Barcelona School of Economics

WORKING PAPERS

Incentive Contracts and Peer Effects in the Workplace

BSE Working Paper 1457| August 2024

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§

August 19, 2024

Latest version.

Abstract

We study the problem of a principal designing wage contracts that simultaneously incentivize and insure workers. Workers' incentives are connected through chains of productivity spillovers, represented by a network of peer-effects. We solve for the optimal linear contract for any network and show that optimal incentives are steeper for more central workers. We link firm profits to organizations' structure via the spectral properties of the co-worker network. When production is modular, the incentive allocation rule is sensitive to the link structure across and within modules. When firms can't write personalized contracts, better connected workers extract rents and total surplus is reduced. In this case, unemployment emerges endogenously because large within-group differences in centrality can decrease firm's profits.

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, Yves Zenou, participants at the 9th European Conference on Networks (Essex), 27th CTN Conference (Budapest), BSE Summer Forum (Barcelona), 2024 Conference on Mechanism and Institution Design (Budapest), and seminar participants at Aix-Marseille School of Economics, UAB and UPF 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.

†Stony Brook University (mclaveriamayol@gmail.com)

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

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

1 Introduction

Wages differ substantially within firms. The distribution of labor earnings tends to skew to the right and displays long right tails. Mean earnings exceed median earnings, and top earners account for a disproportionate share of total earnings. This trend has only heightened in recent years as executive compensation has taken off dramatically, particularly in the form of large stock options, bonuses, and other forms of *variable pay*. The most recent micro data shows that salaries of the top 1% of US employees respond four times more to firm performance than the bottom 1%. At the same time, the structure of organizations has "flattened at the top", with the number of workers reporting directly to upper management increasing over time. Having a larger number of subordinates *amplifies* managerial talent and leads to a *steep earnings profile* within the firm, for most occupational hierarchies.[1]

We currently lack a general framework that connects wages to organizational structure. Building on the theory of incentive design in *network games*, we characterize optimal wage contracts in an agency model with moral hazard and productivity spillovers.[2] In our framework, risk-averse workers collaborate in teams to produce joint output, and effort is not contractible. In the simplest version, output is just the sum of everyone's effort plus a random component. Productivity spillovers enter through workers' marginal cost of effort, which depends on the effort level of their peers. This implies that, in equilibrium, each worker's effort is *amplified* via a complementarity network, which we take as fixed. As a result, better-connected individuals exert a disproportionate influence on overall firm performance. Our theory relates wages to workers' positions in their organizations. Among other things, our model confirms that greater *spans of control* rise performance-pay gaps and contribute to growing earnings disparity.

[1] See Neal and Rosen (2000) for a detailed survey of earnings inequality in firms and existing theories that try to explain it. Jensen, Murphy, and Wruck (2004) note that CEO pay in S&P 500 firms surged from $850,000 in 1970 to over $14 million in 2000, with stock options driving over half of this increase. Wallskog, Bloom, Ohlmacher, and Tello-Trillo (2024) argue that the disparity in pay-response between top and bottom 1% of workers accounts for 40% of the increase in the pay gap between CEOs and median workers from 1980 to 2013. See Bertrand (2009) for a survey of the literature on CEO pay.

In terms of flattening firms, Rajan and Wulf (2006) find that the number of managers reporting directly to the CEO has increased from approximately four in 1986 to more than seven by 1999. Possible explanations for this trend have to do with knowledge hierarchies (Garicano, 2000), or competitive pressures from trade liberalization rendering tall corporate hierarchies obsolete (Guadalupe and Wulf, 2010).

Fox (2009) regresses log wages on a measure of managerial span-of-control. He finds that span-of-control wage gaps increase as you move up the hierarchy. For sales workers, rank-4 workers who supervise three times the number of workers earn 1% more. For rank-7 workers, the span of control wage gap is 3.4%. Smeets and Warzynski (2008) also find that the span of control is positively related to wages and bonuses. Managers supervising a team twice as large as the average implies a salary difference of 2.8%. For middle managers, being indirectly responsible for twice as many workers implies a wage difference of 4.1%.

