Barcelona School of Economics

WORKING PAPERS

Incentive Contracts and Peer Effects in the Workplace

BSE Working Paper 1457

August 2024 (Revised March 2025)

Marc Claveria-Mayol, Pau Milán, Nicolás Oviedo-Dávila

bse.eu/research

Incentive Contracts and Peer Effects in the Workplace*

Marc Claveria-Mayol† Pau Milán‡ Nicolás Oviedo-Dávila§

March 17, 2025

Abstract

We analyze how firms should design wage contracts when workers collaborate in teams and effort costs depend on colleagues through a peer network. Performance-based compensation generates incentives that cascade through the organization, which firms target to boost profits. We analyze optimal incentive design if firms can—and can't—fully discriminate across workers, and when the production technology is separable or complementary across divisions. When workers' effort is substitutable, the most central workers—those who influence most colleagues directly and indirectly—receive the steepest incentives only when output risk is sufficiently large; otherwise firms prioritize workers that are closer to those they influence. We derive a sufficient network statistic that measures the surplus loss when firms must compensate workers of varying centrality equally. Finally, when production technology exhibits complementarity across teams, stronger incentives are assigned to workers who influence colleagues in small teams that receive little influence from others. We apply our findings to organizational design questions, such as optimal firm structure and workforce investments.

Keywords: Incentives, Organizations, Contracts, Networks, Moral Hazard

JEL Codes: D21, D23, D85, D86, L14, L22

*We thank Yann Bramoullé, Hector Chade, Fred Deroian, Pradeep Dubey, Jan Eeckhout, Andrea Galeotti, Ben Golub, Sanjeev Goyal, Ting Liu, Inés Macho-Stadler, Mihai Manea, David Perez-Castrillo, Eran Shmaya, Ran Shorrer, participants at the 9th European Conference on Networks (Essex), 27th CTN Conference (Budapest), 2024 BSE Summer Forum (Barcelona), 2024 Conference on Mechanism and Institution Design (Budapest), Networks Webinar Janeway Institute (Cambridge), 2024 EWMES and 49th SAEe (Palma) and seminar participants at Aix-Marseille School of Economics, UAB, UCL, UPF and TSE for useful comments. Milán acknowledges financial support from the Spanish Ministry of Economy and Competitiveness, through grant ECO2017-83534-P and grant PID2020-116771GB-I00, and from the Severo Ochoa Program for Centres of Excellence in R&D (CEX2019-000915-S). Oviedo-Dávila acknowledges financial support from the Spanish Ministry of Science and Innovation, through grant PID2022-140014NB-100, funded by MCIN/AEI/10.13039/501100011033 and by the FSE+. All errors are ours. Declarations of interests: none.

†Universitat de les Illes Balears (mclaveriamayol@gmail.com)

‡UAB-Barcelona School of Economics (*corresponding author:* pau.milan@gmail.com)

§UAB-Barcelona School of Economics (nicolasoviedod@gmail.com)

1 Introduction

Earnings inequality within firms has increased significantly in recent decades, with top earners capturing disproportionately large shares of total compensation. Consequently, wage distributions have become highly skewed, with increasingly pronounced right tails. One key driver of this trend is the heightened sensitivity of top earners' salaries to firm performance through bonuses, stock options, and other forms of *variable pay*: recent micro-level evidence shows that earnings of the top 1% of U.S. employees respond four times more to firm performance than earnings at the bottom 1%. At the same time organizational structures have flattened, with fewer hierarchical layers and larger spans of control at the top. Empirical findings indicate that greater spans of control correlate with steeper earnings profiles within firms and across occupations, suggesting that an employee's position in an organization shapes compensation.[1]

This paper ties salaries to organizational structure by analyzing how firms design wage contracts when workers influence each other's productivity. We extend a standard moral hazard model by including productivity spillovers. In our model, motivating one worker improves the performance of others throughout the organization. Such spillovers appear in many workplace settings: doctors treat patients more efficiently when nurses monitor symptoms diligently; nurses work harder when matched with diligent peers; and plant managers who streamline protocols reduce subordinates' workloads. As these examples suggest, peer effects reflect assistance, knowledge, delegation, or peer pressure, and they can flow across hierarchical levels (vertical spillovers) or between co-workers at similar ranks (horizontal spillovers). Evidence shows that peer effects significantly impact productivity, contributing nearly half the total effect of performance pay in some settings (Ashraf and Bandiera, 2018).[2]

[1]For a survey of earnings inequality in firms and its theoretical explanations, see Neal and Rosen (2000). Jensen, Murphy, and Wruck (2004) report that CEO pay in S&P 500 firms rose from 850,000 in 1970 to over 14 million in 2000, with stock options driving over half the increase. Wallskog, Bloom, Ohlmacher, and Tello-Trillo (2024) attribute 40% of the CEO-median worker pay gap growth (1980–2013) to the differing pay sensitivity of the top and bottom 1%. See Bertrand (2009) for a survey on CEO pay.

On firm flattening, Rajan and Wulf (2006) show that the number of managers reporting to the CEO rose from about four in 1986 to over seven by 1999. Possible drivers include knowledge hierarchies (Garicano, 2000) and trade liberalization eroding tall corporate structures (Guadalupe and Wulf, 2010).

Fox (2009) finds span-of-control wage gaps widen up the hierarchy: for sales workers, rank-4 supervisors managing three times as many workers earn 1% more, while rank-7 supervisors earn 3.4% more. Smeets and Warzynski (2008) confirm this, showing that managers overseeing twice the average team size earn 2.8% more, with middle managers indirectly responsible for twice as many workers earning 4.1% more.

[2]Mas and Moretti (2009) finds that the pattern of peer effects reflects firms' physical layout, while Bandiera, Barankay, and Rasul (2005) detects informal friendship bonds. Workers' effort has been found to respond to co-workers' effort even when remuneration is independent of output (Falk and Ichino, 2006).

How should firms design monetary incentives to exploit productivity networks?[3] Our framework integrates contract theory with network games in order to characterize optimal compensation.[4] In our model, risk-averse workers collaborate in a team and each worker's effort cost depends on colleagues' effort levels. The firm cannot observe individual effort but can condition wages on total output, which combines workers' efforts plus a random component. Workers receive both fixed and performance-based wages. Variable payments motivate effort but also introduce risk, requiring the firm to provide risk compensation, which reduces profits. Firms must therefore balance incentive provision against the cost of exposing workers to risk—a fundamental trade-off in classical incentive design (Holmstrom and Milgrom, 1987, 1991; Bolton and Dewatripont, 2004). Our theoretical contribution highlights how firms can leverage the structure of peer complementarities to distribute incentives across the organization while optimally managing workers' exposure to risk.

We analyze optimal incentive design while varying two key aspects of our environment: institutional constraints and the production technology of the firm. Specifically we characterize incentive rules for three related cases. We first assume the production technology is linear and we design incentives when the firm can—and cannot—write personalized wage contracts. We then analyze personalized contracts when the firm's technology is "modular", meaning that workers are divided into teams which are complementary in production.[5]

Our first result (Proposition 1) characterizes incentives for the simplest case (linear technology and personalized contracts). We show that each worker's performance-pay is determined as a linear combination of everyone's *Bonacich centrality*—a network statistic that accumulates direct and indirect paths along the organization. The exact contribution of worker j's centrality to worker i's performance-pay depends on the interaction of different forces. When i and j influence the same colleague—directly or indirectly—incentivizing i will raise the marginal cost of incentivizing j, creating a substitution effect. This is because incentivizing i pushes the colleague to work more in equilibrium, making her less responsive to additional incentives from j. Conversely, if i and j influence different colleagues that are themselves connected, incentivizing i lowers the marginal cost of incentivizing j, creating

[3] Previous work has focused on optimizing team composition to leverage social incentives. For instance, Mas and Moretti (2009) find that by maximizing skill diversity in each shift, a supermarket chain could save up to 123,529 hours worked per year which, in 2009 wage costs, amounted to $2.5 million per year.

[4] Network games describe environments where agents best-reply to a "local" subset of other players and the overlapping sets of players define a graph, or network. Ballester, Calvó-Armengol, and Zenou (2006) pioneered the case of strategic complements and Bramoullé and Kranton (2007) the case of strategic substitutes.

[5] There is a fourth possible case: that the firm's technology is modular and the firm cannot write personalized contracts. This complicates the analysis without providing additional insights, so we focus on the other three.

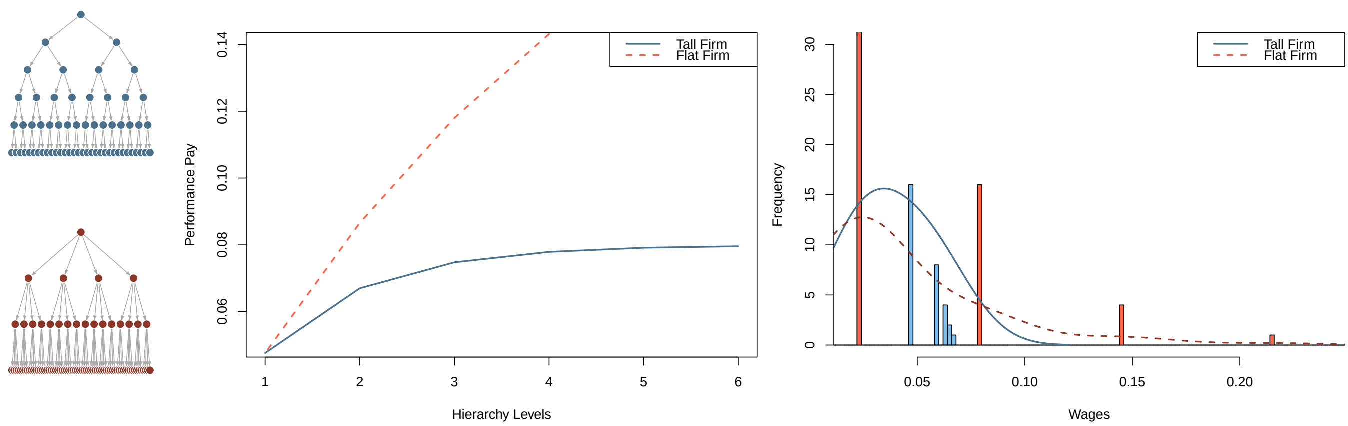

Figure 1: Performance pay and overall wage distribution for two different organizational structures: a tall (blue) firm and a flat (red) firm.

complementarities. This is because incentivizing i pushes i's colleague to work more in equilibrium, making j's colleague more responsive to additional incentives from j. We show how these countervailing forces enter the optimal incentive allocation rule as a function of the firm's level of output risk.[6]

Figure 1 illustrates how our framework can explain the relationship between wage dispersion and organizational structure by plotting the optimal allocation rule for a simple example with two hierarchical firms. To keep things simple, assume that a worker's effort cost is affected only by her direct supervisor. This is drawn as an arrow from a worker to her subordinates.[7] We simulate wages for a *tall firm* with 6 levels and a span of control of 2, and a *flat firm* with 4 levels and a span of control of 4. The left panel shows optimal performance-pay across different hierarchy levels. Consistent with empirical findings described above, performance-pay increases as we move up the organization and, more importantly, the pay-profile is steeper for the flat firm. The right panel shows the distribution of total earnings, which sums performance-pay and fixed salaries. As in the data, the flatter firm has a longer right tail because the wage gap between top earners is much larger. In other words, optimal wage contracts with peer effects can replicate the stylized facts described above because in flatter firms upper management has shorter paths to the rest of the workforce, which makes concentrating incentives on the top more profitable when output risk is significant.[8]

[6] These two forces are explained in more detail below and illustrated in Figure 3 and Figure 4 respectively.

[7] In reality, and in our model, peer effects need not coincide with the firm's formal chain of command. We do this here to link earnings distribution with flat and tall firms.

[8] Alternative explanations for within-firm earnings inequality typically assume an initial distribution of managerial talent which is sorted across job levels and firms over time in order to produce the observed earnings distribution (Roy, 1951; Heckman and Honore, 1990; Gabaix and Landier, 2008; Garicano and

Second, we consider how institutional constraints affect our findings if firms are forced to offer identical contracts to workers in the same job category (Proposition 4). This *wage benchmarking* scenario is motivated by recent policy responses to the negative effects associate with large wage disparities.[9] How should firms design coarse contracts? In this case, incentives are allocated according to an occupational-wide network measure, which aggregates network paths across all workers in an occupation. More importantly, forcing a common contract on workers with different network positions prevents the firm from tailoring incentives and diminishes surplus. We show that the efficiency loss from these constraints is captured by within-group centrality variance (Proposition 5): uniform wages impose no efficiency loss when all members of the same category are equally central (such as in the simplified example in Figure 1), but in more realistic scenarios losses increase linearly with within-group variability in centrality. This finding has significant practical implications, suggesting that standardized pay structures, while equitable, can be economically costly.

Third, we generalize the production function of the firm. Our analysis so far assumed that workers are substitutable in output. To address this limitation, we extend our analysis to modular production where output depends on the minimum performance across essential components. By "modular", we mean production processes where the final output requires successful completion of multiple distinct tasks or modules—like an assembly line where a failure at any stage compromises the entire product.[10] To model this, we divide the workforce into modules, and we assume the production function is substitutable within modules but perfectly complementary across modules. Since we allow for any partition of workers into modules, modular production nests our baseline environment because if the entire workforce belongs to a single module, output is linear.

With modular production, incentives are allocated very differently (Proposition 6). A worker's performance pay depends on the entire organization—even on workers she is not connected to who are in different modules, since all modules depend on each other to generate output. Although the allocation rule is computed analogously to before, the relevant centrality measure now weighs paths by module-specific factors, which reflect how relatively

Rossi-Hansberg, 2006). We complement this perspective by emphasizing that firms tie wages not only to workers' attributes but also to job positions, as evidenced by the fact that talent does not change on promotion day. Having said this, our model is flexible enough to allow for variation in individual skills.

[9] Significant disparities in peers' salaries lead to job dissatisfaction and higher quit rates (Card, Mas, Moretti, and Saez, 2012; Breza, Kaur, and Shamdasani, 2018). As a result, governments and agencies have encouraged wage transparency (Mas, 2017; Obloj and Zenger, 2022; Cullen, 2024), salary benchmarking (Cullen, Li, and Perez-Truglia, 2022), and national wage setting (Hazell, Patterson, Sarsons, and Taska, 2022), all of which compress wages within occupational categories, especially at lower skill levels.

[10] Kremer (1993) named the O-ring theory on the fatal Challenger spacecraft incident in 1986, which malfunctioned due to the failure of one small metal gasket. See also, Matouschek, Powell, and Reich (2025) for a recent analysis on how to organize communication networks in a firm with modular production.

costly it is for modules to match each other's performance. We highlight two special cases. First, we turn off peer effects and show how incentives depend on relative module size: managers who validate the work of many subordinates become critical choke points, similar to senior engineers who must approve all code before deployment. We show that managerial bottlenecks create large pay disparities across hierarchical levels. Second, when every worker is essential, firms prioritize incentivizing workers with fewer incoming links rather than those with greater outgoing influence, since these workers face higher costs and are more tempted to reduce effort. More generally, we show that wage profiles exhibit sharp discontinuities between modules—contrary to the smoother profiles seen in non-modular production—and that these jumps can be much larger than the variation in pay within modules.

Finally, while our primary focus is deriving incentive rules across different production technologies and institutional constraints, our model has broader implications for *organizational design* questions—such as how firms should structure internally, or how they should invest in their workforce. To address these questions, we derive a *spectral decomposition* of profits based on eigenvalues and eigenvectors of the associated network, reducing organizational complexity to its principal components. In Section 5, we leverage well-known spectral properties to establish how profits depend on fundamental structure of the organization. Among other things, we find that profits are greater when centrality is evenly distributed, everything else equal. We also evaluate investment strategies by establishing an average connectivity threshold that determines when firms should prioritize "team-building exercises" that strengthen peer effects over comparable investments in individual human capital.

Related Literature - We contribute to the contract design literature with multiple agents (Mookherjee, 1984; Macho-Stadler and Pérez-Castrillo, 1993; Bolton and Dewatripont, 2004; Winter, 2010), tracing back to Holmstrom (1982) on moral hazard in teams. We link this framework to firms' organizational structure by analyzing pay-for-output incentives when workers interact via productivity spillovers. Under certain conditions, worker behavior aligns with a modified linear-quadratic network game (Ballester et al., 2006). Our key contribution is extending the classic moral hazard problem with multiple agents using network-based peer effects. Unlike prior work on network-targeted incentives, we focus on classical performance-pay schemes (Demange, 2017; Belhaj and Deroïan, 2018; Galeotti, Golub, and Goyal, 2020; Parise and Ozdaglar, 2023).

The main applied contribution is to use our model to speak to questions at the heart of organizational economics. This allows us to explore adjacent questions, such as how optimal incentive design depends on the granularity of the contract and the modular structure of production. There has been significant recent interest connecting contracts to peer networks, focusing on very different topics, from relative performance compensation schemes (DeMarzo and Kaniel, 2023) to endogenous spillovers (Shi, 2024). For a recent contribution seeking to understand non-parametric conditions on incentive optimality on networks see Dasaratha, Golub, and Shah (2024). In independent work recently circulated, Sun and Zhao (2024) has a related framework with a focus on relative status concerns, focusing on peer pressure's psychological costs and the allocation of psycho-therapeutic resources.

Recent work on optimal price discrimination with local network effects shares many conceptual similarities with our approach. In these studies, a monopolist designs a menu of prices or discounts to leverage consumer externalities within an existing network. For example, Bloch and Quérou (2013) identify market conditions under which price discrimination may not be optimal, as the incentive to subsidize central consumers is counterbalanced by their higher willingness to pay. Similarly, Fainmesser and Galeotti (2016) analyze how varying levels of information about network effects can lead to increased price discrimination and even improve overall welfare when price effects are strong. Candogan, Bimpikis, and Ozdaglar (2012) is most similar to our analysis because they relate prices (in our case wages) to Bonacich centrality and they also consider a restricted scenario where the firm cannot fully price discriminate.

Finally, our theory provides a new set of testable predictions that speak to a strand of empirical work estimating peer effects, group composition, and team incentives in organizations (Hamilton, Nickerson, and Owan, 2003; Cornelissen, Dustmann, and Schönberg, 2017; Calvó-Armengol, Patacchini, and Zenou, 2009; Amodio and Martinez-Carrasco, 2018). Most of these papers emphasize how different remuneration schemes or other aspects of the contract—like employment termination decisions—affect productivity spillovers across workers. Our framework takes the peer effect structure as given and solves for the optimal contract.

Roadmap - The rest of the paper is organized as follows. Section 2 presents the baseline model (linear technology and personalized contracts) and solves the optimal contract. We provide comparative statics results and discuss negative spillovers and incomplete information. In Section 3 we consider wage benchmarking and derive its impact on overall surplus. In Section 4 we generalize the firm's technology and derive optimal incentive allocation rules for modular production. In Section 5 we summarize our results on organizational design, which we relegate to Supplementary Appendix G to keep the paper concise. We conclude in Section 6 with a discussion of future lines of research. All proofs are in the Appendix.

2 The Model

2.1 Basic Setup

Consider a risk-neutral firm that hires n workers $N = \{1, 2, \dots, n\}$, to conduct a joint production process.[11] Each worker chooses individual effort $e_i \in \mathbb{R}_+$ and the firm's production is given by

$$X(\mathbf{e}) = \sum_{i \in N} e_i + \varepsilon,$$

where $\mathbf{e}$ is the vector of workers' efforts and $\varepsilon \sim \mathcal{N}(0, \sigma^2)$ is an unobserved random shock to output. Because individual effort is not observable, contractual wage agreements must be based on observable (and verifiable) outcomes, such as output. We focus on the case in which the firm offers linear wage schemes of the form

$$\omega_i(X) = \beta_i + \alpha_i X,$$

where β_i is a fixed payment and α_i captures the contract's variable payment or performance-based compensation.[12] Although linear contracts may seem restrictive, they are parsimonious and resemble most equity payments and bonus schemes typically offered in corporate wage contracts. They are also optimal in some circumstances.[13]

A worker's cost of effort depends on the effort that is exerted by her co-workers. Following the peer effects literature, we assume a linear-quadratic function of own and neighbors' efforts:

$$\psi_i(\mathbf{e}; \mathbf{G}) = \frac{1}{2} e_i^2 - \lambda e_i \sum_{j \in N} g_{ij} e_j. \tag{1}$$

where $g_{ij} = 1$ if j is i's *co-worker* and 0 otherwise. In other words, we define i's co-workers as those members of the firm that can influence i's costs by exerting more or less effort. These co-workers can (but need not) reflect the formal organizational structure of the firm: they may be i's subordinates, managers, or just someone that sits next to i. We allow for

[11] We consider individual production in Supplementary Appendix C and show that the optimal structure of incentives is equivalent under individual and joint production.

[12] Although α_i can be thought of as a form of equity compensation whereby a share of the firm is transferred to the worker, one can also consider cases where $\sum_i \alpha_i > 1$ and $\beta_i < 0$, in which case the contract corresponds to a franchise contractual arrangement.

[13] Holmstrom and Milgrom (1987) show that with continuous efforts in a dynamic setting the optimal contract is linear in the final outcome. Carroll (2015) also demonstrates that linear contracts are optimal with limited liability and risk neutrality, particularly when the principal is uncertain about the agent's available technology.

any co-worker structure and define it by a fixed and exogenous network $\mathbf{G}$.[14] The parameter λ captures the strength of peer effects. If $\lambda > 0$, then actions are strategic complements; if $\lambda < 0$, then actions are strategic substitutes. When $\lambda = 0$ the model reduces to the classical textbook model in Bolton and Dewatripont (2004).

Workers are risk averse with constant absolute risk aversion (CARA) parameter r:

$$u_i(\mathbf{e}, \mathbf{G}, X; \alpha_i, \beta_i) = -\exp\left[-r\left(\omega_i(X; \alpha_i, \beta_i) - \psi_i(\mathbf{e}, \mathbf{G})\right)\right].$$

Since wages are linear and output is normally distributed, expected utility takes a tractable form as

$$\mathbb{E}[u_i(\mathbf{e}; \mathbf{G}; \alpha_i, \beta_i)] \equiv -\exp\left[-r \ \mathrm{CE}_i(\mathbf{e}; \mathbf{G}, \alpha_i, \beta_i)\right],$$

where the certain equivalent of agent i, CE_i, is defined as:

$$\mathrm{CE}_i(\mathbf{e}; \mathbf{G}, \alpha_i, \beta_i) = \beta_i + \alpha_i \sum_{j \in N} e_j - \frac{1}{2} e_i^2 + \lambda e_i \sum_{j \in N} g_{ij} e_j - \alpha_i^2 \frac{r\sigma^2}{2}. \tag{2}$$

The above functional form is conveniently analogous to the utility functions proposed by Ballester et al. (2006) and Calvó-Armengol et al. (2009), with an additional term correcting for risk. The last term captures how adding risk into workers' compensation (through α_i) decreases individual welfare.

If a contract (α_i, β_i) is acceptable, worker i will optimally choose the effort level that maximizes expected utility, taking all other workers' equilibrium effort levels as given,

$$e_i^\star \in \arg\max_{\hat{e}_i \in \mathbb{R}_+} \mathrm{CE}_i(\hat{e}_i, \mathbf{e}_{-i}^\star).$$

A worker accepts the contract only if the certain equivalent in equilibrium is greater than or equal to her reservation utility, U_i,

$$\mathrm{CE}_i(\mathbf{e}) \geq U_i.$$

We take U_i as exogenous and fixed. We consider therefore a situation in which the firm has all bargaining power and essentially makes a take-it-or-leave-it offer to the worker.[15] The firm will select a contract profile (represented by vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$) that maximizes expected

[14] The network is allowed to be weighted and directed. Among other things, a weighted network allows us to capture *congestion effects*, whereby a supervisor with many subordinates might have less influence on each of them. This could be captured by defining the network $\mathbf{G}_D = diag(\mathbf{G}\mathbf{1})^{-1}\mathbf{G}$.

[15] A natural extension considers how the optimal contract looks like when firms compete for workers in different industrial structures. We leave this for future work.

profits subject to these constraints:

$$\max_{\boldsymbol{\alpha},\boldsymbol{\beta}} \mathbb{E}[\pi(\mathbf{e} \mid \boldsymbol{\alpha},\boldsymbol{\beta})] = \sum_{i\in N} e_i - \sum_{i\in N} \omega_i$$

$$\text{subject to: } \mathrm{CE}_i(\mathbf{e}) \geq U_i,\ \forall i \in N \quad \text{(IR)}$$

$$e_i \in \arg\max_{\hat{e}_i\in\mathbb{R}_+} \mathrm{CE}_i(\hat{e}_i, \mathbf{e}_{-i}),\ \forall i \in N \quad \text{(IC)}$$

Main Modeling Assumption

Before solving for the optimal contract, we briefly discuss why we choose to model peer effects through workers' cost structure, as defined in equation (1). Productivity spillovers can also be modeled on the production side by assuming, for instance, that $X(\mathbf{e}) = \sum_i e_i + \lambda \sum_{ij} g_{ij}e_ie_j + \varepsilon$. Although both of these approaches capture incentive spillovers across workers, they differ in important aspects. In the latter case, how much i is influenced by her co-workers' decisions depends on the contract itself because output affects workers' payoffs through the wage. This implies that the firm's contract decision in turn affects peer effect strength across workers, making it difficult to disentangle each effect and precluding closed-form incentive rules.[16] On the other hand, by modeling spillovers in the cost function (1), the strength of co-worker interaction is always modulated by λ, no matter the wage contract chosen by the firm. This allows us to understand more plainly how contracts respond to spillovers and how investments in peer complementarities (i.e. raising λ) affect contracts and profits.

