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 and EIEF*
[3] *Tilburg University*
[4] *University of British Columbia and NBER*
[5] *Luiss University, EIEF and CEPR*

Abstract

We investigate the compensation policies of family and non-family firms using a novel employer-employee matched dataset, comprising virtually the universe of Italian incorporated companies combined with information on their shareholders. Family firms pay significantly lower wages and offer slower careers. Half of their wage gap is accounted for by workers' time invariant characteristics, while the rest cannot be explained by differences in productivity or in non-monetary compensating differentials. The evidence is consistent with a model where family owners protect their control over the firm by creating a "glass ceiling" that limits workers' career progression, thus hindering their human capital accumulation.

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, Lorenzo Cappellari, Pietro Coretto, Sara Formai, Giulia Giupponi, Francis Kramarz, Salvatore Lattanzio, Vincenzo Scrutinio, and seminar participants at LUISS, the UCSD Rady School of Management Brown Bag Finance Seminar, the University of Navarra, the XXXVIII AIEL Conference, and the VI Annual Conference of the Visitinps Scholars Program. 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 and Fabiano Schivardi). The findings and conclusions expressed are solely those of the authors and do not represent the views of INPS.

1 Introduction

Family firms employ a sizable fraction of the total workforce: between 40% and 50% of all European private employment is accounted for by family firms.[1] In this paper we investigate whether the employment policies of these firms differ from that those of non-family firms and, if so, what accounts for these differences.

We address these questions using a rich merged employer-employee dataset for the universe of Italian incorporated companies, based on social security records and detailed information about firms' ownership structure and financial accounts. The data includes almost 900,000 unique firms and over 19 million workers, over a 14-year time period. Such data presents two key advantages over those analyzed in previous work. First, it covers highly heterogeneous companies, ranging from private firms with few employees to large public corporations, whereas most previous research focuses on very large firms. Second, we have information on the careers of individual workers, rather than employment and labor cost data aggregated at the firm level.

A prerequisite for our analysis is to identify family firms in our sample. In most of our analysis, a family firm is defined as one where a single family holds more than 50% of the shares. The family firms in our sample are smaller and substantially less productive than non-family ones. However, the two types of firms have very similar return on assets. The fact that family firms are not less profitable despite their lower productivity can be attributed at least partly to their ability to pay lower wages. We will refer to this wage gap, which is the main focus of our analysis, as the "family firm discount."

The first step in our work is to measure the gap in compensation between family and non-family firms. Controlling for gender and experience, the employees of family firms earn on average 16% less than those of non-family firms, in line with previous evidence by Sraer and Thesmar (2007) and Bassanini et al. (2013) based on French data, and by Ellul

[1] See the data reported by https://europeanfamilybusinesses.eu/about-european-family-businesses/, where family firms are defined as those where the majority of decision rights are in the hands of a family, and at least one family member is involved in the firm's governance.

et al. (2017) based on cross-country data for listed companies.

Second, upon decomposing this wage gap into the portion associated with workers' characteristics and that related to firm characteristics, based on the "AKM" approach introduced by Abowd et al. (1999), it turns out that roughly half of the total wage differential (8 percentage points) is accounted for by individual characteristics of family firms' employees and the remaining half by firm characteristics. Next, we further partition the portion of the wage differential accounted for by firm characteristics in three components, using the Oaxaca-Blinder decomposition (Oaxaca, 1973; Blinder, 1973): (i) a component related to the firm's ownership (family or non-family); (ii) a rent-sharing component, capturing how firms split value added between wages and profits; (iii) a productivity component, which determines the size of the rents to be shared. The component reflecting firm ownership accounts for 10 percent of the wage gap: hence the family firm discount is quite sizeable. In comparison, the productivity shortfall of family firms relative to non-family ones is small: it only accounts for 2 additional points. Instead, the rent-sharing component partly offsets the first two components: workers in family firms appear to have greater bargaining power than those in non-family firms in the wage-setting process, which accounts for a 4-percent reduction in the wage differential.

We then explore the potential sources of the family firm discount, considering two possible explanations. The first is that the discount may reflect compensating differentials: family firms may provide superior non-monetary amenities, such as a less confrontational work environment (Mueller and Philippon, 2011; Kang and Kim, 2020) or may be more likely to uphold "implicit contracts" (Shleifer and Summers, 1988), resulting in greater job stability. Rather than focusing on a specific amenity, we adopt the revealed-preference approach developed by Sorkin (2018), that is, we measure the systematic utility provided to workers by examining their voluntary transitions. We do not find support for this compensating differential hypothesis: family firms appear to provide *lower* utility to workers, even after controlling for the fact that they pay lower wages.

Next, we turn to an alternative potential explanation, which we label the "glass ceiling" hypothesis. Our conjecture is informed by a stark pattern in the data, namely the fact that the family firm discount is larger for higher percentiles of the within-firm earnings distribution. In other words, top earners in family firms earn substantially less than top earners in non-family firms, while the wage gap is much smaller for median wages or for those at the bottom of the within-firm earnings distribution. This suggests that family and non-family firms differ greatly in the career prospects that they can offer to their employees.

Consistent with this hypothesis, family firms have a substantially lower fraction of managers and middle managers, and are less likely to promote their own workers. Not only are their promotions less frequent, but their "returns to promotion" are also lower: the same type of promotion (from white collar to middle manager or from middle manager to manager) leads to lower gains in earnings for workers employed in family firms, controlling for a host of firm and worker characteristics.

Our preferred explanation for this pattern is that family firms feature high "benefits of control": as families do not want to relinquish control to outsiders, they are reluctant to share technical know-how and operational knowledge with other workers. We test one implication of this theory by analyzing firm performance in the wake of CEO deaths. Our hypothesis is that the loss of the top manager should be highly disruptive in family firms, where decision making is extremely concentrated. Conversely, in non-family firms, owing to a higher degree of delegation, more workers should be ready to "step in" and take strategic decisions when the CEO dies. Consistent with our prediction, productivity drops more in family than in non-family firms following the death of the CEO. Moreover, CEO deaths in family firms have significantly worse consequences for workers' wages, earnings, and employment outcomes. Interestingly, we find no effect in family firms whose CEOs do not belong to the controlling family. This dovetails with the idea that in these firms family owners already surrendered control, so that CEOs are more easily replaced upon their demise. In addition, we do not find differential effects on productivity when studying *directors'* deaths. This test is motivated by the possibility that CEO deaths have more negative effects in family firms as they are associated with a a higher "emotional toll." Yet, these events do not result in larger losses in productivity when the departing director belongs to the controlling family, even though the psychological consequences should be analogous to those of a family CEO death.

In the last part of the paper, we rationalize our findings through the lens of a simple model. In our framework, workers can learn new skills on the job by exerting effort. Such effort increases the probability of acquiring skills but is unobservable, so that workers earn information rents. Firm owners promote a fraction of the workers who were successful in acquiring managerial skills. Promoted workers (the "managers") receive higher compensation and produce additional output. We assume that, even though promotions of skilled workers are profit-maximizing, they result in a loss of private benefits of control, which owners trade off against the additional profits they can generate. As family owners place a greater value on the benefits of control, they are less likely to promote their employees, who therefore have weaker incentives to invest in human capital. As a result, family firm employees earn lower average wages, are less likely to receive managerial bonuses, so that their expected lifetime utility is lower that that of non-family firm employees. However, the more attractive career prospects offered by non-family firms lead them to receive more job applications relative to the available vacancies than family firms do. Hence, applicants to jobs in non-family firms face greater probability of unemployment than family firm applicants. This compensates for the greater utility of careers in the former than in the latter, and ensures that in equilibrium applying to either delivers the same expected utility. The model's predictions match many of our empirical results, namely, family firms being characterized by fewer managers and lower wages, returns to promotion, and utility for their employees.

The paper is organized as follows. Section 2 presents a brief review of the relevant literature. Section 3 describes the data. Section 4 presents our key empirical results. Section 7 describes our theoretical model. Section 8 makes some concluding remarks.

2 Literature

A key difference between family and non-family firms is the fact that family owners typically place a great value on retaining control over the firm. As highlighted by Burkart et al. (2003), there are three possible benefits from family control over a company: first, nonpecuniary benefits such as the pleasure from the firm bearing the family name; second, reputational benefits, such as product quality associated with the family or political connections; third, the ability to prevent expropriation by non-family professional managers. The latter may be more skilled at running the company than family members, but may extract resources from it to their own benefit, unless shareholders are adequately protected. As a result, families may delegate control to professional managers only in jurisdictions with sufficiently high shareholder protection, where they may even be willing to sell their controlling stake. These predictions are broadly in line with empirical evidence, as firms in countries with low shareholders protection tend to have more dispersed ownership (LaPorta et al., 1998; Faccio and Lang, 2002).

In principle, private benefits accruing to family owners need not be value-decreasing: the desire to preserve the reputation associated with the family name may motivate a family manager to exert more effort than a non-family manager would; however, dynastic transmission of control may place incapable heirs at the helm of family firms. Anderson and Reeb (2003) provide the first empirical investigation of the relationship between firm performance and family ownership, and find that US family firms belonging to the S&P 500 are smaller but perform better than non-family firms. Sraer and Thesmar (2007) also find family firms to feature higher profitability than non-family ones in French data. Conversely, when studying CEO transitions, Bennedsen et al. (2007) and Pérez-González

(2006) provide causal evidence that family firms tend to exhibit a worse performance in terms of operating profitability and market-to-book ratios after a succession when they appoint a family CEO relative to when they choose an external manager, consistent with wasteful nepotism in family firms. Lippi and Schivardi (2014) provide evidence that family firms tend to select less capable managers, possibly because constrained by the supply of managerial skills within the family. These results dovetail with family ownership being typically associated with *worse* managerial practices: Bloom and Van Reenen (2007), Bloom and Van Reenen (2010), Bandiera et al. (2017), and Bloom et al. (2012) find that family firms tend to be poorly managed. Mullins and Schoar (2016) present survey evidence that family firms, especially if run by founders or family members, have a more hierarchical management. In line with this evidence, which is primarily survey-based, we also find that family firms are characterized by lower productivity and wages. Moreover, our evidence suggests that family firms are characterized by a "glass ceiling" that leads to steep hierarchies and slower career progressions.

While poor management should lead to low productivity and wages, some studies suggest that family firms feature better labor relations and offer greater employment stability to workers, and thus provide valuable nonmonetary amenities to their employees. In implicit contract models along the lines of Baily (1974), Azariadis (1975), and Holmstrom (1983), the profit maximizing strategy for firms consists in absorbing firm shocks and grant risk-averse workers a fixed salary. Under more restrictive conditions, firms may provide workers with insurance against unemployment even in the presence of negative firm shocks. Commitments to honor these implicit contracts may be especially credible when firms are controlled by families, as they have their reputation and prestige at stake (Shleifer and Summers, 1988). Consistent with this hypothesis, family firms appear to pay lower wages and provide more "insurance" to workers vis-à-vis demand shocks (Sraer and Thesmar, 2007; Bassanini et al. (2013); Ellul et al., 2017). Moreover, Mueller and Philippon (2011) show that family ownership is more prevalent in countries in which labor relations are hostile, suggesting that this type of ownership structure may be especially beneficial in these contexts. In addition, Kang and Kim (2020) show that family firms invest more in employee relations than non-family firms. This evidence suggests that family firm employees receive amenities that may compensate for their comparatively low wages. We add to this literature by showing that amenities cannot fully explain the family firm wage discount, as workers on average appear to prefer employment in non-family firms.

This paper also connects with the literature in labor economics that has evaluated the importance of firms for wage inequality (Robinson, 1933; Card et al., 2018). Several papers provide evidence of substantial heterogeneity in how firms might pay the same worker (e.g., Card et al., 2013; Song et al., 2019). This paper adds to this literature by asking *why* that occurs and in particular whether a different ownership structure maps into different firm-wage policies (Bender et al., 2018). In considering this mapping, we quantify how much differences in (i) productivity (Faggio et al., 2010) (ii) rent-sharing practices (Van Reenen, 1996; Card et al., 2015; Kline et al., 2019), and (iii) amenities (Rosen, 1986; Sorkin, 2018) might explain the presence of systematic differences in pay between family and non-family firms.

3 Data

In this section, we illustrate our data sources, explain how we identify family firms, and provide descriptive statistics for the key variables of interest.

3.1 Data Sources

This paper relies on three main data sources. First, we obtain on-site access to social security records of *all* the 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 regarding earnings, type of contract (open ended or temporary), 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*).

Information on firms' ownership and management comes from our second data source, which we refer to as *Cerved-Infocamere*.[2] It is a digitized business registry, comprising three datasets. The first one, *Anagrafica*, includes the universe of Italian limited liability firms for the period 2003–2019, with basic information, such as date of creation, date of cessation (if any), and address. It includes 2.47 million incorporated businesses. The second file, *Soci* has information on firms' owners, which could be either individuals or other firms. For each owner, the dataset includes the name and tax ID, share owned, and address. As the tax ID in Italy is a function of name, city of birth, date of birth, and gender, we can recover this information as well. Coverage in the ownership file is not complete but increases over time, ranging from 41% of the firms in the *Anagrafica* dataset in 2003 to 83% in 2019. In most of our analysis, our time range covers the period 2006–2017.[3] The third file, *Esponenti*, lists the members of companies' board of directors and their top executives, and contains individuals' tax IDs, as well as their addresses and roles, as they appear in the companies' financial statements. Coverage is nearly complete.

The third data source is *Cerved*, a firm-level dataset with accounting information for all the Italian nonfinancial limited liability companies.[4] Importantly, we can easily link all three datasets as both individuals and companies are recorded with their respective tax ID in the three data sources.

[2]Infocamere is the IT company owned by the Italian Chamber of Commerce. We purchased the data through Cerved, a private company that imported and made available to us the dataset in a user-friendly format.

[3]We choose 2006 as the starting year as the coverage of the ownership database becomes fairly high and exceeds 60%. Conversely, we end the sample in 2017 because our administrative data extend up to 2020 and in the event-study evidence presented in Section 6 we use up to three years after the event year (CEO's or director's death or worker's promotion).

[4]The primary data source is once again the business registry managed by Infocamere. All incorporated companies by law must file balance sheets and income statements to the business registry. Data providers, such as Cerved, can then acquire the data, elaborate them and sell them on the market.

3.2 Identifying Family Firms

In most of our analysis, we follow an ownership-based definition of family firm status. For this reason, our first step is to identify the ultimate owners of a firm. This distinction becomes relevant whenever a firm is fully or partially owned by another firm, in which case we track the owning firm's shareholders. We repeat this procedure along the control chain. Each ultimate shareholder is then assigned a share computed using the so-called weakest link principle. Intuitively, 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.[5]

Neither *Cerved-Infocamere* nor *INPS* record family linkages; as a result, we need to make assumptions to identify individuals who are likely to be relatives. First, we assume that, among a firm's shareholders, those with the same last name belong to the same family. Second, we assume that any two shareholders of a firm who differ in gender and last name, but share the same address, are partners.