[2] Network games describe environments were agents best-reply to a "local" subset of other players and the overlapping sets of players defines 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.

There are many reasons why workers may produce the same amount at a lower cost when their peers are more active. Peer effects typically represent "help at work", such as advice, mentorship, or learning, but they may also capture psychological costs like pressure and guilt. Productivity spillovers have been detected in many different kinds of organizations, operating via firms' physical layout (Mas and Moretti, 2009), their corporate hierarchy (Lieber and Skimmyhorn, 2018), or through informal friendship bonds (Bandiera, Barankay, and Rasul, 2005).[3] Most (though not all) studies find that peer effects confer a positive and significant effect on productivity, amounting to nearly half of the total effect of performance pay (Ashraf and Bandiera, 2018).[4] According to recent estimates, firms can realize large productivity gains by designing shift schedules that optimize team composition and thus leverage social incentives.[5] We focus, instead, on how monetary incentives interact with social incentives, and what this means for optimal remuneration as well as optimal team composition.

Workers in our model are paid both fixed and variable wages. Variable payments increase with the firm's output, thus motivating workers but also introducing risk. Higher risk means that a firm must sacrifice larger profits to compensate its workforce. Firms must therefore strike an optimal balance between providing incentives on the one hand and insuring workers on the other. This is a key trade-off in classical incentive design models (Holmstrom and Milgrom, 1987, 1991; Bolton and Dewatripont, 2004). We introduce organizational structure into the textbook theory of contracts and show that, if properly incentivized, *central workers* can distribute incentives across the entire organization *without spreading risk*. This logic pushes firms to concentrate high-powered incentives on specific members of the workforce.

We first examine a scenario where the firm can make a *personalized* take-it-or-leave-it wage offer based on each worker's position in the broader co-worker network. Since firms are able to extract all surplus, the optimal wage contract is efficient in this case. Our first main result (Proposition 1) characterizes the efficient wage contract for any organizational structure and distribution of observable worker characteristics.[6] In Proposition 2, we show that firms effectively concentrate high-powered incentives on workers that are "closest" to the rest of the workforce, as determined by a statistic that aggregates *paths of different lengths* in a specific manner. This new measure of centrality emerges as the natural incentives target in our setup

[3]See Ashraf and Bandiera (2018) for a comprehensive survey of the empirical literature on peer effects in the workplace.

[4]Workers' effort choice has been found to respond to co-workers' effort choice even when remuneration is independent of output (Falk and Ichino, 2006).

[5]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.

[6]While Proposition 1 in the main text focuses on a model with homogeneous workers, the Supplementary Appendix extends the analysis to varying risk aversion, marginal costs of effort, productivity, and bilateral influences.