2.2 Optimal Incentive Contracts

We now characterize the optimal contract. Consider first the optimal effort decision of the worker for any contract (α_i, β_i). Workers play a non-cooperative game similar to that in Ballester et al. (2006). The best-reply function of worker i is given by

$$e_i^\star(\mathbf{e}_{-i}) = \alpha_i + \lambda \sum_{j\in N} g_{ij}e_j, \quad \forall i \in N. \quad (3)$$

Notice that the contract's fixed payment β_i has no effect on workers' effort incentives. A worker is motivated to work only through performance-based compensations α_i, and by the

[16] Recent work by Dasaratha et al. (2024) makes progress in this direction analyzing a different incentive environment.

actions of peers. Any Nash equilibrium effort profile $\mathbf{e}^\star$ satisfies

$$(\mathbf{I} - \lambda\mathbf{G})\,\mathbf{e}^\star = \boldsymbol{\alpha}. \tag{4}$$

As is common in these network games, an interior solution is guaranteed as long as the strength of complementarities is bounded above.

Assumption 1. *The largest eigenvalue of $\lambda\mathbf{G}$ is less than 1.*

Assumption 1 guarantees that equation (4) is a necessary and sufficient condition for best-responses and ensures that the Nash equilibrium is unique. Under this assumption, the unique Nash equilibrium effort profile $\mathbf{e}^\star$ of the game can be characterized by:

$$\mathbf{e}^\star = (\mathbf{I} - \lambda\mathbf{G})^{-1}\boldsymbol{\alpha}.$$

In what follows we let $\mathbf{C} := (\mathbf{I} - \lambda\mathbf{G})^{-1}$, such that $\mathbf{e}^\star = \mathbf{C}\boldsymbol{\alpha}$. Finally, notice that the firm can set fixed payments β_i in order to extract all surplus from workers, such that $\text{CE}_i(\mathbf{e}) = U_i$. We assume that U_i are all sufficiently low such that the firm wants to hire all workers and we normalize U_i to 0.[17] We can therefore rewrite the firm's problem as:

$$\begin{aligned} &\max_{\boldsymbol{\alpha},\boldsymbol{\beta}} \ \mathbb{E}[\pi(\mathbf{e} \mid \boldsymbol{\alpha}, \boldsymbol{\beta})] \\ &\text{subject to:} \quad \text{CE}_i(\mathbf{e}) = 0, \ \forall i \in N \qquad \text{(IR)} \\ &\qquad\qquad\qquad \mathbf{e}^\star = \mathbf{C}\boldsymbol{\alpha} \qquad \text{(IC)} \end{aligned}$$

Solving the firm's problem we obtain an explicit characterization of optimal wage contracts for any peer-network $\mathbf{G}$. To ensure that the firm's problem is a concave optimization problem we must bound peer effects from above, as we did for the worker's problem in Assumption 1. It turns out that the firm's problem requires further restrictions on λ.

Assumption 2. *The spectral radius of $\lambda^2/(1 + r\sigma^2)(\mathbf{GC})'\mathbf{GC}$ is less than 1.*

Recall that, by Assumption 1, $\mathbf{C} := (\mathbf{I} - \lambda\mathbf{G})^{-1} = \sum_{q=0}^{\infty} \lambda^q \mathbf{G}^q$ and therefore $\mathbf{C}'\mathbf{1}$ corresponds to the vector of *Bonacich centralities*, where the i-th component, $b_i(\lambda)$, aggregates all incentive paths (of all lengths) emanating from worker i.[18] The following result shows

[17] In Supplementary Appendix B we solve a fully general model with varying U_i.

[18] Notice that we are taking the column sum of $\mathbf{C}$ (rather than the usual row sum) because we are interested in all "outgoing paths" from j, which precisely capture the agents j can influence. See Ballester et al. (2006) and Jackson (2008) for more details on Katz-Bonacich and related measures of centrality in graphs.

that the optimal incentive rule, $\boldsymbol{\alpha}^\star$, is a linear transformation of this vector.[19]

Proposition 1 (Optimal Incentives Rule)**.** *Under Assumptions 1 and 2, there exists a unique profit-maximizing incentive rule for any peer network* $\mathbf{G}$ *given by*

$$\boldsymbol{\alpha}^\star = \mathbf{W}\mathbf{C}'\mathbf{1} \tag{5}$$

where $\mathbf{W} = \frac{1}{1+r\sigma^2}\left[\mathbf{I} - \frac{\lambda^2}{1+r\sigma^2}(\mathbf{GC})'\mathbf{GC}\right]^{-1}$ *and* $\mathbf{C} = (\mathbf{I} - \lambda\mathbf{G})^{-1}$.

What is the intuition behind Proposition 1? The firm anticipates that, at the margin, worker s will respond to i's incentives in proportion to how many *incentive paths* (of any length) lead to s starting from i:[20]

$$\frac{\partial e_s}{\partial \alpha_i} = \sum_{q=0}^{\infty} \lambda^q \mathbf{G}^q_{si} \geq 0 \tag{6}$$

If $\frac{\partial e_s}{\partial \alpha_i} > 0$ we say that s is an *incentive target* of i. The total marginal benefit of α_i is therefore obtained by summing across all of i's targets. This equals the *Bonacich centrality measure*, which captures the total *amplification potential* of worker i:

$$MB_{\alpha_i} = \sum_s \frac{\partial e_s}{\partial \alpha_i} = \sum_s \sum_{q=0}^{\infty} \lambda^q \mathbf{G}^q_{si} := b_i(\lambda).$$

On the other hand, α_i also affects the total wage costs of the firm because it raises i's targets' effort costs, and it exposes i to additional risk. The marginal costs associated to α_i are:

$$MC_{\alpha_i} = \underbrace{\sum_s e_s \frac{\partial e_s}{\partial \alpha_i} - \lambda \sum_s \sum_\ell g_{s\ell}\left(e_s \frac{\partial e_\ell}{\partial \alpha_i} + e_\ell \frac{\partial e_s}{\partial \alpha_i}\right)}_{\text{effort costs}} + \underbrace{\quad r\sigma^2\alpha_i \quad}_{\text{risk exposure}}. \tag{7}$$

Proposition 1 shows how the firm balances the marginal benefit and marginal cost across all α_i. It turns out that optimal incentives are a weighted average of everyone's centrality:

$$\alpha_i^\star = \sum_{j\in N} w_{ij}(\lambda, \sigma^2)\, b_j(\lambda), \quad \forall i \in N. \tag{8}$$

where the weight w_{ij} captures how much j's centrality matters for i's incentives, and is obtained as the (i, j) element of the symmetric matrix $\mathbf{W}$. This weight reflects how much

[19] Details about $\boldsymbol{\beta}$ are in the Appendix, as it merely ensures workers reach their reservation utility. Its form, shaped by outside options and $\boldsymbol{\alpha}$, is uninformative.

[20] To see this take a derivative of the IC constraint.

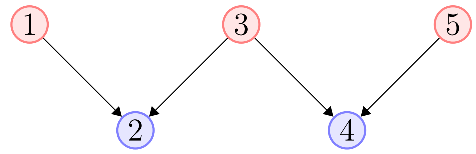

Figure 2: A directed network with 5 agents.

common influence worker i shares with j: $w_{ij} > 0$ if i and j jointly influence a third worker, or if i shares common influence with someone who shares common influence with j, and so on along chains of any length.[21]

Example 1 (Chains of Common Influences)**.** *Consider the network in Figure 2. Worker* 2 *is influenced by workers* 1 *and* 3*, and worker* 4 *is influenced by workers* 3 *and* 5*. Entries* $w_{1,3} = w_{3,1} > 0$ *because workers* 1 *and* 3 *jointly influence worker* 2*. Similarly* $w_{3,5} = w_{5,3} > 0$ *because workers* 3 *and* 5 *jointly influence* 4*. More surprisingly,* $w_{1,5} = w_{5,1} > 0$ *since workers* 1 *and* 5 *exert common influence on workers* 2 *and* 4 *via their common influence with worker* 3*. Finally, notice that workers* 2 *and* 4 *have no common influence with anybody:* $w_{2,i} = w_{4,j} = 0$ *for all* $i \neq 2$ *and* $j \neq 4$ *and* $w_{2,2} = w_{4,4} = 1$*.*

Why do common influences shape the allocation of incentives? Notice that MC_{α_i} in equation (7) doesn't just depend on i's incentive paths to s: $\partial e_s / \partial \alpha_i$. It depends on the interaction term $e_s \cdot \partial e_s / \partial \alpha_i$. At the same time, e_s depends on many workers' incentives since $e_s = \sum_j \alpha_j \frac{\partial e_s}{\partial \alpha_j}$, as seen by differentiating the IR constraint. Substituting this in equation (7), the firm's profit-maximizing condition, $MB_{\alpha_i} = MC_{\alpha_i}$, can be written in terms of incentive paths as follows:

$$b_i(\lambda) = \sum_s \sum_j \alpha_j \frac{\partial e_s}{\partial \alpha_j} \frac{\partial e_s}{\partial \alpha_i} - \lambda \sum_s \sum_\ell g_{s\ell} \sum_j \alpha_j \left(\frac{\partial e_s}{\partial \alpha_j} \frac{\partial e_\ell}{\partial \alpha_i} + \frac{\partial e_\ell}{\partial \alpha_j} \frac{\partial e_s}{\partial \alpha_i} \right) + r\sigma^2 \alpha_i, \quad \forall i \in N. \tag{9}$$

Proposition 1 solves this system of equations and returns $\boldsymbol{\alpha}^\star$.[22] System (9) reveals how incentives complement and oppose each other in the firm's problem. The first term on the right hand side shows that incentivizing any worker j who shares incentive targets with

[21] Note that $(\mathbf{GC})_{ij}$ contains the same information as $\mathbf{C}_{ij}$ but ignores walks of length zero. Thus, $(\mathbf{GC})'\mathbf{GC}$ forms a symmetric matrix where each (i, j) element sums over workers indirectly influenced by both i and j. Under Assumption 2, the weight matrix $\mathbf{W}$ can be expressed as a geometric series of these common-influence matrices:

$$\mathbf{W} = \frac{1}{1 + r\sigma^2} \sum_{k=0}^{\infty} \left(\frac{\lambda^2}{1 + r\sigma^2} \right)^k ((\mathbf{GC})'\mathbf{GC})^k .$$

Each power $((\mathbf{GC})'\mathbf{GC})^k$ tracks the (discounted) common influence of two workers through k intermediaries.

[22] In fact, notice that by stacking up the N equations in (9) we get $\mathbf{C'1} = \mathbf{W^{-1}}\boldsymbol{\alpha}$ which gives the incentive rule in (5).

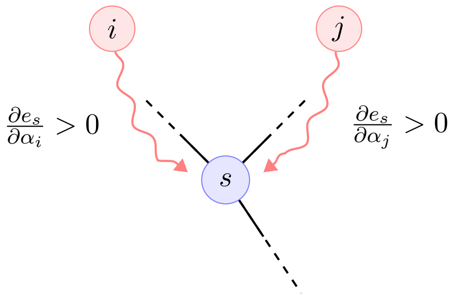

Figure 3: ***Substitutes:*** *Workers i and j both have incentive paths leading to worker s. This means that raising α_i increases the marginal cost of α_j.*

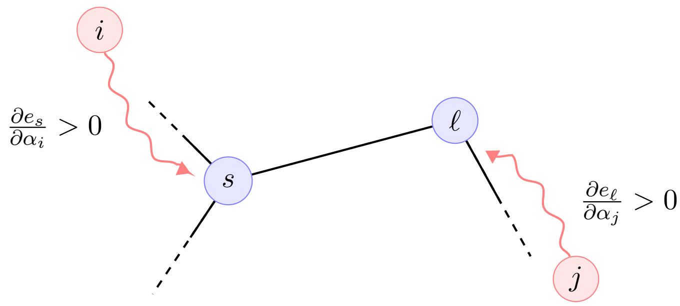

Figure 4: ***Complements:*** *Worker i has an incentive path leading to worker s, who is a neighbor of j's incentive target, ℓ. This means that raising α_i decreases the marginal cost of α_j.*

i raises α_i's marginal costs. This is because incentivizing j pushes j's targets to exert more effort, which means that they become less responsive to additional incentives from i. This implies that α_j and α_i are substitutes (see Figure 3). The second term shows that incentivizing a worker j who influences the *neighbors* of i's targets (rather than the targets themselves) lowers α_i's marginal costs. This is because, incentivizing j again pushes j's targets to exert more effort, which through complementarity links makes their neighbors more responsive to additional incentives. This includes i's target so α_i and α_j are complements (see Figure 4). Workers can simultaneously substitute and complement each other; the total effect depends on the relative length of their incentive paths, as captured by the terms $\partial e_s/\partial \alpha_i$ defined in equation (6). The following example computes these effects and finds α_i in a simple example.

Example 1 Continued. *Consider α_2 in the network in Figure 2. Since worker 2 is her own incentive target, any worker who also targets 2 will raise MC_{α_2}. Using the fact that $\frac{\partial e_i}{\partial \alpha_i} = 1$ for all i, the first term on the right of equation (9) is $\alpha_1 \frac{\partial e_2}{\partial \alpha_1} + \alpha_3 \frac{\partial e_2}{\partial \alpha_3} + \alpha_2$. This captures how others' incentives raise MC_{α_2}. On the other hand, α_1 and α_3 also raise the effort of workers 1 and 3 respectively. These are adjacent to worker 2, so this reduces 2's marginal effort cost through peer effects: the second term on the right of equation (9) is $-\lambda(\alpha_1 + \alpha_3)$. Finally, $b_2(\lambda) = 1$ because nobody "listens" to 2. Therefore, equation (9) can be written as*

$$1 = \alpha_1 \left(\frac{\partial e_2}{\partial \alpha_1} - \lambda \right) + \alpha_3 \left(\frac{\partial e_2}{\partial \alpha_3} - \lambda \right) + \alpha_2 + r\sigma^2 \alpha_2.$$

Looking at the network in Figure 2, it is clear that $\frac{\partial e_2}{\partial \alpha_1} = \frac{\partial e_2}{\partial \alpha_3} = \lambda$ because there is only one incentive path of length 1 between 1 and 2 (and between 3 and 2). Therefore, for this particular example, the substitute effect and the complement effect associated with α_1 and α_3

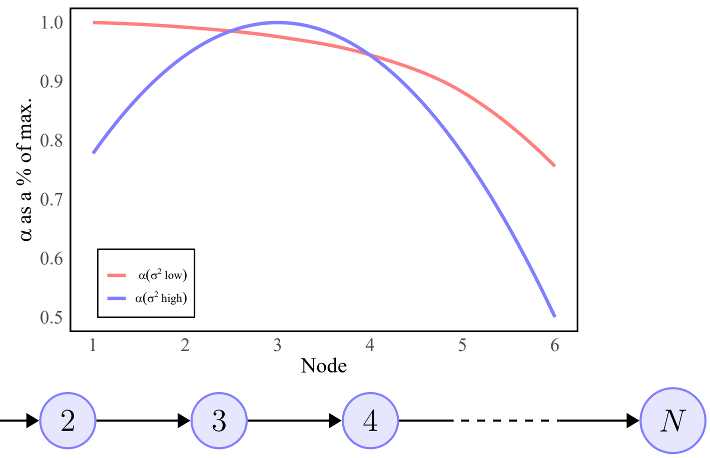

Figure 5: In the supermarket chain network from Mas and Moretti (2009) it can pay off (if risk is low) to incentivize nodes in the middle more than nodes on the left, even though nodes on the left are more central.

cancel each other out, yielding

$$\alpha_2 = \frac{1}{1 + r\sigma^2}.$$

Interestingly, since the network effects exactly cancel, worker 2 ens up receiving the textbook-level of incentives for the standard textbook case with no peer effects, covered in Bolton and Dewatripont (2004) and obtained below as a corollary of Proposition 1

Corollary 1 (No Peer Effects)**.** *In the absence of peer effects (i.e. $\lambda = 0$) incentives are constant across workers and equal to*

$$\alpha_i^\star = \frac{1}{1 + r\sigma^2}, \quad \forall i \in N.$$

We can see this result directly from equation (9). Since no incentive paths exist when $\lambda = 0$ (i.e. $\partial e_s/\partial \alpha_i = 0$ for all $i \neq s$), then no amplification is possible and $b_i(0) = 1 \, \forall i$. Equation (9) boils down to $1 = \alpha_i + r\sigma^2\alpha_i$, which yields the classical result. Notice that, without risk, the principal effectively franchises the firm to each worker (i.e. $\alpha_i = 1$).

We are most interested in understanding how the presence of peer effects ($\lambda > 0$) interacts with fundamental risk ($\sigma^2 > 0$) in distributing incentives. To develop intuition behind equation (9) consider a simple "line network" shown in Figure 5.[23] Notice that $b_1(\lambda) > b_2(\lambda) > \ldots > b_n(\lambda)$ for all $\lambda > 0$, which means that the marginal benefit of α_i always exceeds the marginal benefit of α_j for $j > i$. This might suggest that incentives should decrease as

[23] This network corresponds to the peer effect structure found by Mas and Moretti (2009) in their analysis of check-out clerks in US supermarkets.

we move to the right of the chain. However, when risk is low, the firm finds it optimal to provide worker 3 with larger incentives than 1. The reason is that the marginal costs also differ. Although worker 3 has strictly less incentive-targets than 1, she's "closer" to her targets than 1 is, which implies stronger complementarity effects (see Figure 4). Moreover, because 3's incentive-targets are a subset of 1's, she generates less susbtitutability effects (see Figure 3). We show in Supplementary Appendix E that both of these effects push α_3 to have a lower impact on firms' total wage costs than α_1 (i.e. $MC_{\alpha_3} < MC_{\alpha_1}$) and therefore a higher α.[24] Intuitively, when risk is low, the firm finds it optimal to "spread" incentives more locally in order to exploit shorter incentive paths – which improve complementarities – and avoid central nodes that share many common targets with others.

On the other hand, when risk is high, every extra unit of α_i is expensive, no matter where i is in the network. Now, the firm saves more in risk compensation costs by concentrating incentives on a few workers and therefore prefers to load incentives to those workers with more amplification potential (see red line in Figure 5). As a result, when fundamental risk is large incentives are allocated following a *simple* rule.

Proposition 2 (Monotonicity)**.** *There exists a value* $\bar{\sigma}^2(\mathbf{G})$ *such that if* $\sigma^2 > \bar{\sigma}^2$ *optimal incentives are a monotonic transformation of centrality:* $b_i(\lambda) > b_j(\lambda) \iff \alpha_i^\star > \alpha_j^\star$.

Comparative Statics: The Strength of Connections

How should firms react to changes in the structure of spillovers? One might think that when worker j's influence over i intensifies, the principal should decrease incentives from other workers and concentrate them on i and j's strengthened relationship. We show below that this is not optimal.

Proposition 3 (An Increase in Link Strength)**.** *Every worker's incentive pay* weakly *increases in any link's strength (i.e.* $\partial\alpha_s/\partial g_{ij} \geq 0$ *for all* $i, j, s \in N$*). Moreover, an increase in* g_{ij} strictly *increases the incentive pay of worker* j *and any worker* s *who has a common influence with* j *(i.e.* $w_{j,s} > 0$*).*

Proposition 3 states that incentives and efforts do not decrease as *any* link is strengthened. It turns out the firm recognizes that incentive spillovers travel in all directions, even if the

[24] In fact, Figure 5 also shows that workers 2 and 4 have the second-highest level of incentives, followed by workers 1 and 5. This symmetry around α_3 might seem strange at first. It stems from the fact that, for this particular line network, the additional marginal benefit that you get from 2's greater amplification potential are exactly compensated by 4's stronger complementarity and lower marginal cost. Therefore we get that $\alpha_2 = \alpha_4 < \alpha_3$ and similarly that $\alpha_1 = \alpha_5 < \alpha_2 < \alpha_3$.

network is directed, and thus allocates more incentives to all influential workers that share influence with j. Worker j's incentives will increase more than other workers.

Example 1 Continued. *Consider the network in Figure 2. If worker* 1*'s influence on* 2 *increases (i.e.,* g_{21} *increases), Proposition 3 implies that incentives increase for workers* 1*,* 3*, and* 5*. Intuitively, bolstering* g_{21} *strengthens all incentive paths that use that link. This rises* 1*'s centrality and increases* MB_{α_1} *directly. Moreover, by equation (9), rising* g_{21} *rises* MC_{α_i} *for all* i *that target* 2 *and lowers it for all* k *that target adjacent to* 1 *or* 2*. In both cases these are workers* 1*,* 3*, and* 2*. As discussed above, these effects perfectly cancel for worker* 2*, yielding* $w_{2,1} = 0$ *and no change in* α_2*. Worker* 4 *is also unaffected, because* 4 *does not have any common influence with* 1 *(*$w_{4,1} = 0$*). Finally, these effects further propagate to anyone with a target adjacent to* 3*, which is worker* 5*, due to second-order spillover effects.*

Heterogeneity in Productivity

We have assumed for simplicity that workers only differ in their connections, but are otherwise identical in productivity, outside options, and risk aversion. In Supplementary Appendix B we provide a general characterization of optimal contracts for any distribution of these parameters. One particularly interesting case considers different productivities, with output defined as $X(\mathbf{e}) = \sum_i \theta_i e_i + \varepsilon$. Since workers have different marginal products, the marginal benefit of α_i must now weigh incentive paths by the productivity of their respective targets: $MB_{\alpha_i} = \sum_j \theta_j \frac{\partial e_j}{\partial \alpha_i}$. Therefore, the incentive rule under heterogeneity now aggregates a *productivity-weighted* centrality: $\boldsymbol{\alpha}^\star = \mathbf{W}_\theta \mathbf{C}' \boldsymbol{\theta}$ where $\mathbf{W}_\theta$ is a modified weighting matrix of common influence paths, similarly weighted by $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_n)'$. Interestingly, when $\sigma^2 = 0$ optimal incentives correspond to Proposition 1 because θ_i affects marginal benefits and marginal costs equally. Thus, in the absence of risk, differences in individual worker productivity don't lead to differences in incentives.

Optimal Incentives with Negative Spillovers

We have focused on peer-to-peer complementarities ($\lambda > 0$) but many workplaces experience negative spillovers ($\lambda < 0$), such as free-riding documented by Amodio and Martinez-Carrasco (2018) and Bandiera et al. (2005). Propositions 1 and 2 apply in both cases, though the identity of workers receiving incentives changes with λ. For example, in Figure 6, nodes most central under strategic complements are least central under strategic substitutes, flipping incentive allocations.

Figure 6: ***Panel A:*** *Positive peer effects (i.e., $\lambda > 0$);* ***Panel B:*** *Negative peer effects (i.e., $\lambda < 0$). The size of the node represents their Bonacich centrality and the color represents the allocation of incentives (red being most incentives and blue being least).*

Recent work by Galeotti et al. (2020) shows that when $\lambda < 0$, optimal interventions alternate incentives between adjacent individuals due to convex costs. Their model assumes quadratic costs for adjusting incentives, $\sum_{i \in N}(\alpha_i - \hat{\alpha}_i)^2$. In contrast, our model's costs depend on the wage risk, expressed as $\frac{r\sigma^2}{2}\sum_{i \in N}(\alpha_i^2 - \hat{\alpha}_i^2)$. This difference is significant: lowering incentives in our model reduces costs by decreasing risk, making it unprofitable to exploit alternating patterns. Consequently, the principal's optimal strategy diverges from the alternating structure in Galeotti et al. (2020).

Optimal Incentives with Partial Information

We have assumed that the firm knows the entire network structure and can fully condition on it when designing optimal contracts. There are two main directions in which to relax this assumption. The first is that the principal may know the entire network but may be unable to write contracts that fully discriminate across workers. This is the approach we take in Section 3 where we force firms to offer the same contract to entire sections of the workforce. The second approach is to assume that the firms may have only partial information about the relevant peer-to-peer network.

To address this, in Supplementary Appendix F, we modify the firm's problem and derive the optimal incentive rule as a function of the parameters in a simple random-graph model, rather than a realized network $\mathbf{G}$. Specifically we assume that, when designing the contract, the firm knows the linking probabilities between each pair of workers, but not the realized linking structure. Using a mean-field approximation, we provide a closed-form representation of the incentive rule in this case and show that worker i's incentives depend on her expected connections, but not on the likelihood that i's connections are themselves connected.

3 Wage Benchmarking

In this section, we consider allocating incentives when firms can't write personalized contracts. To do this we divide the workforce into occupational categories, and we assume that the firm must offer the same linear wage contract to every member of that category – everything else remains exactly as in Section 2.

This section is motivated by evidence that recent hiring practices have compressed wage distributions across firms and occupations. Pay transparency tools reduced wage variance by 20% among 100,000 US academics (Obloj and Zenger, 2022) and had a similar effect on California city managers after a 2010 salary disclosure mandate (Mas, 2017).[25] Using national payroll data, Cullen (2024) shows that salary dispersion drops 25% when firms access market pay benchmarks, with most compression (40% vs. 15%) occurring in low-skilled roles. Finally, Hazell et al. (2022) find that 40–50% of job postings within firms offer identical wages across locations despite varying local conditions.