Throughout the paper, we define a firm as being family owned if we can identify an individual, or a group of individuals that we classify as related, who own more than 50% of the shares of the company. Given that the vast majority of our sample is comprised by small, non-publicly listed firms, we chose a cutoff higher than those employed in most of the literature, which tends to focus on large, publicly 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 the same address or last name may not be actually 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 the former as non-family firm. However, the controlling firm may be actually controlled by a single individual or firm, for which we happen not to have ownership information in our

[5]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.

dataset, for example because it is incorporated in a foreign country. This type of error will tend to underestimate the proportion of family firms.

Relative to the previous literature, which tends to focus on small and selected samples of large firms, a key advantage of our data is that our sample is an order of magnitude larger and covers a significant fraction of the private sector of a developed economy, with substantial benefits in terms of statistical power and external validity of the results. These advantages come at the cost of a higher risk of misclassification, which is unavoidable in our case, as manual data collection for each firm in our sample is unfeasible. Fortunately, we can get a sense of the severity of the misclassification. In 2018, *Cerved* developed an in-house classification of firm ownership for a sample of large and medium-sized firms. They strive to accurately identify family relationships by consulting the real estate registry and linking individuals if they share real estate properties. The overlap of the *Cerved* sample with ours covers 68,367 firms and the two classifications coincide for 93% of the observations. This percentage is likely a lower bound to the true accuracy of our classification algorithm, as this sample over-represents relatively large firms, which tend to have more complex and diversified ownership structures, and where our procedure are more likely to deliver incorrect answers.

3.3 Descriptive statistics

Table 1 presents descriptive statistics for the data used in this paper. Our sample includes about 900,000 unique firms and 4.7 million firm-years for the 2006–2017 period, two thirds of which are family firms. For the worker-level analysis, our starting sample will include over 19 million unique workers and about 91 million worker-years. As family firms are smaller on average (14 employees as opposed to 31.8 for non-family firms), the distribution of workers between the two groups of firms is roughly even.

Even though our sample is much larger than those used in previous studies, the data patterns displayed in Table 1 are broadly in line with the existing literature. Family firms are substantially less productive than non-family firms, as measured by operating value added per worker (45.9 versus 63.3 thousand euros per worker), consistent with Bandiera et al. (2017). However, family firms' profitability, as measured by return-on-assets, is slightly higher (4.42 versus 4.37 percentage points). Although the gap is fairly small, this finding is qualitatively in line with Anderson and Reeb (2003) and Sraer and Thesmar (2007). Leverage is also similar between the two groups of firms.[6]

How can family firms roughly match non-family firms' profitability despite their lower productivity? At least part of the answer lies in their ability to pay lower wages. Weekly wages are 64 euros lower in family firms, corresponding to a 14 log-point gap, a "family firm discount" that we decompose in detail in Section 4.2. The gap in the entry wage is of comparable magnitude, suggesting that the family firm discount emerges immediately in a worker's career. But the gap tends to widen over time, as family firms also display a slightly lower wage growth rate.

Interestingly, family firms are also characterized by a lower degree of within-firm wage inequality, which is apparent by looking either at the log-difference between the highest and the median wage, or at the log-difference between wages at the 90^{th} and 10^{th} percentiles. The gaps are 82 and 72 log-points, respectively, as opposed to 1.06 and 82 log-points for non-family firms. Part of this gap is due to organizational differences between the two types of firms: non-family firms have two and a half more workers in managerial positions (1% versus 0.4%) and four times more workers in middle management positions (4% versus 1%). Age, experience and gender do not seem to play a role in explaining the gap, as the two types of firms do not meaningfully differ in these dimensions. Family firms have fewer workers with full-time or open-ended contracts but, again, the difference is fairly small.

[6]As shareholders of a family firm tend to have a high portion of their wealth invested in the firm, they are generally thought as leaning towards more conservative financial choices, and so to less leverage. However, a reluctance to raise equity to prevent a dilution of their stake can lead to a *higher* leverage (and a lower size). In our data, these two channels appear to approximately offset each other.

4 Econometric Framework

This section presents our methodology. Section 4.1 introduces the econometric design used to isolate differences in wage policies between family and non-family firms. Section 4.2 presents a method to decompose differences in wages paid by family and non-family firms so as to take into account the role of worker and firm-level differences. Section 4.3 introduces a test of whether family firms pay lower wages because they offer better amenities.

4.1 Workplace Heterogeneity Across Family and Non-Family Firms

There are two fundamental challenges in interpreting the effect of family firm ownership on the wage structure. The first is that differences between the average wage paid by family and non-family firms can be driven by sorting, e.g., workers with higher skills may be systematically more likely to work for non-family firms. The second challenge is that, even after accounting for workers' selection, differences in wages between family and non-family firms may be driven by differences in productivity (Bloom and Van Reenen, 2007; Sraer and Thesmar, 2007; Card et al., 2018). A key goal of our analysis is assessing whether there are systematic differences in pay between a family and a non-family firm even when these two firms are equally productive.

We address both challenges by studying how a worker's wage changes following a job move. In particular, we estimate the two-way fixed effects specification pioneered by Abowd et al. (1999) (henceforth AKM):

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

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 (dominant) employer of worker i in period t.[7] The vector X_{it} includes a cubic in age and years of labor market experience, as well as year fixed effects; α_i represent the portable component of worker i's wage, i.e., the portion that the worker always receives irrespective of the employer's identity. Finally, ψ_j is the systematic pay premium (or discount) associated with working for firm j after controlling for worker selection. The average (person-year weighted) wage effect of working for a family firm is therefore calculated as follows:

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

where the function $f(\cdot) : (1, \ldots, J) \to (F, NF)$ determines whether j is a family (F) or non-family (NF) firm. Under the identification assumption described below, $\Delta_{\psi,F}$ measures the average wage impact ofrandom switches of workers from non-family to family firms.

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 across 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 occurs 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 several countries appears inconsistent with endogenous mobility.[8] In the Italian context, using the same data available to us, Casarico and Lattanzio (2023) 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.

[7]Following the literature, we define as the dominant employer the firm that paid the worker the most in a given year. 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.

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

Moreover, the wage gains from moving from a lower paying employer to a high-paying one are roughly symmetric to the wage losses when moving from a high-paying employer to a low-paying one. 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 does not represent a first-order concern in Italian data.

Importance of not aggregating moves: Note that we do not estimate the wage effect of family firms via a collapsed version of equation (1) where firm effects are replaced by a dummy for whether the current employer is a family firm or not. As noticed by Card et al. (2023) when studying industry-wage differentials, moves between family and non-family firms might occur within a selected sample of firms (e.g., movers tend to leave highly productive family firms for low-productivity non-family firms). If so, estimates based on a specification with worker fixed effects and a family firm dummy might provide 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 across family firms or across non-family firms—our estimate of the family firm wage effect will not suffer from this bias. Moreover, this approach enables us to assess how much differences in pay between family and non-family firms stem from differences in productivity, as described below.

4.2 Oaxaca Decomposition

If wages are determined by equation (1), then the average gap in 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". It measures how much of the wage gap is due to differences in the unobserved abilities of workers hired by the two types of firms. 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 linked with financial records and ownership structure of firms permits gauging the role of productivity differences in explaining $\Delta_{\psi,F}$. Consider a linear projection of the firm effects, ψ_j, on productivity:

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

where $\theta_{f(j)}$ is the firm wage differential unrelated to productivity, 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 selection of workers (Guiso et al., 2005; Card et al., 2018). Importantly, we allow this coefficient to differ between family and non-family firms, in light of recent evidence that the effect of sales shocks on wages differs between these two types of firms (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 accounted for via an 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 average wage gap between family and non-family firms is driven by differences in productivity. The second component captures how much of it reflects differences in the pass-through coefficient between family and non-family firms.[9] The last term captures systematic wage differences between family and non-family firms that would emerge even if they had the same productivity and pass-through coefficient.

4.3 Compensating Differentials

A standard theory to explain systematic pay differences between firms is that they offer different amenities to workers (Rosen, 1986). Assessing the role of compensating differentials in explaining wage gaps across firms 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), the role of compensating differentials is often assessed via a hedonic approach that typically amounts to a cross-sectional regression of pay on non-pay characteristics. This approach assumes that labor markets are perfectly competitive and thus neglects the possibility that labor market frictions may generate dispersion in utilities across workers. In a frictional labor market, family firms may offer lower pay while also providing *lower* amenities to workers (Mortensen, 2003).[10]

[9]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.

[10]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 5 and with the literature

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

The common value of working for firm j is calculated as follows. Assume that the utility that individual i obtains 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 typically offered by firm j to its employees, i.e., v_j can be identified from the following recursive equation:

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

where $\omega_{j,\ell}$ is the number of workers who voluntarily move from employer ℓ to employer j, scaled by the number of workers who voluntarily leave employer j. Intuitively, "good" employers are unlikely to see workers leaving them and are likely to attract workers from other good employers (just like 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 INPS data provides a reason behind a job separation. A voluntary separation is a job transition 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

surveyed in Section 2.

projections of the page-rank 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 differential motives, 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 whether family firms offer different utility to their employees via their non-pay characteristics.

5 The Family Firm Wage Discount

In this section we describe the main results. Section 5.1 shows results from the decomposition of the raw gap in wages between family and non-family firms presented in equation (5). Section 5.2 assesses whether differences in pay between family firms and non-family firms are driven by differences in the amenities offered by these two types of firms.

5.1 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 (1) to the largest connected set of firms in our data (Card et al., 2013). The raw gap in wages between family and non-family firms is around 14 log points, in line with results from Sraer and Thesmar (2007) and Bassanini et al. (2013) for France and the cross-country evidence of Ellul et al. (2017). After 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 workers with a higher portable component of wages being systematically more likely to sort into non-family firms. This underlines the importance of controlling for workers' composition when trying to estimate the family firm wage discount. The remaining half of the wage gap is explained by systematic differences in the wage policies between family and non-family firms.

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

Figure 1 displays a binscatter plot where for each centile of value added per worker (calculated using both family and non-family firms) we overlay on the $x-$axis the corresponding average 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 to wages is roughly linear, and with a slope not too dissimilar between the two types of firms, it appears that non-family firm wage effects are invariably above the firm-wage effects of family firms by almost a constant amount, except that the gap tends to shrink in the left-tail of the productivity distribution and widen in the right-tail.

To isolate the role of differences in productivity in driving differences in average firm effects between family and non-family firms, we fit a linear specification highlighted in equation (4) to the relationship depicted in Figure 1 and then apply the decomposition of equation (5) (see 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 somewhat smaller (0.13) and statistically different from that found for family firms.

Results from the Oaxaca decomposition of (5) show that the bargaining component would close the family-firm wage gap by around 4 log points, since family firms have a higher rent-sharing coefficient. Differences in productivity account for about 2 log points and thus explain around 25% of the average difference in firm effects found between family and non-family firms. In line with the visual evidence of Figure 1, there is a large component, unexplained by differences in productivity or bargaining power, that drives differences in pay between family and non-family firms. In a counterfactual in which we move each family firm worker to a non-family firm with the same productivity and with the same rent-sharing splitting rule of family firms, their wage would increase on average by around 10 log points according to equation (4).[11]

5.2 The Role of Compensating Differentials

The systematic wage gap between family and non-family firms may be compensated by greater amenities offered by the former. Indeed, previous work notes that family firms might offer greater employment stability as family ownership lowers the incentives of employers to break the implicit contract designed to shield employees from negative shocks (Sraer and Thesmar, 2007; Ellul et al., 2017). The existence of a job security premium among family firms can rationalize why these firms tend to pay less on average. Others argue that family invest more in the relationship with their employees (Mueller and Philippon, 2011; Kang and Kim, 2020), which may also tend to compensate for lower pay.

To assess the general role of compensating differentials in driving the family firm wage discount, we implement the revealed preference approach of Sorkin (2018) described in Section 4.3. That is, we estimate the ranking of employers based on voluntary job moves made by Italian workers. In deriving this ranking, we do *not assume* that the utility from employment has to be equalized across jobs, as instead typically assumed in hedonic-style approaches. However, if compensating differentials account for the existence of a family-firm wage discount, then we should *find* that, on average, the utility of working for a family firm is similar to the utility of working for a non-family firm.

[11] Table 3 shows that the gap in firm effects between family and non-family firms continues to persist after accounting for the possibility that family firms might be concentrated in low-wage sectors or geographical locations. Similarly, these gaps are also present when splitting the analysis of Table 3 by firm size, as shown in Table A1.

Figure 2 displays the common valuation of firms obtained from the recursive formulation of equation (7) by overlaying, for each centile of value added per worker, the Pagerank of family and non-family firms. The pattern of utility that emerges from Figure 2 is very similar to the one displayed in Figure 1. More productive firms tend to be more desirable firms. Moreover, non-family firms tend to be more desirable firms than family ones, even when these two firms have a similar underlying productivity. According to the voluntary job transitions made by Italian workers, therefore, it appears that family firms have a lower Pagerank compared to non-family firms, i.e., they offer *lower* utility from employment compared to family firms.

This is further confirmed by the evidence provided in Table 4, which reports estimates from equation (8). Moving a worker from a family firm to a non-family firm would increase average utility by a margin equal to 48% of the overall standard deviation of the Pagerank index that we find in our data. This large gap in utility is explained partially by the fact that family firms offer lower wages, as measured by the firm effect (Column 2), and tend to be located in sectors and locations where utility from work tends to be lower (Columns 3 and 4). However, as shown in Column 5, even after controlling flexibly for province-by-sector indicators as well as firm effects, it remains true that family firms have a lower ranking than non-family firms. It is therefore hard to reconcile this evidence with a compensating differential interpretation where family firms pay lower wages because they tend to offer higher amenities to workers.

We consider a few robustness checks for this result. First, we re-estimate equation (8) using a version of the Pagerank index that accounts for the possibility that part of the voluntary moves that we observe are driven by family firms differing in terms of offer intensity as well as firm size, along the lines suggested by Sorkin (2018). As shown by (Table A5 in the Appendix, our qualitative conclusions remain unchanged when using this augmented Pagerank index.

Second, we zoom in on potential amenities that can be measured in our data and explore whether the provision of these amenities is somewhat different across types of firms. Following the implicit insurance argument discussed above, we calculate the churn rate and see whether the latter differs between family and non-family firms. Results are displayed in Appendix-Figure A3. Differently from what we see when looking at firm-wage effects or Pageranks, there appears to be little difference in churn rates among similarly productive family and non-family firms. This contrasts with the idea of family firms systematically offering higher employment stability to their workers.