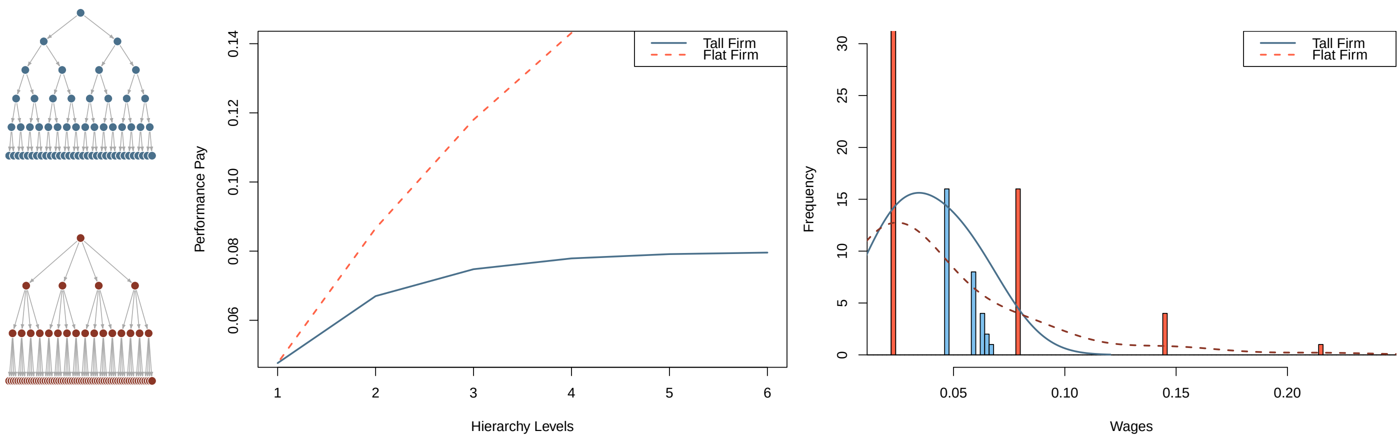


Figure 1: Performance pay and overall wage distribution for two different organizational structures: a tall (blue) firm and a flat (red) firm.

with moral hazard and peer effects.[7]

Consider our theory's implications for the earnings distribution. Figure 1 illustrates the implication of Proposition 1 as organizations become flatter. We simulate our results for two hierarchical organizations: a *tall firm* with 6 levels and a (direct) span of control of 2, and a *flat firm* with 4 levels and a (direct) span of control of 4. Consistent with empirical findings, we observe that variable-pay increases as we move up the organization, and that this pay gap is more pronounced for flatter firms. The second panel shows that the distribution of total earnings (which includes variable and fixed-pay) has a longer right tail in flatter firms, where managers oversee more subordinates. Notice also that the wage gap between top earners is small in tall firms (blue bars) and much larger for flat firms (red bars), resulting in a steeper earnings profile in flatter organizations. Simply put, if productivity is amplified through formal chains of command, and if eliciting effort is costly, jobs with greater "amplification potential" should receive higher pay, particularly in the form of performance-based compensation such as bonuses and stock options. As firms flatten (and spans of control increase), upper management obtains shorter paths to the rest of the workforce, so their amplification potential rises. This makes concentrating incentives on a few *central* workers more advantageous, resulting in the steeper earnings profiles observed in recent data.[8]

[7]Unlike previous work on optimal interventions (Galeotti, Golub, and Goyal, 2020), our incentives target is not proportional to the principal component of the peer network because our objective must also account for workers' risk exposure, which affects the principal's intervention cost differently. We find that our centrality target is essentially a linear transformation of the well-known Katz-Bonacich measures of network centrality, where the precise transformation depends on the pattern of "common influences", as shown below.

[8]Alternative theories on within-firm earnings inequality typically begin with distributional assumptions about managerial talent and, through a process of self-selection (Roy, 1951; Heckman and Honore, 1990), sorting (Gabaix and Landier, 2008), learning (Garicano and Rossi-Hansberg, 2006), and tournaments (Rosen, 1986), talent allocates along job levels to generate the distribution of earnings. We start from the observation that

We also characterize the incentive allocation rule when the firm does not have detailed information on the full structure of productivity spillover across workers. We assume instead the firm holds a *parametric model* in mind which determines the probability that any two workers are linked. We show in Proposition 4 how performance-based compensation should look as a function of the parameters of a very simple mental model of workplace relations. This opens the door to examining more sophisticated models, and, in the spirit of Fainmesser and Galeotti (2016), to explore how wage contracts depend on the information content available to employers.

How should firms organize internally in order to leverage productivity spillovers and allocate incentives most profitably? To tackle this question we analyze how overall firm performance depends on organizational structure. In Proposition 5 we connect optimized profits to the peer effects network through a convenient summary statistic. Specifically, following a technique in Galeotti et al. (2020) we link firms' profits to the eigenvalues and eigenvectors of the associated graph, thus reducing the complexity of an organization to its principal components. We reveal that very different firm-structures are in fact profit-equivalent, and we complement established results on optimal complementarity networks and *nested split graphs* (Belhaj, Bervoets, and Deroïan, 2016).[9] We also use Proposition 5 to compare different investment strategies by providing an average connectivity threshold beyond which firms prefer to invest in "team-building exercises" that strengthen peer effects, over comparable investments in individual human capital.