Since contracts can't perfectly discriminate, the firm no longer extracts all surplus from its workforce. More importantly, the peer-effects network will determine exactly how surplus is distributed, with more "central" workers extracting larger rents from the employer. To see this, we assign N workers into $K \leq N$ groups (or occupational categories) and we assume that the firm must offer the same wage contract to all agents in category $k \in K$:

$$\omega_i = \beta_k + \alpha_k X, \quad \forall i \in k.$$

We allow for any level of granularity in the contract because we make no restrictions on how to partition workers into categories.[26] Define the *group assignment matrix* $\mathbf{T}_{K\times N}$ such that $\mathbf{T}_{ki} = 1$ if worker $i \in N$ is assigned to occupational category $k \in K$, and zero otherwise.[27] Define $\hat{\boldsymbol{\alpha}}$ as the $K \times 1$ vector of incentives chosen by the firm: the k-th term, $\hat{\alpha}_k$, corresponds to the value of α offered to all workers in group k. We can now relate $\hat{\boldsymbol{\alpha}}$ to the full N-vector of incentives by the following simple relation: $\boldsymbol{\alpha} = \mathbf{T}'\hat{\boldsymbol{\alpha}}$ (and similarly for $\hat{\boldsymbol{\beta}}$). It follows that the Incentive Compatibility (IC) constraint can be obtained from the $K \leq N$ contracts as, $\mathbf{e}^\star = \mathbf{C}\mathbf{T}'\hat{\boldsymbol{\alpha}}$.

When contracts are coarse ($K < N$) the Individual Rationality (IR) constraints look differ-

[25] See Cullen (2024) for a review of the wage transparency literature.

[26] As a partition, each worker must be assigned to one and only one category.

[27] For example, if in a law firm workers 1, 2, and 3 are in one occupational category (e.g. paralegals) and workers 4 and 5 are in another (e.g. senior partners), the group assignment matrix is $\mathbf{T}_{2\times 5} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}$.

ent.[28] Firms no longer extract full rents from all workers because if a group contains multiple workers occupying different network positions, then a single contract cannot simultaneously guarantee that everyone is exactly compensated their reservation utility. We assume that the firm sets β_k such that no one in group k rejects the contract. This means that every worker will extract (weakly) positive rents from their contract, and only the "highest-cost worker" will receive her reservation utility. Let $\overline{\psi}_k = \max_{i \in k} \psi_i$ represent the highest effort cost in group k, where $\psi_i = \frac{1}{2}e_i^2 - \lambda e_i \sum_j g_{ij} e_j$ is worker i's cost of effort defined in equation (1), and let $\bar{i}(k)$ represent the worker with highest cost in group k.[29] We can relate β_k to the effort costs of each worker $i \in k$:

$$\beta_k = \frac{r\sigma^2}{2}\alpha_k^2 - \alpha_k \sum_i e_i + \psi_i + \underbrace{(\bar{\psi}_k - \psi_i)}_{\eta_i}, \quad \text{for } i \in k.$$

The last term, η_i, is new and represents the *centrality rents* that worker i now extracts as a result of having lower effort costs than $\bar{i}(k)$. Following equation (2), this implies that $\text{CE}(\boldsymbol{\alpha}, \mathbf{G}, \mathbf{T})_i \geq 0$ for all $i \in k$ and $\text{CE}(\boldsymbol{\alpha}, \mathbf{G}, \mathbf{T})_{\bar{i}} = 0$ for $\bar{i}(k)$. We can now re-write the Principal's problem under coarse contracts using modified (IR) and (IC) constraints as:

$$\begin{aligned} &\max_{\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\beta}}} \mathbb{E}[\pi(\mathbf{e} \mid \boldsymbol{\alpha}, \boldsymbol{\beta})] \\ \text{subject to: } & \text{CE}_i(\boldsymbol{\alpha}, \mathbf{G}, \mathbf{T}) - \eta_i = 0, \quad \forall i \in N \qquad \text{(IR)} \\ & \eta_i \geq 0, \quad \forall i \in N \\ & \mathbf{e} = \mathbf{C}\mathbf{T}'\hat{\boldsymbol{\alpha}} \qquad \text{(IC)} \end{aligned}$$

Proposition 4 (Wage Benchmarking)**.** *The optimal allocation of incentives under wage benchmarking with group assignment* $\mathbf{T}$ *is given by:*

$$\hat{\boldsymbol{\alpha}}^\star = \left(\mathbf{T}\mathbf{W}^{-1}\mathbf{T}'\right)^{-1}\mathbf{T}\mathbf{C}'\mathbf{1}. \tag{10}$$

Equation (10) differs from our original incentive rule in Proposition 1 in a very natural way. Incentives are now allocated at the group-level following an aggregated centrality measure, $\mathbf{TC}'\mathbf{1}$, which simply sums the centrality of all members in each group. The weighting matrix is similarly transformed and aggregates the marginal cost terms MC_{α_k} of all members in each group. In other words, our incentive rule is just like before, if we take the groups as if they

[28]Notice that if $K = N$ (i.e. if each worker is allowed to have a different job-title/contract) then $\mathbf{T} = \mathbf{I}$ and we get the original setup of Section 2

[29]There may be multiple workers with highest cost in group k. It is not important which of these is identified by $\bar{i}(k)$. Notice that, $\bar{\psi}_k = \psi_{\bar{i}(k)}$

were the individual workers.[30] Notice that in the absence of peer effects, coarse contracts should coincide with baseline, because everyone receives the same salary (as seen in Corollary 1). We can confirm this: if $\lambda = 0$, then $\hat{\boldsymbol{\alpha}}^\star = ((1 + r\sigma^2)\mathbf{T}\mathbf{T}')^{-1}\mathbf{T}\mathbf{1}$, and $\alpha_i^* = 1/(1 + r\sigma^2)$, as expected.[31]

3.1 Loss in Surplus due to Wage Benchmarking

Coarse contracts generate a loss in overall surplus akin to the efficiency loss that a monopolist generates when it cannot perfectly price discriminate on all infra-marginal units. In this section, we develop an intuitive way of measuring the surplus-loss generated by any group assignment $\mathbf{T}$ on any peer network $\mathbf{G}$.

Surplus, S, corresponds to the profits collected by the firm and the consumption-equivalent units of expected utility received by agents:

$$S = \mathbb{E}(\pi) + \sum_i \mathrm{CE}_i \,.$$

Notice that in Section 2 the second term was zero because the firm guaranteed each worker their reservation utility. In that case, surplus reduced to profits: $S = \mathbb{E}(\pi) = \frac{1}{2}\sum_i e_i$. With coarse contracts, surplus is split between firms and workers, but we can still find a convenient expression for it.

Lemma 1 (Surplus with Coarse Contracts)**.** *With wage benchmarking, a firm's profits in expectation are maximized at one-half of equilibrium output* ***minus*** *the sum of agents' centrality rents:*

$$\mathbb{E}(\pi^\star(\mathbf{e}|\boldsymbol{\alpha}, \boldsymbol{\beta})) = \frac{1}{2}\sum_i e_i - \sum_i \eta_i$$

Therefore since $\mathrm{CE}_i = \eta_i$ *by the (IR) constraint above, surplus is given by:*

$$S = \frac{1}{2}\sum_i e_i. \tag{11}$$

for any group assignment $\mathbf{T}$ *and any peer network* $\mathbf{G}$

Define $\mathbf{e}_\mathbf{T} := \mathbf{C}\mathbf{T}'\hat{\boldsymbol{\alpha}}$ as the vector of equilibrium effort contributions if contracts are coarse and groups are assigned according to $\mathbf{T}$. Then, the loss in surplus between group assignment

[30] When $\mathbf{T} = \mathbf{I}$ equation (10) is identical to equation (5), as expected.

[31] Notice that $\mathbf{T}\mathbf{T}'$ is a $K \times K$ diagonal matrix with the size of each group along the diagonal. Then, $(\mathbf{T}\mathbf{T}')^{-1}$ is a diagonal matrix with (k, k)th element equal to $1/n_k$, where n_k is group k's size. Thus, $(\mathbf{T}\mathbf{T}')^{-1}\mathbf{T}\mathbf{1} = \mathbf{1}$.

Figure 7: ***Panel A:*** *Within-group variance is zero. No surplus loss.* ***Panel B:*** *Within-group variance is 0.53. Surplus loss, following Proposition 5, is about 0.85.*

$\mathbf{T}$ and group assignment $\widetilde{\mathbf{T}}$ is given by:

$$\Delta S_{\widetilde{\mathbf{T}}-\mathbf{T}} = \frac{1}{2}\,\mathbf{1}'(\mathbf{e}_{\widetilde{\mathbf{T}}} - \mathbf{e}_{\mathbf{T}}).$$

The following result focuses on $\Delta S_{\mathbf{I}-\mathbf{T}}$, which corresponds to the loss of surplus when groups are assigned by $\mathbf{T}$, relative to fully personalized contracts, $\mathbf{I}$. We show that the efficiency loss can be measured easily by computing within-group dispersion in centrality.

Proposition 5 (Loss in Surplus)**.** *The surplus lost due to wage benchmarking with group assignment* $\mathbf{T}$ *is proportional to the sum of within-group variances in Bonacich centrality, weighted by group size:*

$$\lim_{\frac{\lambda}{r\sigma^2}\to 0} \Delta S_{\mathbf{I}-\mathbf{T}} = \frac{1}{1+r\sigma^2}\sum_{k\in K} n_k \operatorname{Var}(\mathbf{b}_k),$$

where $\mathbf{b}_k$ *is the (sub)vector of Bonacich centralities for workers in group* k*.*

Intuitively, if $\operatorname{Var}(\mathbf{b}_k) = 0$ for all $k \in K$, then all workers with the same job title are equally central, making uniform contracts within each group optimal and coarse contracts are efficient. Therefore, *regular networks*—where all individuals are identical—incur zero surplus loss. More importantly, Proposition 5 allows for arbitrary differences in centrality across groups—so long as within-group variability is zero, coarse contracts remain efficient. This underscores that the interaction between the peer structure $\mathbf{G}$ and group assignment $\mathbf{T}$ is what truly matters. For instance, Figure 7 illustrates how the same network under different assignments $\mathbf{T}$ leads to starkly different values of $\Delta S_{\mathbf{I}-\mathbf{T}}$.

Proposition 5 also establishes a limiting result for small λ or large σ^2. While this may seem restrictive, recall that λ is bounded above by the network's spectral radius, meaning it must be small in dense graphs. One might still ask whether surplus loss is generally determined by the sum of within-group variances. Figure 8 presents simulation results from Erdős-Rényi random graphs, showing that $\Delta S_{\mathbf{I}-\mathbf{T}}$ tends to rise with $\sum_k n_k \operatorname{Var}(\mathbf{b}_k)$ and converges to an

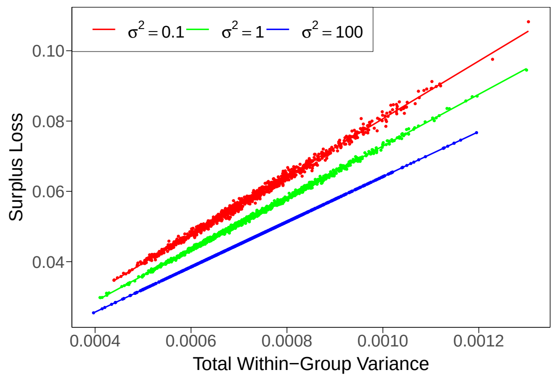

Figure 8: The loss of surplus $\Delta S_{\mathbf{I-T}}$ is increasing in within-group variance in centrality. The relationship is approximately linear and becomes deterministic as σ^2 increases.

exact linear relationship as σ^2 grows. Thus, while proportionality holds strictly only in the limit, the efficiency cost of coarse contracts is well captured by this simple statistic.

4 Modular Production

So far, we have assumed output is a linear function of effort, implying workers' efforts are substitutable despite occupying different positions in the peer network. This allowed us to isolate the impact of peer effects on wages and profits but overlooked how a firm's production function shapes its organization. We now examine how incentive contracts with peer effects are structured when firms have fragmented organizational structures.

We modify our production function to incorporate *modules*, assuming final output depends on the weakest-performing module. Formally, N workers are assigned to K teams, $k_1, k_2, \ldots, k_K$, each responsible for a separate module.[32] Within teams, performance is substitutable, but across teams, it is perfectly complementary.[33] Firm output is now:

$$X(\mathbf{e}) = \min\left\{\sum_{i\in k_1} e_i, \sum_{i\in k_2} e_i, \ldots, \sum_{i\in k_K} e_i\right\} + \varepsilon. \tag{12}$$

This technology is extremely versatile. It nests our original production function as a special

[32] We use K to denote both the number of teams and the set of teams $k \in K$. Moreover, we use $k(i)$ to denote worker i's team.

[33] This assumption simplifies the model while remaining flexible, as we impose no restrictions on module size or composition. Unlike Matouschek et al. (2025), who model modular production as a network of interdependent decisions, our approach avoids introducing an additional "modular production network" alongside the peer effects network. We just have to partition workers into modules.

case in which all workers belong to the same module. However, we can also capture *managerial bottlenecks* within organizations, like when a senior software engineer must approve all code before deployment. Equation (12) can also capture *interdependent production teams*, whereby a small failure in a single critical component of an airplane wing, for instance, compromises the final product.[34]

To find the optimal contract we must re-consider workers' equilibrium effort. In any equilibrium, all modules contribute the same total effort; otherwise, workers in higher-effort modules would benefit by reducing their effort. Thus, $\sum_{i\in k} e_i = \bar{e}$ for all $k \in K$, for some $\bar{e} \geq 0$. The question is which values of $\bar{e}$ constitute a Nash equilibrium? Suppose each module contributes $\bar{e}$ and consider deviations. Increasing effort never benefits a worker, while decreasing it is profitable if the marginal cost (α_i) is less than the marginal benefit ($e_i - \lambda \sum_{j\in N} g_{ij} e_j$). Thus, any equilibrium must satisfy:

$$\alpha_i \geq e_i - \lambda \sum_{j\in N} g_{ij} e_j, \quad \forall i \in N. \tag{13}$$

A contract proposed by the firm allows multiple equilibria. For instance, $e_i = 0, \forall i \in N$ is always an equilibrium.[35] We focus on the *maximal equilibrium* $\hat{e}(\boldsymbol{\alpha})$, where any $\bar{e} > \hat{e}(\boldsymbol{\alpha})$ is not a Nash equilibrium.[36] The firm will never offer a contract $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ leading to a maximal equilibrium $\hat{e}(\boldsymbol{\alpha})$ that does not equate (13).[37] Thus, any optimal wage contract with modular production must satisfy the following two restrictions:

1. $\alpha_i = e_i - \lambda \sum_{j\in N} g_{ij} e_j, \ \forall i \in N.$
2. $\sum_{i\in k} e_i = \hat{e}, \ \forall k \in K.$

Define the *module assignment matrix* $\mathbf{M}_{K\times N}$, such that $\mathbf{M}_{ki} = 1$ if worker $i \in N$ is assigned to module $k \in K$, and zero otherwise.[38] We can now use matrix $\mathbf{M}$ in order to write down

[34] In an extreme and popular example, the *Challenger* NASA spacecraft disaster occurred because a single small metal gasket called the O-ring failed. See Kremer (1993); Garud, Kumaraswamy, and Langlois (2009); Baldwin and Clark (2003) for more details on this and other examples of modular production.

[35] Unilaterally raising e_i never benefits worker i, given everyone else's strategy and the production function.

[36] Firms are assumed to have rational expectations and expect workers to play $\hat{e}(\boldsymbol{\alpha})$. Alternative models with ambiguity (e.g., firms assigning different probabilities to all the equilibria associated with a specific contract) do not change our qualitative results.

[37] For any such wage offer, there exists another $(\tilde{\boldsymbol{\alpha}}, \tilde{\boldsymbol{\beta}})$ with the same $\hat{e}$ that equates (13) and yields higher profits since $\tilde{\boldsymbol{\alpha}} \leq \boldsymbol{\alpha}$ element-wise.

[38] For example, in a firm with workers 1, 2, and 3 in module k_1 and workers 4 and 5 in module k_2, the module assignment matrix is $\mathbf{M}_{2\times 5} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}$, and $\mathbf{Me} = \begin{bmatrix} \sum_{i\in k_1} e_i \\ \sum_{i\in k_2} e_i \end{bmatrix} = \begin{bmatrix} \hat{e} \\ \hat{e} \end{bmatrix} = \hat{e}\mathbf{1}_2$. Notice that matrix $\mathbf{M}$ is defined identically to the group assignment matrix $\mathbf{T}$ in Section 3. We choose to use

the principal's problem as optimizing profits subject to restrictions 1 and 2 above.

$$\max_{\boldsymbol{\alpha}} \mathbb{E}[\pi(\mathbf{e} \mid \boldsymbol{\alpha}, \boldsymbol{\beta})] = \left(\hat{e} - \frac{1}{2}\mathbf{e}'(\mathbf{I} - 2\lambda\mathbf{G})\mathbf{e} - \frac{\sigma^2 r}{2}\boldsymbol{\alpha}'\boldsymbol{\alpha}\right)$$
$$\text{subject to: } \mathbf{e} = \mathbf{C}\boldsymbol{\alpha}$$
$$\mathbf{M}\mathbf{e} = \hat{e}\mathbf{1}_K$$

Proposition 6 (Modular Production)**.** *The optimal allocation of incentives under modular production with module assignment* $\mathbf{M}$ *is given by:*

$$\boldsymbol{\alpha}^\star = \mathbf{W}\mathbf{C}'\boldsymbol{\mu} \tag{14}$$

where $\boldsymbol{\mu} = \mathbf{M}'\mathbf{H}^{-1}\mathbf{1}/(\mathbf{1}'\mathbf{H}^{-1}\mathbf{1})$ *is a* $N\times 1$ *vector of* K *module-specific weights* $(\mu_{k(1)}, \ldots, \mu_{k(n)})'$ *and* $\mathbf{H} = (\mathbf{M}\mathbf{C})\mathbf{W}(\mathbf{M}\mathbf{C})'$.

Proposition 6 fully characterizes optimal incentive allocations under modular production, for any assignment of workers to modules, as captured by matrix $\mathbf{M}$ and any peer-effects network $\mathbf{G}$. In other words, equation (14) generalizes Proposition 1 to any number of modules with arbitrary sizes, and for any structure of incentive links within and across modules. Notice that incentives are aggregated using the same weights $\mathbf{W}$ as before, but the relevant centrality measure now is $\mathbf{C}'\boldsymbol{\mu}$ rather than $\mathbf{C}'\mathbf{1}$.[39]

The intuition behind this incentive rule builds on our earlier discussion surrounding equation (9). Notice that, no matter how one partitions workers into modules, MC_{α_i} remains unchanged; it only depends on the network structure through $\mathbf{W}^{-1}$, as discussed before (see footnote 22). On the other hand, the marginal benefit, MB_{α_i}, changes because the firm's expected revenues correspond to the weakest-performing module. Given that each module contributes $\hat{e}$ in equilibrium, the marginal benefit of efforts are interdependent—raising one worker's effort lowers another's within the same module. Thus, a worker's effort no longer increases revenue 1-to-1; it now depends on which module she belongs to. Specifically, worker j's effort now increases revenue 1-to-$\mu_{k(j)}$, where $\mu_{k(j)}$ is a *modular factor* associated with j's module. Therefore, MB_{α_i} now weighs each of i's incentive paths by the module the target

different notation across sections to emphasize that the K modules in this section have no relation to the K occupational categories of Section 3. Analyzing modular production and wage benchmarking simultaneously is beyond the scope of this paper, and we leave it for future research.

[39] If there is only one module (i.e. if $\mathbf{M} = \mathbf{1}'_N$) then $\boldsymbol{\mu} = \mathbf{1}$ and equation (14) corresponds to the incentive rule from Proposition 1, as expected.

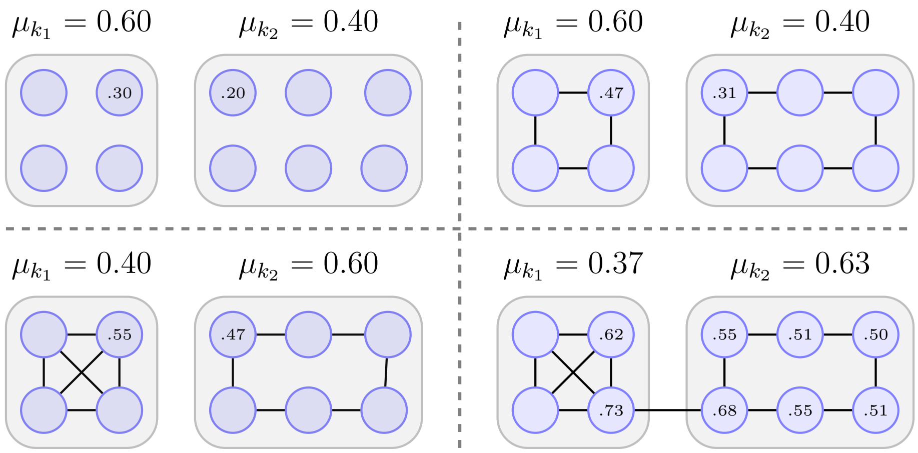

Figure 9: Modular shares, μ_k, and incentives, α_i, in four different firm configurations. Each firm has a 4-worker and 6-worker module. (parameters: $\lambda = 0.15$ and $r\sigma^2 = 1$).

belongs to:

$$MB_{\alpha_i} = \sum_j \frac{\partial e_j}{\partial \alpha_i} \mu_{k(j)}.$$

Stacking these up across all workers, you get the vector of marginal benefits which is precisely the new centrality measure $\mathbf{C}'\boldsymbol{\mu}$ in equation (14).

The values of μ_k (which must sum to 1) are determined by module size and link structure, as captured by $\mathbf{H}^{-1}$ in Proposition 6. As shown in the Appendix, μ_k is the shadow value (Lagrange multiplier) of the constraint ensuring all modules contribute $\hat{e}$ in equilibrium. It reflects the average per-worker cost of implementing $\hat{e}$ in module k, relative to other modules. For instance, with two equally-sized, symmetrically-linked modules, then $\mu_{k_1} = \mu_{k_2} = 0.5$ and incentives are half those in Proposition 1. This is intuitive since each worker relies on half of the workforce to generate marginal revenue.

For asymmetric cases, μ_k varies inversely with module k's aggregate centrality. Figure 9 reports (μ_{k_1}, μ_{k_2}) for four firms with two modules each. The top-left panel shows that, without links, size differences drive μ_k: workers in smaller modules require stronger incentives (relative to larger groups) to achieve a given $\hat{e}$, a result formalized in Corollary 2 below. The top-right panel confirms that symmetric connections preserve μ_k's since the (relative) incentives required to achieve $\hat{e}$ have not changed. However, relative to the previous case, α_i's increase by 56% for everyone, due to uniform centrality gains.[40] The bottom-left panel shows that workers in small but well-connected modules receive lower μ_k than those in larger but poorly connected ones, as it becomes relatively easier for them to generate $\hat{e}$. Conse-

[40]This is computed as the top right α_i's over the top left α_i's, $0.47/0.3 = 0.31/0.2 = 1.56$

quently, incentives rise less (17%) for these workers than for workers in the larger module (52%). Lastly, the bottom-right panel shows that cross-module links have asymmetric effects, boosting incentives more for loosely connected groups.

Finally, this example highlights another key feature of modular production: whereas in the linear technology of Section 2, network components were independent, here incentives depend on the entire network, including links in separate components. To see this, note that incentives for the large module in Figure 9 increase from 0.31 (top-right) to 0.47 (bottom-left), even though the only additional links are formed in a separate component of the peer network. Before proceeding, we examine the incentive rule in detail for two special cases that illustrate how modular technology (12) can represent distinct organizational environments.

Modular Production without Peer Effects

The case without peer effects (i.e. $\lambda = 0$) teaches us a lot already about how modular production drives earnings disparities, and connects with some of the lessons from the literature on knowledge hierarchies (Garicano, 2000; Garicano and Rossi-Hansberg, 2006). More specifically, a technology like (12) can capture managerial bottlenecks and sequential hierarchies, such as when a set of 100 medical officers at the FDA (module 1) prepare reports that must be processed and approved by a senior official (module 2) before drugs can go to market. Even abstaining from peer-to-peer complementarities, this 100-to-1 span-of-control ratio in the firm's production function yields large differences in incentive pay.

Recall that, without peer effects, there are no incentive paths (i.e. $\partial e_i/\partial \alpha_j = 0$ for $i \neq j$ and $\partial e_i/\partial \alpha_i = 1$). Therefore, the marginal benefit from incentivizing worker i simplifies to her own module's productivity weight: i.e. $MB_{\alpha_i} = \mu_{k(i)}$. Following equation (9) the profit-maximizing incentive rules must satisfy $\mu_{k(i)} = (1 + r\sigma^2)\alpha_i$, for some constants $\mu_{k(i)}$ such that $\sum_k \mu_k = 1$. Since the manager alone must validate 100 workers' worth of output, the firm must weigh the manager with a productivity share, μ_k, 100 times larger than that of workers. Thus, relative to Corollary 1, incentive rules are now scaled to capture differences in module size. The following corollary generalizes this intuition.

Corollary 2 (No Peer Effects)**.** *Take $\lambda = 0$. Consider a firm with K modules with sizes $n_1, n_2, \ldots, n_K$ and let $k(i)$ represent worker i's module. Incentives are allocated according to the following rule:*

$$\alpha_i^\star = \frac{1}{1 + r\sigma^2} \frac{1}{n_{k(i)}} \left(\sum_{s \in K} \frac{1}{n_s} \right)^{-1}, \quad \forall\, i \in N. \tag{15}$$

If modules are of equal size, $\alpha_i^\star = \frac{1}{1+r\sigma^2}\frac{1}{K}$, and if all workers form part of one single module,

α_i simplifies to Corollary 1, as expected.

