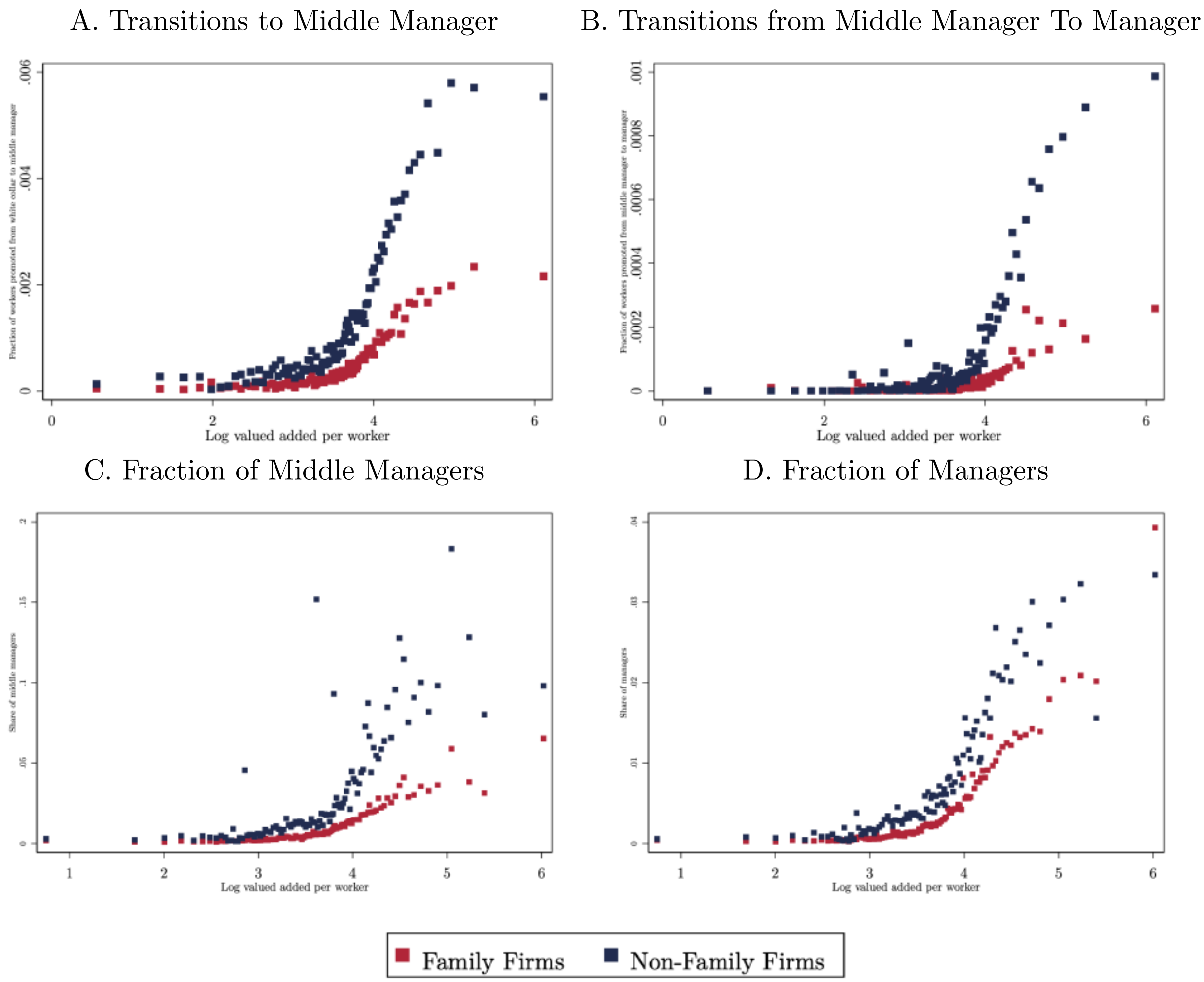
Notes: Each point represents the average of the estimated Pagerank (Page et al., 1998; Sorkin, 2018) for family and non-family firms belonging to a specific percentile of the overall log value added per worker distribution. Log value added per worker of a given firm is calculated by averaging this measure across years. Family firms are defined as those in which an individual or a group of family members own more than 50% of the shares. The Pagerank is calculated by recursive formulation of equation (7), where we leverage the fact that the INPS data separate voluntary from involuntary job moves.

Figure 6: Within-Firm Distribution of Wages across Family and Non-Family Firms



Notes: Each panel displays the corresponding average within-firm percentile of the logarithm of weekly wages for family and non-family firms in a given centile of the distribution of log value added per worker.

Figure 7: Promotion Rates and Fraction of Workers in Managerial Positions in Family and Non-Family Firms

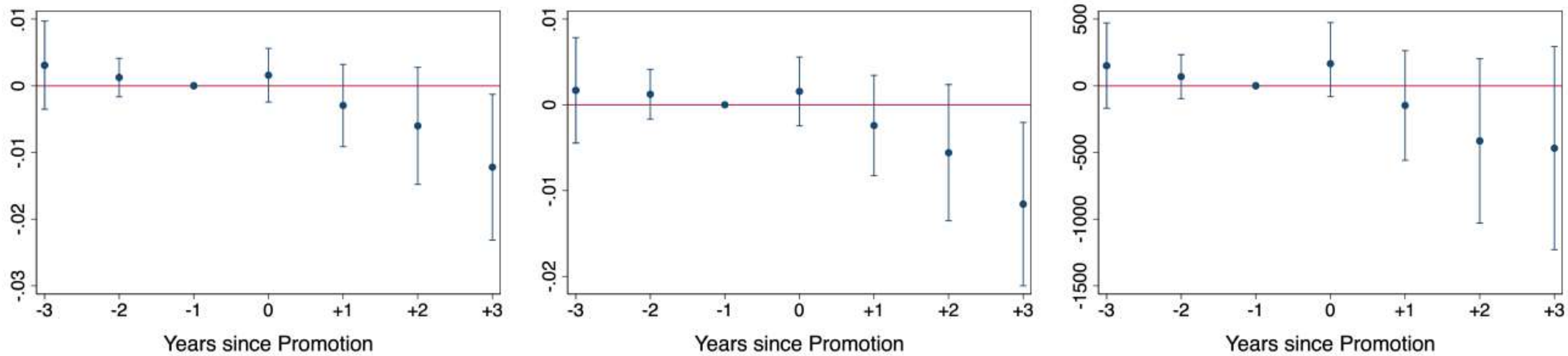


Notes: Panel (a) displays the fraction of workers promoted to middle-managment ("quadro") in family and non-family firms in a given centile of the distribution of log value added per worker. These promotions are measured by instances in which a given worker is not employed either as a manager or as a middle manager in year t and in year $t+1$ is employed by the same firm as a middle manager. Panel (b) is similar but shows the fraction of workers promoted from middle to upper management. Panel (c) and Panel (d) report the fractions of middle managers and managers in family and non-family firms in a given centile of the distribution of log value added per worker.