Production technologies also shape how organizations distribute incentives. The results discussed so far assume that each worker's effort is substitutable. In Section 3 we extend our results to *modular production*, where firms assemble their final product from essential components (Kremer, 1993; Baldwin and Clark, 2003; Garud, Kumaraswamy, and Langlois, 2009; Matouschek, Powell, and Reich, 2023). While versatile, these firms are vulnerable to the failure of any single component.[10] To model this, we divide the workforce into teams, and we assume the production function is substitutable within teams but complementary across teams. We allow for any peer-to-peer link structure within and across teams and we derive optimal contracts for any partition of workers to teams. The special case of a single module recovers the basic model we have been discussing, but for any other modular structure incentives are allocated very differently. For instance, if each worker is in a separate module then every worker is *essential* and firms no longer allocate incentives to those workers with most amplification potential

talent does not change simply due to a promotion, which means that firms tie wages to job positions as well as individuals. Highly skewed wage distributions within firms may in fact represent an *incentive device* (Neal and Rosen, 2000).

[9]Nested-split graphs are networks where agents' neighborhoods are ordered by set inclusion. That is, if i has more friends than j, then i knows everyone j knows. Belhaj et al. (2016) show that NSGs maximize total utility under linear linking costs. König, Tessone, and Zenou (2014) show NSGs are stochastically stable in a preferential attachment process of network formation.

[10]Kremer (1993) popularized the O-ring theory of economic development, based on the fatal Challenger spacecraft incident in 1986, which malfunctioned due to the failure of one small metal gasket.

(i.e. more outgoing paths). Instead, firms prioritize the least susceptible workers (i.e. those with less incoming links), because they are most tempted to lower their effort and hurt overall output. More generally, we find that incentives don't propagate as well across modules and that performance-pay profiles – such as those in Figure 1 – are very sensitive to the precise modular structure, exhibiting sharp jumps across modules.

In the final section of the paper (Section 4) we analyze how incentives are allocated when firms must offer a *one-size-fits-all* contract to all workers with the same job title. This is motivated by recent "fairness norms" in wage determination, which are designed to address undesirable effects that have been linked to large pay disparities within seemingly similar jobs.[11] To understand the impact of wage benchmarking on incentive allocation, we assign workers into occupational categories and restrict firms to offer the same wage contract to everyone in a category. We find that, when it interacts with our forces on incentive allocation and peer complementarities, wage compression can lead to unintended welfare effects – including lower surplus and higher unemployment – that are not apparent in a model without these ingredients.

We begin by characterizing the incentive allocation rule under *wage benchmarking* for any assignment of workers into categories and any arbitrary peer network. We show in Proposition 8 that incentives are now distributed based on an occupational-wide network measure that aggregates the centrality of all members within a job category. More importantly, the inability to provide personalised wage contracts diminishes surplus, similar to the dead-weight loss generated by a monopolist unable to implement perfect price discrimination. The link pattern within and across job categories determines how salary benchmarking affects welfare. We show in Proposition 10 that the loss in overall surplus can be captured by a simple statistic: within-group variance in centrality. Indeed, if all members of a job category are equally central (i.e. if within-group variance is 0), assigning them the same wage incurs no cost in our model. However, efficiency decreases (almost) linearly with aggregate within-variability, while differences in centrality across job categories do not affect welfare. This implies that, when benchmarking salaries, firms with seemingly identical workers might actually face large efficiency costs coming from underlying differences in productivity spillovers.