Another margin that we check is the presence of worker-friendly policies. INPS data allow us to measure the number of days off (n_{it}) that a worker can obtain for sickness (while being fully compensated) 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 the one used in Lachowska et al. (2022) to isolate the importance of firms in unemployment insurance take-ups or in Bana et al. (2018) to isolate the importance of firms in maternity leaves take-ups. Figure A4 shows the distribution of this firm-level effect for centiles of value added per worker. The evidence does not suggest that family firms are systematically associated with higher propensity to offer days off when a worker is sick.[12]

All in all, the evidence provided here is at odds with the idea of family firms paying lower wages because they offer better amenities. When constructing a ranking of employers based on voluntary transitions made by workers, we find that family firms offer lower utility than non-family firms, even after controlling for differences in pay across these two types of firms. Therefore, on top of lower wages, family firms also seem to offer lower amenities. In other words, family firms are systematically worse employers that

[12] Appendix-Figure A1 addresses the possibility that the family firm discount is accounted for by workers' geographical preferences. For example, working in a non-family firm may entail a higher compensation but require a longer commuting time, as non-family firms are fewer than family firms. We regress the firm wage fixed effect on the logarithm of productivity and 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.

amplify inequalities not only in pay but also in utility. This suggests that these firms are able to hire and retain workers because frictions (such as asymmetric information, search frictions, congestion, horizontal differentiation) prevent workers from leaving these firms to join employers offering higher levels of utility. The next sections will present theoretical rationale and empirical evidence for an economic mechanism that can explain why family firms might be able to operate while offering lower pay and utility.

6 The Glass Ceiling of Family Firms

Having established that, even after controlling for sorting, family firms consistently pay lower salaries, it is worth asking whether *all* workers are equally subject to a family firm discount, or there is heterogeneity in the family wage discount across them. In particular, is the discount uniform across the wage distribution or does it mostly affect a specific group of workers, such as low-pay or high-pay employees? Evidence that the discount is larger at the top of the wage distribution would suggest that family firms offer worse career prospects, as top wages are typically achieved upon promotion.

In Section 6.1, we present three pieces of evidence that, taken together, strongly suggest that workers in family firms are subject to a "glass ceiling", leading to a compression of the right tail of the wage distribution. First, the wage gap between family and non-family firms is mainly accounted by a gap in top wages; second, family firms feature significantly less frequent promotions. Thirdly, their promotions are associated with lower wage increases than those of non-family firm employees. In Section 6.2, we show that the loss of top managers is more harmful to family-firms (and especially family-managed firms) than to non-family ones, consistently with their lower propensity to promote non-family members and delegate managerial power to them.

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

We start by plotting, in Figure 3, the within-firm log-wage distribution by log-value added per worker for different wage percentiles. We show the 5th, 10th, 25th, 75th, 90th, and 95th percentiles of the within-firm distribution of the logarithm of wages in Panels A, B, C, D, E, and F, respectively. The log-wage percentiles for family and non-family firms are fairly close for the 5th and 10th percentile, and a clearly visible gap starts emerging for the 25th percentile. The gap becomes more apparent, and economically large, once we focus on the 75th, 90th, and 95th percentiles, especially for high values of log-value added per worker, hence for the most productive firms. Although this evidence should be taken as merely descriptive, as it does not account for sorting, it shows that the gap we observe in average wages can be explained primarily by the differences in compensation of workers at the high end of the within-firm wage distribution.

Next, to further investigate this point, we look at whether and how family and non-family firms differ in their promotion practices. We consider promotions from white collar to middle managers (*quadri*), and from middle managers to top managers (*dirigenti*). We consider all the workers employed with the same firm for two consecutive years.

For each firm and year, Panels A and B of Figure 4 display the fraction of promoted workers, either from white collar to middle management, or from middle to top management, 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 family and non-family firms. For more productive firms, however, they become more frequent. Moreover, a stark gap between promotion rates appears for more productive firms: zooming on the top decile of the productivity distribution, the propensity to promote in non-family firms is two-to-three times higher than in family firms. For example, in the highest productivity percentile, the fraction of workers promoted every year is close to 0.6% in non-family firms and 0.2% in family firms.

Results are qualitatively similar in Panel B, where we plot the average promotion rate from middle to top management. The average promotion rates are, not surprisingly, lower, but the difference in the promotion rates is even higher: in the top percentile of the productivity distribution, the average promotion rates for family and non-family firms are 0.02% and 0.1%, respectively.

While non-family firms promote more intensively to managerial positions, this might be due to a higher turnover of workers in such positions, due to firings or quits. In this case, it would be questionable to interpret the higher promotion rate as an indication of better career opportunities. Panels C and D of Figure 4 plot the fraction of middle and top managers over the total workforce. They mirror those of promotions: at higher levels of productivity, non-family firms have a larger share of workers in managerial positions. Hence, family firms appear to provide significantly lower career opportunities to their employees.

In principle, family firms might use promotions more selectively but reward promoted employees more generously. From the employees' perspective, the greater boost in compensation could compensate the lower chance of promotion. Hence, in our third piece of evidence we compare the returns to promotion of workers employed in family and non-family firms via a matching approach. Intuitively, we select workers who have similar characteristics, including the type of promotion, but who differ in the type of employer. Every worker promoted in a family firm is paired with a "twin" worker in a non-family firm, matching on the logarithm of firm size, the logarithm of value added per worker, wage growth, and age, as well as (exactly) on promotion type, industry (using the 2-digit NACE classification), contract type, full-time status, and gender.[13] Upon imposing a two-year tenure restriction and following workers over a $(-3, +3)$-year window, we are left with 29,632 promotions from white collar to middle management and 1,416 promotions from

[13]We implement a caliper matching algorithm without replacement. For the logarithm of firm size, the logarithm of value added per worker, wage growth, and age (the variables for which the match is approximate) the caliper widths are 1, 5, 0.5, and 0.05, respectively.

middle management to manager. We then estimate the following model:

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

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, that is, either the logarithm of the wage, a dummy equal to one if the worker is employed, or raw earnings. FF is a family firm dummy, γ and δ are vectors of worker and year fixed effects, respectively. We are interested in the estimated β_k coefficients, which track the evolution of the dependent variable prior and after the promotion event. As β_{-1} is normalized to 0, our baseline year is the year prior to the event.

The results are plotted in Figure 5. Panel A shows that promotions to middle management result in a 1.1% lower increase in wages in family firms three year after the event: as to the level of earnings, workers promoted in family firms earn 600 euros less than their matched counterparts three years after the event. The results for employment are less precise and the estimated coefficients are relatively small, although they point out to a lower retention rate for workers promoted in family firms.

Panel B, which focuses on promotions from middle to top management, reveals qualitatively similar results, although with lower precision, due to the smaller sample. Workers promoted in a family firm earn 2% less three year after the promotion than those promoted in non-family firms, although this effect is only marginally statistically significant. Although there is no effect on the employment rate, the difference in the earnings trajectory is economically large: three years after event, a €4,000 difference in the earnings emerges, relative to the baseline year.

6.2 Evidence from CEO Deaths

A potential explanation for the evidence presented in the previous sections is that members of the controlling family firms wish to retain most, if not all, managerial positions in the firms that they control, and thus retain 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-family members.

This narrative can explain why family firms tend to promote fewer workers to managerial positions, as well as reward them less upon promotion. It is also consistent with the family wage discount being especially pronounced for top earners: it likely reflects family firms' unwillingness to compete keenly to attract or retain non-family employees to decision-making roles at the top of the firm's hierarchy.

An implication of this hypothesis is that losing the top decision maker should have different effects depending on whether the firm hit by this loss is a family firm or not. In a family firm, most of the technical and organizational knowledge should be concentrated in a few members of the controlling family; as a result, losing the most important member of the management team – someone who is difficult to replace within the family – is likely to be very disruptive. Conversely, in non-family firms, the ability to take key managerial decisions is shared by several incumbents or, failing that, similar managers can be hired externally. As a result, in these firms even losing the top manager should be less detrimental to performance, as there is a larger pool of workers who should be able to "step in" and effectively replace the deceased CEO.

To test the implication of this hypothesis, we make use of the *Esponenti* file in our data, which lists the top executives and the members of the board of directors 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 chairman of the board (*presidente*). For brevity, in what follows both will be referred to as CEOs. INPS records the death year 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 death event.[14]

[14]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.

In the sample of firms subject to CEO death event, we then distinguish between family and non-family firms, defined as treatment and control firms, respectively. For every treated firm, we select a control firm in the same industry (defined using the 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.[15] This matching procedure leaves us with a sample of 1,268 events, half of which involve the death of a family CEO. We then estimate the following regression to investigate whether the loss of the firm's CEO affects the productivity of family firms differently from that of non-family firms:

$$\log(Productivity)_{it} = \sum_{k=-3}^{3} \alpha_k D_{it}^k + \sum_{k=-3}^{3} \beta_k D_{it}^k \times FF_i + \gamma_i + \delta_t + \varepsilon_{it}, \qquad (10)$$

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

The β coefficients obtained upon estimating equation 10, together with 95% confidence intervals, are displayed in Figure 6, Panel A.i. (Appendix-Table A3 lists coefficients and associated standard errors.) After the CEO deaths, productivity falls more in family firms than in non-family firms. While the effect is initially limited, two and three years after the event family firms experience productivity losses 7.6 and 5.7 percent higher than their non-family counterparts, respectively, although the coefficient β_3 is not very precisely estimated.

Our detailed board and shareholding data allow us to further refine the analysis. So far, family firms have been identified solely based on firm ownership, ignoring management

[15] 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.

characteristics. However, some family firms may not be "family managed," having hired an external manager to run their operations. In these firms the controlling family can be expected to be more inclined to delegate key decisions than in family-managed ones. For example, an incumbent managers can normally be replaced with other non-family managers, without significant losses in the firm's technical know-how. Indeed, Mullins and Schoar (2016) provide survey evidence that professional CEOs of family firms (i.e., managers not related to the controlling family) behave very similarly to managers of non-family firms in terms of propensity to delegate decision-making. Specifically, they have a relatively high number of managers reporting directly to them, as opposed to family CEOs, and especially founder CEOs. In contrast, losses of technical know-how can be expected in "family managed" firms, where the owner-manager is likely harder to replace with an equally able family member. Hence, one can expect the differential effect of CEO deaths on productivity to be driven by family-managed firms.

Panel A.ii displays event-study evidence that only refers to family-managed firms, together with the corresponding matched control firms. The evidence is qualitatively similar to that of Panel A.i, but the point estimates of β_2 and β_3 are larger in magnitude. Conversely, Panel A.iii reveals no detectable differential effect of CEO deaths in non-family-managed firms. For completeness, Appendix-Figure A2 in the Appendix displays the average productivity in the seven years surrounding the CEO death event. Both groups of firms experience substantial drops in productivity in the year of the CEO's death. However, these effects are much more persistent for family firms: even three years after the event, their productivity does not recover the level featured in the year prior to the CEO death. Conversely, non-family firms recover relatively quickly and, three years after the event, have a productivity level that is *higher* than that displayed prior to the CEO death.[16]

[16]This improvement in productivity could stem from the new management's ability to eliminate or reduce agency problems created by the previous management: for instance, the incoming CEO may divest low-quality assets acquired by the previous CEO (Pan et al., 2016) or breaking entrenchment created by them (Taylor, 2010).

A potential explanation for this finding is that the drop in productivity is not driven by a loss in know-how but rather by the "emotional" impact that the loss of the owner-manager of a company can have on its workers. This explanation would be in line with the fact that results are driven by family-managed firms, where the departing CEO is a member of the family controlling the company.

To address this concern, 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 that of the CEO. The intuition for this test is simple. The death of a family member should be an emotionally disruptive event independent of her role; however, we would expect a substantial loss in technical and operational know-how only for the loss of a family member with an executive role, namely the CEO. This is precisely what we find in Panel B: there is no detectable evidence of a differential productivity response for director death events, whether we focus on all firms (Panel B.i), on family-managed firms (Panel B.ii), or on non-family-managed firms (Panel B.ii).

In Figure 7, we adopt the same design but now examine outcomes at the worker level. (Appendix-Table A4 reports coefficients and standard errors in tabular form.) Again, we adopt a matching approach, pairing every worker in a family firm with a worker in a non-family firm, matching on the same variables used when looking at promotions. Importantly, this matching is performed within pairs of firms matched in the firm-level analysis. The final sample includes 8,306 workers, 5,802 of which are employed either in family-managed firms or in the corresponding matched firms. For brevity, and given the evidence on firm outcomes, we focus on the latter subsample.

Panel A shows that employees of family-managed firms experience larger wage losses after the CEO death, relative to employees of non-family firms. There is little effect in the short run but, three years after the event, their wages are 3.4 percent lower. Their employment probability is also lower (Panel B), but the effect emerges with the opposite timing: a large short-run effect, namely, a 2 percentage point lower employment probability for family-firm employees one year after the event, but no significant difference three years after the CEO death.

Earnings of family firm workers are driven down by the lower employment probability in the short run, and by the lower wages in the long run. As a result, the negative effect of a family firm CEO death on the level of labor earnings is immediate and fairly stable, with coefficients implying losses between 650 and 950 euros (Panel C).

Finally, one can expect that, after the death of the CEO, firms may promote other employees to higher positions in the corporate ladder, as they may be asked to perform more managerial tasks to make up for the CEO's loss. However, these career advancement opportunities may differ in family and non-family firms. This is indeed the case, as shown in Panel D, where the dependent variable is a dummy equal to 1 if a worker is a manager or middle manager after the death of the CEO. Family firms remain disproportionately more averse to promote their employees to managerial positions: 3 year after the CEO death, the probability of being either a manager or a middle manager is 0.65 percentage points lower for workers of family firms. Hence, the larger drop in productivity suffered by family firms following CEO deaths may be due, at least partly, to their reluctance to delegate strategically important decisions to new managers, consistent with our glass ceiling hypothesis.

7 A Model of Career Paths in Family Firms

The evidence presented so far can be broadly summarized as follows. First, careers in family firms are slower than in non-family ones, the frequency of promotions being significantly smaller in the former. Second, family firms pay lower average compensation to their employees, at each stage of their careers, even after accounting for sorting of workers into firms with different characteristics. Third, workers appear to prefer working in non-family firms, implying that family firms do not offer greater amenities to make up for the lower attractiveness of their careers. In what follows, we provide a simple theoretical framework to understand these empirical regularities.

In the model, firm owners determine their promotion policies by balancing the efficiency gains from promoting skilled workers against the personal advantages of retaining control over the firm. Specifically, in family-owned firms the owners' preference for retaining decision-making power is assumed to be greater than in non-family firms, consistent with the evidence in Lippi and Schivardi (2014). Consequently, they are less likely to promote their employees. Workers exert unobservable effort to accumulate human capital and develop productive skills. Only workers who are successful at developing skills are considered for promotion to managerial positions. The lower career prospects offered by family firms induce weaker incentives to invest in human capital, leading workers to accumulate less skills and earn commensurately lower wages and managerial bonuses. Given the better economic conditions that non-family firms offer, in equilibrium they receive more job applications relative to the available vacancies, increasing the probability of applicants being rejected and remaining unemployed. This higher probability of unemployment balances the greater expected utility of being employed in a non-family firm: in equilibrium, applying for a job to family and non-family firms delivers the same expected utility.