Corollary 2 implies that workers in small modules are compensated more in performance-pay when they collaborate with larger modules, and that the ratio in incentives is inversely proportional to the ratio in module size: $\alpha_i/\alpha_j = n_{k(j)}/n_{k(i)}$. This hereto unknown result in the theory of optimal contracts sheds light on how non-linear technologies can drive large wage disparities (even in the absence of peer effects). It captures the simple fact that workers in small modules are individually valuable, since they are essential in order to validate the output of larger groups.

When Every Worker is Essential

Now imagine a "weakest-link" production function, where every worker is *essential* (i.e. each worker belongs to a separate module). This captures extremely precise production processes that consist of many indispensable steps, whereby a single error halts the entire process. For instance, a small defect rate in extreme ultraviolet lithography (EUV) can cause massive yield losses in semiconductor fabrication – one worker miscalibrating a machine in a cleanroom can disrupt production for days.

Although all modules have the same size (i.e. size of 1) workers differ in how they are connected to each other. Since everyone contributes $\hat{e}$ in equilibrium, differences in the length and targets of their influence paths determine the productivity factor μ_k to weigh each module. It turns out that $\mathbf{H} = \mathbf{CWC}'$ and therefore, following equation (14), $\boldsymbol{\alpha}^\star \propto \mathbf{C}^{-1}\mathbf{1}$. We thus find that incentives are allocated very differently. Firms no longer concentrate incentives on workers with more outgoing paths (i.e. higher Bonacich centrality). Rather, firms prioritize workers with fewer incoming links.

Corollary 3. *When every worker is essential (i.e. every module is of size 1), incentives are allocated inversely to workers' in-degree. Formally, optimal incentives are allocated following:*

$$\alpha_i^\star = \frac{1 - \lambda d_i}{\xi}, \quad \forall i \in N,$$

where d_i is worker i's in-degree and $\xi = \sum_{j \in N} (1 - 2\lambda d_j) + \sigma^2 r (1 - \lambda d_j)^2$ is common across all workers.

Intuitively, all workers exert the same effort in equilibrium so paying central workers a large α_i won't raise others' contributions (beyond their own value of α). On the other hand, those with few incoming links have large effort costs and therefore stand more to gain by

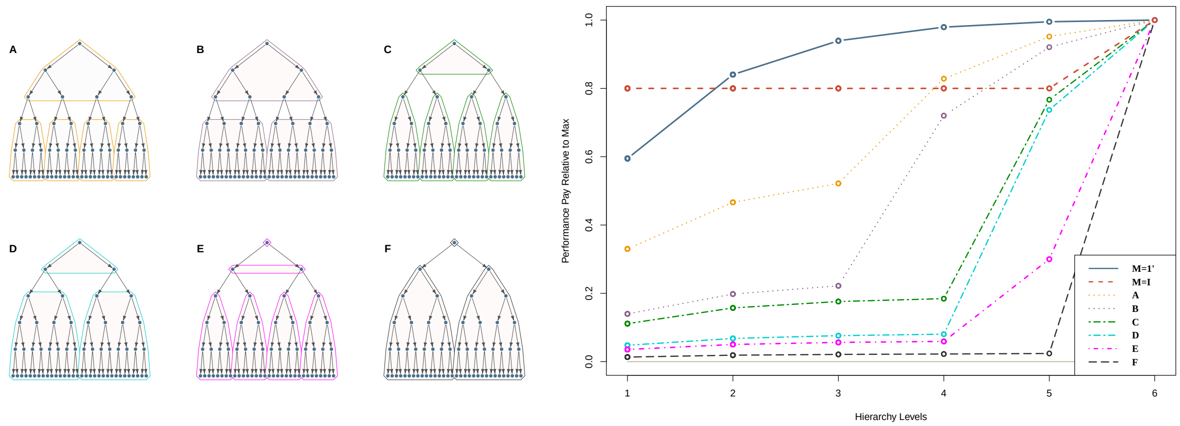

Figure 10: Performance-pay for Different Modular Configurations. Simulations are run for $\lambda = 0.2$, $\sigma^2 = 2$ and $r = 5$.

lowering their effort, unless variable pay is set sufficiently high.[41] These two forces imply that firms ensure a sufficiently large level of output by disregarding Bonacich centrality and allocating incentives based on incoming links instead.[42]

Modules in Hierarchies

Before moving on, let's consider how modular structure impacts earnings profiles by revisiting the hierarchy example from Figure 1. The right panel of Figure 10 tracks performance-pay along the firm's hierarchy for different module configurations. There are various things to note from this figure. First, the $\mathbf{M} = \mathbf{1}'$ case replicates the curve in Figure 1, as expected. Second, the $\mathbf{M} = \mathbf{I}$ case shows everyone receiving 80% of the CEO's performance compensation. This is in line with Corollary 3 since the CEO has an in-degree of 0 while everyone else has an in-degree of 1, and $1 - 0.2 * 1 = 0.8$. Now consider other modular configurations, which are drawn on the left panel. Notice that performance-pay jumps up as we cross to a higher module, which occurs, for instance, in level 3 for firms A and B, and level 4 for firms C and D. Notice also that the jump is more pronounced the larger are the lower modules. Within a module, performance pay tracks network centrality (as in the $\mathbf{M} = \mathbf{1}'$ curve), but the network effect is dampened by the large jump across modules (for instance in firm F). In fact, we decompose total variance in performance pay and find that *within*-module variation only accounts somewhere from 3% (firm D) to 35% percent (firm A) of the total variation:

[41]Recall that with modular technology it never pays to increase effort, given that everyone is doing the same effort. It only (sometimes) pays to lower effort.

[42]On some networks (See Figure 1) the worker with least incoming links might also be the most central, but this is generally not the case.

the rest is driven by variation in pay *between* modules. In fact, firm E shows that with multiple module jumps performance-pay profile can even look convex.

5 Implications for Organizational Design

While our focus throughout has been on deriving incentive rules across different production technologies and institutional constraints, our model also has broad implications for *organizational design*. We can use our framework to ask how firms should structure their organizations – or how they should invest in their workforce – based on the relationship between wages and the underlying network of bilateral ties. We explore these questions in detail in Supplementary Appendix G and provide some intuition for our findings below.

We first show that firm's overall profits in equilibrium are proportional to aggregate output, a result that extends the canonical intuition on team production to our setting with peer-to-peer network of complementarities. More importantly, using a decomposition method as in Galeotti et al. (2020), Proposition 9 derives an explicit formula linking expected profits to network structure through its *principal components*, for both directed and undirected networks. The general expressions are in Supplementary Appendix G but a lot of intuition can be obtained by focusing on undirected networks and taking a first-order approximation.[43] Letting μ_1 and $\mathbf{u}_1$ represent the leading eigenvalue and unit-eigenvector of $\mathbf{G}$ respectively, we can write profits as:

$$\mathbb{E}\left(\pi^{\star}\right) \approx \frac{n}{2} \frac{1 - n \operatorname{Var}\left(\mathbf{u}_1\right)}{\left(1 + r\sigma^2\right)\left(1 - \lambda\mu_1\right)^2 - \left(\lambda\mu_1\right)^2} \tag{16}$$

This expression highlights four key features. First, profits are decreasing in σ^2 because, everything else equal, higher risk implies larger compensation packages for workers. Second, profits are increasing in μ_1 (if $\lambda > 0$), meaning that networks with larger leading eigenvalues—i.e. denser and more expansive graphs—generate larger profits. This makes sense since these networks are better at amplifying incentives (notice the effect is reversed if $\lambda < 0$). Third, and perhaps most interestingly, profits decrease with $\operatorname{Var}(\mathbf{u}_1)$. In other words, profits are larger when eigenvector centrality is evenly distributed across workers.[44] Finally, keeping everything else constant, optimal firm size is given by $n^{\star} = \frac{1}{2\operatorname{Var}(\mathbf{u}_1)}$, reflect-

[43] In Section G we argue that this approximation, akin to principal component analysis, is particularly good for networks with large spectral gaps.

[44] For instance, an undirected ring and star for $n = 4$ have the same leading eigenvalue, but because the ring has a lower centrality variance, it generates larger profits according to (16).

ing that firms scale better when centrality is evenly distributed than when there are incentive bottlenecks.[45]

To showcase these effects, in Section G we compare networks whose degree distribution follows a *power law*, $P(d) \sim d^{-\gamma}$. As γ decreases, both μ_1 and $\text{Var}(\mathbf{u}_1)$ increase since there is more weight on the tail of the degree distribution (creating dominant hubs with many links). We show that there is a threshold $\gamma^\star$ such that profits decrease with γ above the threshold (because the μ_1 effect dominates) and increase below it (because the $\text{Var}(\mathbf{u}_1)$ effect dominates). This implies that profits are maximized at $\gamma^\star$.

Our profit decomposition result also enables an exact comparison of different organizational structures using well-known spectral properties for different families of networks. For instance, we can show that, among all complete bipartite networks, profits are maximized when divisions are of equal size (Corollary 4). Similarly, in networks where each worker has the same number of connections, expected profits are independent of how the organization is partitioned (Corollary 5). We also consider how community structure might influence profits, and find that it doesn't (Corollary 6).

Our analysis also highlights how firms should allocate resources to maximize profits. Imagine firms must decide whether to enhance workers' human capital by investing in a general training program, or whether to strengthen peer effects by investing in team-building exercises. If network connectivity is sufficiently sparse, investing in human capital is (obviously) preferable, but the advantage shifts towards strengthening peer effects if workers are sufficiently connected. Proposition 10 establishes a sharp threshold in Erdős-Rényi random graphs, showing that, once each worker has at least one expected connection, investments in team strength yield higher profits than investments in individual skills.[46]

These findings underscore a broader lesson: organizational design decisions should be guided by workers' structure of interactions. While standard economic models often focus on individual incentives, this analysis demonstrates that network structure fundamentally shapes firm performance. This first set of results suggest that this is a very promising line of future research.

[45] In Section G we show how these four features generalize to directed networks. The numerator to equation (16) becomes $1 - n\,\text{Cov}(\mathbf{u}_1, \mathbf{v}_1)$, where $\mathbf{v}_1$ is the leading *left*-eigenvector of $\mathbf{G}$. Intuitively, profits are larger if workers that have most outward influence are not most influenced themselves: hierarchical networks like those in Figure 1 yield larger profits than more reciprocal networks with comparable overall connectivity.

[46] This threshold coincides with the threshold for the emergence of a giant component in Erdos-Renyi random graphs.

6 Concluding Remarks

This paper examines optimal wage contracts in firms with productivity spillovers. We show that firms can leverage network-based incentives to boost productivity, with optimal contract design depending on co-worker externalities, production technology, and institutional constraints. Our framework provides a rationale for observed trends in the steepening earnings profile within firms that doesn't rely on assuming different endowments of managerial talent or intricate market forces that elevate talent up firm's hierarchy, to larger teams, or to more valuable organizations. Neal and Rosen (2000), for instance, note that the shape of the earnings distribution cannot be explained by the "super-star" CEO phenomenon because scale economies imply there are not enough of them to make a dent in the upper tail of the earnings distribution. By introducing peer effects into standard contract theory, we connect salaries to workers' span of control and rationalize why variable-pay varies so much within organizations.

Empirical validation using co-worker network data could test whether performance-based pay tracks worker centrality and identify untapped productivity gains. Future research could also explore firm competition (Chade and Eeckhout, 2023) and the impact of common ownership on managerial incentives and firm strategies (Garud et al., 2009). We can also provide quantitative measures of the welfare effects associated to benchmarking salaries when peer effects are taken into account. We expect the impact to depend on organizational structure in interesting ways. These extensions would allow us to understand how variations in network structures across firms might lead to different competitive outcomes and the sorting patterns that ensue (Gabaix and Landier, 2008).

Finally, our framework offers a foundation to study how artificial intelligence (AI) and automation may disrupt organizational structures. As AI systems automate cognitive tasks and relocate humans to routing knowledge work, the dynamics of employee interactions and organizational structure are likely to shift (Ide and Talamas, 2024). The integration of AI is poised to transform production processes, potentially leading to more modular configurations and certainly disrupting the structure of peer-to-peer complementarities in the workplace. Firms will need to reevaluate incentive structures and contract designs moving forward.

References

Amodio, F. and M. A. Martinez-Carrasco (2018): "Input allocation, workforce management and productivity spillovers: Evidence from personnel data," *The Review of Economic Studies*, 85, 1937–1970.

Ashraf, N. and O. Bandiera (2018): "Social incentives in organizations," *Annual Review of Economics*, 10, 439–463.

Avella-Medina, M., F. Parise, M. T. Schaub, and S. Segarra (2018): "Centrality measures for graphons: Accounting for uncertainty in networks," *IEEE Transactions on Network Science and Engineering*, 7, 520–537.

Baldwin, C. Y. and K. B. Clark (2003): "Managing in an age of modularity," *Managing in the modular age: Architectures, networks, and organizations*, 149, 84–93.

Ballester, C., A. Calvó-Armengol, and Y. Zenou (2006): "Who's who in networks. Wanted: The key player," *Econometrica*, 74, 1403–1417.

Bandiera, O., I. Barankay, and I. Rasul (2005): "Social Preferences and the Response to Incentives: Evidence from Personnel Data," *The Quarterly Journal of Economics*, 120, 917–962.

Belhaj, M., S. Bervoets, and F. Deroïan (2016): "Efficient networks in games with local complementarities," *Theoretical Economics*, 11, 357–380.

Belhaj, M. and F. Deroïan (2018): "Targeting the key player: An incentive-based approach," *Journal of Mathematical Economics*, 79, 57–64.

Bertrand, M. (2009): "CEOs," *Annual Review of Econonomics*, 1, 121–150.

Bloch, F. and N. Quérou (2013): "Pricing in social networks," *Games and Economic Behavior*, 80, 243–261.

Bollobás, B., S. Janson, and O. Riordan (2007): "The phase transition in inhomogeneous random graphs," *Random Structures & Algorithms*, 31, 3–122.

Bolton, P. and M. Dewatripont (2004): *Contract theory*, MIT press.

Bramoullé, Y. and R. Kranton (2007): "Public goods in networks," *Journal of Economic Theory*, 135, 478–494.

Breza, E., S. Kaur, and Y. Shamdasani (2018): "The morale effects of pay inequality," *The Quarterly Journal of Economics*, 133, 611–663.

Calvó-Armengol, A., E. Patacchini, and Y. Zenou (2009): "Peer effects and social networks in education," *The Review of Economic Studies*, 76, 1239–1267.

Candogan, O., K. Bimpikis, and A. Ozdaglar (2012): "Optimal pricing in networks with externalities," *Operations Research*, 60, 883–905.

Card, D., A. Mas, E. Moretti, and E. Saez (2012): "Inequality at work: The effect of peer salaries on job satisfaction," *American Economic Review*, 102, 2981–3003.

Carroll, G. (2015): "Robustness and Linear Contracts," *The American Economic Review*, 105, 536–563.

Chade, H. and J. Eeckhout (2023): "Do incentives or competition determine managers' wages?" *Working Paper.*

Cornelissen, T., C. Dustmann, and U. Schönberg (2017): "Peer Effects in the Workplace," *American Economic Review*, 107, 425–456.

Cullen, Z. (2024): "Is pay transparency good?" *Journal of Economic Perspectives*, 38, 153–180.

Cullen, Z. B., S. Li, and R. Perez-Truglia (2022): "What's My Employee Worth? The Effects of Salary Benchmarking," Tech. rep., National Bureau of Economic Research.

Dasaratha, K. (2020): "Distributions of centrality on networks," *Games and Economic Behavior*, 122, 1–27.

Dasaratha, K., B. Golub, and A. Shah (2024): "Incentive design with spillovers," *Available at SSRN 4853054*.

Demange, G. (2017): "Optimal targeting strategies in a network under complementarities," *Games and Economic Behavior*, 105, 84–103.

DeMarzo, P. M. and R. Kaniel (2023): "Contracting in peer networks," *The Journal of Finance*, 78, 2725–2778.

Fainmesser, I. P. and A. Galeotti (2016): "Pricing network effects," *The Review of Economic Studies*, 83, 165–198.

Falk, A. and A. Ichino (2006): "Clean evidence on peer effects," *Journal of Labor Economics*, 24, 39–57.

Fox, J. T. (2009): "Firm-size wage gaps, job responsibility, and hierarchical matching," *Journal of Labor Economics*, 27, 83–126.

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

Galeotti, A., B. Golub, and S. Goyal (2020): "Targeting interventions in networks," *Econometrica*, 88, 2445–2471.

Galeotti, A., S. Goyal, M. O. Jackson, F. Vega-Redondo, and L. Yariv (2010): "Network games," *The Review of Economic Studies*, 77, 218–244.

Garicano, L. (2000): "Hierarchies and the Organization of Knowledge in Production," *Journal of Political Economy*, 108, 874–904.

Garicano, L. and E. Rossi-Hansberg (2006): "Organization and inequality in a knowledge economy," *The Quarterly Journal of Economics*, 121, 1383–1435.

Garud, R., A. Kumaraswamy, and R. Langlois (2009): *Managing in the modular age: architectures, networks, and organizations*, John Wiley & Sons.

Guadalupe, M. and J. Wulf (2010): "The flattening firm and product market competition: The effect of trade liberalization on corporate hierarchies," *American Economic Journal: Applied Economics*, 2, 105–127.

Hamilton, B., J. A. Nickerson, and H. Owan (2003): "Team Incentives and Worker Heterogeneity: An Empirical Analysis of the Impact of Teams on Productivity and Participation," *Journal of Political Economy*, 111, 465–497.

Hazell, J., C. Patterson, H. Sarsons, and B. Taska (2022): "National wage setting," *University of Chicago, Becker Friedman Institute for Economics Working Paper*.

Heckman, J. J. and B. E. Honore (1990): "The empirical content of the Roy model," *Econometrica: Journal of the Econometric Society*, 1121–1149.

Hiller, T. (2017): "Peer effects in endogenous networks," *Games and Economic Behavior*, 105, 349–367.

Holmstrom, B. (1982): "Moral hazard in teams," *The Bell Journal of Economics*, 324–340.

Holmstrom, B. and P. Milgrom (1987): "Aggregation and Linearity in the Provision of Intertemporal Incentives," *Econometrica*, 55, 303–328.

——— (1991): "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design," *Journal of Law, Economics, & Organization*, 7, 24–52.

Ide, E. and E. Talamas (2024): "Artificial Intelligence in the Knowledge Economy," in *Proceedings of the 25th ACM Conference on Economics and Computation*, 834–836.

Jackson, M. O. (2008): *Social and economic networks*, vol. 3, Princeton University Press.

Jensen, M. C., K. J. Murphy, and E. G. Wruck (2004): "Remuneration: Where we've been, how we got to here, what are the problems, and how to fix them," *Harvard NOM Working Paper*.

Kremer, M. (1993): "The O-ring theory of economic development," *The Quarterly Journal of Economics*, 108, 551–575.

Macho-Stadler, I. and D. Pérez-Castrillo (1993): "Moral hazard with several agents: The gains from cooperation," *International Journal of Industrial Organization*, 11, 73–100.

Mas, A. (2017): "Does transparency lead to pay compression?" *Journal of Political Economy*, 125, 1683–1721.

Mas, A. and E. Moretti (2009): "Peers at work," *American Economic Review*, 99, 112–145.

Matouschek, N., M. Powell, and B. Reich (2025): "Organizing Modular Production," *Journal of Political Economy*, 133, 986–1046.

MOOKHERJEE, D. (1984): "Optimal Incentive Schemes with Many Agents," *The Review of Economic Studies*, 51, 433–446.

NEAL, D. AND S. ROSEN (2000): "Theories of the Distribution of Earnings," *Handbook of Income Distribution*, 1, 379–427.

OBLOJ, T. AND T. ZENGER (2022): "The influence of pay transparency on (gender) inequity, inequality and the performance basis of pay," *Nature Human Behaviour*, 6, 646–655.

PARISE, F. AND A. OZDAGLAR (2023): "Graphon games: A statistical framework for network games and interventions," *Econometrica*, 91, 191–225.

RAJAN, R. G. AND J. WULF (2006): "The flattening firm: Evidence from panel data on the changing nature of corporate hierarchies," *The Review of Economics and Statistics*, 88, 759–773.

ROY, A. D. (1951): "Some thoughts on the distribution of earnings," *Oxford Economic Papers*, 3, 135–146.

SHI, X. (2024): "Optimal compensation schemes in organizations with interpersonal networks," *Available at SSRN 4023019*.

SMEETS, V. AND F. WARZYNSKI (2008): "Too many theories, too few facts? What the data tell us about the link between span of control, compensation and career dynamics," *Labour Economics*, 15, 687–703.

SUN, Y. AND W. ZHAO (2024): "Contracts on peer pressure networks," *Available at SSRN 4881522*.

SUNDARARAJAN, A. (2008): "Local network effects and complex network structure," *The BE Journal of Theoretical Economics*, 7, pp. 0000102202193517041319.

WALLSKOG, M., N. BLOOM, S. W. OHLMACHER, AND C. TELLO-TRILLO (2024): "Within-Firm Pay Inequality and Productivity," Tech. rep., National Bureau of Economic Research.

WINTER, E. (2010): "Transparency and incentives among peers," *RAND Journal of Economics*, 41, 504–523.

A Proofs

A.1 Proof of Proposition 1

The firm's problem can be written in matrix form as:

$$\max_{\boldsymbol{\alpha}} \mathbb{E}[\pi(\mathbf{e} \mid \boldsymbol{\alpha}, \boldsymbol{\beta})] = \boldsymbol{\alpha}'\mathbf{C}'\mathbf{1} - \frac{1}{2}\boldsymbol{\alpha}'\mathbf{C}'\mathbf{C}\boldsymbol{\alpha} + \lambda\boldsymbol{\alpha}'\mathbf{C}'\mathbf{G}\mathbf{C}\boldsymbol{\alpha} - \frac{1}{2}\sigma^2 r\boldsymbol{\alpha}'\boldsymbol{\alpha}.$$

Taking the first order condition and solving for $\boldsymbol{\alpha}$ we get:

$$\begin{aligned}
&\frac{\partial E(\pi)}{\partial \boldsymbol{\alpha}} = \mathbf{C}'\mathbf{1} - \mathbf{C}'\mathbf{C}\boldsymbol{\alpha}^\star + \lambda\mathbf{C}'(\mathbf{G}+\mathbf{G}')\mathbf{C}\boldsymbol{\alpha}^\star - \sigma^2 r\boldsymbol{\alpha}^\star = 0 \\
\implies &\boldsymbol{\alpha}^\star = \left(\mathbf{C}'(\mathbf{I} - \lambda(\mathbf{G}+\mathbf{G}'))\mathbf{C} + \sigma^2 r\mathbf{I}\right)^{-1}\mathbf{C}'\mathbf{1} = \left(\mathbf{C} - \mathbf{I} - \lambda\mathbf{C}'\mathbf{G}\mathbf{C} + (1+\sigma^2 r)\mathbf{I})\right)^{-1}\mathbf{C}'\mathbf{1} \\
&\quad = \left(\lambda\mathbf{C}\mathbf{G} - \lambda\mathbf{C}'\mathbf{C}\mathbf{G} + (1+\sigma^2 r)\mathbf{I}\right)^{-1}\mathbf{C}'\mathbf{1} = \left((1+\sigma^2 r)\mathbf{I} - \lambda(\mathbf{C}'-\mathbf{I})\mathbf{C}\mathbf{G}\right)^{-1}\mathbf{C}'\mathbf{1} \\
&\quad = \left((1+\sigma^2 r)\mathbf{I} - \lambda^2(\mathbf{C}\mathbf{G})'\mathbf{C}\mathbf{G}\right)^{-1}\mathbf{C}'\mathbf{1} = \frac{1}{1+r\sigma^2}\left[\mathbf{I} - \frac{\lambda^2}{1+r\sigma^2}(\mathbf{G}\mathbf{C})'\mathbf{G}\mathbf{C}\right]^{-1}\mathbf{C}'\mathbf{1},
\end{aligned}$$

where the third equality after the implication follows from $\mathbf{G}\mathbf{C} = \mathbf{G} + \lambda\mathbf{G}^2 + ... = \frac{1}{\lambda}(\mathbf{C} - \mathbf{I})$ and the fifth equality uses the fact that $(\mathbf{G}\mathbf{C})' = (\mathbf{C}\mathbf{G})' = \frac{1}{\lambda}(\mathbf{C}' - \mathbf{I})$. Finally, writing (IR) in matrix form and solving for $\boldsymbol{\beta}^*$ we have:

$$\boldsymbol{\beta}^*(\boldsymbol{\alpha}^\star, \mathbf{e}^\star) = \frac{1}{2}(\mathbf{e}^\star \circ \mathbf{e}^\star) - (\mathbf{e}^\star \circ \lambda\mathbf{G}\mathbf{e}^\star) + \frac{1}{2}(r\sigma^2)(\boldsymbol{\alpha}^\star \circ \boldsymbol{\alpha}^\star) - \boldsymbol{\alpha}^\star \circ \mathbf{1}(\mathbf{1}'\mathbf{e}^\star),$$

where $\circ$ denotes the Hadamard product. Recalling that $\mathbf{e}^\star = \mathbf{C}\boldsymbol{\alpha}^\star$ completes the proof. ■