Finally, one-size-fits-all wage policies not only reduce surplus, they also affect rent sharing between workers and their employers because firms cannot ensure each worker their reservation utility. In Section 4.2, we show that firms extract all surplus only from the "poorest connected" workers in each job category, while others receive *centrality rents* based on their network connec-

[11] 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.

tions. This may lead firms to exclude the least connected workers: it lowers output but improves the firm's share of remaining surplus. As in the literature on efficiency wage models (Shapiro and Stiglitz, 1984; Akerlof and Yellen, 1986), unemployment emerges endogenously from firms' inability to personalize contracts, albeit for very different reasons. Unemployment here does not serve as a worker discipline device, because firms can't fire workers upon observing low performance (in fact, we assume the firm observes joint output only). Instead, unemployment is voluntary here. Simply put, firms optimally design wage contracts that don't satisfy everyone's participation constraint. In Proposition 11, we characterize the ensuing level of unemployment as a function of the distribution of centrality rents. We show that workers remain unemployed if the distribution within each occupational category is sufficiently disperse.

Related Literature - We contribute to the theoretical literature on contract design with multiple agents (Mookherjee, 1984; Macho-Stadler and Pérez-Castrillo, 1993; Bolton and Dewatripont, 2004; Winter, 2010) going all the way back to the original contribution on moral hazard in teams due to Holmstrom (1982). We connect this canonical framework to firms' organizational structure by analyzing familiar pay-for-output incentive schemes when workers interact through productivity spillovers. We show that, under certain specifications, workers' behavior can be cast as a modified version of the linear-quadratic network game pioneered by Ballester et al. (2006). Thus our key theoretical contribution is to extend the canonical moral hazard problem with one principal and many agents using the tools from the networks literature on peer effects. Prior work on targeting incentives on networks have analyzed more abstract non-monetary notions of incentives different from the classical performance-pay elements in standard contract theory. (Demange, 2017; Belhaj and Deroïan, 2018; Galeotti et al., 2020; Parise and Ozdaglar, 2023).

The main applied contribution is to use that model to speak to questions at the heart of organizational economics. This allows us to explore adjacent questions, such as as how optimal incentive design depends on the granularity of the contract and the modular structure of production. There's been a lot of very 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.

A more distant but still relevant recent literature on optimal price discrimination with (local) network effects shares some conceptual aspects with our approach. In all these papers a monopolist must design a menu of prices/discounts to exploit an existing consumer network of externalities. Bloch and Quérou (2013) find market conditions under which price discrimination is not optimal because the tendency to subsidize central consumers (in order to spur demand) is exactly offset by their higher willingness to pay.Fainmesser and Galeotti (2016) vary the information the firm has about the existence and quality of network effects. Interestingly, they find that improving information leads to more price discrimination and can improve overall welfare when price effects are sufficiently 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 labor 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 instead.

Roadmap- The rest of the paper is organized as follows. Section 2 presents the baseline model and solves for the optimal contract when the firm knows and doesn't know the full network. We provide comparative statics results and discuss the case of negative spillovers. In Section 2.3 we relate profits to the spectral properties of the co-worker network. In Section 3 we generalize the firm's technology and consider how contracts change under modular production. In Section 4 we restrict firms to offer a one-size-fits-all contract to everyone in an occupational category and we derive a new theory of unemployment. We conclude in Section 5 with a brief discussion of how our findings open 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 $N = \{1, 2, \ldots, n\}$, to conduct a joint production process.[12] Each worker chooses individual effort $e_i \in \mathbb{R}_+$ and the firm's production is given by

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

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

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

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

where β_i is a fixed payment and α_i captures the contract's variable payment or performance-based compensation.[13] 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.[14]

We assume that workers are embedded in a fixed and exogenous network of peers, represented by the adjacency matrix $\mathbf{G}$.[15] Depending on the application in mind, the network might reflect the firm's organizational structure – i.e. how roles and responsibilities are assigned– but it can also capture the informal bonds of assistance that are forged between workers. Effort costs are assumed to be a linear-quadratic function of own and neighbors' efforts following the standard form in the peer effects literature:

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

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 specifies to the classical textbook model in Bolton and Dewatripont (2004).