7.1 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 costly effort. Effort is unobservable and its cost is $c_w e_w^2/2$, but skills can be observed once they are acquired. 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. Promoted workers 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 $c_m e_m^2/2$.[17] Firm j chooses its promotion rate ϕ_j by trading off its expected profits with 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 larger for family than for non-family firms ($\beta_F > \beta_{NF}$) 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 8, comprises four stages:

- **Matching and hiring**. At $t = 0$, each worker decides if to apply for a job in a randomly chosen family or a non-family firm, taking into account the expected lifetime income respectively offered by them. Firms hire workers by randomly picking them from their applicant pool, with unsuccessful applicants remaining unemployed. Upon hiring workers, each firm j offers them 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 $t = 2$.
- **Workers' 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.

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

- **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 to offering a certain bonus contract to promoted workers at $t = 2$.

7.2 Main Results

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

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

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 are paid their reservation wage, which we normalize to zero with no loss of generality. The efficiency wage (11) is not only increasing in the firm 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 raise the surplus that skilled workers are able to generate. Upon substituting for the optimal promotion probability, one sees that non-family firms, which feature a lower β_j, pay higher wages than family firms to their best-paid employees: in setting them, they take into account that skill acquisition by their employees generates a larger expected surplus, due to their greater probability of promotion (as we show next).

Moreover, non-family firm employees also feature a greater probability of acquiring productive skills than family firm ones, 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 non-family and family firm employees stems both from the difference in the efficiency wages paid to their respective best-paid workers and also from the the difference in their chances of achieving high-pay positions in their respective within-firm wage distribution. Moreover, on both accounts the wage gap between non-family and family firms widens for 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 Tables 2 and 3 and in Figure 1, which show that, even after controlling for workers' sorting, family firms pay lower wages, the more so for high productivity levels. They are also consistent with the findings in Figure 3, which documents that the wage gap is absent for the low percentiles of the within-firm wage distribution, who, in the model, represent the workers who do not develop skills, and therefore earn the reservation wage in both family and non-family firms.

Result 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 family firms, which feature higher 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 non-family and family firms is greater for more productive firms. This is in line with the evidence of Figure 4, which characterizes the promotion policies of family and non-family firms for different classes of value added per worker, and show that family firms have lower promotion rates.[18]

Result 3. *Promotions in family firms lead to a lower increase in compensation.*

This result holds under a mild regularity assumption on the moments of the distribution of y_j that is formally stated in the appendix, and is satisfied by most common distributions with positive support, namely, uniform, beta, gamma and log-normal.[19] First, notice that the expected bonus b for a worker who is promoted, conditional on y_j, is equal for family and non-family firms, and given by:

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

However, *unconditionally*, the expected bonus paid by non-family firms is higher. This result, formally proven in Appendix A.2, has a simple intuition. Recall from Result 1 that non-family firms are more likely to promote workers than family firms, and that the gap in the promotion rate increases with productivity y_j. Hence, workers promoted in non-family

[18] The figure plots the fraction of promotions against the value added the unobserved y_j but only against the observed value added per worker, and we do so in Figure 4. 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 4, making the difference between promotion rates in family and non-family firms even larger.

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

firms 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 non-family firms. This prediction is consistent with the event study evidence in Figure 5, which shows that the return to promotion (i.e., the bonus paid to promoted workers) is significantly lower in family firms.

Result 4. *Family firm employees have lower expected utility than non-family firm ones, but job applicants in family firms have 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 family and non-family firms, the average $\mathbb{E}(U_{1j}^*)$, computed over y_j, is larger for non-family firms than for family ones: by offering more rewarding careers, on average the former foster greater human capital accumulation and offer higher lifetime expected utility than the latter. 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 set of non-family and family firms.[20] This prediction is consistent with our evidence in Figure 2, based on the methodology proposed by Sorkin (2018), which shows that workers attach lower expected utility to being employed in a family firm.

The model helps understand that this fact can be reconciled with labor market equilibrium: even though family firm employees enjoy on average lower utility than their non-

[20]Since our evidence indicates that family firms feature lower productivity than non-family ones (see Table 1), it is worth pointing out that this result generalizes to the case where non-family firms are more productive than family ones, in the sense that the productivity of the former first-order stochastically dominates that of the latter, 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 non-family firms were more productive than family ones.

family counterparts, at the job application stage the expected utility of the two groups of workers can be the same. This is because those who apply to jobs in non-family firms run a greater risk of unemployment, as such jobs are more eagerly sought for and thus have a greater number of applicants than jobs in family firms. The reason why non-family firms do not cut their wages is that they are optimally set to elicit effort from employees.

8 Conclusion

This paper asks how and why the compensation and promotion policies of family and non-family firms differ, and what do these differences imply for their performance. We confirm, on a much larger sample of Italian firms, the finding of the existing literature that there is a meaningful gap in average wages, with family firms paying substantially lower wages. However, our comprehensive employer-employee matched data enable us to explore the possible reasons for this gap. We find that about half of the gap, corresponding to 8 log-points, is due to workers' time invariant characteristics, but the remaining difference cannot be explained by sorting. Accounting for differences in rent sharing, productivity, industry, and location does not fully eliminate the gap.

Moreover, the gap is more pronounced in high-productivity firms and is mainly driven by differences in compensation for relatively high-pay workers, i.e., those in the highest percentiles of the within-firm earnings distribution. In addition, family firms feature fewer top and middle management positions, fewer promotions and lower returns to promotions, controlling for observable worker and firm characteristics. 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 other employees. To test this hypothesis, we conduct event studies around CEO deaths. It turns out that productivity, workers' earnings and employment are more negatively affected by CEO deaths in family firms, and especially in family-managed ones, than in non-family firms, matching for a host of observable characteristics. This finding corroborates our glass ceiling hypothesis: if in family firms the organizational and technical know-how is not shared with other employees, part of it is irremediably lost absent the top decision maker.

We also investigate an alternative "compensating differential" hypothesis, namely the possibility that the lower wage paid by family firms may be compensated by superior non-monetary amenities, such as a less confrontational work environment (Mueller and Philippon, 2011) or greater job security (Ellul et al., 2017). We measure the systematic utility associated with family and non-family firms following Sorkin (2018)'s revealed-preference approach, but find that family firms appear to be inferior employers. Moreover, we do not detect meaningful differences in observable amenities, such as the type of contract offered to workers.

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 shareholders of the former suffer a "disutility" from promoting workers, consistent with the presence of higher benefits from control in family firms. Despite its simplicity, our model rationalizes most of our empirical findings, namely that family firms display lower wages, have fewer managers, deliver lower wage gains upon promotions, and lead to lower expected utility for their workers. The model also suggests that, while offering lower wages and slower careers, family firms may still be able to compete for job applicants with non-family firms, by offering a greater likelihood of getting a job upon applying for it.

References

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

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.

Azariadis, Costas, 1975, "Implicit Contracts and Underemployment Equilibria", *Journal of Political Economy* 83, 1183–1202.

Baily, Martin Neil, 1974, "Wages and Employment under Uncertain Demand", *Review of Economic Studies* 41, 37–50.

Bana, Sarah, Kelly Bedard, and Maya Rossin-Slater, 2018, "The impacts of paid family leave benefits: Regression kink evidence from california administrative data", Working Paper 24438, National Bureau of Economic Research.

Bandiera, Oriana, Renata Lemos, Andrea Prat, and Raffaella Sadun, 2017, "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.

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.

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

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, 2023, “Industry wage differentials: A firm-based approach”, Technical report, National Bureau of Economic Research.

Casarico, Alessandra, and Salvatore Lattanzio, 2023, “What Firms Do: Gender Inequality in Linked Employer-Employee Data”, *Journal of Labor Economics* Forthcoming.

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.

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.

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, CESifo.

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.

Faggio, Giulia, Kjell G Salvanes, and John Van Reenen, 2010, "The Evolution of Inequality in Productivity and Wages: Panel Data Evidence", *Industrial and Corporate Change* 19, 1919–1951.

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.

Holmstrom, Bengt, 1983, "Equilibrium Long-Term Labor Contracts", *The Quarterly Journal of Economics* 98, 23–54.

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

Kline, Patrick, Neviana Petkova, Heidi Williams, and Owen Zidar, 2019, "Who Profits from Patents? Rent-sharing at Innovative Firms", *Quarterly Journal of Economics* 134, 1343–1404.

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", Technical report, 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", Working Paper 25954, National Bureau of Economic Research.

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

Lavetti, Kurt, and Ian M Schmutte, 2016, "Estimating compensating wage differentials with endogenous job mobility" .

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

Mortensen, Dale T., 2003, *Wage dispersion: why are similar workers paid differently?* (MIT press, Boston, MA).

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.

Pan, Yihui, Tracy Yue Wang, and Michael S. Weisbach, 2016, "CEO Investment Cycles", *The Review of Financial Studies* 29, 2955–2999.

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

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, 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.

Taylor, Lucian A, 2010, "Why Are CEOs Rarely Fired? Evidence from Structural Estimation", *The Journal of Finance* 65, 2051–2087.

Van Reenen, John, 1996, "The Creation and Capture of Rents: Wages and Innovation in a Panel of UK Companies", *Quarterly Journal of Economics* 111, 195–226.

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-Level Component of Wages and Value Added per Worker

Figure 1 presents the average "AKM" firm fixed effect ($y-$axis) visà-vis the percentile of the logarithm of value added per worker ($x-$axis) . We regress the logarithm of worker's wage on year fixed effects, cubic polynomials of age and years of experience, worker fixed effects, and firm fixed effects. We then compute the average firm fixed effect for each percentile of the logarithm of value added per worker, separately for family and non-family firms. Red and blue squares refer to family and non-family firms, respectively.

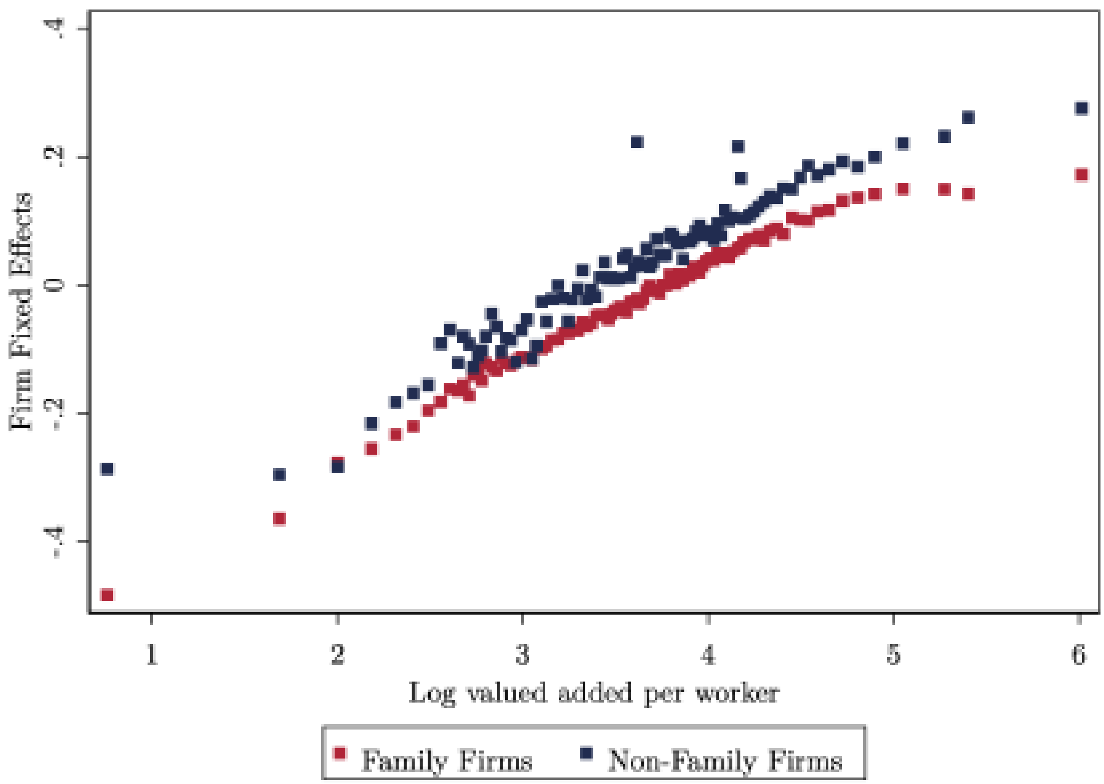


Figure 2
Worker's Utility and Labor Productivity

Figure 2 shows the average worker's typical utility from being employed in the firm visà-vis the percentile of the logarithm of value added per worker, computed using Sorkin (2018)'s approach.

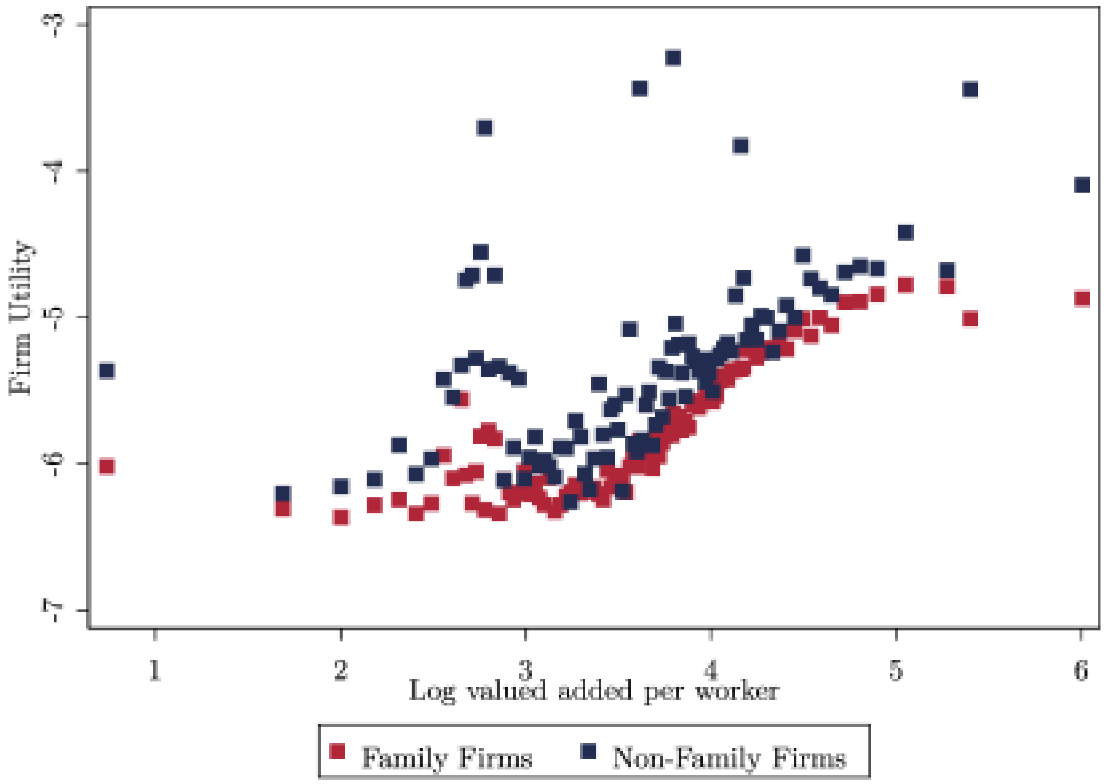


Figure 3
Within-Firm Wage Distribution

Figure 3 presents averages of the logarithm of weekly wages vis-à-vis the percentile of the logarithm of value added per worker. The variables on the $y-$axis are 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. Red and blue squares refer to family and non-family firms, respectively.