A.2 Proof of Proposition 2

It is easy to see that as σ^2 increases, the symmetric matrix $\mathbf{W}$ decreases. This implies that there is a finite σ^2, call it $\bar{\sigma}^2$, such that for all $\sigma^2 \geq \bar{\sigma}^2$ optimal incentives, as defined in equation 5, are arbitrarily small. Then, for all $\sigma^2 \geq \bar{\sigma}^2$, equation 9 becomes $b_k(\lambda) = r\sigma^2\alpha_k$ for all $k \in N$, which implies that optimal incentives are a monotonic transformation of Katz-Bonacich centrality. ■

A.3 Proof of Proposition 3

Using the chain rule, we can write the derivative of optimal incetives (equation (5)0 with respect to g_{ij} as the sum of two terms:

$$\frac{\partial \boldsymbol{\alpha}^\star}{\partial g_{ij}} = \frac{1}{1+r\sigma^2}\left[\frac{\partial \mathbf{W}}{\partial g_{ij}} \cdot \mathbf{C}'\mathbf{1} + \mathbf{W} \cdot \frac{\partial \mathbf{C}'\mathbf{1}}{\partial g_{ij}}\right]. \tag{17}$$

Since $\mathbf{W}$ is an inverted matrix, the derivative of the first term in the brackets is

$$\frac{\partial \mathbf{W}}{\partial g_{ij}} = -\mathbf{W}\frac{\partial[\mathbf{I} - \frac{\lambda^2}{1+r\sigma^2}(\mathbf{GC})'\mathbf{GC}]}{\partial g_{ij}}\mathbf{W} = \frac{\lambda^2}{1+r\sigma^2}\mathbf{W}\frac{\partial \mathbf{C}'\mathbf{G}'\mathbf{GC}}{\partial g_{ij}}\mathbf{W}.$$

Next, notice that the derivatives of $\mathbf{G}$ and $\mathbf{C}$ are $\mathbf{E_{ij}}$ and $\lambda\mathbf{C}\mathbf{E_{ij}}\mathbf{C}$, respectively, where $\mathbf{E_{ij}}$ is a matrix of all zeros except element (i, j), which is equal to one. Thus, applying the chain rule a few more times we can write the first term in the brackets of (17) as

$$\frac{\partial \mathbf{W}}{\partial g_{ij}} \cdot \mathbf{C}'\mathbf{1} = \frac{\lambda^2}{1+r\sigma^2}\mathbf{W}\left[\mathbf{C}'\mathbf{E_{ji}}\mathbf{GC} + \lambda\mathbf{C}'(\mathbf{E_{ji}}\mathbf{C}'\mathbf{G}'\mathbf{G} + \mathbf{G}'\mathbf{GCE_{ij}})\mathbf{C} + \mathbf{C}'\mathbf{G}'\mathbf{E_{ij}}\mathbf{C}\right]\mathbf{W} \cdot \mathbf{C}'\mathbf{1} \geq 0,$$

which is non-negative because all matrices are non-negative. Next, focus on the second term in brackets of (17). Entry s of the derivative vector of $\mathbf{C}'\mathbf{1}$ with respect to g_{ij} is given by

$$\frac{\partial(\mathbf{C}'\mathbf{1})_s}{\partial g_{ij}} = \frac{\partial \sum_{t=1}^n c_{ts}}{\partial g_{ij}} = \sum_{t=1}^n \frac{\partial c_{ts}}{\partial g_{ij}} = \sum_{t=1}^n \lambda c_{ti} c_{js} = \lambda c_{js} \sum_{t=1}^n c_{ti},$$

where the second equality follows from taking the (t, s) element of the derivative of $\mathbf{C}$ with respect to g_{ij}. Note that the derivative above is positive if $c_{js} > 0$, i.e., if worker j is influenced by worker s. Moreover, when $s = j$ this derivative is always strictly greater than zero. Thus, the second term in the brackets of (17) is

$$\mathbf{W} \cdot \frac{\partial \mathbf{C}'\mathbf{1}}{\partial g_{ij}} = \left(\lambda \sum_{t=1}^n c_{ti}\right)\mathbf{W} \cdot \begin{pmatrix} c_{j1} & \cdots & c_{jn} \end{pmatrix}' \geq 0.$$

Thus, we have established that (17) is non-negative. For the second part of the proposition, consider the derivative of $\alpha_s^\star$ with respect to the link g_{ij}:

$$\frac{\partial \alpha_s^\star}{\partial g_{ij}} = \frac{1}{1+r\sigma^2}\sum_{r=1}^n \left[\frac{\partial w_{sr}}{\partial g_{ij}} \cdot (\mathbf{C}'\mathbf{1})_r + w_{sr} \cdot \frac{\partial (\mathbf{C}'\mathbf{1})_r}{\partial g_{ij}}\right].$$

Let $W(j) \subseteq N$ be the set of workers that share common influence with worker j. First, consider a worker $s \notin W(j)$. Then, for all $r \in W(j)$, it must be that $w_{sr} = 0$. Next, if s has no influence on any worker then $\alpha_s^\star$ is given by Corollary 1, in which case the increase in g_{ij} has no effect on s's incentives. If s has influence on some other worker $s' \notin W(j)$, then it must be that s is not in the same network component as worker j. In this case, an increase in g_{ij} does not change the centrality of any worker in s's component and so $\delta w_{ss'}/\delta g_{ij}$ must be equal to zero. This implies that (17) is equal to zero for all $s \notin W(j)$.

Next, consider a worker $s \in W(j)$. Since s and j share common influence, we know that $w_{sj} = w_{js} > 0$. Thus, in the summation above when $r = j$ we have

$$\frac{\partial w_{sj}}{\partial g_{ij}} \cdot b_j + w_{sj} \cdot \frac{\partial b_j}{\partial g_{ij}} > 0,$$

since an increase in g_{ij} leads to an increase in j's influence over i, i.e., $\partial b_j / \partial g_{ij} > 0$, which in turn increases any common influence that j has with others, including w_{sj}, i.e, $\partial w_{sj}/\partial g_{ij} > 0$. In other words, there is at least one positive term in the summation above. Lastly, recalling that $w_{jj} \geq 1$, notice that in the summation above when $s = j$ and $r = j$, the term $w_{jj}(\partial b_j / \partial g_{ij})$ must be positive. Thus, (17) is strictly positive for all $s \in W(j) \cup \{j\}$. ■

A.4 Proof of Proposition 4

We can write the principal's problem under wage benchmarking in matrix form:

$$\max_{\hat{\alpha}} \mathbb{E}(\pi(\mathbf{e}|\boldsymbol{\alpha}, \boldsymbol{\beta})) = \mathbf{1}'\mathbf{e} - \mathbf{1}'\mathbf{w} = (\mathbf{1} - \boldsymbol{\alpha})'\mathbf{e} - \mathbf{1}'\boldsymbol{\beta}$$

subject to

$$\mathbf{1}'\boldsymbol{\beta} = \frac{1}{2}\mathbf{e}'\mathbf{e} - \lambda \mathbf{e}'\mathbf{G}\mathbf{e} - \boldsymbol{\alpha}'\mathbf{e} + \frac{1}{2} r\sigma^2 \boldsymbol{\alpha}'\boldsymbol{\alpha} + \mathbf{1}'\boldsymbol{\eta} \qquad \text{(IR)}$$

$$\mathbf{e} = \mathbf{C}\boldsymbol{\alpha} = \mathbf{C}(\mathbf{T}'\hat{\boldsymbol{\alpha}}) \qquad \text{(IC)}$$

Substituting in the (IR) and (IC) constraints, the principal's problem as a function of $\hat{\boldsymbol{\alpha}}$ is given by:

$$\max_{\hat{\alpha}} (\mathbf{C}\mathbf{T}'\hat{\boldsymbol{\alpha}})'\mathbf{1} + \lambda(\mathbf{C}\mathbf{T}'\hat{\boldsymbol{\alpha}})'\mathbf{G}(\mathbf{C}\mathbf{T}'\hat{\boldsymbol{\alpha}}) - \frac{1}{2}(\mathbf{C}\mathbf{T}'\hat{\boldsymbol{\alpha}})'(\mathbf{C}\mathbf{T}'\hat{\boldsymbol{\alpha}}) - \frac{1}{2} r\sigma^2 (\mathbf{T}'\hat{\boldsymbol{\alpha}})'(\mathbf{T}'\hat{\boldsymbol{\alpha}}) - \mathbf{1}'\boldsymbol{\eta}$$

Taking the first order condition and solving for $\hat{\boldsymbol{\alpha}}$, we get:

$$\frac{\partial E[\pi]}{\partial \hat{\boldsymbol{\alpha}}} = \mathbf{TC}'\mathbf{1} + \lambda \mathbf{TC}'(\mathbf{G}+\mathbf{G}')\mathbf{CT}'\hat{\boldsymbol{\alpha}} - \mathbf{TC}'\mathbf{CT}'\hat{\boldsymbol{\alpha}} - r\sigma^2\mathbf{TT}'\hat{\boldsymbol{\alpha}} = 0$$
$$\Rightarrow \hat{\boldsymbol{\alpha}}^{\star} = (\mathbf{T}(r\sigma^2\mathbf{I} + \mathbf{C}'(\mathbf{I} - \lambda(\mathbf{G}+\mathbf{G}'))\mathbf{C})\mathbf{T}')^{-1}\mathbf{TC}'\mathbf{1}.$$

Similar steps as in the proof of Proposition 1 lead to the desired expression. ■

A.5 Proof of Lemma 1

We can write the principal's expected profits in matrix form as:

$$\mathbb{E}(\pi(\mathbf{e}|\boldsymbol{\alpha},\boldsymbol{\beta})) = \mathbf{1}'\mathbf{e} - \frac{r\sigma^2}{2}\boldsymbol{\alpha}'\boldsymbol{\alpha} - \frac{1}{2}\mathbf{e}'\left(\mathbf{I} - 2\lambda\mathbf{G}\right)\mathbf{e} - \mathbf{1}'\boldsymbol{\eta},$$

Using $\mathbf{e}^{\star} = \mathbf{C}\boldsymbol{\alpha}^{\star}$ and $\boldsymbol{\alpha}^{\star} = \mathbf{T}'\hat{\boldsymbol{\alpha}} = \mathbf{T}'(\sigma^2 r\mathbf{TT}' + \mathbf{TC}'(\mathbf{I}-2\lambda\mathbf{G})\mathbf{CT}')^{-1}\mathbf{TC}'\mathbf{1}$, focus on the second and third terms above to get:

$$-\frac{r\sigma^2}{2}(\mathbf{T}'\hat{\boldsymbol{\alpha}})'(\mathbf{T}'\hat{\boldsymbol{\alpha}}) - \frac{1}{2}(\mathbf{CT}'\hat{\boldsymbol{\alpha}})'\left(\mathbf{I}-2\lambda\mathbf{G}\right)\mathbf{CT}'\hat{\boldsymbol{\alpha}} = -\frac{1}{2}\hat{\boldsymbol{\alpha}}'\left(r\sigma^2\mathbf{TT}' + \mathbf{TC}'\left(\mathbf{I}-2\lambda\mathbf{G}\right)\mathbf{CT}'\right)\hat{\boldsymbol{\alpha}}$$
$$= -\frac{1}{2}\hat{\boldsymbol{\alpha}}'\left(r\sigma^2\mathbf{TT}' + \mathbf{TC}'\left(\mathbf{I}-2\lambda\mathbf{G}\right)\mathbf{CT}'\right)\underbrace{(\sigma^2 r\mathbf{TT}' + \mathbf{TC}'(\mathbf{I}-2\lambda\mathbf{G})\mathbf{CT}')^{-1}\mathbf{TC}'\mathbf{1}}_{=\hat{\boldsymbol{\alpha}}}$$
$$= -\frac{1}{2}\hat{\boldsymbol{\alpha}}'\underbrace{\left(r\sigma^2\mathbf{TT}' + \mathbf{TC}'\left(\mathbf{I}-2\lambda\mathbf{G}\right)\mathbf{CT}'\right)(\sigma^2 r\mathbf{TT}' + \mathbf{TC}'(\mathbf{I}-2\lambda\mathbf{G})\mathbf{CT}')^{-1}}_{=\mathbf{I}}\mathbf{TC}'\mathbf{1}$$
$$= -\frac{1}{2}\hat{\boldsymbol{\alpha}}'\mathbf{TC}'\mathbf{1}$$

Reincorporating the middle terms into the expected profits of the firm, and noticing that $\mathbf{1}'\mathbf{e} = \mathbf{e}'\mathbf{1}$, we get:

$$E(\pi^{\star}) = \mathbf{1}'\mathbf{e} - \frac{1}{2}\hat{\boldsymbol{\alpha}}'\mathbf{TC}'\mathbf{1} - \mathbf{1}'\boldsymbol{\eta} = \mathbf{1}'\mathbf{e} - \frac{1}{2}\boldsymbol{\alpha}'\mathbf{C}'\mathbf{1} - \mathbf{1}'\boldsymbol{\eta} = \mathbf{1}'\mathbf{e} - \frac{1}{2}\mathbf{e}'\mathbf{1} - \mathbf{1}'\boldsymbol{\eta} = \frac{1}{2}\mathbf{1}'\mathbf{e} - \mathbf{1}'\boldsymbol{\eta}. \quad ■$$

A.6 Proof of Proposition 5

Let $\gamma = 1/(1+r\sigma^2)$. Following proposition 1 we can write the optimal efforts for personalized contracts (P) compactly as

$$\mathbf{e}^P = \gamma\mathbf{C}\boldsymbol{\alpha}^P = \gamma\mathbf{CWC}'\mathbf{1};$$

and, by proposition 4, we can do the analogous for the wage benchmarking case (WB) as

$$\mathbf{e}^{WB} = \gamma \mathbf{C}\boldsymbol{\alpha}^{WB} = \gamma \mathbf{C}\mathbf{T}'\hat{\boldsymbol{\alpha}}^{WB} = \gamma \mathbf{C}\mathbf{T}'(\mathbf{T}\mathbf{W}^{-1}\mathbf{T}')^{-1}\mathbf{T}\mathbf{C}'\mathbf{1},$$

where $W = [\mathbf{I} - \lambda^2/(1+r\sigma^2)(\mathbf{CG})'\mathbf{CG}]^{-1}$. Thus, the difference in outputs can be written as

$$\mathbf{1}'(\mathbf{e}^P - \mathbf{e}^{WB}) = \gamma \mathbf{1}'(\mathbf{C}\mathbf{W}\mathbf{C}'\mathbf{1} - \mathbf{C}\mathbf{T}'(\mathbf{T}\mathbf{W}^{-1}\mathbf{T}')^{-1}\mathbf{T}\mathbf{C}'\mathbf{1}) = \gamma \mathbf{b}'[\mathbf{W} - \mathbf{T}'(\mathbf{T}\mathbf{W}^{-1}\mathbf{T}')^{-1}\mathbf{T}]\mathbf{b}. \tag{18}$$

where $\mathbf{b} = \mathbf{C}'\mathbf{1}$ is the vector of outward Bonacich centralities. We want to show that there exists a δ such that:

$$\Delta(X^G - X^C) = X = \sum_{i \in A}(\mathbf{e}_i^P - \mathbf{e}_i^{WB}) = \delta \frac{1}{n} \sum_k n_k \mathrm{Var}(\mathbf{b}_k)$$

where $\mathrm{Var}(\mathbf{b}_k)$ is the variance within group k and $\sum_k$ sums over all k groups[47]. The average within-group variance, weighted by group size, is given by:

$$\frac{1}{n}\sum_k n_k \sum_{i \in k}(b_i - \bar{b})^2 = \frac{1}{n}\sum_k n_k \left(\sum_{i \in k} b_i^2 - \frac{(\sum_{i \in k} b_i)^2}{n_k} \right) = \frac{1}{n}\left[\sum_{i \in A} b_i^2 - \sum_k \frac{1}{n_k} (\sum_{i \in k} b_i)^2 \right],$$

where the last equality follows from the fact that summing over workers in group k and then over all groups k is the same as summing over *all* workers. Next, since the matrix $(\mathbf{TT}')^{-1}$ is a $K \times K$ diagonal matrix with (k, k) element equal one over group k's size, we can write the expression above in matrix form as:

$$\frac{1}{n}\sum_k n_k \mathrm{Var}(\mathbf{b}_k) = \frac{1}{n}\left[\mathbf{b}'\mathbf{b} - (\mathbf{Tb})')(\mathbf{TT}')^{-1}\mathbf{Tb}\right] = \frac{1}{n}\left(\mathbf{b}'\left[\mathbf{I} - \mathbf{T}'(\mathbf{TT}')^{-1}\mathbf{T}\right]\mathbf{b}\right) \tag{19}$$

Putting together (18) and (19) we have

$$\gamma \mathbf{b}'[\mathbf{W} - \mathbf{T}'(\mathbf{T}\mathbf{W}^{-1}\mathbf{T}')^{-1}\mathbf{T}]\mathbf{b} = \delta \frac{1}{n}\mathbf{b}'(\mathbf{I} - \mathbf{T}'(\mathbf{TT}')^{-1}\mathbf{T})\mathbf{b}.$$

Finally, notice that when $\lambda^2/(r\sigma^2) \to 0$ it is true that $\mathbf{W} \to \mathbf{I}$ as well as $(\mathbf{T}\mathbf{W}^{-1}\mathbf{T}')^{-1} \to (\mathbf{TT}')^{-1}$. Thus, in the limit, as $\lambda^2/(r\sigma^2) \to 0$, we have that (18) and (19) are the same when $\delta = n\gamma = n/(1 + r\sigma^2)$. That is,

$$\frac{1}{1+r\sigma^2}\mathbf{b}'[\mathbf{I} - \mathbf{T}'(\mathbf{TT}')^{-1}\mathbf{T}]\mathbf{b} = \delta \frac{1}{n}\mathbf{b}'(\mathbf{I} - \mathbf{T}'(\mathbf{TT}')^{-1}\mathbf{T})\mathbf{b}. \quad \blacksquare$$

[47] Notice we are abusing notation by using k as a group as well as the total number of groups in the firm.

A.7 Proof of Proposition 6

Notice that we can replace $\boldsymbol{\alpha} = (\mathbf{I} - \lambda\mathbf{G})\mathbf{e}$ and using the auxiliary variable $\hat{e}$ solve for the equivalent dual problem

$$\max_{(\hat{e},\mathbf{e})} \left(\hat{e} - \frac{1}{2}\mathbf{e}'[(\mathbf{I} - 2\lambda\mathbf{G}) + \sigma^2 r(\mathbf{I} - \lambda\mathbf{G}')(\mathbf{I} - \lambda\mathbf{G})]\mathbf{e} \right)$$

subject to

$$\mathbf{M}\mathbf{e} = \hat{e}\mathbf{1}_K,$$

and retrieve $\boldsymbol{\alpha}^\star$ using $\boldsymbol{\alpha}^\star = (\mathbf{I} - \lambda\mathbf{G})\mathbf{e}^\star$. Let $\breve{\boldsymbol{\Sigma}} \equiv (\mathbf{I} - 2\lambda\mathbf{G}) + \sigma^2 r(\mathbf{I} - \lambda\mathbf{G}')(\mathbf{I} - \lambda\mathbf{G})$. Considering the $k \times 1$ vector of Lagrangian multipliers $\boldsymbol{\mu}_k$, the Lagrangian of the above problem can be expressed as:

$$\mathcal{L}(\mathbf{e}, \hat{e}, \boldsymbol{\mu}) = \hat{e} - \frac{1}{2}\mathbf{e}'\breve{\boldsymbol{\Sigma}}\mathbf{e} - \boldsymbol{\mu}_k'(\hat{e}\mathbf{1}_{\boldsymbol{k}} - \mathbf{M}\mathbf{e}).$$

We have the following system of first order conditions:

$$\frac{\partial \mathcal{L}}{\partial \mathbf{e}} = -\frac{1}{2}\left(\breve{\boldsymbol{\Sigma}} + \breve{\boldsymbol{\Sigma}}'\right)\mathbf{e} + \mathbf{M}'\boldsymbol{\mu} = -\boldsymbol{\Sigma}\mathbf{e} + \mathbf{M}'\boldsymbol{\mu} = \mathbf{0}_{\boldsymbol{n}} \tag{20}$$

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{\mu}} = \mathbf{M}\mathbf{e} - \hat{e}\mathbf{1}_{\boldsymbol{k}} = \mathbf{0}_{\boldsymbol{k}} \tag{21}$$

$$\frac{\partial \mathcal{L}}{\partial \hat{e}} = 1 - \mathbf{1}_{\boldsymbol{k}}{}'\boldsymbol{\mu}_{\boldsymbol{k}} = 0 \tag{22}$$

Take $\boldsymbol{\Sigma} = \frac{1}{2}\left(\breve{\boldsymbol{\Sigma}} + \breve{\boldsymbol{\Sigma}}'\right) = (1 + \sigma^2 r)(\mathbf{I} - \lambda(\mathbf{G} + \mathbf{G}')) + \sigma^2 r(\lambda\mathbf{G})'(\lambda\mathbf{G})$.[48] From (20) we obtain $\mathbf{e} = \boldsymbol{\Sigma}^{-1}\mathbf{M}'\boldsymbol{\mu}_k$. Plugging $\mathbf{e}$ into (21) and defining the $K \times K$ matrix $\mathbf{H} \equiv \mathbf{M}\boldsymbol{\Sigma}^{-1}\mathbf{M}'$, we can solve for $\boldsymbol{\mu}_k$: $\mathbf{M}\boldsymbol{\Sigma}^{-1}\mathbf{M}'\boldsymbol{\mu}_k = \hat{e}\mathbf{1}_k \Rightarrow \boldsymbol{\mu}_k = \hat{e}\mathbf{H}^{-1}\mathbf{1}_k$, and use $\boldsymbol{\mu}_k$ in (22) to solve for $\hat{e}$:

$$\mathbf{1}_k'\boldsymbol{\mu}_k = \hat{e}\mathbf{1}_k'\mathbf{H}^{-1}\mathbf{1}_k = 1 \quad \Rightarrow \quad \hat{e} = \frac{1}{\mathbf{1}_k'\mathbf{H}^{-1}\mathbf{1}_k}.$$

Finally, putting everything together and using the fact that $\boldsymbol{\alpha} = (\mathbf{I} - \lambda\mathbf{G})\mathbf{e}$, and defining the vector of worker-specific lagrangian multipliera $\boldsymbol{\mu} = \mathbf{M}'\boldsymbol{\mu}_k$, we get:

$$\mathbf{e} = \frac{1}{\mathbf{1}_k'\mathbf{H}^{-1}\mathbf{1}_k}\boldsymbol{\Sigma}^{-1}\mathbf{M}'\mathbf{H}^{-1}\mathbf{1}_k \quad \Longrightarrow \quad \boldsymbol{\alpha} = \frac{1}{\mathbf{1}_k'\mathbf{H}^{-1}\mathbf{1}_k}(\mathbf{I} - \lambda\mathbf{G})\boldsymbol{\Sigma}^{-1}\mathbf{M}'\mathbf{H}^{-1}\mathbf{1}_k = (\boldsymbol{\Sigma}\mathbf{C})^{-1}\boldsymbol{\mu}.$$

[48] In the case of undirected graphs, $\mathbf{G} = \mathbf{G}'$, and $\boldsymbol{\Sigma}$ simplifies to $(\mathbf{I} - 2\lambda\mathbf{G})(1 + \sigma^2 r) + \sigma^2 r(\lambda\mathbf{G})^2$ and $\boldsymbol{\Sigma} = \breve{\boldsymbol{\Sigma}}$.