Workers are assumed to be risk averse with constant absolute risk aversion (CARA) parameter r, and preferences given by:

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

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

[13] 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.

[14] 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.

[15] The network is allowed to be weighted and directed. A link from i to j is represented by $g_{ij} \neq 0$ and $g_{ij} = 0$ in the absence of such a link. We define a worker's *in-degree* as $d_i = \mathbf{G1}$, which counts those workers that influence i.

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 uncertainty. The last term captures how adding risk into workers' compensation (through α_i) decreases individual welfare. Indeed, contractual arrangements with larger contingent payments, more risk averse agents, or highly volatile production processes will deliver lower utility to workers.

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.[16]

The firm will select a contract for each worker in order to maximize expected profits. Contracts (represented by vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$) must be individually rational and incentive compatible. Formally we have,

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

subject to

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

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

The solution to this problem characterizes optimal linear contracts as a function of the existing peer network ($\boldsymbol{\alpha}(\mathbf{G})$, $\boldsymbol{\beta}(\mathbf{G})$).

[16] 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.

First-Best Contract

As is typical in principal-agent models, we begin by presenting the optimal contract under symmetric information, to establish a starting point for comparison. In this scenario, 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

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

Lemma 1 (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.*

2.2 Optimal Wage Contracts with Moral Hazard

To solve the model, consider first the optimal effort decision of the worker for any contract (α_i, β_i). Recall that e_i maximizes worker i's certain equivalent consumption as defined in (2). Workers play a non-cooperative game similar to that in Ballester et al. (2006). The best-reply function of worker i is thus given by

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

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

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

We now make an assumption about the strength of strategic spillovers. Recall that the spectral radius of a matrix is the maximum of its eigenvalues' absolute values.

Assumption 1. *The spectral radius 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 can therefore rewrite the firm's problem as:

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

From the (IR) constraint we can obtain an expression for the fixed payments as a function of equilibrium actions $\mathbf{e}^\star$ and incentives payments $\boldsymbol{\alpha}$. To simplify notation, we normalize outside options to zero for everyone, i.e. $U_i = 0$ for all $i \in N$.[17] Therefore, we have that,

$$\beta_i(\boldsymbol{\alpha}, \mathbf{e}) = -\alpha_i \sum_k e_k + \frac{1}{2}e_i^2 - \lambda e_i \sum_{j \in N} g_{ji} e_k + \alpha_i^2 \frac{r\sigma^2}{2}. \tag{5}$$

We can use this expression to rewrite profits only as a function of $\boldsymbol{\alpha}$ as,

$$\begin{aligned} \mathbb{E}[\pi(\mathbf{e} \mid \boldsymbol{\alpha}, \boldsymbol{\beta})] &= \sum_i e_i - \sum_i w_i \\ &= \left(1 - \sum_i \alpha_i\right) \sum_i e_i - \sum_i \beta_i \\ &= \sum_i e_i - \frac{1}{2}\sum_i e_i^2 + \frac{r\sigma^2}{2}\sum_i \alpha_i^2 + \lambda \sum_{ji} g_{ji} e_i e_j. \end{aligned}$$

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.

[17] In Supplementary Appendix B we solve a more general model with heterogeneous outside options, risk aversion and productivity parameters.

Assumption 2. *The spectral radius of* $\lambda^2/(1+r\sigma^2)(\mathbf{GC})'\mathbf{GC}$ *is less than 1.*[18]

We are now ready to characterize optimal contracts.