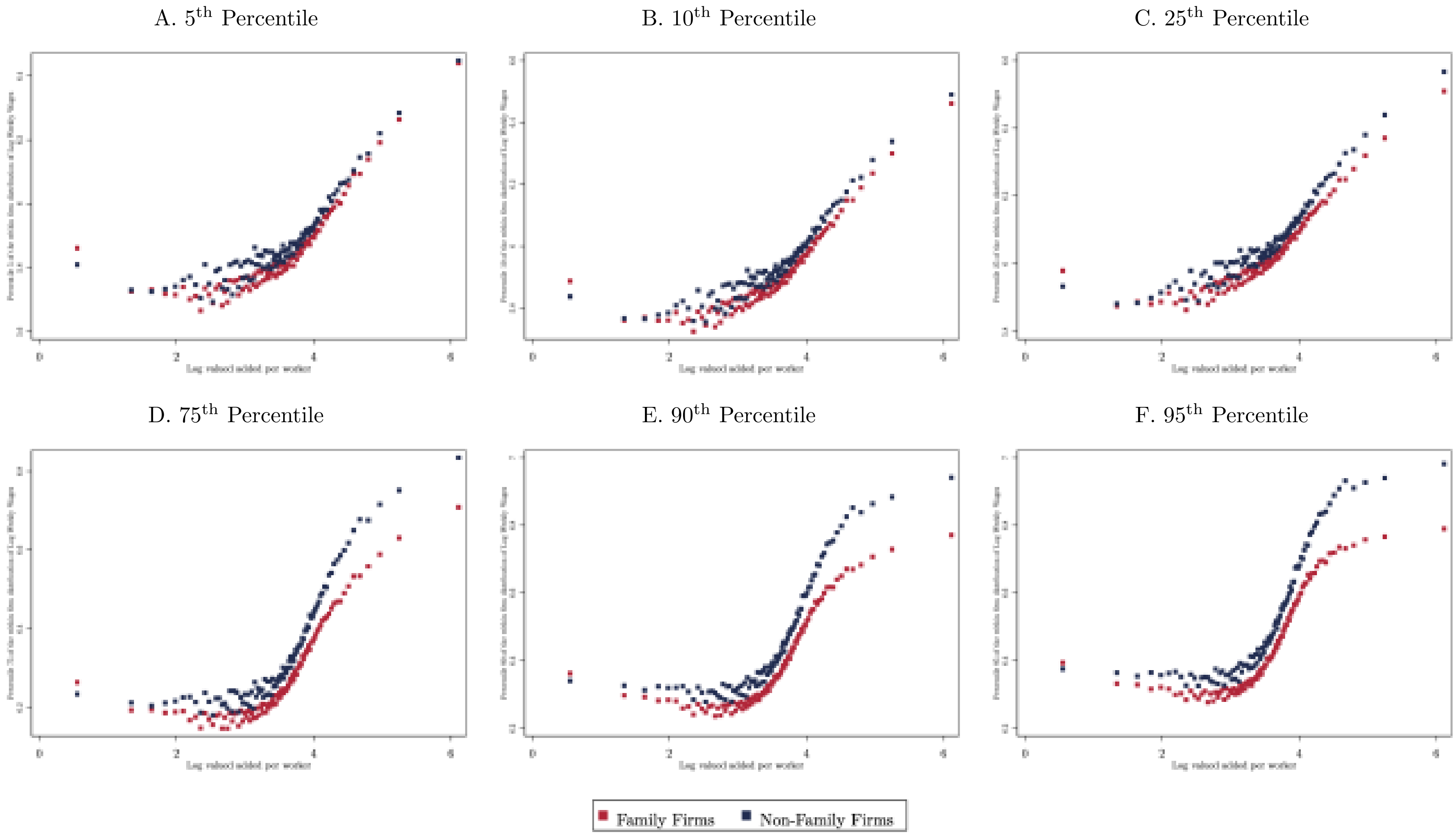


Figure 4
Promotion Rates and Fraction of Managerial Positions

Figure 4 presents averages of the promotion rates and of the fractions of managerial positions in family and non-family firms, by percentile of the logarithm of value added per worker. Panels A and B display the fractions of workers who are promoted (i.e., transitions) from a white-collar job to a middle-management position and from a middle-management position to a managerial position, respectively. Panels C and D display the fraction of middle managers and managers, respectively.

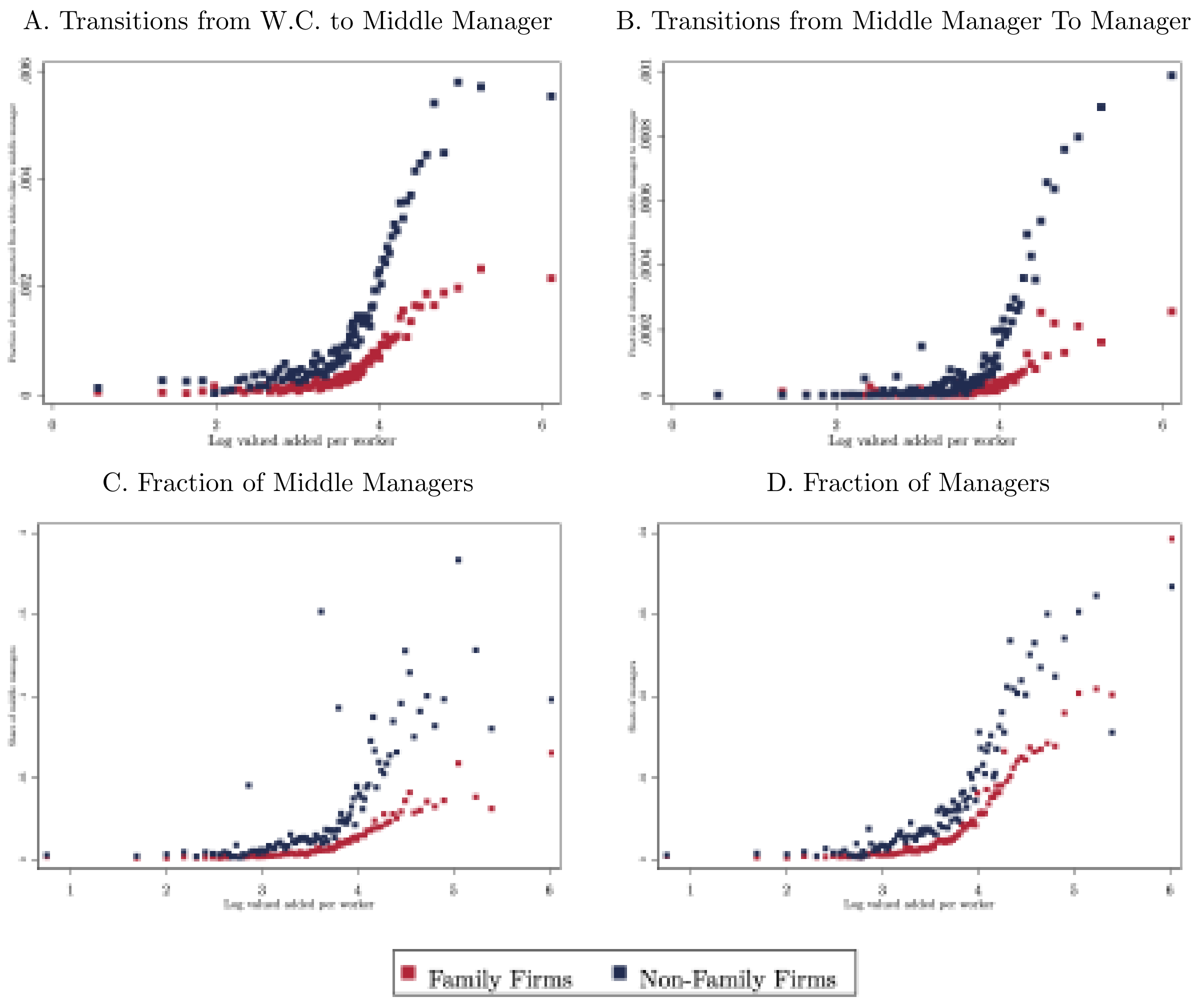


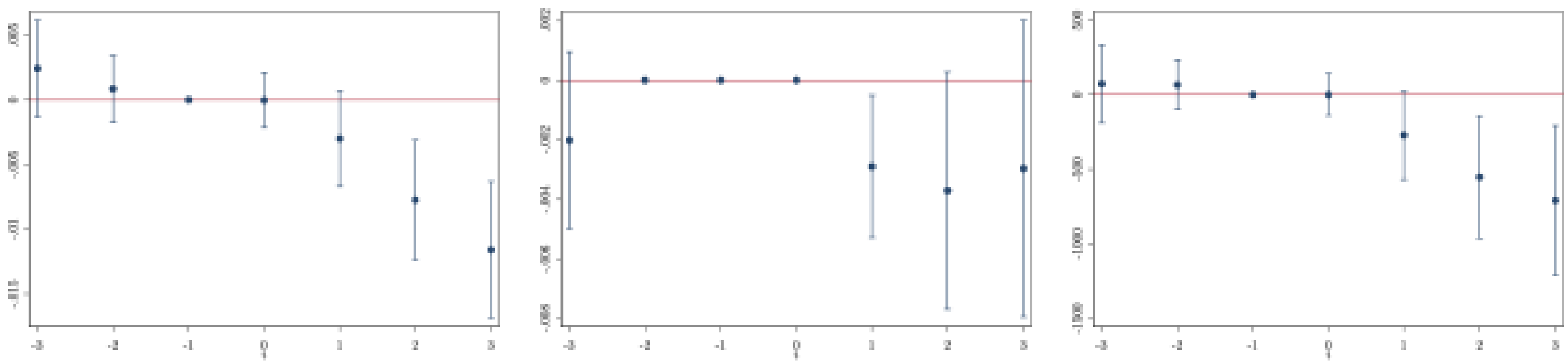
Figure 5

Returns to Promotions: Event-Study Evidence

Figure 5 presents event-study evidence of the differential effects of a promotion depending on whether the worker was promoted by a family vs. non-family firm. We consider two types of promotion. Panel A includes workers promoted from white collar to middle management positions. Panel B includes workers promoted from middle management to managerial positions. For every worker promoted in a family firm, we find a worker who obtained the same type of promotion in a non-family firm, matching on the following characteristics: logarithm of firm size, the logarithm of value added per worker, and 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 event. 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 terms with the associated 95% confidence intervals based on standard errors clustered at the firm level.

A. From White Collar to Middle Management

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



B. From Middle Management to Manager

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

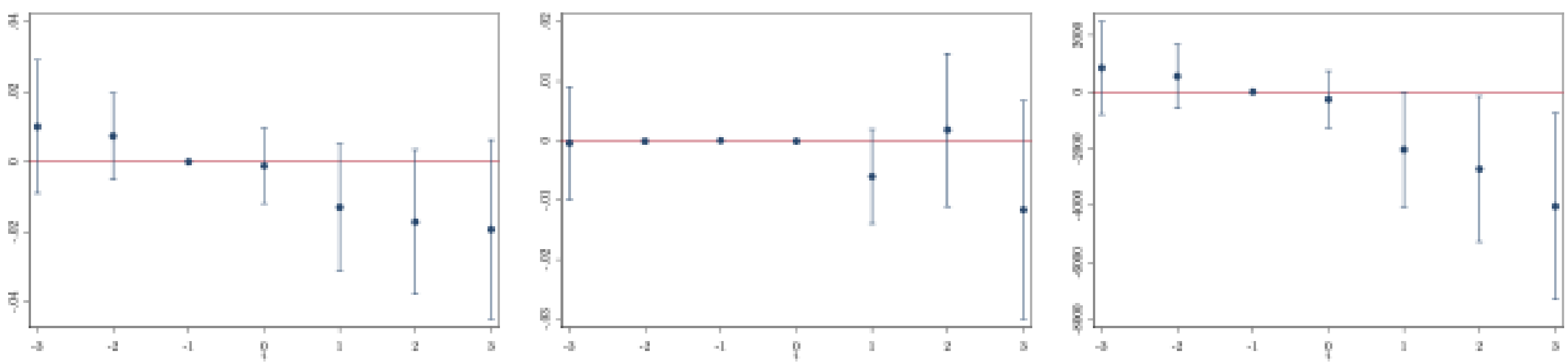
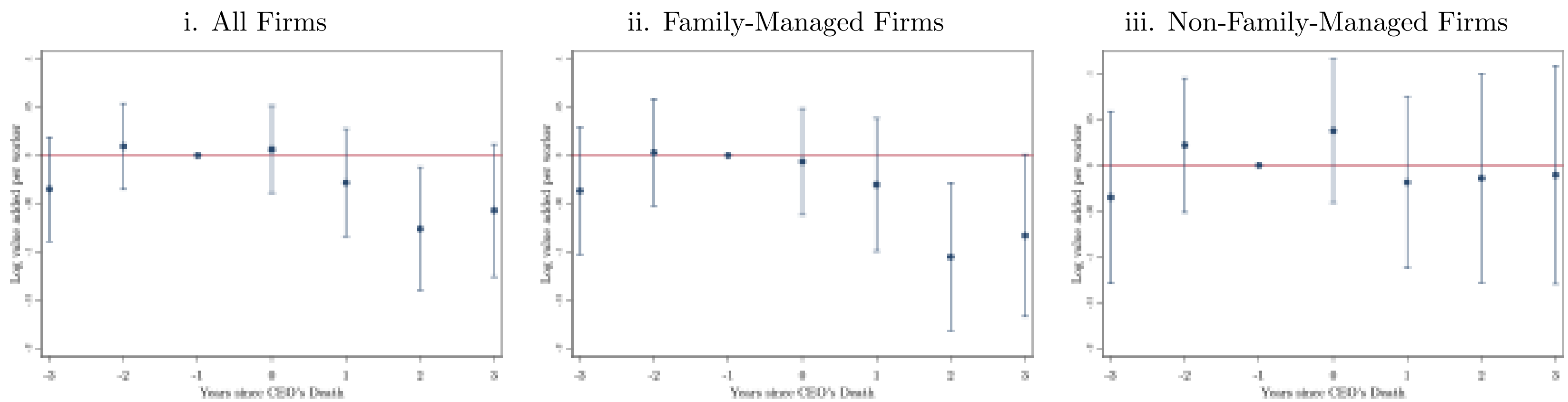


Figure 6
Productivity around CEO and Director Deaths

Figure 6 presents event-study evidence of the differential effects of a CEO (Panel A) or director (Panel B) death depending on whether the firm is a family or a non-family firm. For every family firm in which the CEO/director dies, we find a non-family firm matching 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 CEO/director 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 figures plot the coefficients on these interaction terms with the associated 95% confidence intervals, based on standard errors clustered at the firm level. Subpanels A.i and B.i include all firms. Subpanels A.ii and B.ii include 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 include only treated firms in which the deceased CEO/director was not a member of the controlling family, together with the respective matched control firms.