Figure 8: Returns to Promotions: Event-Study Evidence

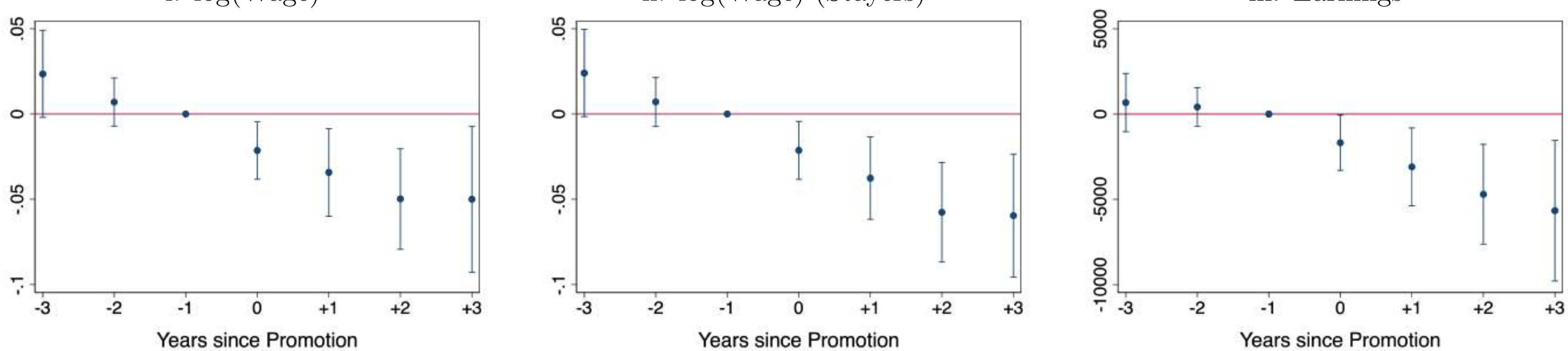
A. To Middle Management

i. log(Wage) ii. log(Wage) (Stayers) iii. Earnings



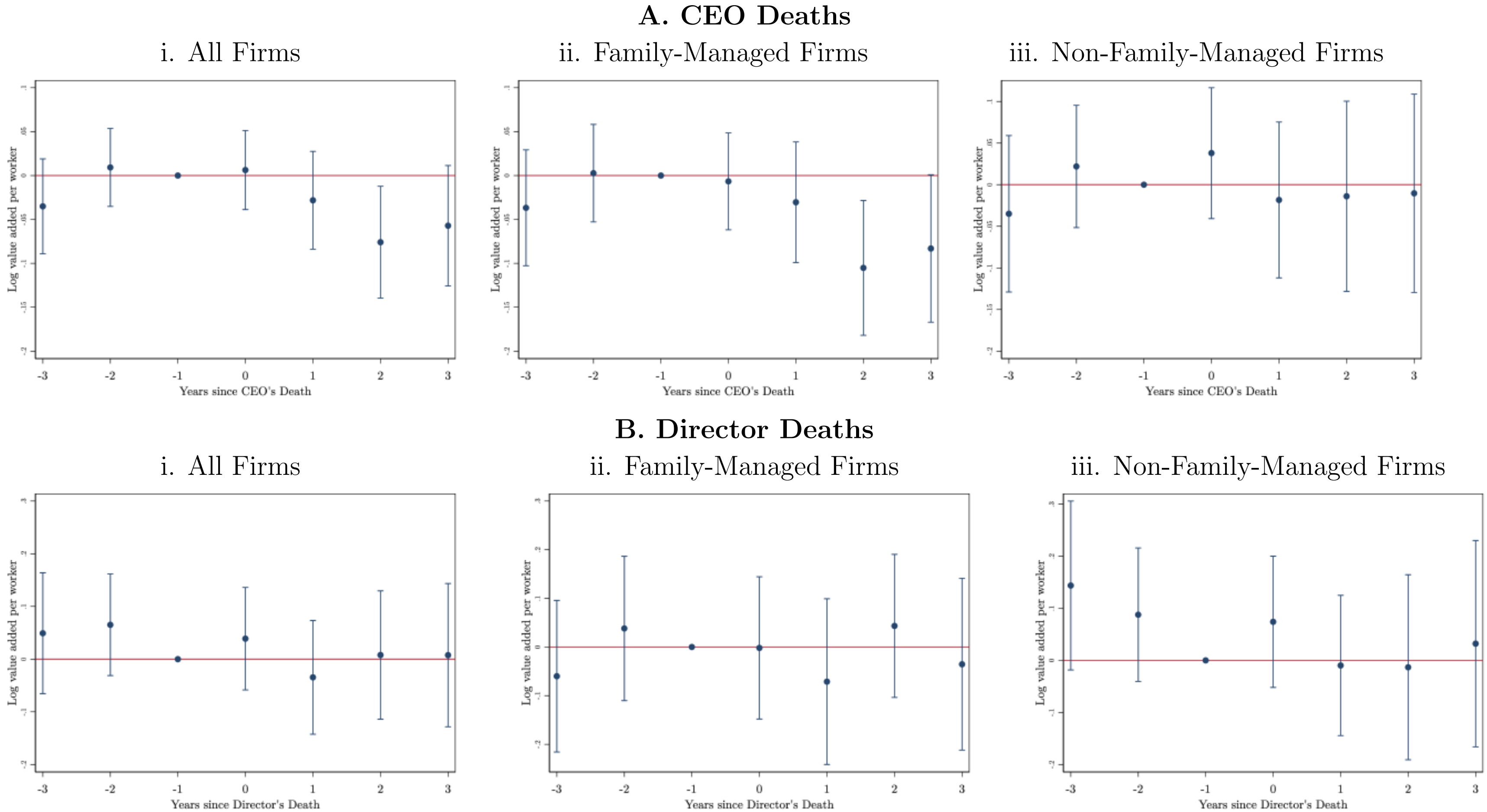
B. From Middle Management to Manager

i. log(Wage) ii. log(Wage) (Stayers) iii. Earnings



Notes: This figure presents event-study evidence of the differential effects of promotion between family and non-family firms. Panel A is for workers promoted to middle management positions, Panel B those promoted from middle to upper managerial positions. Every worker promoted in a family firm is matched with one who obtained the same type of promotion in a non-family firm, based on the following set of characteristics: logarithm of firm size, logarithm of value added per worker, two lags of log weekly wage, age, and (exactly), industry, contract type, full time status, and gender. All the characteristics are measured in the year prior to the promotion. Workers are kept in the sample for a window of $(-3, +3)$ years, where 0 is the year of the promotion. The dependent variable is then regressed on year fixed effects, worker fixed effects, event-year dummies, and event-year dummies interacted with a family firm dummy. The figures plot the coefficients on these interaction coefficients with the associated 95% confidence intervals based on standard errors clustered at the firm level.

Figure 9: Productivity around CEO and Director Deaths



Notes: This figure presents event-study evidence of the differential effects of the death of a CEO (Panel A) or member of the board of director(Panel B) between family and non-family firms. For every family firm in which the CEO/director dies, we find a similar non-family firm matching on industry, the logarithm of firm size, and the logarithm of value added per worker, all measured in the year before the death. Firms are kept in the sample for a $(-3, +3)-$year window, where 0 is the death event year. The logarithm of value added per worker is then regressed on year fixed effects, firm fixed effects, event-year dummies, and event-year dummies interacted with a family firm dummy. The figures plot the coefficients of these interaction terms with the associated 95% confidence intervals, based on standard errors clustered at the firm level. Subpanels A.i and B.i comprise all firms. Subpanels A.ii and B.ii cover only treated firms in which the deceased CEO/director was a member of the controlling family, together with the respective matched control firms. Subpanels A.iii and B.iii cover only treated firms in which the deceased CEO/director was not a member of the controlling family, together with the respective matched control firms.