Proposition 1 (Optimal Contracts)**.** *Under Assumptions 1 and 2, there exists a unique profit-maximizing linear wage function* $w_i(\mathbf{G}) = \beta_i + \alpha_i X$*, for each worker* $i \in N$*, with*

$$\boldsymbol{\alpha}^\star = \frac{1}{1+r\sigma^2}\left[\mathbf{I} - \frac{\lambda^2}{1+r\sigma^2}(\mathbf{GC})'\mathbf{GC}\right]^{-1}\mathbf{C}'\mathbf{1} \tag{6}$$

and

$$\boldsymbol{\beta}^\star = \frac{1}{2}\left[\mathbf{C}\boldsymbol{\alpha}^\star \circ (\mathbf{I} - 2\lambda\mathbf{G})\,\mathbf{C}\boldsymbol{\alpha}^\star + \boldsymbol{\alpha}^\star \circ \left(r\sigma^2\mathbf{I} - 2\mathbf{1}\mathbf{1}'\mathbf{C}\right)\boldsymbol{\alpha}^\star\right] \tag{7}$$

with $\mathbf{C} := (\mathbf{I} - \lambda\mathbf{G})^{-1}$ *and where* $\circ$ *denotes the Hadamard (element-wise) product.*

The intuition behind Proposition 1 is that firms optimally concentrate high-powered incentives on workers who are "closest" to the rest of the workforce, as determined by aggregating all directed (outward) paths on the peer network. Technically speaking, this measure is obtained by a weighted average of the canonical Katz-Bonacich measure of centrality, $(\mathbf{C}'\mathbf{1})$.[19] In fact, the (i,j) element of $\mathbf{W} := \left[\mathbf{I} - \frac{\lambda^2}{1+r\sigma^2}(\mathbf{GC})'\mathbf{GC}\right]^{-1}$, which we call w_{ij}, determines how much j's centrality, $b_j(\lambda) := (\mathbf{C}'\mathbf{1})_j$, matters for i's incentive provision, α_i. More concretely, we can write

$$\alpha_i^\star = \frac{1}{1+r\sigma^2}\sum_{j\in N} w_{ij}(\lambda,\sigma^2)\, b_j(\lambda), \quad \forall i \in N. \tag{8}$$

where we drop the explicit dependence on $\mathbf{G}$ to ease notation. The weight w_{ij} captures how much *common influence* worker i shares with j: $w_{ij} > 0$ if i and j jointly influence a third worker, even if they do not influence each other. To see this, notice that $(\mathbf{GC})_{ij}$ captures the same information as $\mathbf{C}_{ij}$, except it ignores the walks of length zero. Therefore, the product $(\mathbf{GC})'\mathbf{GC}$ defines a symmetric matrix where the (i,j) element equals $\sum_{r=1}^n (\mathbf{GC})_{ri}(\mathbf{GC})_{rj}$ and thus only counts workers that are indirectly influenced by both i and j. Finally, notice that under Assumption 2 we can express the weight matrix $\mathbf{W}$ as a geometric series of these common-

[18] If $\mathbf{G}$ is symmetric then assumption 2 can be written as:

$$\lambda\mu_1(\mathbf{G}) \leq \frac{1 + r\sigma^2 - \sqrt{1+r\sigma^2}}{r\sigma^2} \leq 1,$$

where $\mu_1(\mathbf{G})$ is the largest eigenvalue of $\mathbf{G}$. That is, the strength of peer effects is bounded above. Obviously, in this case, Assumption 1 is implied by Assumption 2.

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

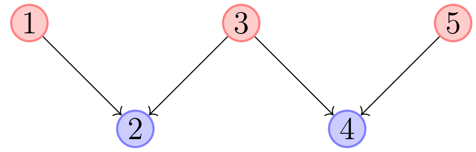


Figure 2: A directed network with 5 agents.

influence matrices,

$$\mathbf{W} = \sum_{k=0}^{\infty} \left(\frac{\lambda^2}{1 + r\sigma^2} \right)^k ((\mathbf{GC})'\mathbf{GC})^k . \tag{9}$$

The kth power of the matrix $((\mathbf{GC})'\mathbf{GC})^k$ keeps track of the (discounted) common influence of two workers through a series of k other workers with whom they have common influence. Therefore, w_{ij} captures the total weight of the common influence of agents i and j.