A. CEO Deaths



B. Director Deaths

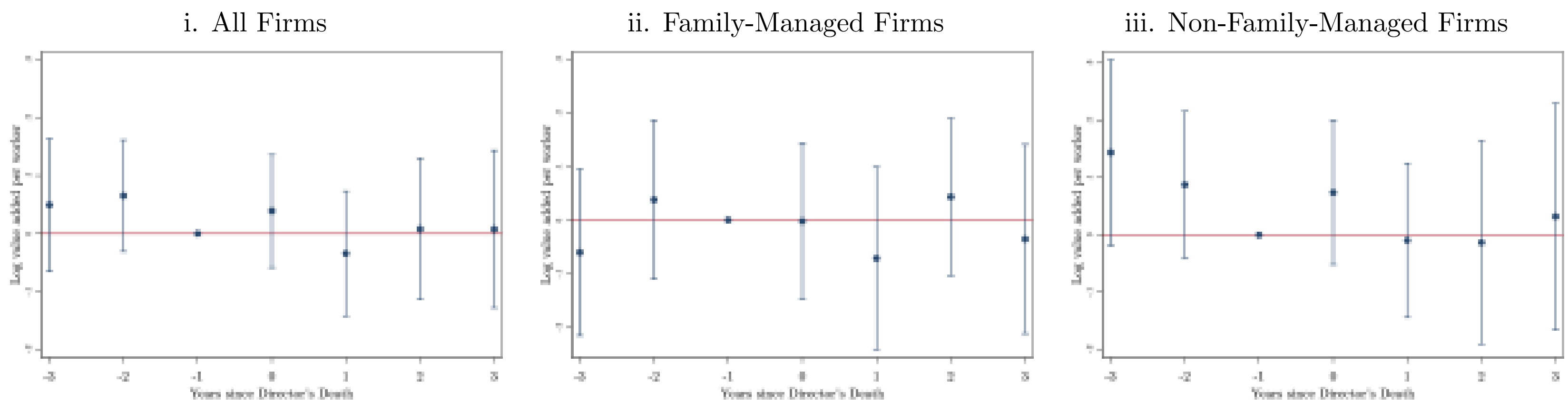


Figure 7

Worker-Level Outcomes around CEO Deaths

Figure 7 presents event-study evidence of the differential effects of a CEO death depending on whether the firm is a family or a non-family firm. For every family-managed family firm in which the CEO dies, we find a non-family firm matching on industry, the logarithm of firm size, and the logarithm of value added per worker. Within these paired firms, we match each worker in a family firm with a worker in a non-family firm based on wage growth, age, and (exactly), 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 event. The dependent variable, indicated at the top of each panel, is then regressed on year fixed effects, worker fixed effects, event-year dummies, and event-year dummies interacted with a family firm dummy. Employment is a dummy equal to 1 if the individual is employed, and Middle Mananger or Manager is a dummy equal to 1 if the worker is a middle manager or a manager. The figures plot the coefficients on these interaction terms with the associated 95% confidence intervals based on standard errors clustered at the firm level.

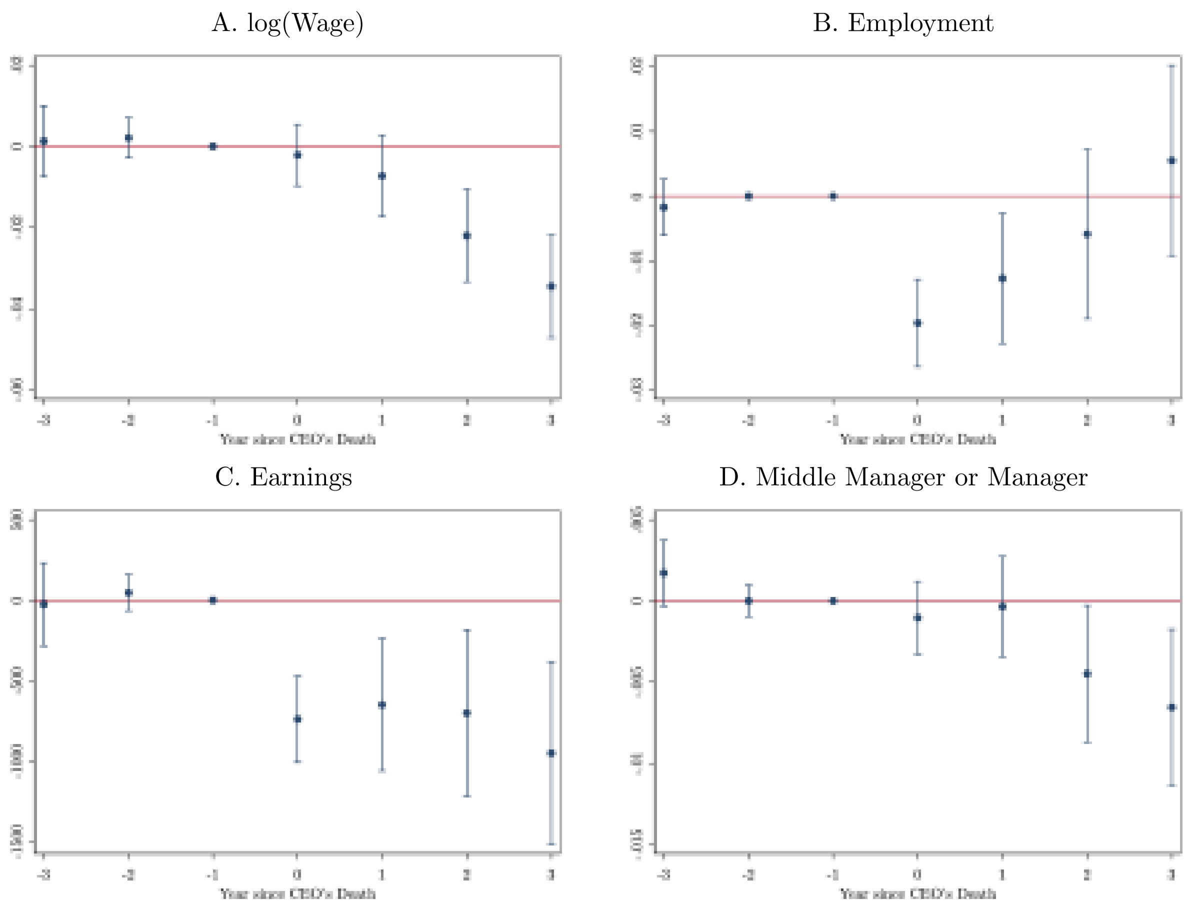


Figure 8
Timeline of the Model

$t = 0$	$t = 1$	$t = 2$	$t = 3$
Matching and hiring:	**On-the-job learning**:	**Workers' promotion**:	**Production and compensation**:
(i) workers apply for jobs in family or non-family firms; (ii) each firm j offers a wage contract to successful applicants (unsuccessful applicants stay unemployed)	(i) workers pick effort level e_w; (ii) acquire skills with probability e_w, anticipating promotion probability ϕ_j and bonus contract for promoted workers	(i) firm j promotes a fraction ϕ_j of skilled workers; (ii) offers them a bonus contract; (iii) managers pick effort level e_m; (iv) acquire skills with probability e_m	(i) skilled workers and managers produce y_j, (ii) skilled managers also produce δy_j; (iii) firms pay agreed wages and bonuses

Table 1
Descriptive Statistics

Table 1 presents means of some of the firm-level variables observed in the sample, broken down between family and non-family firms. The time period covers the years 2006–2017. Firm size is the number of workers employed in the firm. Value added per worker is operating value added, in thousand euros, scaled by the number of workers. ROA is earnings before interests and taxes scaled by total assets, all multiplied by one hundred. 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 (that is, the wage earned 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$^{\text{th}}$ percentile minus the logarithm of the wage measured at the 10$^{\text{th}}$ percentile. Share managers and share middle managers are the fractions of managers and middle maangers, respectively. Age is the average age of the firm's workforce. Experience is the number of years the worker has been employed in the current firm. Share males, part time, and temporary, are the fractions of male workers, part-time workers, and workers with temporary 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 the worker-year level.

	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
Experience	18.345	18.585
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

Table 2
The Anatomy of the Family Firm Wage Discount

Table 2 shows the average characteristics across family and non-family firms for the universe of employers with available financial information. The observation period comprises the years 2006–2017. The first row presents the average log weekly wage across family and non-family firms, weighted by the number of person-year observations. We then show the average person effects and firm effects across family and non family firms after fitting a two-way AKM model on the log weekly wage in these data. In this AKM model, the employer fixed effect is represented by the unrestricted interaction between the unique tax identifier of the employer and a dummy for whether the employer is under family ownership or not. The AKM model also controls for year fixed effects, as well as cubic polynomials in age and experience. The fourth row shows the average log value added per worker while the fifth row shows the rent-sharing coefficient obtained after regressing the firm effect on a constant and log value added per worker separately across family and non-family firms. The last three rows then performs the Oaxaca decomposition where the difference in the firm effects across family and non-family firms is divided into a productivity component (differences in average log value added per worker across family and non-family firms, weighted by the rent-sharing coefficient of non-family firms) and a bargaining component (differences in rent-sharing coefficient weighted using the distribution of log value added per worker of family firms). The last row reports the constant obtained when projecting the firm onto log value added per worker separately for family and non-family firms.

	Family Firm	**Non-Family Firm**	**Difference**
log(Wage)	6.03	6.18	-0.14
Adjusted log(Wage)	5.90	6.07	-0.16
Person Effects	-0.03	0.05	-0.08
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
Constant	-0.54	-0.44	-0.10

Table 3

Family Firm Wage Discount after Controlling for Productivity, Location and Sector

This table shows the results from a regression where the outcome variable is the firm effect estimated from a two-way (AKM) fixed effects decomposition on log wages. The regression is run on the micro, person-year, data and always controls for time fixed effects. Productivity is value added per worker. Industry is defined as 2-digit Ateco code. Standard errors are clustered at the firm level.

	(1)	(2)	(3)	(4)	(5)
Family Firm	-0.082***	-0.059***	-0.047***	-0.042***	-0.038***
	(0.006)	(0.004)	(0.003)	(0.002)	(0.001)
log(Productivity)		0.137***	0.134***	0.113***	0.108***
		(0.002)	(0.002)	(0.002)	(0.002)
Std Dep. Variable	0.20	0.20	0.20	0.20	0.20
# of Observations	89,507,579	89,507,579	89,507,579	89,507,579	89,507,579
# of Firms	753,154	753,154	753,154	753,154	753,154
Industry FE			X	X	
Province FE				X	
LLM FE					X

Table 4
Systematic Utility and Firm-level Wages

Table 4 shows the results from a regression where the dependent variable is page-rank utility index of Page et al. (1998) popularized in economics by Sorkin (2018). This index ranks employers on the basis of voluntary transitions made by workers. To define a voluntary transition, we make use of the fact that the Italian data provides information on whether a job was terminated because the worker has voluntarily resigned from the job. The page-rank utiility is then regressed on a dummy for whether the firm is a family firm and (from column 2) on the AKM firm-wage effect described in Table 2. All regression results are estimated in the micro person-year data and are thus person-year weighted. In all regression we control for year fixed effects as well the additional fixed effects listed at the bottom of the table. Standard errors are reported in parathesis and are clustered at the firm-level.

	(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 effect		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

A Appendix

A.1 Additional Results

Figure A1
The Family Firm Discount
and Workers' Places of Birth, Residence, and Location

Figure A1 shows regression coefficients obtained after regressing the firm "AKM" fixed effect on a family firm dummy, the logarithm of value added per worker, year fixed effects, and different sets of fixed effects. Every horizontal bar corresponds to the coefficient estimated on the family firm dummy in a different specification, with the associated 95% confidence interval in red. Going from the top bar to the bottom bar, the specifications include: no additional fixed effects, city of birth fixed effects, city of residence fixed effects, city of work location fixed effects, city of birth × city of residence fixed effects, city of residence × city of work location fixed effects, city of birth × city of work location fixed effects, and city of birth × city of residence × city of location fixed effects.

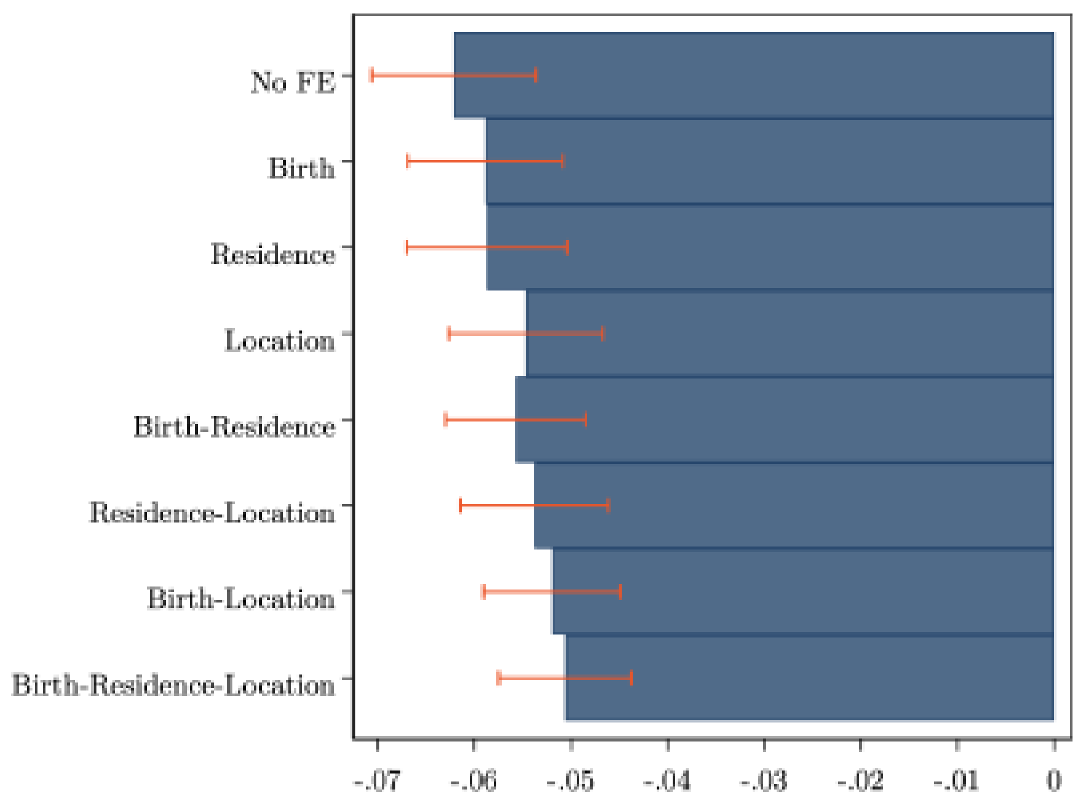


Table A1
Family Firm Wage Discount after Controlling for Productivity, Location and Sector: Firm Size Split

This table shows the results from a regression where the outcome variable is the firm effect estimated from a two-way (AKM) fixed effects decomposition on log wages. The regression is run on the micro, person-year, data and always controls for time fixed effects. Log value Added per worker is the average value added per worker observed for a given firm. Sector is defined as 2-digit Ateco code. Standard errors are clustered at the firm level. Panels A, B, and C display results for small (1 to 49-employees), medium (49 to 249 employees), and large (more than 249 employees) firms.