Figure 10: Timeline of the Model

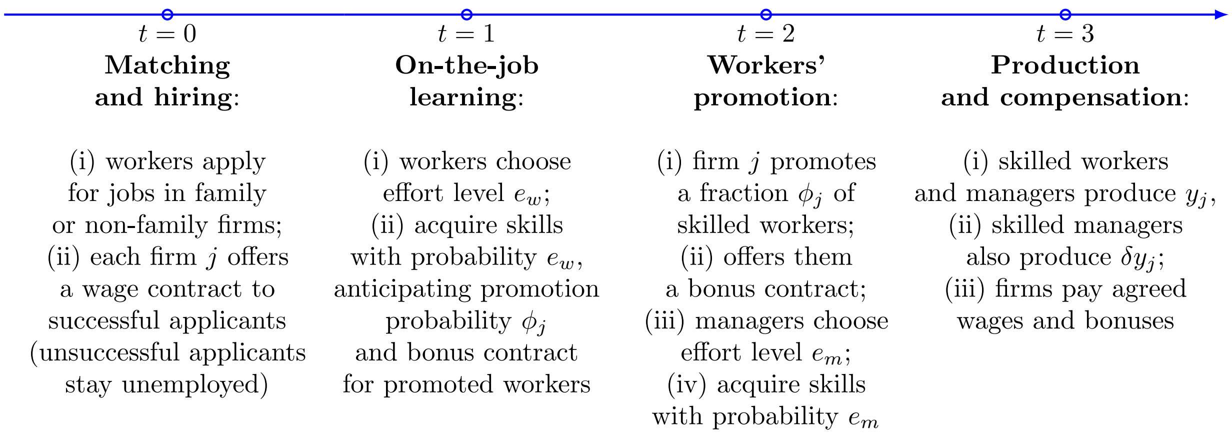


Table 1: Descriptive Statistics

	Family Firms	**Non-Family Firms**
Firm Size	14.056	31.782
Value Added per Worker	45.919	63.303
ROA	4.421	4.372
Leverage	0.788	0.762
Wage	436.833	510.696
Log(Wage)	6.034	6.177
Entry Log(Wage)	5.914	6.06
Log(Wage) Growth	0.042	0.046
Log(Wage) (Max - Median)	0.821	1.061
Log(Wage) (90 - 10 Perc.)	0.719	0.817
Share Managers	0.004	0.01
Share Middle Managers	0.01	0.04
Age	39.555	39.839
Share Males	0.651	0.634
Share Part Time	0.184	0.168
Share Temporary	0.210	0.221
# Firms	607,200	292,213
# Workers	9,972,643	9,421,821
# Firm-Years	3,319,919	1,394,195
# Worker-Year	46,664,781	44,310,305

Notes: This table presents average characteristics of family and non-family firms, for the period 2006–2017. Firm size is number of workers. Value added per worker is operating value added, in thousand euros, scaled by number of workers. ROA is earnings before interest and taxes scaled by total assets, all multiplied by 100. Leverage is one minus common equity scaled by total assets. Wage is the weekly wage. Entry Log(Wage) is the logarithm of the entry wage (i.e., the wage in the first year the worker has been employed with the firm). Log(Wage) (Max - Median) is the logarithm of the highest wage minus the logarithm of the median wage. Log(Wage) (90 - 10 Perc.) is the logarithm of the wage measured at the 90^{th} percentile minus the logarithm of the wage measured at the 10^{th} percentile. Share managers and share middle managers are the fractions of managers and middle managers. Age is the average age of the firm's workforce. Share males, part time, and temporary, are the fractions of male workers, part-time workers, and workers with fixed-term contracts, respectively. All the means, except for firm size, are weighted by the number of workers. The last four rows report the number of unique firms and workers, and the number of observations, at the firm-year level and at the worker-year level present in our sample.

Table 2: Decomposing the Wage Gap between Family and Non-Family Firms

	Family Firm	**Non-Family Firm**	**Difference**
Log(Wage) (Adjusted)	5.90	6.07	-0.16
Person Effects	-0.03	0.05	-0.08
Firm Effects	-0.03	0.05	-0.08
Log VA/L	3.56	3.73	-0.17
Rent Sharing Coefficient	0.14	0.13	0.01
Decomposing the Difference in Firm Effects into...			
Bargaining Channel	0.51	0.47	0.04
Productivity Channel	0.47	0.49	-0.02
Systematic component	-0.54	-0.44	-0.10

Notes: This table decomposes the average wage gaps between family and non-family firms after fitting the AKM specification described in equation (1) on the sample of Italian workers employed by a firm with non-missing value added data for the years 2005 and 2017. For each variable listed in a given row, we report the average for family firms (Column 1), that for non-family firms (Column 2) and the corresponding difference (Column 3). Log Wage (adjusted) refers to the average log wage after netting out the covariates in the AKM specification of (1). Person effect refers to the worker fixed effects from the AKM specification in equation (1). Firm effect refers to the firm fixed effects from the same AKM regression. Rent-sharing coefficient refers to the coefficient from a regression of the firm effects on a constant (denoted as θ in equation (4)) and value-added per worker, estimated separately for family and non-family firms. Using these two regressions, we then decompose average differences between firm effects in family and non-family firms into a productivity channel (differences in productivity in firm effects between family and non-family firms weighted by the non-family rent-sharing coefficient) and a bargaining component (differences in rent-sharing weighted by average log value added per worker by family firms)—see equation (5). All reported statistics are person-year weighted.

Table 3: PageRank of Family and Non-Family Firms

	(1)	(2)	(3)	(4)	(5)
Family Firm	-0.703***	-0.581***	-0.299***	-0.282***	-0.231***
	(0.076)	(0.070)	(0.020)	(0.018)	(0.012)
Firm Wage Premium		1.511***	1.318***	1.169***	0.935***
		(0.208)	(0.101)	(0.085)	(0.049)
Std. of Dep. Var.	1.45	1.45	1.45	1.45	1.45
# of Firms	313,375	310,584	310,584	310,584	310,584
Industry FE			X	X	
Province FE				X	
LLM FE					X

Notes: This table shows the results from a regression where the dependent variable is the PageRank utility index of Page et al. (1998), popularized in economics by Sorkin (2018). The index ranks employers on the basis of voluntary transitions of workers. To identify voluntary transitions, we exploit the fact that the Italian data specify whether a job was terminated owing to voluntary resignation or a different cause. The PageRank utility is then regressed on a dummy for family firm and (starting from column 2) on the AKM firm-wage effect described in Table 2. All regression results are estimated on the micro person-year data and are thus person-year weighted. All regressions control for year fixed effects as well the additional fixed effects listed at the bottom of the table. The local labor market (LLM) is defined as an industry-province combination. Standard errors are reported in parentheses and are clustered at the firm-level. ***, **, and * indicate the 1%, 5%, and 10% level of significance, respectively.

A Appendix

Appendix A.1 includes additional results and robustness checks.

Table A1 replicates the Oaxaca decomposition reported in Table 2 in the main text, but excluding shareholders from the sample.

Table A2 reports, in tabular form, coefficients for the event-study regressions used to estimate how the "returns to promotion" (to middle manager or from middle manager to upper manager) differ between family and non-family firms (see Figure 8 in the main text for the corresponding plots).