Thus, $\boldsymbol{\mu} = \mathbf{M}'\mathbf{H}^{-1}\mathbf{1}/(\mathbf{1}'\mathbf{H}^{-1}\mathbf{1})$. Notice that, as $\lambda\mathbf{G}\mathbf{C} = \mathbf{C} - \mathbf{I}$, then:

$$\begin{aligned}\mathbf{C}'\boldsymbol{\Sigma}\mathbf{C} &= (1+r\sigma^2)\mathbf{C}'\Big(\mathbf{I} - \lambda(\mathbf{G}' + \mathbf{G})\Big)\mathbf{C} + r\sigma^2\Big(\lambda\mathbf{G}\mathbf{C}\Big)'\lambda\mathbf{G}\mathbf{C} \\ &= (1+r\sigma^2)\Big(\mathbf{C}'\mathbf{C} - \mathbf{C}'(\lambda\mathbf{G}' + \lambda\mathbf{G})\mathbf{C}\Big) + r\sigma^2\Big(\mathbf{C}'\mathbf{C} - \mathbf{C}' - \mathbf{C} + \mathbf{I}\Big) \\ &= \mathbf{C}'\Big(\mathbf{I} - \lambda(\mathbf{G}' + \mathbf{G})\Big)\mathbf{C} + r\sigma^2\mathbf{I} = \mathbf{W}^{-1}.\end{aligned}$$

Where the last equality follows directly from Proposition 1. As $\mathbf{W} = (\mathbf{C}'\boldsymbol{\Sigma}\mathbf{C})^{-1} = (\boldsymbol{\Sigma}\mathbf{C})^{-1}\mathbf{C}'^{-1}$, then $\mathbf{W}\mathbf{C}' = (\boldsymbol{\Sigma}\mathbf{C})^{-1}$. Moreover, $\mathbf{H} = \mathbf{M}\boldsymbol{\Sigma}^{-1}\mathbf{M}' = (\mathbf{M}\mathbf{C})\mathbf{W}(\mathbf{M}\mathbf{C})'$ and $\boldsymbol{\alpha} = \mathbf{W}\mathbf{C}'\boldsymbol{\mu}$. ■

A.8 Proof of Corollary 2

If $\lambda = 0$ then $\boldsymbol{\Sigma} = (1+r\sigma^2)\mathbf{I}$ and $\mathbf{H} = \frac{1}{1+r\sigma^2}\mathbf{M}\mathbf{M}'$. Substituting into equation (14), we get

$$\boldsymbol{\alpha}^\star = \frac{1}{1+r\sigma^2}\frac{\mathbf{M}'(\mathbf{M}\mathbf{M}')^{-1}\mathbf{1}}{\mathbf{1}'(\mathbf{M}\mathbf{M}')^{-1}\mathbf{1}}$$

notice that $\mathbf{M}\mathbf{M}' = \text{diag}(n_1, n_2, \dots, n_K)$ so the denominator of the second term above can be written as:

$$\mathbf{1}'(\mathbf{M}\mathbf{M}')^{-1}\mathbf{1} = \sum_{k=1}^{K}\frac{1}{n_k} = \frac{\sum_r \prod_{k\in K\setminus r} n_k}{\prod_{k\in K} n_k}.$$

Notice that the numerator is simply an $N \times 1$ vector where the i-th position is $1/n_{k(i)}$. Putting everything together we get the desired expression. Finally, if $n_1 = n_2 = \ldots = n_K = \tilde{n}$ then the expression in Corollary 2 becomes

$$\alpha_i^\star = \frac{1}{1+r\sigma^2}\frac{1/\tilde{n}}{K\tilde{n}^{K-1}/\tilde{n}^K} = \frac{1}{1+r\sigma^2}\frac{1}{K}. \qquad ■$$

A.9 Proof of Corollary 3

Notice from Proposition 6 that for $\mathbf{M} = \mathbf{I}$, we get $\mathbf{H} = \mathbf{M}\boldsymbol{\Sigma}^{-1}\mathbf{M}' = \boldsymbol{\Sigma}^{-1}$, and:

$$\boldsymbol{\alpha}^* = (1 - \lambda\mathbf{G})\boldsymbol{\Sigma}^{-1}\frac{\mathbf{M}'\mathbf{H}^{-1}\mathbf{1}}{\mathbf{1}'\mathbf{H}^{-1}\mathbf{1}} = (1 - \lambda\mathbf{G})\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}\mathbf{1}\frac{1}{\mathbf{1}'\boldsymbol{\Sigma}\mathbf{1}} = \frac{1}{\mathbf{1}'\boldsymbol{\Sigma}\mathbf{1}}(1 - \lambda\mathbf{G})\mathbf{1}.$$

Next, notice that $\mathbf{G1} = \mathbf{d}$ where $\mathbf{d} := (d_1, d_2, \ldots, d_N)'$. Then,

$$\begin{aligned}
\mathbf{1}'\boldsymbol{\Sigma}\mathbf{1} &= \mathbf{1}'\left[(1+\sigma^2 r)(\mathbf{I} - \lambda(\mathbf{G}+\mathbf{G}')) + \sigma^2 r(\lambda\mathbf{G})'(\lambda\mathbf{G})\right]\mathbf{1} \\
&= (1+r\sigma^2)\mathbf{1}'(\mathbf{I} - \lambda(\mathbf{G}+\mathbf{G}'))\mathbf{1} + r\sigma^2(\lambda\mathbf{G1})'(\lambda\mathbf{G1}) \\
&= \sum_{j \in N}\left[(1+r\sigma^2)(1-2\lambda d_j) + \lambda^2 r\sigma^2 d_j^2\right]
\end{aligned}$$

Therefore, $\boldsymbol{\alpha}^* = \frac{1}{\mathbf{1}'\boldsymbol{\Sigma}\mathbf{1}}(1-\lambda\mathbf{G})\mathbf{1} = \frac{1}{\xi}(1-\lambda\mathbf{d})$. And thus, $\alpha_i^* = \frac{1-\lambda d_i}{\xi}$, as desired. ■

Supplementary Appendix

B Optimal Incentive Contracts with Heterogeneous Agents

Consider the model in section 2 when workers are heterogeneous in their productivity per unit of effort (θ_i), risk aversion (r_i), and reservation utility ($-\exp[-r_i(U_i)]$). The firm's production is now given by:

$$X(\mathbf{e}) = \sum_{i=1}^{n} \theta_i e_i + \varepsilon,$$

and worker i's certain equivalent is now defined as:

$$\text{CE}_i(\mathbf{e}, \mathbf{\Lambda}) = \beta_i + \alpha_i \sum_{i=1}^{n} \theta_i e_i - \frac{1}{2} e_i^2 + \lambda e_i \sum_{j \in N} g_{ij} e_j - \alpha_i^2 \frac{r_i \sigma^2}{2}.$$

For any contract (α_i, β_i), the best reply-function of worker i is given by:

$$e_i^\star = \theta_i \alpha_i + \lambda \sum_{j \in N} g_{ij} e_j$$

Define $\mathbf{\Theta} = \text{diag}(\theta)$ and $\mathbf{R} = \text{diag}(r)$. Then, the vector of best responses is:

$$\mathbf{e} = \mathbf{\Theta} \alpha + \lambda \mathbf{G} \mathbf{e}.$$

Under Assumption 1, the unique Nash equilibrium effort profile $e^\star$ of the game can be characterized:

$$\mathbf{e}^\star = [\mathbf{I} - \lambda \mathbf{G}]^{-1} \mathbf{\Theta} \boldsymbol{\alpha}.$$

As in the main text, the firm can set fixed payments β_i in order to extract all surplus from the workers, such that $\text{CE}_i(\mathbf{e}) = U_i$. We can write the firm's problem as:

$$\max_{\alpha,\beta} \mathbb{E}[\pi(\mathbf{e}|\boldsymbol{\alpha}, \boldsymbol{\beta})] = \sum_{i}^{n} \theta_i e_i - \sum_{i}^{n} S_i(X)$$

subject to

$$\text{CE}_i(\mathbf{e}) = U_i, \ \forall i \in N \quad \text{(IR)}$$

$$\mathbf{e}^\star = [\mathbf{I} - \lambda \mathbf{G}]^{-1} \mathbf{\Theta} \boldsymbol{\alpha} \quad \text{(IC)}$$

The (IR) condition under heterogeneity for worker $i \in N$ can be written as:

$$\beta_i + \alpha_i \sum_{j \in N} \theta_i e_j - \frac{1}{2} e_i^2 + \lambda e_i \sum_{j \in N} g_{ij} e_j - \alpha_i^2 r_i \frac{\sigma^2}{2} = U_i \tag{23}$$

Substituting for β_i into the objective function and writing the firm's expected profits in matrix form we have: Which obviating the $\sum_k^n U_k$ constant terms, can be expressed in matrix form as:

$$\max_{\alpha} \mathbb{E}[\pi(\mathbf{e}|\boldsymbol{\alpha}, \boldsymbol{\beta})] = \{\mathbf{e}'\boldsymbol{\Theta}1 - \frac{1}{2}\mathbf{e}'\mathbf{e} - \frac{\sigma^2}{2}\boldsymbol{\alpha}'\mathbf{R}\boldsymbol{\alpha} + \lambda \mathbf{e}'\mathbf{G}\mathbf{e}\}$$

subject to

$$\mathbf{e}^\star = [\mathbf{I} - \lambda \mathbf{G}]^{-1} \boldsymbol{\Theta}\boldsymbol{\alpha} \tag{IC}$$

where we omit the constant term $\sum_{i \in N} U_i$. Taking $\tilde{\mathbf{C}} \equiv [\mathbf{I} - \lambda \mathbf{G}]^{-1} \boldsymbol{\Theta}$ and replacing $\mathbf{e} = \tilde{\mathbf{C}}\boldsymbol{\alpha}$ and $\mathbf{e}' = \boldsymbol{\alpha}'\tilde{\mathbf{C}}'$, the above maximization problem becomes:

$$\max_{\boldsymbol{\alpha}} \mathbb{E}[\pi(\mathbf{e}|\boldsymbol{\alpha})] = \{\boldsymbol{\alpha}'\tilde{\mathbf{C}}'\boldsymbol{\Theta}\mathbf{1} - \frac{1}{2}\boldsymbol{\alpha}'\tilde{\mathbf{C}}'\tilde{\mathbf{C}}\boldsymbol{\alpha} - \frac{\sigma^2}{2}\boldsymbol{\alpha}'\mathbf{R}\boldsymbol{\alpha} + \lambda \boldsymbol{\alpha}'\tilde{\mathbf{C}}'\mathbf{G}\tilde{\mathbf{C}}\boldsymbol{\alpha}\}$$

Using matrix calculus, the first order conditions with respect to $\boldsymbol{\alpha}$ imply:

$$\mathbf{0} = \tilde{\mathbf{C}}'\boldsymbol{\Theta}\mathbf{1} - \left[\left(\tilde{\mathbf{C}}'\left(\mathbf{I} - \lambda(\mathbf{G}' + \mathbf{G})\right)\tilde{\mathbf{C}}\right) + \sigma^2\mathbf{R}\right]\boldsymbol{\alpha}^\star$$
$$\Rightarrow \boldsymbol{\alpha}^\star = \left[\left(\tilde{\mathbf{C}}'\left(\mathbf{I} - \lambda(\mathbf{G}' + \mathbf{G})\right)\tilde{\mathbf{C}}\right) + \sigma^2\mathbf{R}\right]^{-1}\tilde{\mathbf{C}}'\boldsymbol{\Theta}\mathbf{1}.$$

The optimal induced effort in this case is given by $\mathbf{e}^\star = \tilde{\mathbf{C}}\boldsymbol{\alpha}^\star$ and, defining $\mathbf{u} = (U_1, U_2. \ldots, U_N)$, the vector of optimal fixed payments $\boldsymbol{\beta}^\star$ can be recovered using (IR) in matrix form:

$$\beta^\star(\boldsymbol{\alpha}^\star) = \frac{1}{2}\left[\mathbf{u} + \tilde{\mathbf{C}}\boldsymbol{\alpha}^\star \circ (\mathbf{I} - 2\lambda\mathbf{G})\,\tilde{\mathbf{C}}\boldsymbol{\alpha}^\star + \boldsymbol{\alpha}^\star \circ \left(\sigma^2\mathbf{R} - 2\mathbf{1}\mathbf{1}'\boldsymbol{\Theta}\tilde{\mathbf{C}}'\right)\boldsymbol{\alpha}^\star\right].$$

C Incentive Provision Under Individual Production

Consider a version of the model in which each worker's effort results in a noisy individual production (IP) according to

$$q_i = e_i + \varepsilon_i. \tag{24}$$

The random variables $(\varepsilon_i)_{i\in N}$ are assumed to be independently and normally distributed with mean zero and variance σ^2.[49] This means that worker i's compensation is now conditional on individual output rather than the joint output $X(\mathbf{e})$. That is,

$$w_i = \beta_i^{IP} + \alpha_i^{IP} q_i, \tag{25}$$

where β_i^{IP} is a fixed term of the compensation, and α_i^{IP} is a variable or performance-related compensation coefficient under individual production. In this case, worker i's certainty equivalent is given by

$$CE_i^{IP}(\mathbf{e}, \mathbf{G}; \alpha_i^{IP}, \beta_i^{IP}) = \beta_i^{IP} + \alpha_i^{IP} e_i^{IP} - \frac{1}{2}(e_i^{IP})^2 + \lambda e_i^{IP} \sum_{j\in A} g_{ij} e_j^{IP} - (\alpha_i^{IP})^2 \frac{r_i \sigma^2}{2}. \tag{26}$$

Notice that the effort-provision problem of worker i is the same when maximizing (2) and (26). However the former expression has an extra term: $\alpha_i \sum_{j\neq i} e_j$. This implies that when the principal sets the fixed part of the compensation to guarantee that the participation constraint is satisfied for each worker we have the following equivalence:

$$\beta_i^{IP} = \beta_i + \alpha_i \sum_{j\neq i} e_j.$$

However, looking at the principal's reduced maximization problems in both cases we confirm that $\alpha_i = \alpha_i^{IP}$. Choosing performance-related compensations under joint production the principal maximizes

$$Max_{\boldsymbol{\alpha}} \ \sum_{i\in A} e_i - \sum_{i\in A} w_i = \sum_{i\in A} e_i - \sum_{i\in A} \beta_i - \sum_{i\in A} \alpha_i \sum_{k\in A} e_k;$$

while under individual production she maximizes

$$Max_{\boldsymbol{\alpha}} \ \sum_{i\in A} e_i - \sum_{i\in A} w_i = \sum_{i\in A} e_i - \sum_{i\in A} \beta_i^{IP} - \sum_{i\in A} \alpha_i e_i.$$

Using the fact that $\beta_i^{IP} = \beta_i + \alpha_i \sum_{j\neq i} e_j$ we can write the latter optimization problem as

$$Max_{\boldsymbol{\alpha}} \ \sum_{i\in A} e_i - \sum_{i\in A} \beta_i - \sum_{i\in A} \alpha_i \sum_{j\neq i} e_j - \sum_{i\in A} \alpha_i e_i = \sum_{i\in A} e_i - \sum_{i\in A} \beta_i - \sum_{i\in A} \alpha_i \sum_{k\in A} e_k.$$

[49]The model by Holmstrom and Milgrom (1987) has also been extended to situations with individual production and correlated outputs (see Bolton and Dewatripont (2004)).

D First-Best Contract

Under symmetric information, effort is both observable and contractible. As a result, there is no need for incentive compatibility, and the Principal's problem can be stated as follows:

$$\max_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{e}} \mathbb{E}[\pi(\mathbf{e} \mid \boldsymbol{\alpha}, \boldsymbol{\beta})]$$

subject to

$$\text{CE}_i(\mathbf{e}) = U_i, \ \forall i \in N \tag{IR}$$

Proposition 7 (First Best)**.** *Under symmetric information, agents are fully insured. Wages and principal's profits are increasing in peer effects. The optimal contract implies* $\boldsymbol{\alpha}^\star = \mathbf{0}$*,* $\mathbf{e}^\star = (\mathbf{I} - \lambda(\mathbf{G} + \mathbf{G}'))^{-1}\mathbf{1}$*, and in the case of* $\mathbf{G} = \mathbf{G}'$*,* $\boldsymbol{\pi}^\star = \frac{1}{2}\mathbf{1}'\mathbf{e}^\star$*, with* $\mathbf{1}$ *representing a vector of ones.*

Proof of Proposition 7. To simplify notation, we normalize outside options to zero for everyone, i.e. $U_i = 0$ for all $i \in N$. Notice that there is no incentive to pay the agents above their reservation utility, i.e., the (IR) constraint is binding. Thus, we have:

$$\beta_i = -\alpha_i \sum_{i \in N} e_i + \frac{1}{2} e_i^2 - \lambda e_i \sum_{j \in N} g_{ij} e_j + \alpha_i^2 \frac{r\sigma^2}{2}, \quad \forall i \in N.$$

Substituting into the objective function of the principal we have:

$$\max_{\mathbf{e},\boldsymbol{\alpha},\boldsymbol{\beta}} E[\pi] = \sum_{i \in N} e_i - \sum_{i \in N} \left[\alpha_i \sum_{j \in N} e_j + \beta_i \right] = \sum_{i \in N} e_i - \sum_{i \in N} \left[\frac{1}{2} e_i^2 - \lambda e_i \sum_{j \in N} g_{ij} e_j + \alpha_i^2 \frac{r\sigma^2}{2} \right].$$

Notice the utility of the principal is globally decreasing in α_i^2, which implies that the optimal contract involves $\alpha_i^2 = 0$ for all $i \in N$. Therefore, in matrix form we have:

$$\max_{\mathbf{e}} E[\pi] = \mathbf{e}'\mathbf{1} - \frac{1}{2}\mathbf{e}'\mathbf{e} + \lambda \mathbf{e}'\mathbf{G}\mathbf{e}$$

and writing the first order conditions to solve for $\mathbf{e}$ we get:

$$\mathbf{1} - \mathbf{e} + \lambda(\mathbf{G} + \mathbf{G}')\mathbf{e} = 0 \quad \Rightarrow \quad \mathbf{e}^\star = (\mathbf{I} - \lambda(\mathbf{G} + \mathbf{G}'))^{-1}\mathbf{1}.$$

Finally, replacing $\mathbf{e}^{\star}$ back into $E[\pi]$, we get:

$$\pi^{\star} = \mathbf{e}'\left(\mathbf{1} - \frac{1}{2}(\mathbf{I} - 2\lambda\mathbf{G})\mathbf{e}\right) = \mathbf{e}'\left(\mathbf{1} - \frac{1}{2}(\mathbf{I} - 2\lambda\mathbf{G})(\mathbf{I} - \lambda(\mathbf{G} + \mathbf{G}'))^{-1}\mathbf{1}\right).$$

Notice that if $\mathbf{G}$ is symmetric, $\mathbf{G} = \mathbf{G}'$ and we get:

$$\pi^{\star} = \mathbf{e}'\left(\mathbf{1} - \frac{1}{2}(\mathbf{I} - 2\lambda\mathbf{G})(\mathbf{I} - 2\lambda\mathbf{G})^{-1}\mathbf{1}\right) = \mathbf{e}'\left(\mathbf{1} - \frac{1}{2}\mathbf{1}\right) = \frac{1}{2}\mathbf{e}'\mathbf{1}.$$

■

E Optimal Incentives and Bonacich Centrality in a Line Network

In this section, we use equation (9) to show that, without risk ($\sigma^2 = 0$), the marginal costs associated with the incentives to worker 3, MC_{α_3} are smaller than the marginal costs of incentives provided to worker 1, MC_{α_1}, for the line network example with $N = 7$ in Figure 11.

First, recall that the substitution effects associated with the incentives to worker k, the first term on the right-hand side of equation (9) for worker $k \in N$, are given by

$$\sum_i \sum_j \alpha_j \frac{\partial e_i}{\partial \alpha_j} \frac{\partial e_i}{\partial \alpha_k},$$

and notice that in the case of a line network, worker k influences the effort of all workers "to the right." Moreover, the influence on worker $j > k$ falls by a factor of λ with the distance between k and j. That is, for each $i \geq k$,

$$\frac{\partial e_i}{\partial \alpha_k} = \lambda^{(k-i)}.$$

Thus, we can rewrite the first term in MC_{α_k} as

$$\sum_{j=1}^{N} \alpha_j \sum_{i \geq \max(j,k)} \lambda^{(i-k)} \lambda^{(i-j)}.$$

We can now give the substitution effect on the marginal costs of the incentives to workers

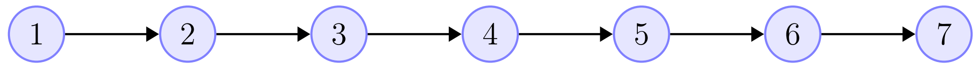

Figure 11: A line network with 7 workers.

1 and 3, i.e., SE_{α_1} and SE_{α_3}:

$$\begin{aligned} SE_{\alpha_1} &= \alpha_1 \left(1+\lambda^2+\lambda^4+\lambda^6+\lambda^8+\lambda^{10}+\lambda^{12}\right)+\alpha_2\left(\lambda+\lambda^3+\lambda^5+\lambda^7+\lambda^9+\lambda^{11}\right) \\ &+\alpha_3\left(\lambda^2+\lambda^4+\lambda^6+\lambda^8+\lambda^{10}\right)+\alpha_4\left(\lambda^3+\lambda^5+\lambda^7+\lambda^9\right)+\alpha_5\left(\lambda^4+\lambda^6+\lambda^8\right) \\ &+\alpha_6\left(\lambda^5+\lambda^7\right)+\alpha_7\left(\lambda^6\right). \\ SE_{\alpha_3} &= \alpha_1\left(\lambda^2+\lambda^4+\lambda^6+\lambda^8+\lambda^{10}\right)+\alpha_2\left(\lambda+\lambda^3+\lambda^5+\lambda^7+\lambda^9\right) \\ &+\alpha_3\left(1+\lambda^2+\lambda^4+\lambda^6+\lambda^8\right)+\alpha_4\left(\lambda+\lambda^3+\lambda^5+\lambda^7\right)+\alpha_5\left(\lambda^2+\lambda^4+\lambda^6\right) \\ &+\alpha_6\left(\lambda^3+\lambda^5\right)+\alpha_7\left(\lambda^4\right). \end{aligned}$$

Notice that α_1 generates stronger substitution effects with α_1 and α_2 while α_3 generates stronger substitution effects with α_3, α_4 , α_5, α_6, and α_7.[50] Taking the difference we get:

$$\begin{aligned} SE_{\alpha_1}-SE_{\alpha_3} &= \alpha_1(1+\lambda^{12})+\alpha_2(\lambda^{11})-\alpha_3(1-\lambda^{10})-\alpha_4(\lambda-\lambda^9) \\ &-\alpha^5(\lambda^2-\lambda^8)-\alpha^6(\lambda^3-\lambda^7)-\alpha_7(\lambda^4-\lambda^6) \end{aligned}$$

Next, focus on the substitution effect associated with the incentives to worker k, i.e., the second term on the right-hand side of equation (9), given by

$$-\lambda\sum_i\sum_j g_{ij}\sum_s\alpha_s\left(\frac{\partial e_j}{\partial\alpha_s}\frac{\partial e_i}{\partial\alpha_k}+\frac{\partial e_i}{\partial\alpha_s}\frac{\partial e_j}{\partial\alpha_k}\right)$$

Notice that now we are looking for positive terms such that worker k influences either side of the link g_{ij}. For a line network, this means that either worker i or worker j or both are "to the right" of worker k. Next, we need to look for workers s that influence the other side of that same link. Whenever worker s influences both sides of the link, we will have that the two terms in the parenthesis are positive. However, if s does not influence one side of the link, we only have one positive term in the parenthesis. Thus, for a line network, we can

[50] Generally, for a line network, we have the following relationship between the substitution effects in the marginal cost of incentives:

$$SE_{\alpha_k}=\lambda^{|k-i|}SE_{\alpha_i}+\sum_{i>j\geq k}\lambda^{|k-j|}\alpha_i.$$

write the substitution effect more concisely as

$$-\lambda \sum_s \alpha_s \left[\sum_{i,j \geq \max(k,s)} g_{ij} \left(\lambda^{(j-s)} \lambda^{(i-k)} + \lambda^{(i-s)} \lambda^{(j-k)} \right) + \sum_{\substack{\max(k,s) > \min(i,j) \geq \min(k,s) \\ \max(i,j) = \max(k,s)}} g_{ij} (\lambda^{(\min(i,j) - \min(k,s))}) \right].$$

Using the above, we can write the complementarity effect on the marginal costs of the incentives to workers 1 and 3, i.e., CE_{α_1} and CE_{α_3}:[51]

$$\begin{aligned} CE_{\alpha_1} = & -\lambda \big[2\alpha_1 \left(\lambda + \lambda^3 + \lambda^5 + \lambda^7 + \lambda^9 + \lambda^{11} \right) + \alpha_2 \left(1 + 2\lambda^2 + 2\lambda^4 + 2\lambda^6 + 2\lambda^8 + 2\lambda^{10} \right) \\ & + \alpha_3 \left(\lambda + 2\lambda^3 + 2\lambda^5 + 2\lambda^7 + 2\lambda^9 \right) + \alpha_4 \left(\lambda^2 + 2\lambda^4 + 2\lambda^6 + 2\lambda^8 \right) \\ & + \alpha_5 \left(\lambda^3 + 2\lambda^5 + 2\lambda^7 \right) + \alpha_6 \left(\lambda^4 + 2\lambda^6 \right) + \alpha_7 \left(\lambda^5 \right) \big]. \\ CE_{\alpha_3} = & -\lambda \big[\alpha_1 \left(\lambda + 2\lambda^3 + 2\lambda^5 + 2\lambda^7 + 2\lambda^9 \right) + \alpha_2 \left(1 + 2\lambda^2 + 2\lambda^4 + 2\lambda^6 + 2\lambda^8 \right) \\ & + 2\alpha_3 \left(\lambda + \lambda^3 + \lambda^5 + \lambda^7 \right) + \alpha_4 \left(1 + 2\lambda^2 + 2\lambda^4 + 2\lambda^6 \right) + \alpha_5 \left(\lambda + 2\lambda^3 + 2\lambda^5 \right) \\ & + \alpha_6 \left(\lambda^2 + 2\lambda^4 \right) + \alpha_7 \left(\lambda^3 \right) \big]. \end{aligned}$$

Notice that α_3 generates stronger complementarity effects with $\alpha_4, \alpha_5, \alpha_6$, and α_7 (as well as α_3) since worker 3 is "less steps away" from workers 4 to 7 than worker 1. On the other hand, worker 1 only generates stronger complementarity effects with α_2 (and α_1 itself). Taking the difference, we get:

$$\begin{aligned} CE_{\alpha_1} - CE_{\alpha_3} = & \alpha_1(\lambda^2 + 2\lambda^{12}) + \alpha_2(2\lambda^{11}) - \alpha_3(\lambda^2 - 2\lambda^{10}) - \alpha_4(\lambda^3 - 2\lambda^9) \\ & - \alpha_5(\lambda^2 + \lambda^4 - 2\lambda^8) - \alpha_6(\lambda^3 + \lambda^5 - 2\lambda^7) - \alpha_7(\lambda^4 - \lambda^6). \end{aligned}$$

Comparing the impact on all α_j for $j \in N$, i.e., the coefficients for each α_j, we can see that, for small λ, incentives for worker 3 generate greater complementarity effects than incentives for worker 1: $CE_{\alpha_3} > CE_{\alpha_1}$.