A. Small Firms

	(1)	(2)	(3)	(4)	(5)
Family Firm	-0.037***	-0.022***	-0.018***	-0.021***	-0.021***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
log(Productivity)		0.129***	0.123***	0.010***	0.096***
		(0.001)	(0.001)	(0.001)	(0.001)
Std Dep. Variable	0.22	0.22	0.22	0.22	0.22
# of Observations	38,546,175	38,338,488	38,338,488	38,338,488	38,338,488
# of Firms	720,714	701,128	701,128	701,128	701,128

B. Medium Firms

	(1)	(2)	(3)	(4)	(5)
Family Firm	-0.063***	-0.035***	-0.030***	-0.030***	-0.028***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
log(Productivity)		0.136***	0.130***	0.111***	0.103***
		(0.002)	(0.002)	(0.003)	(0.003)
Std Dep. Variable	0.17	0.17	0.17	0.17	0.17
# of Observations	20,827,828	20,786,330	20,786,330	20,786,330	20,786,330
# of Firms	30,874	30,642	30,642	30,642	30,642

C. Large Firms

	(1)	(2)	(3)	(4)	(5)
Family Firm	-0.068***	-0.049***	-0.041***	-0.027***	-0.020***
	(0.013)	(0.008)	(0.006)	(0.005)	(0.004)
log(Productivity)		0.140***	0.137***	0.096***	0.094***
		(0.005)	(0.005)	(0.007)	(0.007)
Std Dep. Variable	0.16	0.16	0.16	0.16	0.16
# of Observations	30,133,576	30,099,836	30,099,836	30,099,836	30,099,836
# of Firms	4,332	4,332	4,332	4,332	4,332
Industry FE			X	X	
Province FE				X	
LLM FE					X

Table A2
Returns to Promotions: Event-Study Coefficients

Table A2 presents event-study evidence of the differential effects of a promotion depending on whether the worker was promoted by a family versus a non-family firm. We consider two types of promotion. The sample used in the first three columns includes workers promoted from white collar to middle management positions. The sample used in the columns 4-6 includes workers promoted from middle management to managerial positions. For every worker promoted in a family firm, we find a worker who obtained the same type of promotion in a non-family firm, matching on the following characteristics: logarithm of firm size, the logarithm of value added per worker, and 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 event. 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 table reports coefficients on these interaction terms with the associated standard errors, clustered at the firm level.

Sample:	White Collar to Middle Manager			Middle Manager to Manager		
Dep. Var.	log(Wage) (1)	Employed (2)	Earnings (3)	log(Wage) (4)	Employed (5)	Earnings (6)
β_{-3}	0.002	-0.002	72.512	0.010	0.000	846.121
	(0.002)	(0.002)	(133.786)	(0.010)	(0.005)	(833.488)
β_{-2}	0.001	0.000	65.003	0.001	0.000	550.845
	(0.001)	(0.000)	(85.020)	(0.006)	(0.000)	(573.785)
β_0	0.000	0.000	-0.245	-0.001	0.000	-262.414
	(0.001)	(0.000)	(71.106)	(0.005)	(0.000)	(508.524)
β_1	-0.003	-0.003	-273.708	-0.013	-0.003	-2,029.774
	(0.002)	(0.001)	(151.368)	(0.009)	(0.004)	(1,024.820)
β_2	-0.008	-0.004	-554.126	-0.017	0.002	-2,711.946
	(0.002)	(0.002)	(210.045)	(0.011)	(0.007)	(1,305.083)
β_3	-0.012	-0.003	-710.200	-0.019	-0.012	-4,024.537
	(0.003)	(0.003)	(255.933)	(0.013)	(0.009)	(1,669.233)
R^2	0.848	0.274	0.739	0.838	0.289	0.744
Obs.	203,898	207,424	207,424	9,818	9,912	9,912

Table A3

Productivity around CEO and Director Deaths: Event-Study Coefficients

Table A3 presents event-study evidence of the differential effects of a CEO (first three columns) or director (columns 4-6) death depending on whether the firm is a family or a non-family firm. For every family firm in which the CEO/director dies, we find a non-family firm matching 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 CEO/director 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 with standard errors clustered at the firm level in parentheses. The samples used in columns 1 and 4 include all firms. The samples used in columns 2 and 5 (family-managed, "FM") include only treated firms in which the deceased CEO/director was a member of the controlling family, together with the respective matched control firms. The samples used in 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 respective matched control firms.

Death Events:	CEOs			Directors		
Sample	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

Table A4

Worker-Level Outcomes around CEO Deaths: Event-Study Coefficients

Table A4 displays event-study evidence of the differential effects of a CEO death depending on whether the firm is a family or a non-family firm. For every family firm in which the CEO dies, we find a non-family firm matching on industry, the logarithm of firm size, and the logarithm of value added per worker. Within these paired firms, we match each worker in a family firm with a worker in a non-family firm based on wage growth, age, and (exactly), 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 event. 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 displays the coefficients on these interaction terms with the associated standard errors clustered at the firm level in parentheses.

Dep. Var.	log(Wage)	Employed	Earnings	Manager or Mid. Man.
	(1)	(2)	(3)	(4)
β_{-3}	0.001	-0.002	-24.616	0.002
	(0.004)	(0.002)	(131.832)	(0.001)
β_{-2}	0.002	-0.000	46.861	0.000
	(0.003)	(0.000)	(57.343)	(0.000)
β_0	-0.002	-0.020	-739.353	-0.001
	(0.004)	(0.003)	(136.612)	(0.001)
β_1	-0.007	-0.013	-649.119	-0.000
	(0.005)	(0.005)	(209.385)	(0.002)
β_2	-0.022	-0.006	-700.304	-0.004
	(0.006)	(0.007)	(263.981)	(0.002)
β_3	-0.034	0.006	-949.492	-0.007
	(0.006)	(0.008)	(289.617)	(0.002)
R^2	0.850	0.351	0.803	0.681
Obs.	39,243	40,614	40,614	40,614

Table A5
Systematic Utility and Firm-level Wages

Table A5 shows the results from a regression where the dependent variable is page-rank utility index of Page et al. (1998) popularized in economics by Sorkin (2018), employing a correction for offer intensity and firm size. This index ranks employers on the basis of voluntary transitions made by workers. To define a voluntary transition, we make use of the fact that the Italian data provides information on whether a job was terminated because the worker has voluntarily resigned from the job. The page-rank utiility is then regressed on a dummy for whether the firm is a family firm and (from column 2) on the AKM firm-wage effect described in Table 2. All regression results are estimated in the micro person-year data and are thus person-year weighted. In all regression we control for year fixed effects as well the additional fixed effects listed at the bottom of the table. Standard errors are reported in parathesis and are clustered at the firm-level.

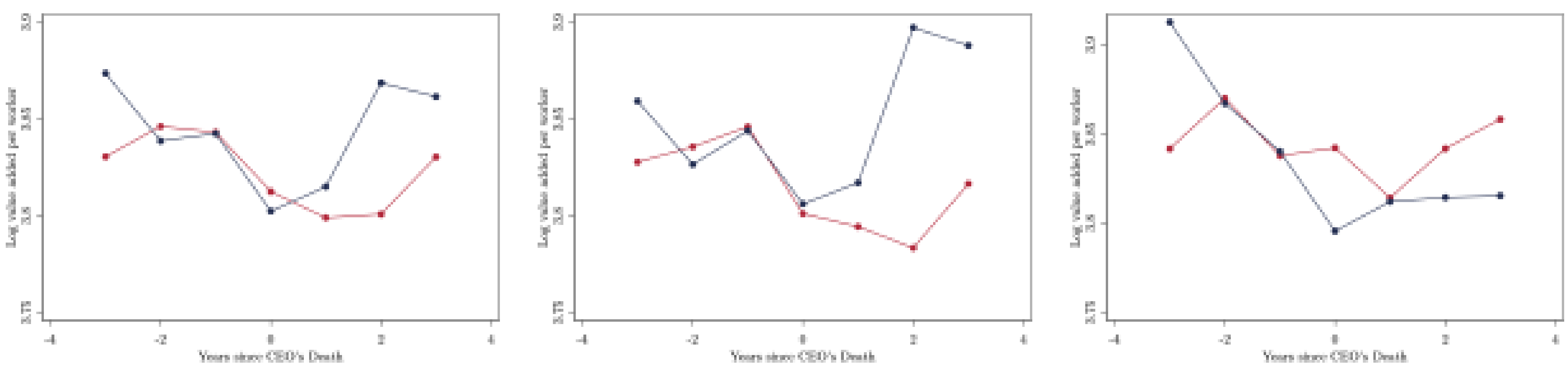
	(1)	(2)	(3)	(4)	(5)
Family Firm	-2.918***	-2.660***	-1.521***	-1.374***	-1.105***
	(0.273)	(0.276)	(0.074)	(0.060)	(0.034)
Firm effect		3.102***	4.064***	3.472***	2.978***
		(0.633)	(0.316)	(0.252)	(0.113)
Std. of Dep. Var.	3.36	3.35	3.35	3.35	3.35
# of Firms	309,395	307,259	307,259	307,259	307,259
Industry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

Figure A2
Productivity around CEO and Director Deaths: Sample Averages

Figure A2 plots the average productivity in the seven years surrounding CEO (Panel A) or director (Panel B) deaths, depending on whether the firm is a family or a non-family firm. For every family firm in which the CEO dies, we find a non-family firm matching 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 CEO / director event. Productivity is defined as the logarithm of value added per worker. Subpanels A.i and B.i include all firms. Subpanels A.ii and B.ii include 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 include only treated firms in which the deceased CEO / director was not a member of the controlling family, together with the respective matched control firms. Average productivity for family and non-family firms is plotted in red and blue, respectively.

A. CEO Deaths

i. All Firms | ii. Family-Managed Firms | iii. Non-Family-Managed Firms



B. Director Deaths

i. All Firms | ii. Family-Managed Firms | iii. Non-Family-Managed Firms

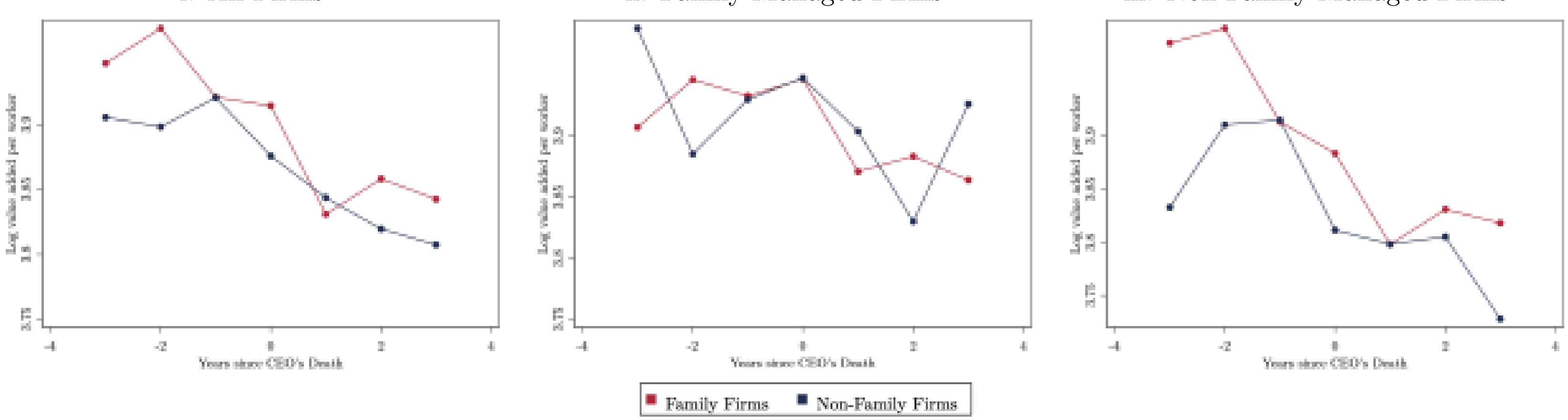


Figure A3
Churn Rates of Family and Non-Family Firms

This figure shows the average log value added per worker and churn rate among family and non-family firms for each centile of log value added per worker. The churn rate is calculated in the micro matched employer-employee data as an indicator equal to one if worker i experiences a separation from their dominant employer from year t to year $t+1$. Red and blue squares refer to family and non-family firms, respectively.