Table A3 reports, in tabular form, coefficients for the event-study regressions used to estimate how the productivity effect of the death of a CEO or director differs between family and non-family firms (see Figure 9 in the main text for the corresponding plots).

In Table A4 we replicate the analysis of Table 3 in the main text, but correcting for offer intensity and firm size to estimate the systematic utility associated with the firms in our sample (see Section IV.B in Sorkin, 2018).

Figure A1 displays the coefficients from regressing the firm wage premium on a family firm dummy. The regressions further control for several combinations of fixed effects at the level of city of birth, city of residence, and city of work.

Appendix A.2 sets out model derivations and proofs omitted from the main text for brevity.

Figure A3 presents the results of a simple simulation exercise. Without attempting to match the data quantitatively, we simply assume that the productivity parameter in the model is drawn from a uniform distribution. With some additional parametric assumptions regarding β_i, δ, c_M, and c_w, we obtain the equilibrium wages and promotion rates for family and non-family firms.[22] We find: (i) wages and promotion rates are higher in non-family firms; (ii) they are both increasing in productivity; and (iii) the gap in wages and

[22]We set $\beta_F = 1.5$, $\beta_{NF} = 0.8$, $\delta = 0.5$, $c_M = 1$, and $c_w = 0.6$. The productivity parameter y_i is uniformly distributed between 0.5 and 1 and the range of value added per worker shown in the figures is $(-1.2, -.2)$.

promotion rates increases with productivity.

A.1 Additional Results

Table A1: Decomposing the Wage Gap between Family and Non-Family Firms (Excluding Shareholders)

	Family Firm	**Non-Family Firm**	**Difference**
log(Wage)	6.03	6.17	-0.15
Person Effects	-0.03	0.04	-0.07
Firm Effects	-0.03	0.05	-0.08
log(Value Added / Worker)	3.56	3.73	-0.17
Rent Sharing Coefficient	0.14	0.13	0.01
Decomposing the Difference in Firm Effects into...			
Bargaining Channel	0.51	0.47	0.04
Productivity Channel	0.47	0.49	-0.02
Systematic component	-0.54	-0.44	-0.10

Notes: This table decomposes the average wage gaps between family and non-family firms after fitting the AKM specification described in equation (1) to the sample of Italian workers employed by a firm with non-missing data on value added for the years 2006–2017, with the exclusion of employees who are also shareholders in the firm. For each variable in a given row, we report the averages for family firms (Column 1) and non-family firms (Column 2) with the corresponding difference (Column 3). Log Wage (adjusted) refers to the average log wage after netting out the covariates in the AKM specification of equation (1). Person effect refers to the worker fixed effects from the AKM specification in equation (1). Firm effect refers to the firm fixed effects from the same AKM regression. Rent-sharing coefficient refers to the coefficient from a regression of the firm effects on a constant (denoted as θ in the equation (4)) and value-added per worker, estimated separately for family and non-family firms. Using these two regressions, we then decompose the average differences between firm effects in family and non-family firms into a productivity channel (differences in productivity in firm effects weighted by the non-family rent-sharing coefficient) and a bargaining component (differences in rent-sharing weighted by average log value added per worker by family firms)—see equation (5). All the statistics reported here are person-year weighted.

Table A2: Returns to Promotion: Event-Study Coefficients

Sample:	To Middle Manager			Middle Manager to Manager		
Dep. Var.	log(Wage)	log(Wage) (Stayers)	Earnings	log(Wage)	log(Wage) (Stayers)	Earnings
	(1)	(2)	(3)	(4)	(5)	(6)
β_{-3}	0.003	0.002	150.68	0.024	0.024	674.02
	(0.003)	(0.003)	(163.05)	(0.013)	(0.013)	(869.07)
β_{-2}	0.001	0.001	68.02	0.007	0.007	412.05
	(0.001)	(0.001)	(84.09)	(0.007)	(0.007)	(577.44)
β_0	0.002	0.002	165.95	-0.021**	-0.021**	-1,681.65**
	(0.002)	(0.002)	(126.06)	(0.009)	(0.009)	(825.94)
β_1	-0.003	-0.002	-147.89	-0.034***	-0.038***	-3,099.76***
	(0.003)	(0.003)	(209.56)	(0.013)	(0.012)	(1,162.10)
β_2	-0.006	-0.006	-413.63	-0.050***	-0.058***	-4,706.67***
	(0.004)	(0.004)	(314.63)	(0.015)	(0.015)	(1,489.06)
β_3	-0.012**	-0.012**	-467.65	-0.050**	-0.060***	-5,666.77***
	(0.006)	(0.005)	(388.62)	(0.022)	(0.018)	(2,099.79)
R^2	0.815	0.845	0.686	0.839	0.866	0.737
Obs.	151,406	141,499	154,014	6,499	6,111	6,566

Notes: Table A2 presents event-study evidence of the differential effects of a promotion depending on whether the worker was promoted by a family or a non-family firm. We consider two types of promotion. The first three columns comprise workers promoted to middle management positions. Columns 4-6 cover workers promoted from middle management to upper managerial positions. Every worker in a family firm is matched with a worker who got the same type of promotion in a non-family firm, based on the following characteristics: logarithm of firm size, logarithm of value added per worker, wage growth, age, and (exactly), industry, contract type, full time status, and gender. Workers are kept in the sample for a $(-3, +3)-$year window, where 0 is the year of the promotion. The dependent variable (indicated at the top of each column) is then regressed on year fixed effects, worker fixed effects, event-year dummies, and event-year dummies interacted with a family firm dummy. The table reports coefficients on these interaction terms with the corresponding standard errors, clustered at the level of the original employer (that is, the firm where the employee was promoted). ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

Table A3: Productivity around CEO and Director Deaths: Event-Study Coefficients

Death Events: Sample	CEOs			Directors		
	All	FM	NFM	All	FM	NFM
	(1)	(2)	(3)	(4)	(5)	(6)
β_{-3}	-0.035	-0.037	-0.035	0.049	-0.060	0.144
	(0.028)	(0.034)	(0.048)	(0.059)	(0.079)	(0.083)
β_{-2}	0.009	0.003	0.022	0.065	0.038	0.088
	(0.023)	(0.028)	(0.037)	(0.049)	(0.076)	(0.065)
β_0	0.006	-0.007	0.038	0.039	-0.002	0.074
	(0.023)	(0.028)	(0.040)	(0.050)	(0.075)	(0.064)
β_1	-0.028	-0.030	-0.018	-0.035	-0.071	-0.010
	(0.028)	(0.035)	(0.048)	(0.055)	(0.087)	(0.069)
β_2	-0.076**	-0.105***	-0.014	0.008	0.043	-0.013
	(0.032)	(0.039)	(0.058)	(0.062)	(0.075)	(0.091)
β_3	-0.057	-0.083*	-0.010	0.008	-0.035	0.032
	(0.035)	(0.043)	(0.061)	(0.069)	(0.090)	(0.101)
R^2	0.668	0.650	0.709	0.685	0.682	0.705
Obs.	7,459	5,018	2,441	1,809	807	1,002