We can now compute the marginal costs of α_1 and α_3 by adding up the substitutability and complementarity terms above for each worker, along with the risk term, the last term in equation (9):

$$\begin{aligned} MC_{\alpha_1} = & \alpha_1(1 - \lambda^2 - \lambda^4 - \lambda^6 - \lambda^8 - \lambda^{10} - \lambda^{12}) - \alpha_2(\lambda^3 + \lambda^5 + \lambda^7 + \lambda^9 + \lambda^{11}) \\ & - \alpha_3(\lambda^4 + \lambda^6 + \lambda^8 + \lambda^{10}) - \alpha_4(\lambda^5 + \lambda^7 + \lambda^9) - \alpha_5(\lambda^6 + \lambda^8) - \alpha_6(\lambda^7) - r\sigma^2\alpha_1, \\ MC_{\alpha_3} = & -\alpha_1(\lambda^4 + \lambda^6 + \lambda^8 + \lambda^{10}) - \alpha_2(\lambda^3 + \lambda^5 + \lambda^7 + \lambda^9) + \alpha_3(1 - \lambda^2 - \lambda^4 - \lambda^6 - \lambda^8) \\ & - \alpha_4(\lambda^3 + \lambda^5 + \lambda^7) - \alpha_5(\lambda^4 + \lambda^6) - \alpha_6(\lambda^5) - r\sigma^2\alpha_3. \end{aligned}$$

[51] We let $g_{ij} \in \{0, 1\}$ for all $i, j \in N$.

Letting $\sigma^2 = 0$ and taking the difference we get:

$$MC_{\alpha_1} - MC_{\alpha_3} = \alpha_1(1-\lambda^2-\lambda^{12}) + \alpha_4(\lambda^3-\lambda^9) + \alpha_5(\lambda^4-\lambda^8) + \alpha_6(\lambda^5-\lambda^7) \\ - \alpha_2(\lambda^{11}) - \alpha_3(1-\lambda^2+\lambda^{10}).$$

It can be checked numerically that the sum of all coefficients in this difference is positive for small strength of peer effects λ, as imposed by assumptions 1 and 2. For example, for $\lambda = 0.5$ the difference is 0.142822.

Hence, despite worker 1 being more bonacich central than worker 3,[52] with no risk ($\sigma^2 = 0$), the marginal costs associated with α_1 are greater than those associated with α_3, pushing the principal to provide the highest incentives to worker 3.

F Incentives with Random Networks

In this section, we derive optimal incentive rules as a function of the parameters of a statistical process that generates a random network. First, we give the optimal allocation of incentives under the linear technology from the baseline model (Section 2) and, second, under the modular production from Section 4.

Random Networks and Linear Technology

We now proceed by relaxing the full information assumption on the network structure. Imagine that, instead of knowing the realized network, the principal only knows the parameters of a data-generating process of random networks. Examples of such a statistical model include the stochastic block model (SBM) of random graphs and its special cases like the planted partition and Erdős-Rényi models. We incorporate this possibility by considering that the optimal contract maps from a parameterized model of linking probabilities to a function of aggregate output. Employing a mean-field approximation, the firm could anticipate workers' behaviors as a function of the expected adjacency matrix $\bar{\mathbf{G}}$, rather than the realized network, and compute optimal contracts.[53]

[52] For $\lambda = 0.5$, the vector of bonacich centralities is $\mathbf{b}(\lambda) = (1.984, 1.969, 1.937, 1.875, 1.75, 1.5, 1)$.

[53] In fact, recent work using graphons as the underlying stochastic generative process of a network shows that Nash equilibria in finite sampled network games converge to Nash equilibria of the graphon game with high probability (Parise and Ozdaglar (2023)). In such large and dense networks, it has also been shown that measures of centrality, such as Katz-Bonacich, of the expected adjacency matrix is very close to that of a realized network for stochastic block models (Dasaratha (2020)) and, more generally, for graphons (Avella-Medina, Parise, Schaub, and Segarra (2018)) as stochastic generative models of network formation. Because

A flexible family of random networks is the IRN model proposed by Bollobás, Janson, and Riordan (2007) where each agent has a specific "type" from a finite set and an agent of type i is linked to an agent of type j with independent probability p_{ij}. Although solving the optimal contract for a general class of models is beyond the scope of this paper, we consider a very natural special case that has been used extensively in the literature to capture *homophily* in a parsimonious framework (Jackson, 2008).

Proposition 8. *Consider a special case of the IRN model with two types in equally-sized groups (planted partition model). Let p represent the within-type linking probability while q represents the across-type linking probability. Then the optimal allocation of incentives is given by*

$$\alpha_i^\star(p,q) = \frac{1-\lambda\frac{p+q}{2}N}{(1+r\sigma^2)(1-\lambda\frac{p+q}{2}N)^2-(\lambda\frac{p+q}{2}N)^2}, \quad \forall i \in N.$$

Proof of Proposition 8. Consider two groups of size M such that $N = 2M$. Let $\mathbf{J}_M$ be an $M \times M$ matrix of ones. Then, the expected adjacency matrix for the IRN model in this case is given by:

$$\bar{\mathbf{G}} = \begin{pmatrix} p\mathbf{J}_M & q\mathbf{J}_M \\ q\mathbf{J}_M & p\mathbf{J}_M. \end{pmatrix}$$

Notice that the row sum of $\bar{\mathbf{G}}$ is equal to $M(p+q)$. Moreover, the row sum of $\bar{\mathbf{C}} = \mathbf{I} - \lambda\bar{\mathbf{G}}$ is $1 - \lambda M(p+q)$. Because the row and column sums of $\mathbf{I} - \lambda\bar{\mathbf{G}}$ are all the same, we have that the row and column sums of $(\mathbf{I} - \lambda\bar{\mathbf{G}})^{-1}$ are equal to the reciprocal, i.e., $1/(1 - \lambda M(p+q))$. This means we can write:

$$\bar{\mathbf{C}}'\mathbf{1} = \left(\frac{1}{1-N\lambda\frac{p+q}{2}}\right)\mathbf{1}.$$

The firm anticipates workers' behaviors as a function of the expected adjacency matrix $\bar{\mathbf{G}}$, rather than the realized network, and computes optimal contracts. Thus, we have that $\boldsymbol{\alpha} = \alpha_i\mathbf{1}$, i.e., workers are symmetric. Then, the firm solves:

$$\max_{\alpha_i} \quad \alpha_i\mathbf{1}'\bar{\mathbf{C}}'\mathbf{1} - \frac{1}{2}\alpha_i^2\mathbf{1}'\bar{\mathbf{C}}'\bar{\mathbf{C}}\mathbf{1} + \lambda\alpha_i^2\mathbf{1}'\bar{\mathbf{C}}'\bar{\mathbf{G}}\bar{\mathbf{C}}\mathbf{1} - \frac{r\sigma^2}{2}\alpha_i^2 n.$$

Defining $\chi = 1/(1 - N\lambda((p+q)/2))$, substituting in for $\bar{\mathbf{C}}$, and using the fact that $\mathbf{1}\bar{\mathbf{G}}\mathbf{1} =$

large (and dense) firms are precisely the firms that are more likely to not have precise information about the peer network, recent work on random graph theory suggests that designing optimal contracts using a mean-field approximation could prove a valid approach. Moreover, as shown in Parise and Ozdaglar (2023), such an intervention in the space of the graphon is asymptotically optimal and tractable.

$\frac{p+q}{2}N^2$, the firm's problem can be rewritten as:

$$\max_{\alpha_i} \quad \alpha_i \chi - \frac{1}{2}\alpha_i^2(\chi^2(1-\lambda N(p+q)) + r\sigma^2).$$

Taking the first order condition we get:

$$\chi = \alpha_i(\chi^2(1-\lambda N(p+q)) + r\sigma^2) \implies \alpha_i = \frac{\chi}{(\chi^2(1-\lambda N(p+q)) + r\sigma^2)}$$

Finally, going back to the original notation and rearranging we get:

$$\alpha_i^\star(p,q) = \frac{1-\lambda N\frac{p+q}{2}}{(1+r\sigma^2)(1-\lambda N\frac{p+q}{2})^2 - (\lambda N\frac{p+q}{2})^2}.$$

■

Consider how hompohily – – affects optimal incentives in this generative *planted partition random graph model* with two groups. Notice that incentive allocations in Proposition 8 depend on the expected degree of the network (parametrized by $p+q$), but not on the level of expected homophily (parameterized by p/q).

In Proposition 8 we only restrict the information available to the principal. We continue to assume that worker best-reply on the realized structure as in previous sections.[54] Notice that by comparing across different random network models we are able to vary the firm's level of informativeness. A very exciting line of future research follows the approach of Fainmesser and Galeotti (2016) and considers how raising the information content (i.e. higher certainty on the true relationships within the firm) of the firm may affect the optimal contract, and how this may also depend on the information workers have on the network.[55]

Random Networks and Modular Production

We now turn to investigate how homophily alters workers' optimal incentives as well as each module's productivity shares μ_k under the modular production in Section 4 when the network is generate according to the planted partition random graph model.

[54] Previous work has analyzed the equilibrium of general network games when agents only have partial information on the interaction structure (Galeotti, Goyal, Jackson, Vega-Redondo, and Yariv, 2010; Sundararajan, 2008; Fainmesser and Galeotti, 2016). In these models agents know some sufficient statistics of the network rather than the entire structure.

[55] Fainmesser and Galeotti (2016) consider a monopolist pricing a good with network effects and analyze how varying the information available to the monopolist affects the pricing strategy.

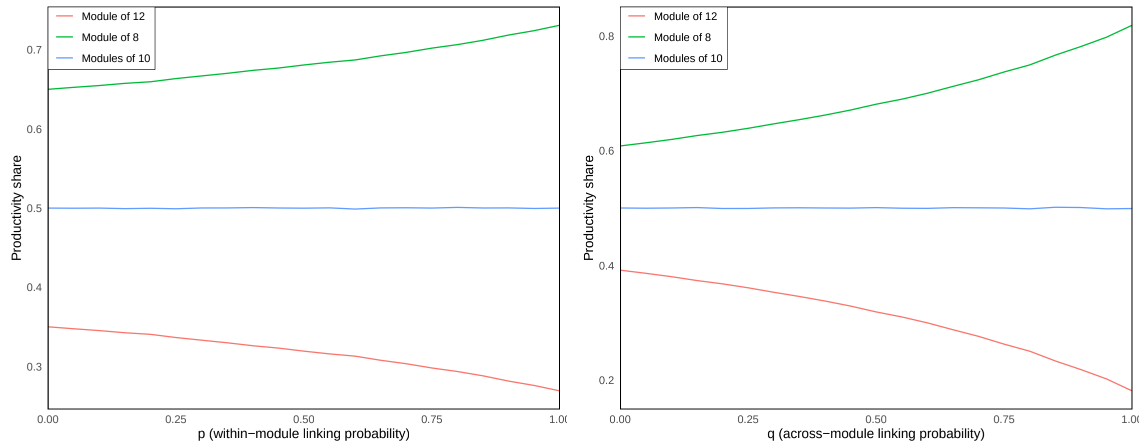

Figure 12: Modular share, μ_k, for the planted partition model with linking parameters p and q. Two equal-sized modules (blue line) and the asymmetric case with two modules of sizes 8 workers (green line) and 12 workers (red line) (parameters: $\lambda = 0.025$, $\sigma^2 = 1$ and $r = 1$).

We consider the example of a firm with $N = 20$ workers and $K = 2$ modules and look at two different cases: equal module sizes (i.e., 10 workers in each module) and asymmetric module sizes with 8 workers in one module and 12 workers in the other. In Figure 12, we plot for the two cases, the mean productivity share of each module, μ_k, across many random networks generated by fixing across-module linking probability $q = 0.5$ and letting within-module linking probability p increase (left) and by fixing $p = 0.5$ and letting q increase (right). This allows us to isolate the impact of each type of link on the productivity share.

Consider first the case with symmetric modules. Because modules are of equal size, the module-specific productivity share is equal to $1/2$ for both modules. In this case, adding more links, no matter if they are within or across modules has no impact on these productivity shares, as illustrated by the flat lines in Figure 12.

Next, consider the case with asymmetric modules of 8 and 12 workers respectively. First, notice that even for $p = 0$ or $q = 0$, workers in the small module are assigned a larger productivity share. That is because these workers have higher costs of effort relative to the workers in the larger module. In fact, without peer effects, i.e., $\lambda = 0$, the productivity shares are exactly 0.6 for the small module and 0.4 for the large one. However, as p or q increases, this gap widens. This is because, the additional links always generate higher complementarities for workers in the large module, no matter if the links are within or across modules, and, hence, workers in the small module have even higher effort costs relative to those in the large module. Lastly, notice that this effect is larger for across-modules links than for within-modules links.

G Organizational Design

This section develops the model's implications for organizational design, which are summarized in Section 5. First, we link firm profits—under the optimal contract from Proposition 1—to the peer network's structure via its principal components. Next, using spectral properties of specific graph families, we classify and compare networks based on their equilibrium profits. Finally, we identify for which networks firms prefer strengthening peer effects over equivalent investments in human capital.

G.1 A Profit Decomposition Result

How do profits under the optimal wage contract depend on network structure? For $\lambda > 0$, additional links enhance incentives and boost firm profits, making the complete graph optimal.[56] When link formation is costly, prior work by Belhaj, Bervoets, and Deroïan (2016) and Hiller (2017) shows that constrained-efficient networks in games with strategic complements belong to the class of nested-split graphs.[57] However, existing research does not identify the most efficient network within this broad class. Moreover, nested-split graphs may be impractical to implement. Firms need to understand which networks yield equivalent profits and which outperform others. To address this, we establish a direct connection between profits and the network's structural (i.e., spectral) properties for any λ.

Proposition 9 (Network Structure and Profits)**.** *In expectation, a firm's profits are maximized at one-half of equilibrium output for any network* $\mathbf{G}$*, any level of peer effects* λ*, and any level of fundamental risk* σ^2*:*

$$\mathbb{E}(\pi^\star(\mathbf{e}|\boldsymbol{\alpha}, \boldsymbol{\beta})) = \frac{1}{2}\mathbb{E}\left(X(\mathbf{e}^\star)\right).$$

Moreover, let $\mathbf{u}_\ell$ *and* $\mathbf{v}_\ell$ *represent the right and left unit-eigenvectors of* $\mathbf{G}$ *associated to eigenvalues* μ_ℓ*. Expected profits are given by:*

$$\mathbb{E}(\pi^\star(\mathbf{e}|\boldsymbol{\alpha}, \boldsymbol{\beta})) = \frac{n}{2}\sum_{\ell} \frac{1 - n\,\mathrm{Cov}\left(\mathbf{u}_\ell, \mathbf{v}_\ell\right)}{(1 + r\sigma^2)(1 - \lambda\mu_\ell)^2 - (\lambda\mu_\ell)^2}, \tag{27}$$

The first part of Proposition 9 extends a well-known result from the team production literature to our setting with bilateral spillovers: optimized profits scale one-to-one with output

[56] For $\lambda < 0$ the efficient network is empty.

[57] These are graphs where each agent's neighbors form a subset of the neighbors of any higher-degree agent. This result does not hold for strategic substitutes ($\lambda < 0$).

(for any network). The second part of Proposition 9 tells us how network structure drives profits by decomposing the network effects into the *principal components* of the underlying graph.

To develop intuition, it is useful to to approximate (27) by focusing on the leading term of the sum above.[58] In this case, we can write

$$\mathbb{E}(\pi^\star) \approx \frac{n}{2} \frac{1 - n \operatorname{Cov}(\mathbf{u}_1, \mathbf{v}_1)}{(1 + r\sigma^2)(1 - \lambda\mu_1)^2 - (\lambda\mu_1)^2} \tag{28}$$

1. Profits decrease with risk: higher σ^2 implies the firm must pay larger compensation packages to employees

2. Profits increase with leading eigenvalue μ_1 (if $\lambda > 0$): Networks with more links and/or more expansive link structures amplify incentives best (opposite is true if $\lambda < 0$).

3. Profits decrease with $\operatorname{Cov}(\mathbf{u}_1, \mathbf{v}_1)$: Profits are larger if workers that have most outward influence on others (high $\mathbf{u}_1$) are NOT the workers that are most influenced themselves (high $\mathbf{v}_1$). For instance, a hierarchical network like those in Figure 1 yield larger profits than more reciprocal networks with comparable overall connectivity, μ_1.

4. Optimal firm size: Keeping everything else constant, $n^\star = \frac{1}{2\operatorname{Cov}(\mathbf{u}_1, \mathbf{v}_1)}$ because firms scale better when an extra influential worker is not heavily influenced by others. Otherwise adding workers does not introduce enough independent contributions and new employees simply reinforce existing dependencies, leading to diminishing returns.

In Section 5 we summarize these points for the special case in which the network is undirected. In this case, the only difference is that the numerator of equation (28) becomes $1 - n\operatorname{Var}(\mathbf{u}_1)$, which modifies the intuition behind points 3 and 4 above slightly: now, profits (and firms) are larger when eigenvector centrality is evenly distributed across workers.

G.2 Comparing Network Structures

The profit characterization above allows us to compare organizational structures based on their expected profits. We begin by focusing on the approximation in equation (16) and later show exact comparisons using spectral properties for certain families of graphs.

[58] This is a particularly good approximation when the spectral gap is large so that eigenvalues drop off quickly.

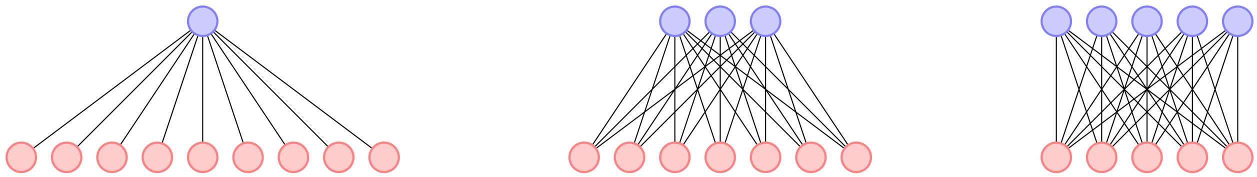

Figure 13: Complete Bipartite graphs with $N = 10$. An asymmetric bipartite graph (left panel) generates lower profits than a symmetric one (right panel).

Consider all networks with a degree distribution that follows a *power law* $P(d) \sim d^{-\gamma}$. Intuitively, as γ increases both μ_1 and $\text{Var}(\mathbf{u}_1)$ in equation (16) increase because the distribution puts more weight on the right tail, creating dominant workers with much larger centrality but also increasing the total number of links. There is a threshold $\gamma^\star$ such that profits decrease with γ above the threshold (because the μ_1 effect dominates) and increase below it (because the $\text{Var}(\mathbf{u}_1)$ effect dominates). This implies that profits are maximized at $\gamma^\star$. A well-known approximation is $\mu_1 \approx \langle d^2\rangle/\langle d\rangle$ and $\mathbf{u}_1(i) \approx \frac{d_i}{\sqrt{\sum_j d_j^2}}$, which implies $\text{Var}(\mathbf{u}_1) = \frac{1}{N}\left(1 - \frac{\langle d\rangle^2}{\langle d^2\rangle}\right)$. Using the fact that $\langle d\rangle = \frac{\gamma-1}{\gamma-2}k_{\min}$ and $\langle d^2\rangle = \frac{\gamma-1}{\gamma-3}k_{\min}^2$, we have $\frac{\langle k\rangle^2}{\langle k^2\rangle} = \frac{(\gamma-1)(\gamma-3)}{(\gamma-2)^2}$. Plugging this in to (16) and taking $k_{\min} = 1$ we obtain

$$\mathbb{E}\left(\pi^\star\right) \approx \frac{n}{2}\frac{1 - \frac{1}{(\gamma-2)^2}}{\left(1 + r\sigma^2\right)\left[1 - \lambda\left(\frac{\gamma-2}{\gamma-3}\right)\right]^2 - \left[\lambda\left(\frac{\gamma-2}{\gamma-3}\right)\right]^2}$$

This expression approximates firm's profits as a function of the decay parameter of the power-law distribution. The exact value of $\gamma^\star$ depends on parameters and can't be obtained analytically..

Exact Comparison of Networks

In what follows we look at classes of networks for which we can solve equation (27) directly and compute profits exactly. Consider a firm deciding how to delegate responsibilities, specifically determining the relative size of two interacting divisions. Suppose all relevant spillovers occur across divisions. The firm must then choose among all complete bipartite graphs of size N, where members are split into groups of size n and m (Figure 13). Using the spectral properties of these graphs,[59] we can express expected profits in terms of n and

[59]For complete bipartite graphs, the eigenvalues are well known: $\mu_1 = \sqrt{mn}$, $\mu_2 = \mu_3 = \cdots = \mu_{n-1} = 0$, and $\lambda_n = -\sqrt{mn}$. The corresponding unit eigenvectors satisfy:

$$\mathbf{u}_{1,i} = \begin{cases} \frac{1}{\sqrt{2m}}, & \text{if } i \text{ is in group } m, \\ \frac{1}{\sqrt{2n}}, & \text{if } i \text{ is in group } n \end{cases}, \quad \mathbf{u}_{n,i} = \begin{cases} \frac{1}{\sqrt{2m}}, & \text{if } i \text{ is in group } m, \\ -\frac{1}{\sqrt{2n}}, & \text{if } i \text{ is in group } n \end{cases}.$$

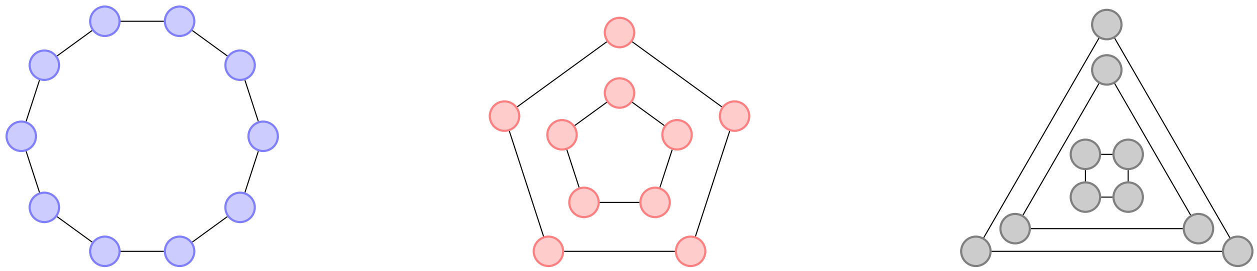

Figure 14: All 2-regular graphs with $N = 10$ give the same profits.

m:

$$2\mathbb{E}(\pi^\star)_{\text{bipartite}} = \frac{\left(\sqrt{m}+\sqrt{n}\right)^2/2}{(1+r\sigma^2)(1-\lambda\sqrt{nm})^2-(\lambda\sqrt{nm})^2} + \frac{\left(\sqrt{m}-\sqrt{n}\right)^2/2}{(1+r\sigma^2)(1+\lambda\sqrt{nm})^2-(\lambda\sqrt{nm})^2}.$$

From this, we identify the profit-maximizing structure among all complete bipartite graphs:

Corollary 4 (**Complete Bipartite Networks**). *Among all complete bipartite graphs with n nodes in group A and m nodes in group B, expected profits are maximized when the two groups are of equal size, i.e., $n = m$.*

Now, consider a different scenario where a homogeneous organization — where each worker is influenced by the same number of peers — decides whether to split into K divisions with sizes $C_1, \ldots, C_k$. Specifically, the CEO must choose between keeping all N workers in a single d-regular structure or dividing them into smaller d-regular components (Figure 14). Again, Proposition 9 helps solve this design problem. Given the spectral properties of d-regular graphs,[60] we express profits as:

$$\mathbb{E}(\pi^\star)_{\text{regular}} = \frac{1}{2}\sum_{i=1}^{k}\frac{C_i}{(1+r\sigma^2)(1-d\lambda)^2-(d\lambda)^2} = \frac{1}{2}\frac{n}{(1+r\sigma^2)(1-d\lambda)^2-(d\lambda)^2}.$$

Expected profits depend only on local spillover structure, not on whether the organization is split into multiple divisions.

Corollary 5 (**Regular Networks**). *All d-regular graphs of size N yield the same expected profits.*

Next, consider the role of homophily – the tendency of individuals to form connections within their group. We analyze this using the planted partition random graph model from

Since $(\mathbf{u}_1'\mathbf{1})^2 + (\mathbf{u}_n'\mathbf{1})^2 = N$, only the first and last terms in Proposition 9 contribute to profits.

[60] For a d-regular graph with k components of sizes $C_1, C_2, \ldots, C_k$, the eigenvalues are $\mu_1 = \mu_2 = \cdots = \mu_k = d$. The corresponding eigenvectors satisfy $\mathbf{u}_i'\mathbf{1} = C_i/\sqrt{C_i}$ for $i = 1, \ldots, k$, while all other eigenvectors are orthogonal to $\mathbf{1}$.