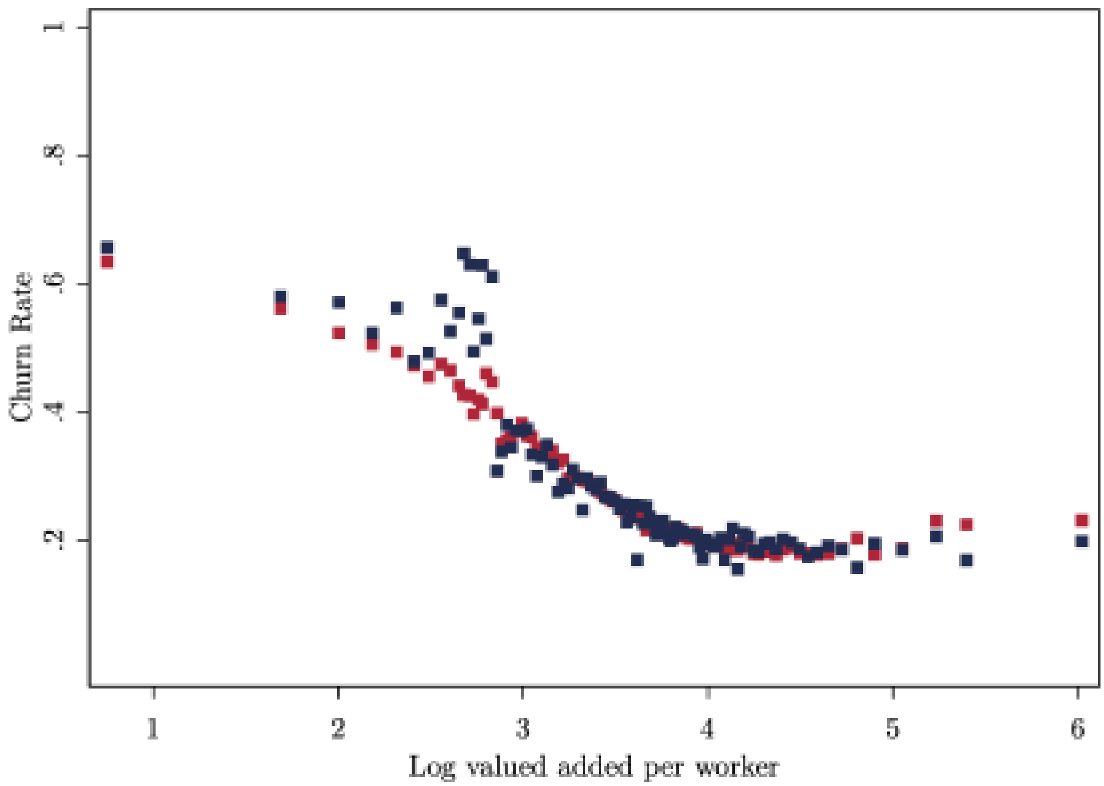
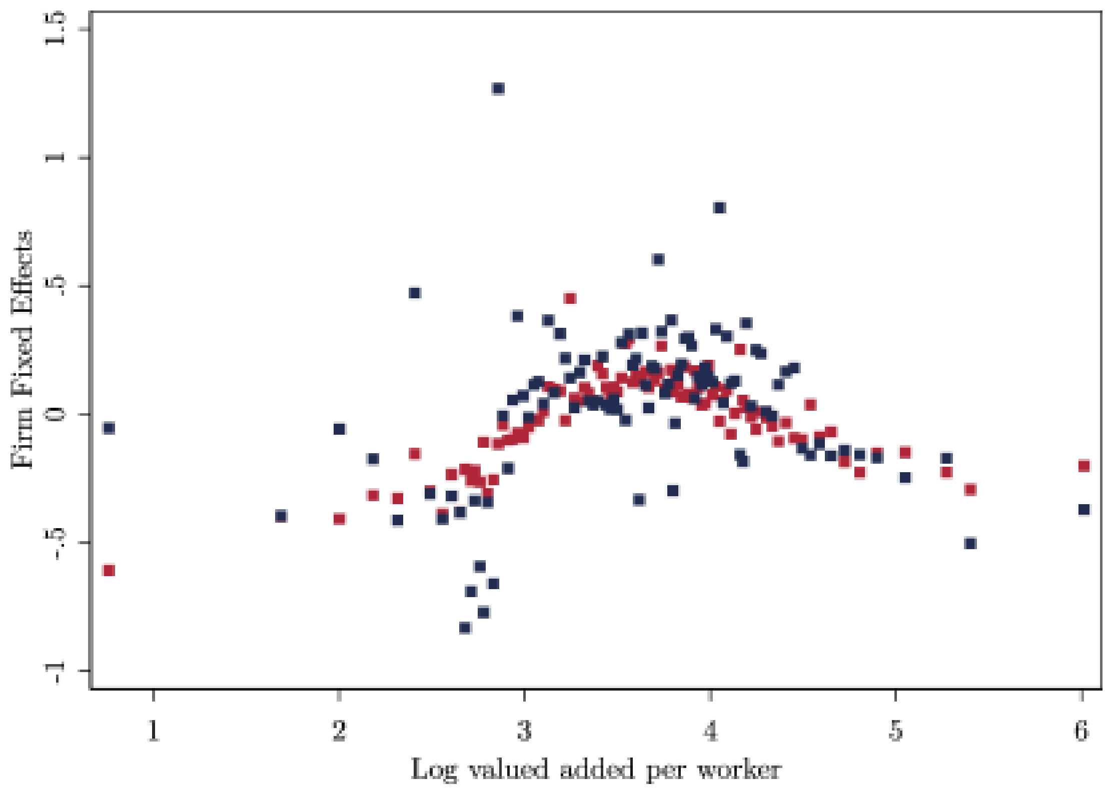


Figure A4
Number of Days Off due to Sickness if Employer by Family vs Non-Family Firms

We begin by estimating the AKM specification in equation 1 using as outcome $\log n_{it}$, i.e. the logarithm in the number of sick paid days of worker i in year t. This figure then computes the average log value added per worker and the firm effects in number of sicks days obtained from this specification for each centile of log value added per worker. Red and blue squares refer to family and non-family firms, respectively.



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 offering a certain bonus contract to promoted workers at $t = 2$. Hence, the model is solved by backward induction.

Stage *t*=2. The optimal compensation scheme that any firm j can offer to its newly appointed managers is one that pays a bonus x_j if the managers succeed in producing the additional output δy_j at $t = 3$ and 0 otherwise (as limited liability prevents a negative bonus). Given this state-contingent contract, workers promoted to managerial positions in firm j 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 $e^*_m = x_j/c_m$. Firm j will pick the bonus x_j so as to maximize the incremental profits that 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 in firm j 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 that managers are expected to generate for firm j 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 will choose the promotion probability ϕ_j of qualified workers by trading off the incremental profits π^*_{2i} against the loss of benefit of control, i.e., by 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 firm 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.[21] 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 are assumed to have a higher taste for corporate control, that is, $\beta_F > \beta_{NF}$, this implies that $\phi^*_F - \phi^*_{NF} < 0$ and decreases in y_j, which establishes **Result 2**.

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

$$\mathbb{E}\left[b|y_j\right] = x^* e^*_m = \frac{\delta y_j}{2}\frac{\delta y_j}{2c_m} = \frac{1}{c_m}\left(\frac{\delta y_j}{2}\right)^2. \tag{A.7}$$

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

By the law of iterated expectations, the unconditional expected bonus is:

$$\mathbb{E}[b] = \mathbb{E}_y\left[\mathbb{E}\left[b|y_j\right]\right] = \frac{\int \frac{1}{c_m}\left(\frac{\delta y_j}{2}\right)^4 \left[y_j + \frac{3}{2\beta c_m^2}\left(\frac{\delta y_j}{2}\right)^4\right] dF(y)}{\int \left(\frac{\delta y_j}{2}\right)^2 \left[y_j + \frac{3}{2\beta c_m^2}\left(\frac{\delta y_j}{2}\right)^4\right] dF(y)}. \tag{A.8}$$

Noting that the raw moment $E[y_j^k] \equiv \int y_j^k dF(y)$, expression (A.8) can be rewritten as:

$$\mathbb{E}[b] = \frac{\frac{1}{c_m}\left(\frac{\delta}{2}\right)^4 \mathbb{E}[y_j^5] + \left(\frac{\delta}{2}\right)^8 \frac{3}{2\beta c_m^3}\,\mathbb{E}[y_j^8]}{\left(\frac{\delta}{2}\right)^2 \mathbb{E}[y_j^3] + \left(\frac{\delta}{2}\right)^6 \frac{3}{2\beta c_m^2}\,\mathbb{E}[y_j^6]} \equiv \frac{\mathbb{N}}{\mathbb{D}}. \tag{A.9}$$

Showing that the expected bonus is lower in family firms than in non-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.9), respectively. Observing that

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

and grouping the common terms, we obtain:

$$\text{sign}\left(\frac{\partial\mathbb{N}}{\partial\beta}\mathbb{D} - \frac{\partial\mathbb{D}}{\partial\beta}\mathbb{N}\right) = \text{sign}\left(\mathbb{E}[y_j^5]\,\mathbb{E}[y_j^6] - \mathbb{E}[y_j^3]\,\mathbb{E}[y_j^8]\right) = \text{sign}\left(\frac{\mathbb{E}[y_j^6]}{\mathbb{E}[y_j^3]} - \frac{\mathbb{E}[y_j^8]}{\mathbb{E}[y_j^5]}\right), \tag{A.11}$$

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

A sufficient condition for the sign (A.11) to be negative is the following monotonicity assumption on the raw moments of the distribution $F(y)$:

$$\frac{\mathbb{E}[y_j^{k-1}]}{\mathbb{E}[y_j^{k-2}]} < \frac{\mathbb{E}[y_j^k]}{\mathbb{E}[y_k^{k-1}]}, \tag{A.12}$$

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

$$\frac{E[y_j^{k-1}]}{E[y_j^{k-h-1}]} < \frac{E[y_j^k]}{E[y_j^{k-h}]}, \tag{A.13}$$

where $h < k - 2$, thanks to a simple induction argument.[22]

Setting $h = 3$ and $k = 7$, inequality (A.13) becomes:

$$\frac{\mathbb{E}[y_j^6]}{\mathbb{E}[y_j^3]} < \frac{\mathbb{E}[y_j^7]}{\mathbb{E}[y_j^4]}, \tag{A.17}$$

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

$$\frac{\mathbb{E}[y_j^7]}{\mathbb{E}[y_j^4]} < \frac{\mathbb{E}[y_j^8]}{\mathbb{E}\, y_j^5}. \tag{A.18}$$

Inequalities (A.19) and (A.18) jointly imply:

$$\frac{\mathbb{E}[y_j^6]}{\mathbb{E}[y_j^3]} < \frac{\mathbb{E}[y_j^8]}{\mathbb{E}[y_j^5]}. \tag{A.19}$$

so that the sign (A.11) is negative. This proves **Result 3**.

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 of firm j choose effort e_w so as to maximize their stage 1 expected utility:

$$\max_{e_w}\ \mathbb{E}(U_1) = e_w(w_j + \phi_j^* e_m^* x_j^*) - \frac{c_w}{2} e_w^2 - \phi_j^* e_w \frac{c_m}{2} e_m^{*2}. \tag{A.20}$$

The first term of expression (A.20) indicates that, when choosing their effort e_w, workers rationally anticipate that, if successful, they not only gain the wage w_j but are also given a chance ϕ_j^* of being promoted to a managerial position and earn the corresponding bonus x_j^*, while bearing the cost of acquiring the necessary managerial skills (the last term).

[22]For $h = 1$, inequality (A.13) reduces to assumption (A.12). Assume that it also holds for a generic $h = n < k - 2$. Then:

$$\frac{\mathbb{E}[y_j^{k-1}]}{\mathbb{E}[y_j^{k-n-1}]} < \frac{\mathbb{E}[y_j^k]}{\mathbb{E}[y_j^{k-n}]}. \tag{A.14}$$

Rewrite assumption (A.12) for the $(k - n)^{th}$ moment as:

$$\frac{\mathbb{E}[y_j^{k-n-1}]}{\mathbb{E}[y_j^{k-n-2}]} < \frac{\mathbb{E}[y_j^{k-n}]}{\mathbb{E}[y_j^{k-n-1}]}. \tag{A.15}$$

Multiplying inequalities (A.14) and (A.15) side by side yields:

$$\frac{\mathbb{E}[y_j^{k-1}]}{\mathbb{E}[y_j^{k-n-2}]} < \frac{\mathbb{E}[y_j^k]}{\mathbb{E}[y_j^{k-n-1}]}, \tag{A.16}$$

showing that inequality (A.13) also holds for $h = n + 1$ and completing the proof.

Substituting from previous expressions for x_j^* and e_m^* and computing the first-order condition with respect to e_w, one finds that the workers' optimal effort is increasing both in the wage and in the probability of promotion:

$$e_w^*(w_j) = \frac{1}{c_w}\left[w_j + \frac{\phi_j^*}{2c_m}\left(\frac{\delta y_j}{2}\right)^2\right] = \frac{1}{c_w}\left[w_j + \frac{1}{2\beta_j c_m^2}\left(\frac{\delta y_j}{2}\right)^4\right]. \tag{A.21}$$

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

$$\begin{aligned}\max_{w_j}\ \mathbb{E}(\pi_{1j}) &= e_w^*[(y_j - w_j) + \phi_j^* e_m^*(\delta y_j - x_j^*)] \\ &= \frac{1}{c_w}\left[w_j + \frac{\phi_j^*}{2c_m}\left(\frac{\delta y_j}{2}\right)^2\right]\left[y_j - w_j + \frac{\phi_j^*}{c_m}\left(\frac{\delta y_j}{2}\right)^2\right],\end{aligned} \tag{A.22}$$

yielding the optimal efficiency wage paid to employees who manage to acquire productive skills at stage $t = 1$:

$$w_j^* = \frac{1}{2}\left[y_j + \frac{\phi_j^*}{2c_m}\left(\frac{\delta y_j}{2}\right)^2\right] = \frac{1}{2}\left[y_j + \frac{1}{2\beta_j c_m^2}\left(\frac{\delta y_j}{2}\right)^4\right]. \tag{A.23}$$

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.23) in (A.21) yields workers' equilibrium effort at stage $t = 1$:

$$e_w^*(w_j^*) = \frac{1}{2c_w}\left[y_j + \frac{3\phi_j^*}{2c_m}\left(\frac{\delta y_j}{2}\right)^2\right] = \frac{1}{2c_w}\left[y_j + \frac{3}{2\beta_j c_m^2}\left(\frac{\delta y_j}{2}\right)^4\right]. \tag{A.24}$$

Effort decreases in β_j, and is therefore higher in non-family firms. Taken together, equations (A.23) and (A.24) imply all the findings stated in **Result 1**.

Replacing the equilibrium effort (A.24) into expression (A.22), it is immediate to see that expected profits are always positive. By replacing expression (A.24) into (A.20) we obtain the equilibrium expected utility of workers from working in firm j:

$$\mathbb{E}(U_1^*) = \frac{1}{8c_w}\left[y_j + \frac{3}{2}\frac{\phi^*}{c_m}\left(\frac{\delta y_j}{2}\right)^2\right]^2 = \frac{1}{8c_w}\left[y_j + \frac{3}{2\beta_j c_m^2}\left(\frac{\delta y_j}{2}\right)^4\right]^2. \tag{A.25}$$

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

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

In equilibrium, workers will allocate their applications so as to equalize their expected utility from applying to family or non-family firms. Denoting the mass of applicants to family and non-family firms' jobs by A_F and A_{NF} respectively, and recalling that both types of firms hire workers by randomly drawing from their respective applicant pools, applicants obtain 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 of firm yields a positive expected utility by equation (A.25), 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 in either set of firms if

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

Since in equilibrium careers in non-family firms have been shown to offer higher expected utility than those 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 low enough compared to that of being hired upon applying to a family firm. Otherwise said, in equilibrium applying for a job in non-family firms entails greater unemployment risk than applying for a job in family firms.