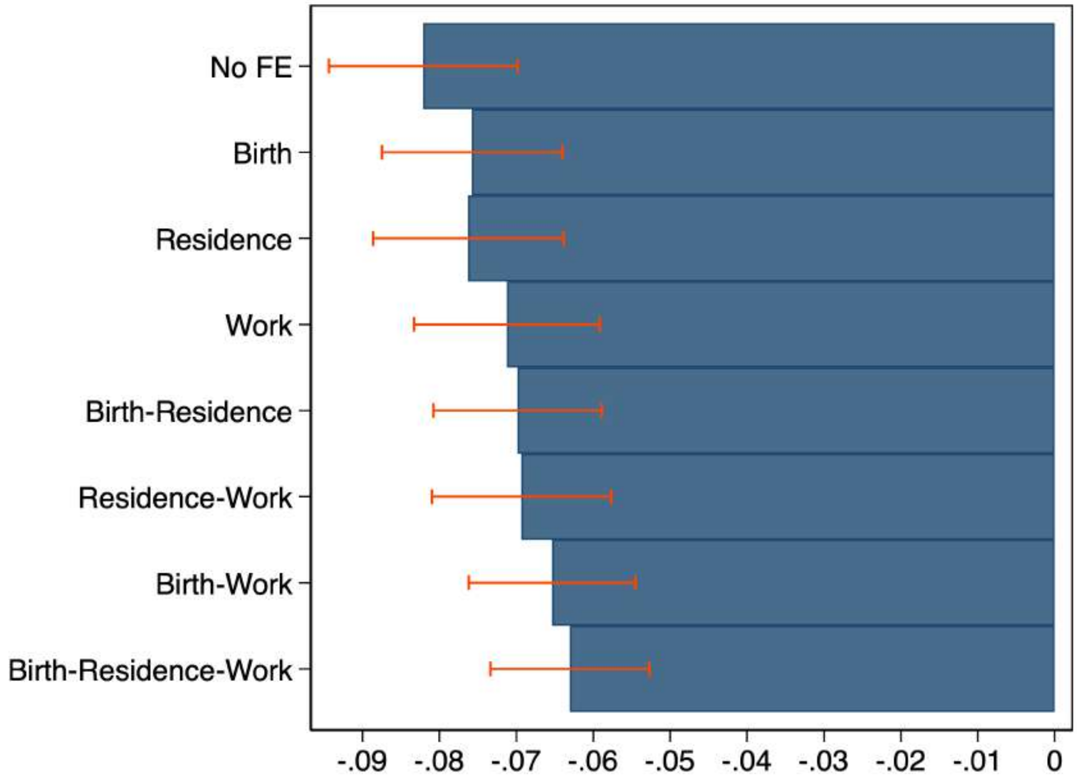
Notes: Table A3 presents event-study evidence on the differential effects of the death of a CEO (columns 1-3) or director (columns 4-6) in family and non-family firms. Every family firm is matched with a non-family firm based on industry, the logarithm of firm size, and the logarithm of value added per worker. Firms are kept in the sample for a $(-3, +3)-$year window, where 0 is the year of the death event. The logarithm of value added per worker is then regressed on year fixed effects, firm fixed effects, event-year dummies, and event-year dummies interacted with a family firm dummy. The table displays the coefficients on these interaction terms (standard errors clustered at the firm level in parentheses). The samples in columns 1 and 4 include all firms. Those in columns 2 and 5 (family-managed, "FM") comprise only treated firms in which the deceased CEO/director was a member of the controlling family, together with the matched control firms. Columns 3 and 6 (non-family-managed, "NFM") include only treated firms in which the deceased CEO/director was not a member of the controlling family, together with the matched control firms. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

Table A4: PageRank of Family and Non-Family firms after adjusting for Offer Intensity and Firm Size

	(1)	(2)	(3)	(4)	(5)
Family Firm	-2.918***	-2.660***	-1.521***	-1.374***	-1.105***
	(0.273)	(0.276)	(0.074)	(0.060)	(0.034)
Firm effect		3.102***	4.064***	3.472***	2.978***
		(0.633)	(0.316)	(0.252)	(0.113)
Std. of Dep. Var.	3.36	3.35	3.35	3.35	3.35
# of Firms	309,395	307,259	307,259	307,259	307,259
Industry FE			X	X	
Province FE				X	
LLM FE					X

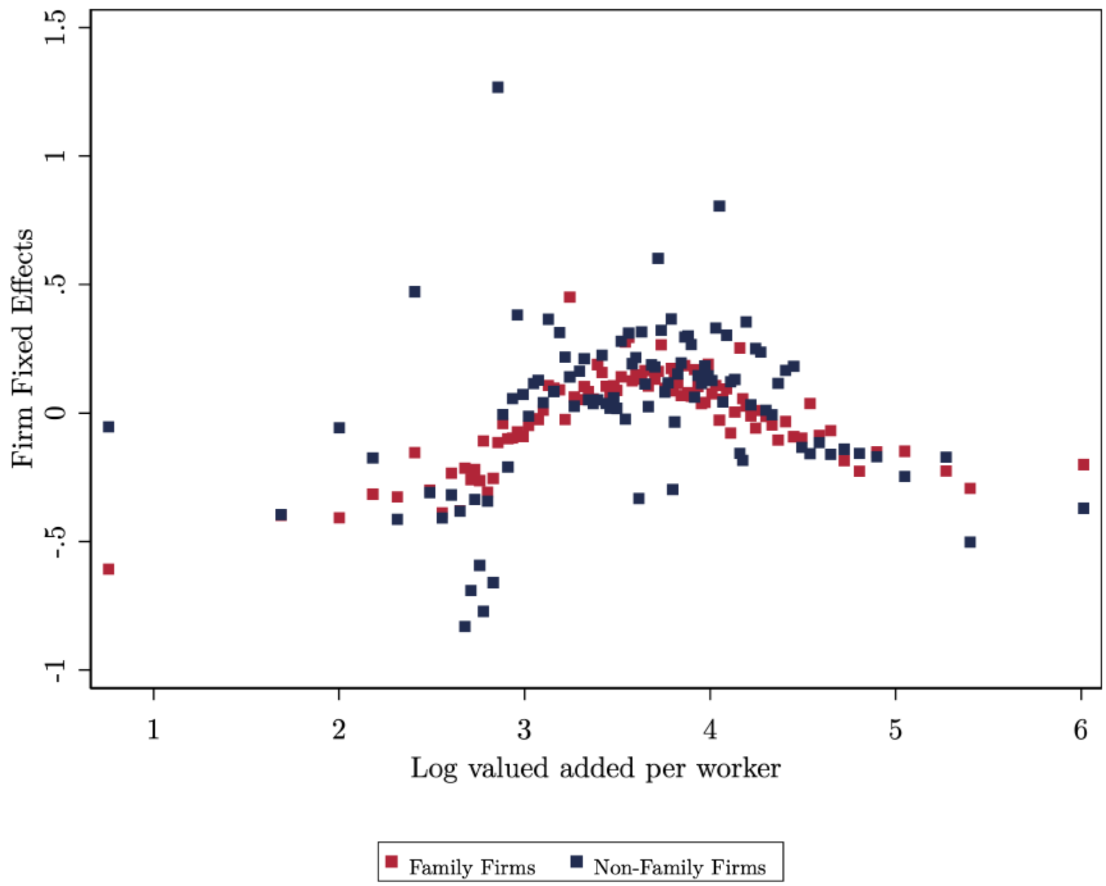
Notes: Table A4 shows the results of a regression where the dependent variable is the PageRank utility index of Page et al. (1998) popularized in economics by Sorkin (2018), but correcting for offer intensity and firm size. The index ranks employers on the basis of voluntary transitions of workers. To identify voluntary transitions, we exploit the fact that the Italian data specify whether a job termination was due to voluntary resignation or a different cause. The PageRank utility is then regressed on a family firm dummy and (from column 2) on the AKM firm-wage effect described in Table 2. All regression results are estimated on the micro person-year data and are thus person-year weighted. In all the regressions we control for year fixed effects as well the additional fixed effects listed at the bottom of the table. Standard errors (reported in parentheses) are clustered at the firm-level. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