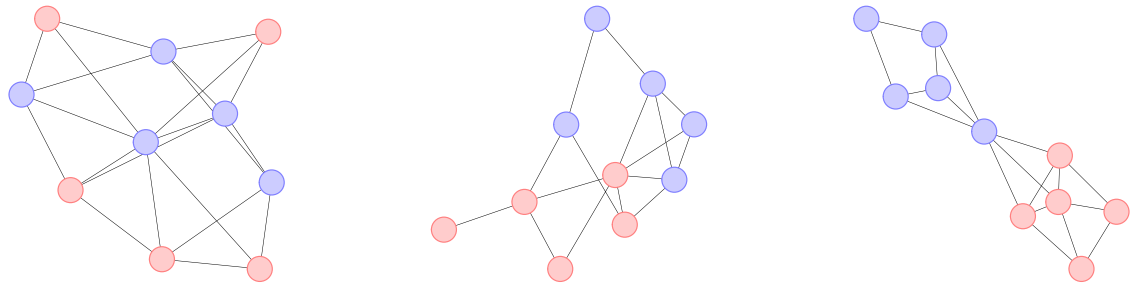

Figure 15: Planted Partition model with $n = 10$ and $p + q = 0.8$.
Panel A: $p = q = 0.4$. *Panel B: $p = 0.6, q = 0.2$.* *Panel C: $p = 0.75, q = 0.05$.*

Proposition 8, where p represents the probability of within-group connections and q the probability of cross-group connections. While $p + q$ determines overall connectivity, homophily is captured by the ratio p/q. Using Proposition 9 and the expected adjacency matrix $\bar{\mathbf{G}} \in \{p, q\}^{n\times n}$ we obtain:[61]

$$\mathbb{E}(\pi^{\star})_{\text{planted partition}} = \frac{1}{2}\frac{n}{(1 + r\sigma^2)(1 - \lambda n\frac{p+q}{2})^2 - (\lambda n\frac{p+q}{2})^2}.$$

Corollary 6 (Community Structure and Profits). *In a planted partition model with connection probabilities p and q, expected profits depend only on average degree (i.e. on $p + q$), and not on expected homophily (i.e. on p/q).*

Directed Networks

We can also use Proposition 9 to compute profits for directed networks. For instance, we could think of an assembly line production process in which each worker directly influences the next co-worker in a closed loop. Then, a *directed cycle* of N nodes (see the left network

[61]The only nonzero eigenvalues of $\bar{\mathbf{G}}$ are $\mu_1 = n(p + q)/2$ and $\mu_2 = n(p - q)/2$, with $\mathbf{u}_1'\mathbf{1} = n/\sqrt{n}$ and $\mathbf{u}_i'\mathbf{1} = 0$ for $i \geq 2$.

Figure 16: Two directed networks with 5 workers: (a) a directed cycle and (b) a regular tournament graph.

in Figure 16) can be shown to have expected profits equal to:[62]

$$\mathbb{E}\left(\pi^{\star}\right)_{\text{directed cycle}} = \frac{1}{2}\frac{n}{\left(1+r\sigma^{2}\right)(1-\lambda)^{2}-\lambda^{2}}$$

Here, the feedback loop means that when a worker takes an action, it has a spillover effect on the entire cycle. In other words, the presence of a cycle allows positive complementarities (when $\lambda > 0$) to be amplified across the network.

For another example, consider a *regular tournament graph* (see the right network in Figure 16).[63] While these graphs have been used to interpret the outcome of a round-robin tournament, we could interpret them in our case as a sales team in which each salesperson is influenced by only one other salesperson, a perceived direct competitor. We can leverage Proposition 9 to express expected maximized profits as:[64]

$$\mathbb{E}\left(\pi^{\star}\right)_{\text{tournament}} = \frac{1}{2}\frac{n}{\left(1+r\sigma^{2}\right)\left(1-\lambda\frac{n-1}{2}\right)^{2}-\lambda^{2}\left(\frac{n-1}{2}\right)^{2}}.$$

In this case, each worker is, on average, similarly influential. Thus, every worker's influence is aggregated over $\lambda\frac{N-1}{2}$ links, while in a directed cycle every node influences only one other node, so the effective amplification of a worker's action is governed by the factor λ. As a

[62] For a directed cycle with homogeneous weights, the adjacency matrix is a permutation matrix. Its eigenvalues and eigenvectors are well-known. In particular, let $\omega = e^{2\pi i/n}$. Then the eigenvalues are $\mu_k = \omega^k = e^{2\pi i k/n}$, $k = 0, 1, \ldots, n-1$ and a standard orthonormal basis of eigenvectors is $\mathbf{u}_k = \frac{1}{\sqrt{n}}\left(1, \omega^k, \omega^{2k}, \ldots, \omega^{(n-1)k}\right)^{\top}$. Now, for $k = 0$ we have $\omega^0 = 1$ so $\mathbf{u}_0'\mathbf{1} = \frac{1}{\sqrt{n}}\sum_{j=0}^{n-1} 1 = \frac{n}{\sqrt{n}} = \sqrt{n}$ while for $k \neq 0$ it can be easily shown that $\mathbf{u}_k'\mathbf{1} = 0$.

[63] A tournament graph is a complete directed graph in which every pair of nodes has exactly one directed link. A regular tournament graph is a tournament graph in which every node has the same in-degree and out-degree. Regular tournament graphs are guaranteed to exist when N is odd.

[64] For regular tournament graphs, each row adds up to the same constant $(n-1)/2$. This implies that $(n-1)/2$ is the first eigenvalue with an eigenvector equal to $\mathbf{1}$. Thus, by standardizing the left and right eigenvectors associated with the first eigenvalue as $\mathbf{v}_1 = \frac{1}{\sqrt{n}}\mathbf{1} = \mathbf{w}_1$ we are guaranteed to satisfy the bi-orthogonality condition. Moreover, we have $\mathbf{v}_1'\mathbf{1} = \mathbf{w}_1'\mathbf{1} = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} 1 = \sqrt{n}$ while for $\ell \geq 2$ it can be shown that either $\mathbf{v}_\ell'\mathbf{1} = 0$ or $\mathbf{w}_\ell'\mathbf{1} = 0$ or both.

result, the denominator in the tournament case is smaller (for $\lambda > 0$) than in the directed cycle case, leading to higher expected profits.

Finally, for directed tree structures, similar to the hierarchical firms shown in Figure 1, the situation is more involved. Since a directed tree is acyclic, its adjacency matrix is nilpotent and has all eigenvalues equal to zero. This means that $\mathbf{G}$ is not diagonalizable in this case. Therefore, one cannot directly apply Proposition 9 as in the previous cases. To obtain a similar expression as for the cases of directed cycles and tournament graphs, one could use the nilpotent property to express the *finite* geometric series of $(\mathbf{I} - \lambda\mathbf{G})^{-1}$ which would depend on the number of levels in the hierarchy, i.e., the depth of the tree.

G.3 How to Invest in Workers

Firms typically invest in their workforce through training programs that enhance either individual skills or teamwork. Should a firm focus on improving workers' human capital or strengthening peer complementarities through team-building exercises? The profit decomposition in Proposition 9 provides a framework to address such human resource decisions in large and complex organizations.

Consider an extension of the baseline model where the marginal cost of effort is influenced by a human-capital parameter $\nu \geq 1$. We can rewrite (1) as

$$\psi_i(\mathbf{e}) = \nu\frac{e_i^2}{2} - \lambda e_i \sum_{j\in N} g_{ji}e_j.$$

Suppose a firm can invest one dollar to either *decrease* ν (reducing effort costs) or *increase* λ (amplifying peer effects). To maintain comparability, we assume both investments have the same per-unit cost. Allowing for different costs would not alter the fundamental insight. The firm's decision reduces to evaluating:

$$\frac{\partial\mathbb{E}(\pi)}{\partial\lambda} \lesseqgtr \left|\frac{\partial\mathbb{E}(\pi)}{\partial v}\right|.$$

Using Proposition 9, we can compute these marginal effects. We show that investing in team strength is preferably to investing in human capital only for those networks with:

$$(1-\mu_\ell)\left(1 + r\sigma^2\left(v - \lambda\mu_\ell\right)\right) < \ 1/2, \quad \forall\ell \text{ with } \mathbf{u}_\ell'\mathbf{1} \neq 0.$$

The first thing to notice is that, as $r\sigma^2$ grows, it is less profitable to invest in team-building exercises, everything else equal. Intuitively, when the cost associated to providing risky incentives increases – either because the firm is very risky or the workforce is very risk averse – performance-based compensation is costly, so investing in peer effects has little impact.

Network structure also plays a crucial role. For instance, the equation above tells us that the only *regular network* for which investing in human capital dominates is the empty network.[65] *Scale-free networks*, which are characterized by a power-law degree distribution, and *small-worlds networks* will typically favor investments in peer strength.[66] When connections are sparse investing in human capital is (obviously) preferable. However, as network density increases, the advantage shifts toward strengthening peer effects. One might wonder where this shift happens, and whether most networks of N nodes favor one investment over the other. A structured way to increase network density is by raising the linking probability p in an *Erdős-Rényi random graph.* As p increases from 0 (isolated nodes) to 1 (complete network), there is a threshold beyond which investing in human capital ceases to be optimal. It turns out that this threshold aligns with the emergence of a giant component.

Proposition 10. *In the Erdős-Rényi Random Graph model, investing in team strength outperforms investing in human capital if and only if* $np \geq 1$.

This surprising result implies that as long as each worker interacts with at least one other worker (in expectation), then investing in team strength is superior to uniformly enhancing worker productivity. Therefore, investments in team-building exercises dominate investments in human capital for most real-world networks.

G.4 Proofs of Section G

Proof of Proposition 9. Recall that the firm's problem can be written in matrix form as:

$$\max_{\boldsymbol{\alpha}} \mathbb{E}[\pi(\mathbf{e} \mid \boldsymbol{\alpha}, \boldsymbol{\beta})] = \boldsymbol{\alpha}'\mathbf{C}'\mathbf{1} - \frac{1}{2}\boldsymbol{\alpha}' \underbrace{\left[\mathbf{C}'\mathbf{C} - 2\lambda\mathbf{C}'\mathbf{G}\mathbf{C} + \sigma^2 r\mathbf{I}\right]}_{P} \boldsymbol{\alpha}$$

Notice that if $\mathbf{G}$ is not symmetric, then the term $\mathbf{C}'\mathbf{G}\mathbf{C}$ is also non-symmetric. However, the matrices $\mathbf{C}'\mathbf{C}$ and I are both symmetric, even if $\mathbf{G}$ is not. Now, we use the fact that the

[65] To see this notice that Assumption 1 requires that $\nu > \lambda\mu_\ell$. Therefore since $\mu_\ell = d$ for all eigenvalues with $\mathbf{u}_\ell'\mathbf{1} \neq 0$ the result follows.

[66] These arguments require checking how quickly the dominant eigenvalue exceeds 1 for different parameter values.

quadratic form for a non-symmetric matrix P only depends on the symmetric part of P:

$$\mathbf{x}'\mathbf{P}\mathbf{x} = \mathbf{x}'\left(\frac{\mathbf{P}+\mathbf{P}'}{2}\right)\mathbf{x}. \tag{29}$$

This means we can replace $\mathbf{C}'\mathbf{G}\mathbf{C}$ by its symmetric part $(1/2)(\mathbf{C}'\mathbf{G}\mathbf{C} + \mathbf{C}'\mathbf{G}'\mathbf{C})$ without changing the value of the quadratic form $\boldsymbol{\alpha}'\mathbf{P}\boldsymbol{\alpha}$ and, hence, without changing the optimization problem, which we can write now as:

$$\max_{\boldsymbol{\alpha}} \mathbb{E}[\pi(\mathbf{e} \mid \boldsymbol{\alpha}, \boldsymbol{\beta})] = \boldsymbol{\alpha}'\mathbf{C}'\mathbf{1} - \frac{1}{2}\boldsymbol{\alpha}' \underbrace{\left[\mathbf{C}'\left(\mathbf{I} - \lambda(\mathbf{G}' + \mathbf{G})\right)\mathbf{C} + \sigma^2 r\mathbf{I}\right]}_{P} \boldsymbol{\alpha}.$$

This problem yields our solution for optimal incentives in Proposition 5:

$$\boldsymbol{\alpha}^* = \left(\mathbf{C}'(\mathbf{I} - \lambda(\mathbf{G} + \mathbf{G}'))\mathbf{C} + \sigma^2 r\mathbf{I}\right)^{-1}\mathbf{C}'\mathbf{1} = \mathbf{P}^{-1}\mathbf{C}'\mathbf{1}.$$

Substituting this into the objective function we see that, at the optimum, maximum expected profits are:

$$\mathbb{E}[\pi^*] = (\mathbf{C}'\mathbf{1})'\mathbf{P}^{-1}\mathbf{C}'\mathbf{1} - \frac{1}{2}(\mathbf{C}'\mathbf{1})'\mathbf{P}^{-1}\mathbf{P}\mathbf{P}^{-1}\mathbf{C}'\mathbf{1} = \frac{1}{2}(\mathbf{C}'\mathbf{1})'\mathbf{P}^{-1}\mathbf{C}'\mathbf{1} = \frac{1}{2}\mathbf{1}'\mathbf{C}\boldsymbol{\alpha}^* = \frac{1}{2}\mathbf{1}'\mathbf{e}^* = \frac{1}{2}\mathbb{E}[X(\mathbf{e})],$$

where we have used the expression for equilibrium efforts $\mathbf{e} = \mathbf{C}\boldsymbol{\alpha}$. Thus, we have shown that, in expectation, the firm's profits are maximized at one-half of the equilibrium output for any network $\mathbf{G}$.

We now prove the second part of the proposition. We can write the worker's equilibrium condition as:

$$\left[(\mathbf{I} - \lambda(\mathbf{G} + \mathbf{G}')) + \sigma^2 r(\mathbf{I} - \lambda\mathbf{G})(\mathbf{I} - \lambda\mathbf{G}')\right]\mathbf{e}^* = \mathbf{1}.$$

We assume that the non-symmetric matrix $\mathbf{G}$ is diagonalizable with the right unit-eigenvectors $\mathbf{u}_\ell$ and the left unit-eigenvectors $\mathbf{v}_\ell$ satisfying

$$\mathbf{G}\mathbf{u}_\ell = \mu_\ell\mathbf{u}_\ell, \quad \text{and} \quad \mathbf{v}_\ell'\mathbf{G} = \mu_\ell\mathbf{v}_\ell'$$

for corresponding eigenvalues μ_ℓ, which may be complex but will appear in conjugate pairs if $\mathbf{G}$ is real. Moreover, we also assume the bi-orthogonality condition:

$$\mathbf{v}_\ell'\mathbf{u}_k = \delta_{\ell k},$$

where $\delta_{\ell k}$ is the Kronecker delta.[67] Then, defining the matrices $\mathbf{U}$ with columns $\mathbf{u}_\ell$ and $\mathbf{V}$ with columns v_ℓ we also have that $\mathbf{V}'\mathbf{U} = \mathbf{I}$.

We expand $\mathbf{e}^*$ in the basis of the right eigenvectors: $\mathbf{e}^* = \sum_\ell c_\ell \mathbf{u}_\ell$. We multiply the equilibrium effort equation on the left by $\mathbf{v}_i'$, for any arbitrary fixed i:

$$\mathbf{v}_i' \left[(\mathbf{I} - \lambda (\mathbf{G} + \mathbf{G}')) + \sigma^2 r (\mathbf{I} - \lambda \mathbf{G}) (\mathbf{I} - \lambda \mathbf{G}') \right] \mathbf{e}^* = \mathbf{v}_i' \mathbf{1}.$$

Substitute the expansion $\mathbf{e}^* = \sum_\ell c_\ell \mathbf{u}_\ell$ into the equation above:

$$c_i \mathbf{v}_i' \left[(\mathbf{I} - \lambda (\mathbf{G} + \mathbf{G}')) + \sigma^2 r (\mathbf{I} - \lambda \mathbf{G}) (\mathbf{I} - \lambda \mathbf{G}') \right] \mathbf{u}_i = \mathbf{v}_i' \mathbf{1},$$

where, by linearity and the bi-orthogonality condition, only the ith term in the expansion survives. Next, notice that, again by the bi-orthogonality condition, the following is true:

$$\mathbf{v}_i' \mathbf{G} \mathbf{u}_i = \mu_i \mathbf{v}_i' \mathbf{u}_i = \mu_i, \quad \text{and} \quad \mathbf{v}_i' \mathbf{G}' \mathbf{u}_i = (\mathbf{G} \mathbf{v}_i)' \mathbf{u}_i = \mu_i \mathbf{v}_i' \mathbf{u}_i = \mu_i.$$

Thus, using these two equations we can re-write the terms in brackets in the equilibrium effort condition. The first term is:

$$\mathbf{v}_i' \left[(\mathbf{I} - \lambda (\mathbf{G} + \mathbf{G}')) \right] \mathbf{u}_i = 1 - 2\lambda \mu_i,$$

and the second term in the brackets, ignoring the constant $\sigma^2 r$, is:

$$\mathbf{v}_i' \left[(\mathbf{I} - \lambda \mathbf{G}) (\mathbf{I} - \lambda \mathbf{G}') \right] \mathbf{u}_i = (1 - \lambda \mu_i) \mathbf{v}_i' (\mathbf{I} - \lambda \mathbf{G}') \mathbf{u}_i = (1 - \lambda \mu_i)(\mathbf{v}_i' \mathbf{u}_i - \lambda \mathbf{v}_i' \mathbf{G}' \mathbf{u}_i) = (1 - \lambda \mu_i)^2$$

where in the first equality we have used the fact that $(\mathbf{I} - \lambda \mathbf{G}') \mathbf{u}_i = (1 - \lambda \mu_i) \mathbf{u}_i$, since $\mathbf{u}_i$ is a right unit-eigenvector of $\mathbf{G}'$ and $\mathbf{v}_i'$ is a left unit-eigenvector of $\mathbf{G}$, and in the third equality we used the bi-orthogonality condition and the fact that $\mathbf{v}_i' \mathbf{G}' \mathbf{u}_i = \mu_i$, as shown above. Hence, putting everything together we can re-write the equilibrium effort condition as:

$$c_i \left[(1 - 2\lambda \mu_i) + \sigma^2 r (1 - \lambda \mu_i)^2 \right] = \mathbf{v}_i' \mathbf{1} \iff c_i = \frac{\mathbf{v}_i' \mathbf{1}}{(1 - 2\lambda \mu_i) + \sigma^2 r (1 - \lambda \mu_i)^2}.$$

[67] The Kronecker delta is defined as:

$$\delta_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j. \end{cases}$$

Plugging c_i into the ith component of the expansion of $\mathbf{e}^*$ we obtain:

$$e_i^* = \sum_\ell \frac{u_{\ell,i}(\mathbf{v}_\ell'\mathbf{1})}{(1-2\lambda\mu_\ell)+\sigma^2 r(1-\lambda\mu_\ell)^2},$$

where $u_{l,i}$ is the ith entry of the right unit-eigenvector $\mathbf{u}_\ell$. Since maximized profits are equal to one-half of the sum of equilibrium efforts, we have:

$$\mathbb{E}(\pi^\star(\mathbf{e}|\boldsymbol{\alpha},\boldsymbol{\beta})) = \frac{1}{2}\sum_i\sum_\ell \frac{u_{\ell,i}(\mathbf{v}_\ell'\mathbf{1})}{(1-2\lambda\mu_\ell)+\sigma^2 r(1-\lambda\mu_\ell)^2} = \frac{1}{2}\sum_\ell \frac{(\mathbf{v}_\ell'\mathbf{1})(\mathbf{u}_\ell'\mathbf{1})}{(1-2\lambda\mu_\ell)+\sigma^2 r(1-\lambda\mu_\ell)^2},$$

where $\sum_i u_{\ell,i} = \mathbf{u}_\ell'\mathbf{1}$. We can re-write the numerator because we have that:

$$\text{Cov}(\mathbf{v}_\ell,\mathbf{u}_\ell) = \frac{1}{n}\sum_{i=1}^n [(\mathbf{u}_\ell)_i - \bar{u}_\ell]\,[(\mathbf{v}_\ell)_i - \bar{v}_\ell] = \frac{1}{n}\mathbf{v}_\ell'\mathbf{u}_\ell - \bar{u}_\ell\bar{v}_\ell.$$

Then, we have that $\mathbf{v}_\ell'\mathbf{u}_\ell = 1$ by the bi-orthogonality condition. Thus, multiplying both sides of the equation by n, we get:

$$n\,\text{Cov}(\mathbf{v}_\ell,\mathbf{u}_\ell) = 1 - n\bar{u}_\ell\bar{v}_\ell \iff 1 - n\,\text{Cov}(\mathbf{v}_\ell,\mathbf{u}_\ell) = n\bar{u}_\ell\bar{v}_\ell.$$

Thus, since $n\bar{u}_\ell = (\mathbf{u}_\ell'\mathbf{1})$ and $n\bar{v}_\ell = (\mathbf{v}_\ell'\mathbf{1})$, we can re-write the equation above as:

$$n(1 - n\,\text{Cov}(\mathbf{v}_\ell,\mathbf{u}_\ell)) = (\mathbf{v}_\ell'\mathbf{1})(\mathbf{u}_\ell'\mathbf{1}),$$

which gives the numerator in equation (27). Finally, notice that if $\mathbf{G}$ is a symmetric matrix, then the left and right eigenvectors coincide, $\mathbf{u}_1 = \mathbf{v}_1$, which implies that $\text{Cov}(\mathbf{u}_1,\mathbf{v}_1) = \text{Var}(\mathbf{u}_1)$ and everything else remains the same.

■

Proof of Corollary 10. Applying Proposition 9 to the augmented model with the human-capital parameter ν, we can write the firm's profits as:

$$\mathbb{E}(\pi^\star(\mathbf{e}|\boldsymbol{\alpha},\boldsymbol{\beta})) = \frac{1}{2v}\sum_\ell \frac{(\mathbf{1}'\mathbf{u}_\ell)^2}{(1+v\sigma^2 r)(1-\frac{\lambda}{\nu}\mu_\ell)^2 - \left(\frac{\lambda}{\nu}\mu_\ell\right)^2}$$

We want to compute the partial derivatives with respect to the human-capital parameter ν and with respect to the peer effects parameter λ. To simplify notation, let $(\mathbf{1}'\mathbf{u}_\ell)^2/2 = a$,

$\sigma^2 r = b$. Then, we have:

$$\frac{\partial}{\partial \nu}\left(\frac{a}{\nu\left[(1+\nu b)(1-\frac{\lambda}{\nu}\mu_\ell)^2 - \left(\frac{\lambda}{\nu}\mu_\ell\right)^2\right]}\right) = -\frac{a\left(2b(\nu-\mu_\ell\lambda)+1\right)}{\left(-2b\mu_\ell\nu\lambda + \mu_\ell\lambda(b\mu_\ell\lambda - 2) + b\nu^2 + \nu\right)^2}$$

$$\frac{\partial}{\partial \lambda}\left(\frac{a}{\nu\left[(1+\nu b)\left(1-\left(\frac{\lambda\mu_\ell}{\nu}\right)^2\right) - \left(\left(\frac{\lambda\mu_\ell}{\nu}\right)^2\right)\right]}\right) = \frac{2a\mu_\ell\left(b(\nu-\mu_\ell\lambda)+1\right)}{\left(-2b\mu_\ell\nu\lambda + \mu_\ell\lambda(b\mu_\ell\lambda - 2) + b\nu^2 + \nu\right)^2}$$

Thus, as long as $a = \mathbf{u}_\ell'\mathbf{1} \neq 0$, the condition for $\frac{\partial}{\partial\lambda}(\cdot) > -\frac{\partial}{\partial\nu}(\cdot)$, which, if true for all ℓ, implies $\frac{\partial\pi}{\partial\lambda} > -\frac{\partial\pi}{\partial v}$, is:

$$2a\mu_\ell(b(\nu-\mu_\ell\lambda)+1) > a(2b(\nu-\mu_\ell\lambda)+1) \iff 2\mu_\ell b(\nu-\mu_\ell\lambda) - 2b(\nu-\mu_\ell\lambda) + 2\mu_\ell > 1$$
$$\iff 2b(\nu-\mu_\ell\lambda)(\mu_\ell - 1) + 2\mu_\ell - 2 > -1 \iff (b(\nu-\mu_\ell\lambda)+1)(\mu_\ell - 1) > -\frac{1}{2}$$

Going back to the original notation, we have that a sufficient condition for $\frac{\partial\pi}{\partial\lambda} > -\frac{\partial\pi}{\partial v}$ is that, for all ℓ with $\mathbf{u}_\ell'\mathbf{1} \neq 0$, we have:

$$(1-\mu_\ell)(1 + r\sigma^2(v - \mu_\ell\lambda)) < \frac{1}{2} \quad \forall \ell \text{ with } \mathbf{u}_\ell'\mathbf{1} \neq 0.$$

Finally, notice that in special case of a planted partition model (IRN with 2 types), it is well known that the first and second eigenvalues and their corresponding eigenvectors are of the random adjacency matrix are:

$$\mu_1 = \frac{p+q}{2}\cdot n, \qquad \mathbf{u}_1 = \frac{1}{\sqrt{n}}\mathbf{1} \qquad \Longrightarrow \qquad (\mathbf{u}_1'\mathbf{1})^2 = n,$$

$$\mu_2 = \frac{p-q}{2}\cdot n, \qquad \mathbf{u}_{2,i} = \begin{cases} \frac{1}{\sqrt{n}}, & \text{if } i \text{ is in group 1,} \\ -\frac{1}{\sqrt{n}}, & \text{if } i \text{ is in group 2.} \end{cases} \qquad \Longrightarrow \qquad (\mathbf{u}_2'\mathbf{1})^2 = 0$$

In the Erdős-Rényi Random Graph model ($p = q$) we therefore have that $\mu_1 = np$, $(\mathbf{u}_1'\mathbf{1})^2 = n$, and $(\mathbf{u}_\ell'\mathbf{1})^2 = 0, \forall \ell > 1$. Therefore, a sufficient condition for $\frac{\partial\pi}{\partial\lambda} > -\frac{\partial\pi}{\partial v}$ is that $\mu_1 = np = 1$. That is, if workers have at least one peer in expectation, it is profitable to invest in strengthening complementarities over investing in straightening human capital.

■