Figure A1: The Family Firm Discount: Place of Birth, Residence, Work



Notes: Figure A1 shows the coefficients obtained from regressing the firm "AKM" fixed effect on a family firm dummy, the logarithm of value added per worker, year fixed effects, and several sets of fixed effects. Each horizontal bar corresponds to the coefficient for the family firm dummy in a given specification, with the 95% confidence interval in red. From top to bottom, the specifications are: no additional fixed effects, city of birth fixed effects, city of residence fixed effects, city of work fixed effects, city of birth × city of residence fixed effects, city of residence × city of work fixed effects, city of birth × city of work fixed effects, and city of birth × city of residence × city of work fixed effects.

Figure A2: Number of Sick Days in Family and Non-Family Firms



Notes: Figure A2 plots average firm-specific "paid sick days" fixed effects against centiles of productivity. We first estimate the AKM specification in equation (1) using as outcome $\log n_{it}$, i.e., the logarithm of the number of sick days of worker i in year t. We then compute the average log value added per worker and the firm effects in number of sick days obtained from this specification for each centile of log value added per worker.

A.2 Model Derivations

In the model, firms cannot commit to future actions: at $t = 0$ they cannot commit either to a specific promotion probability or to a specified bonus contract to promoted workers at $t = 2$. Hence, the model is solved by backward induction.

Stage *t*=2. The optimal compensation scheme that firm j can offer to newly appointed managers pays a bonus x_j if the manager succeeds in producing the additional output δy_j at $t = 3$ and 0 if not (given limited liability, negative bonuses are precluded). Given this state-contingent contract, workers promoted to managerial positions choose effort e_m to acquire managerial skills, maximizing their incremental utility as of $t = 2$:

$$\max_{e_m} \ \mathbb{E}(U_2) = e_m x_j - \frac{c_m}{2} e_m^2. \tag{A.1}$$

Hence, their optimal effort is $e_m^* = x_j/c_m$. The firm chooses the bonus x_j so as to maximize the incremental profits it can extract from the worker's effort e_m^*:

$$\max_{x_j} \ \mathbb{E}(\pi_{2j}) = \phi_j e_m^*(\delta y_j - x_j) = \frac{\phi_j x}{c_m}(\delta y_j - x_j), \tag{A.2}$$

where the expression takes into account that at $t = 2$ firm j promotes a fraction ϕ_j of its successful employees. Hence, the optimal bonus is $x_j^* = \delta y_j/2$. The resulting equilibrium managerial effort is increasing in the firm's productivity y_j:

$$e_m^*(x_j^*) = \frac{\delta y_j}{2c_m}, \tag{A.3}$$

where it is assumed that $\delta y_j \leq 2c_m$, so that the probability $e_m^*(x_j^*) \leq 1$. Thus, the incremental profit managers are expected to generate in equilibrium at stage $t = 2$ is

$$\mathbb{E}(\pi_{2j}^*) = \phi_j e_m^*(\delta y_j - x_j^*) = \frac{\phi_j}{c_m}\left(\frac{\delta y_j}{2}\right)^2. \tag{A.4}$$

Each firm j chooses the promotion probability ϕ_j of qualified workers trading the incremental profits π^*_{2i} off against the loss of benefit of control, i.e., solving:

$$\max_{\phi_j} \ \mathbb{E}(\pi^*_{2j}) - \frac{\beta_j}{2}\phi_j^2. \tag{A.5}$$

Firm j's optimal promotion probability of skilled workers is decreasing in the owners' taste for the benefits of control:

$$\phi^*_j = \frac{1}{\beta_j c_m}\left(\frac{\delta y_j}{2}\right)^2, \tag{A.6}$$

which, being a probability, is assumed not to exceed 1.[23] It is immediate to show that $\frac{\partial \phi^*_j}{\partial \beta_j} < 0$ and $\frac{\partial^2 \phi^*_j}{\partial \beta_j \partial y_j} < 0$. Recalling that family firms have a stronger taste for corporate control, that is, $\beta_F > \beta_{NF}$, this implies that $\phi^*_F - \phi^*_{NF}$ is negative and decreasing in y_j, which establishes **Prediction 2**.

Stage t=1. Firm j finds it optimal to offer workers a two-valued state-contingent compensation: w_j if the worker succeeds in acquiring productive skills on the job, and 0 otherwise. The employees choose effort e_w so as to maximize their stage 1 expected utility:

$$\max_{e_w} \ \mathbb{E}(U_1) = e_w(w_j + \phi^*_j e^*_m x^*_j) - \frac{c_w}{2}e_w^2 - \phi^*_j e_w \frac{c_m}{2} e^{*2}_m. \tag{A.7}$$

The first term of expression (A.7) indicates that when choosing their effort, workers rationally anticipate that if successful they not only make wage w_j but also gain a chance ϕ^*_j of being promoted to a managerial position and earning bonus x^*_j, while bearing the cost of acquiring the necessary managerial skills (the last term). Substituting from previous expressions for x^*_j and e^*_m and computing the first-order condition with respect to e_w, we get that the workers' optimal effort is increasing both in the wage and in the probability

[23]As $\delta y_j \leq 2c_m$, a sufficient condition for $\phi^*_j \leq 1$ is that $\beta_j \geq \delta y_j/2$.

of promotion:

$$e_w^*(w_j) = \frac{1}{c_w}\left[w_j + \frac{\phi_j^*}{2c_m}\left(\frac{\delta y_j}{2}\right)^2\right] = \frac{1}{c_w}\left[w_j + \frac{1}{2\beta_j c_m^2}\left(\frac{\delta y_j}{2}\right)^4\right]. \tag{A.8}$$

At stage $t = 1$, each firm j sets the wage so as to maximize its overall profit, anticipating its optimal future choice of the bonus x_j^* and the managers' future optimal choice of effort e_m^*:

$$\begin{aligned}\max_{w_j}\ \mathbb{E}(\pi_{1j}) &= e_w^*[(y_j - w_j) + \phi_j^* e_m^*(\delta y_j - x_j^*)] \\ &= \frac{1}{c_w}\left[w_j + \frac{\phi_j^*}{2c_m}\left(\frac{\delta y_j}{2}\right)^2\right]\left[y_j - w_j + \frac{\phi_j^*}{c_m}\left(\frac{\delta y_j}{2}\right)^2\right],\end{aligned} \tag{A.9}$$

yielding the optimal efficiency wage paid to employees who succeed in acquiring productive skills at stage $t = 1$:

$$w_j^* = \frac{1}{2}\left[y_j + \frac{\phi_j^*}{2c_m}\left(\frac{\delta y_j}{2}\right)^2\right] = \frac{1}{2}\left[y_j + \frac{1}{2\beta_j c_m^2}\left(\frac{\delta y_j}{2}\right)^4\right]. \tag{A.10}$$

It is straightforward to show that $\frac{\partial w_j^*}{\partial \beta_j} < 0$ and $\frac{\partial^2 w_j^*}{\partial \beta_j \partial y_j} < 0$. Recalling that $\beta_F > \beta_{NF}$, this implies that $w_F^* - w_{NF}^* < 0$ and decreases in y_j.

Replacing the firm's optimal efficiency wage (A.10) in expression (A.8) yields workers' equilibrium effort at stage $t = 1$:

$$e_w^*(w_j^*) = \frac{1}{2c_w}\left[y_j + \frac{3\phi_j^*}{2c_m}\left(\frac{\delta y_j}{2}\right)^2\right] = \frac{1}{2c_w}\left[y_j + \frac{3}{2\beta_j c_m^2}\left(\frac{\delta y_j}{2}\right)^4\right]. \tag{A.11}$$

Effort decreases in β_j, and is therefore greater in non-family firms. Taken together, equations (A.10) and (A.11) imply all the findings stated in **Prediction 1**.

Substituting the equilibrium effort (A.11) into expression (A.9), it is immediate to see that expected profits are always positive. Substituting expression (A.11) into (A.7) we obtain the equilibrium expected utility of workers from working in firm j:

$$\mathbb{E}(U_1^*) = \frac{1}{8c_w}\left[y_j + \frac{3}{2}\frac{\phi^*}{c_m}\left(\frac{\delta y_j}{2}\right)^2\right]^2 = \frac{1}{8c_w}\left[y_j + \frac{3}{2\beta_j c_m^2}\left(\frac{\delta y_j}{2}\right)^4\right]^2. \tag{A.12}$$

As this expression is decreasing in β_j, the expected utility is lower for employees of family firms, proving **Prediction 4**.

Next, we prove that the expected bonus from a promotion is greater in non-family than family firms. The expected bonus b for a worker who is promoted, conditional on y_j, is:

$$\mathbb{E}\left[b|y_j\right] = x^* e_m^* = \frac{\delta y_j}{2}\frac{\delta y_j}{2c_m} = \frac{1}{c_m}\left(\frac{\delta y_j}{2}\right)^2. \tag{A.13}$$

To obtain the unconditional expected bonus, we apply the law of iterated expectations. Importantly, we take the expectation only over the population of managers. As the promotion probability, conditional on y_j, is $\phi^* e_w^*$, the relevant probability distribution function with respect to which the expectation needs to be taken is $\phi^* e_w^* / \int \phi^* e_w^* dF(y)$. By substituting expressions (A.6), (A.11), and (A.13) for ϕ^*, e_w^*, and $\mathbb{E}\left[b|y_j\right]$, we have:

$$\mathbb{E}[b] = \mathbb{E}_y\left[\mathbb{E}\left[b|y_j\right]\right] = \frac{\int \frac{1}{c_m}\left(\frac{\delta y_j}{2}\right)^4\left[y_j + \frac{3}{2\beta c_m^2}\left(\frac{\delta y_j}{2}\right)^4\right] dF(y)}{\int \left(\frac{\delta y_j}{2}\right)^2\left[y_j + \frac{3}{2\beta c_m^2}\left(\frac{\delta y_j}{2}\right)^4\right] dF(y)}. \tag{A.14}$$

Noting that the raw moment $E[y_j^k] \equiv \int y_j^k dF(y)$, expression (A.14) can be rewritten as:

$$\mathbb{E}[b] = \frac{\frac{1}{c_m}\left(\frac{\delta}{2}\right)^4 \mathbb{E}[y_j^5] + \left(\frac{\delta}{2}\right)^8 \frac{3}{2\beta c_m^3}\mathbb{E}[y_j^8]}{\left(\frac{\delta}{2}\right)^2 \mathbb{E}[y_j^3] + \left(\frac{\delta}{2}\right)^6 \frac{3}{2\beta c_m^2}\mathbb{E}[y_j^6]} \equiv \frac{\mathbb{N}}{\mathbb{D}}. \tag{A.15}$$

Showing that the expected bonus is smaller in family firms is equivalent to showing that $\frac{\partial \mathbb{E}[b]}{\partial \beta} < 0$, namely that $\frac{\partial \mathbb{N}}{\partial \beta}\mathbb{D} - \frac{\partial \mathbb{D}}{\partial \beta}\mathbb{N} < 0$, where $\mathbb{N}$ and $\mathbb{D}$ are the numerator and denominator of expression (A.15), respectively. Observing that

$$\frac{\partial \mathbb{N}}{\partial \beta}\mathbb{D} - \frac{\partial \mathbb{D}}{\partial \beta}\mathbb{N} = -\left(\frac{\delta}{2}\right)^{10} \frac{3}{2\beta^2 c_m^3}\,\mathbb{E}[y_j^3]\,\mathbb{E}[y_j^8] - \left(\frac{\delta}{2}\right)^{14} \frac{9}{4\beta^3 c_m^5}\,\mathbb{E}[y_j^6]\,\mathbb{E}[y_j^8] + \left(\frac{\delta}{2}\right)^{10} \frac{3}{2\beta^2 c_m^3}\,\mathbb{E}[y_j^5]\,\mathbb{E}[y_j^6] + \left(\frac{\delta}{2}\right)^{14} \frac{9}{4\beta^3 c_m^5}\,\mathbb{E}[y_j^6]\,\mathbb{E}[y_j^8] \tag{A.16}$$

and grouping the common terms, we obtain:

$$\text{sign}\left(\frac{\partial \mathbb{N}}{\partial \beta}\mathbb{D} - \frac{\partial \mathbb{D}}{\partial \beta}\mathbb{N}\right) = \text{sign}\left(\mathbb{E}[y_j^5]\,\mathbb{E}[y_j^6] - \mathbb{E}[y_j^3]\,\mathbb{E}[y_j^8]\right) = \text{sign}\left(\frac{\mathbb{E}[y_j^6]}{\mathbb{E}[y_j^3]} - \frac{\mathbb{E}[y_j^8]}{\mathbb{E}[y_j^5]}\right), \tag{A.17}$$

where the last equality relies on all the moments of y being strictly positive.

A sufficient condition for the sign of expression (A.16) to be negative is the following monotonicity assumption on the raw moments of the distribution $F(y)$:

$$\frac{\mathbb{E}[y_j^{k-1}]}{\mathbb{E}[y_j^{k-2}]} < \frac{\mathbb{E}[y_j^{k}]}{\mathbb{E}[y_j^{k-1}]}, \tag{A.18}$$

where $\mathbb{E}[y_j^k]$ is assumed to exist for all $k = 1, 2, ..., K$. This inequality implies that:

$$\frac{E[y_j^{k-1}]}{E[y_j^{k-h-1}]} < \frac{E[y_j^{k}]}{E[y_j^{k-h}]}, \tag{A.19}$$

where $h < k-2$, thanks to a simple induction argument.[24]

[24]For $h = 1$, inequality (A.19) reduces to assumption (A.18). Assume that it also holds for a generic $h = n < k-2$. Then:

$$\frac{\mathbb{E}[y_j^{k-1}]}{\mathbb{E}[y_j^{k-n-1}]} < \frac{\mathbb{E}[y_j^{k}]}{\mathbb{E}[y_j^{k-n}]}. \tag{A.20}$$

Rewrite assumption (A.18) for the $(k-n)^{th}$ moment as:

$$\frac{\mathbb{E}[y_j^{k-n-1}]}{\mathbb{E}[y_j^{k-n-2}]} < \frac{\mathbb{E}[y_j^{k-n}]}{\mathbb{E}[y_j^{k-n-1}]}. \tag{A.21}$$

Multiplying inequalities (A.20) and (A.21) side by side yields:

$$\frac{\mathbb{E}[y_j^{k-1}]}{\mathbb{E}[y_j^{k-n-2}]} < \frac{\mathbb{E}[y_j^{k}]}{\mathbb{E}[y_j^{k-n-1}]}, \tag{A.22}$$

showing that inequality (A.19) also holds for $h = n+1$ and completing the proof.

Setting $h=3$ and $k=7$, inequality (A.19) becomes:

$$\frac{\mathbb{E}[y_j^6]}{\mathbb{E}[y_j^3]} < \frac{\mathbb{E}[y_j^7]}{\mathbb{E}[y_j^4]}, \tag{A.23}$$

and for $h=3$ and $k=8$, it becomes:

$$\frac{\mathbb{E}[y_j^7]}{\mathbb{E}[y_j^4]} < \frac{\mathbb{E}[y_j^8]}{\mathbb{E}\, y_j^5}. \tag{A.24}$$

Inequalities (A.23) and (A.24) jointly imply:

$$\frac{\mathbb{E}[y_j^6]}{\mathbb{E}[y_j^3]} < \frac{\mathbb{E}[y_j^8]}{\mathbb{E}[y_j^5]}. \tag{A.25}$$

so that the sign of expression (A.16) is negative. This proves **Prediction 3**.

Stage *t*=0. In the initial stage, the labor market allocates workers between firms. Each worker chooses whether to apply to a family or a non-family firm, conditioning on expectations of future compensation and promotion prospects, as computed above.

In equilibrium, workers allocate their applications so as to equalize their expected utility from applying to family and non-family firms. Denoting the mass of applicants to the two types of firm by A_F and A_{NF} respectively, and recalling that both types hire workers randomly from their respective applicant pools, applicants get a job in a family firm with probability $\lambda_F \equiv N_F/A_F$ and in a non-family firm with probability $\lambda_{NF} \equiv N_{NF}/A_{NF}$. As employment in either type yields a positive expected utility by equation (A.12), while unemployment entails zero utility, anyone in the labor force L applies for a job ($L = A_F + A_{NF}$), but not all applicants are hired, i.e., $\lambda_F < 1$ and/or $\lambda_F < 1$.

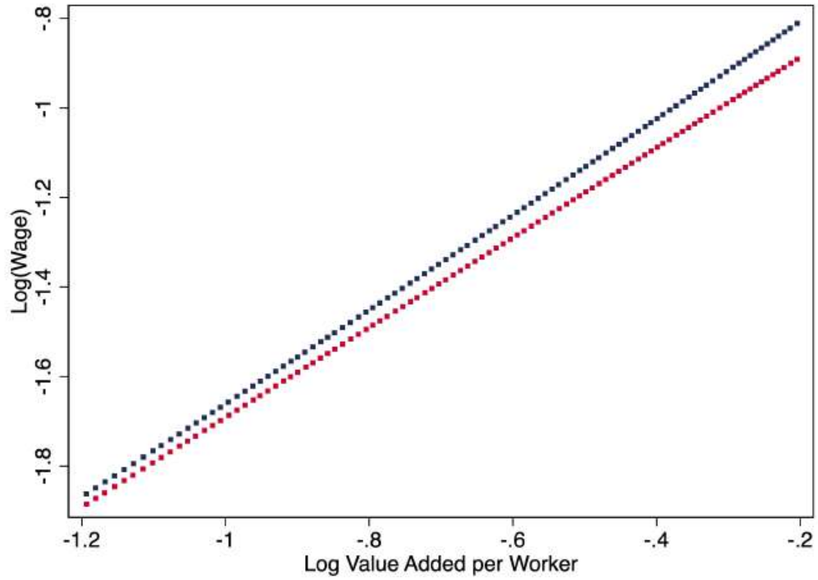
Applicants will be indifferent between applying to the two types of firm if

$$\lambda_F \cdot \mathbb{E}(U_F^*) = \lambda_{NF} \cdot \mathbb{E}(U_{NF}^*). \tag{A.26}$$

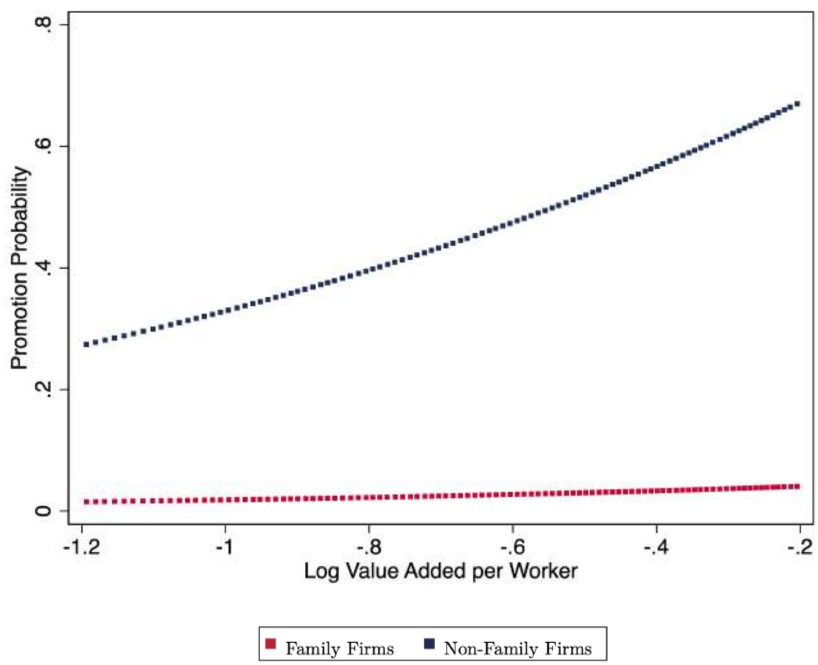
Since in equilibrium careers have been shown to offer greater expected utility in non-family than in family firms, the indifference condition (A.26) implies that $\lambda_{NF} < \lambda_F$: intuitively, the probability of being hired upon applying to a non-family firm must be sufficiently low compared with the probability of being hired upon applying to a family firm. In other words, in equilibrium applying to a non-family firm entails greater unemployment risk than applying to a family firm.

Figure A3: Model Simulation

A. log(Wage)



B. Promotion Probability



Notes: Figure A3 plots the average logarithm of wage (Panel A) and the average promotion probability (Panel B) against the percentile of the logarithm of value added per worker, as produced by the model of Section 6. The parameters are: $\beta_F = 1.5$, $\beta_{NF} = 0.8$, $\delta = 0.5$, $c_M = 1$, and $c_w = 0.6$. The productivity parameter y_i is uniformly distributed between 0.5 and 1 and the range of value added per worker shown in the figures is $(-1.2, -.2)$. We produce 10,000 values of y_i and simulate the model based on these parameters; then compute averages of the variables on the $y-$axes for each percentile of value added per worker separately for family and non-family firms.