Example 1 (Common Influences)**.** *Consider the* 5*-worker network with* 4 *directed links in Figure* 2*. Worker* 2 *is influenced by workers* 1 *and* 3*, and worker* 4 *is influenced by workers* 3 *and* 5*. Entries $w_{13} = w_{31}$ of* (9) *are positive due to workers* 1 *and* 3*'s common influence on worker* 2*, as reflected by the product $(\mathbf{GC})'\mathbf{GC}$ (similarly for $w_{3,5} = w_{5,3} > 0$). More surprisingly, entries $w_{1,5} = w_{5,1}$ are also positive since workers* 1 *and* 5 *exert common influence on workers* 2 *and* 4 *via their (direct) common influence with worker* 3*. This is reflected on the $(1, 5)$ and $(5, 1)$ elements of the matrix $((\mathbf{GC})'\mathbf{GC})^2$, which can be expressed as $(\mathbf{GC})_{2,1}(\mathbf{GC})_{2,3}(\mathbf{GC})_{4,3}(\mathbf{GC})_{4,5}$. Finally, notice that workers* 2 *and* 4 *have no influence on others and receive incentives according to the textbook case with no peer effects described in Corollary* 1 *below.*

Proposition 1 allows us to easily analyze optimal contracts for important benchmarks. For instance, consider what happens in the absence of peer effects (i.e. when $\lambda = 0$). In this case, it is obvious that all agents are minimally-central (i.e. $b_i(0) = 1$ for all $i \in N$) and $\mathbf{C} = \mathbf{W} = \mathbf{I}$. Applying this to Proposition 1 we show that our model specifies to the textbook benchmark as a special case (Bolton and Dewatripont, 2004).

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

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

Since there are no spillovers, the firm finds it optimal to treat each worker separately.[20] Notice that, as is well-known, in the absence of risk (i.e. $\sigma^2 = 0$) the principal finds it optimal to "sell"

[20] Although the contract is assumed to be a linear function of aggregate output, $X = \sum_i e_i + \varepsilon$, workers' problems are separable in the absence of spillovers (i.e. when $\lambda = 0$) because CARA utility has no wealth effects, so at the margin workers' incentives are as in the individual problem. That's why the solution for $\lambda = 0$ corresponds to the classical principal-agent solution.

the firm to the worker (i.e. $\alpha_i = 1$) and that the presence of risk decreases incentives uniformly for all $i \in N$.

We are most interested in understanding how the presence of fundamental risk ($\sigma^2 > 0$) twists the distribution of incentives away from the benchmark model. To see this consider first the case with no risk (i.e. $\sigma^2 = 0$). From Proposition 1 we have that, in this case,

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

With a bit of algebra, one arrives at the following corollary.

Corollary 2 (No Risk)**.** *In the absence of fundamental risk (i.e. $\sigma^2 = 0$), workers' incentives correspond to an affine transformation of Bonacich-centrality*

$$\alpha_i^\star = \frac{1}{2}\left[1 + b_i(2\lambda)\right], \quad \forall i \in N,$$

where $b_i(2\lambda)$ is worker i's Bonacich centrality with parameter 2λ.[21]

Next, consider the case when $\sigma^2 > 0$. The presence of fundamental risk modifies how incentives must be distributed in two important ways. First, $\boldsymbol{\alpha}^\star$ in equation (6) scales with $1/(1 + r\sigma^2)$. Second, the term in brackets converges to $\mathbf{I}$ as $r\sigma^2$ grows. This implies that, although firms decrease incentive provisions when risk is high, the way in which incentives are distributed becomes proportional to standard Bonacich centrality ($\mathbf{C}'\mathbf{1}$).

Corollary 3 (Increasing Fundamental Risk)**.** *Performance-based compensation decreases monotonically with σ^2. However, it does not decrease uniformly for all workers. In the limit, incentives are proportional to Bonacich centrality. Formally, $\boldsymbol{\alpha}^\star \underset{\sigma^2 \to \infty}{\longrightarrow} q\, \mathbf{C}'\mathbf{1}$ for q vanishingly small.*

Therefore, allocating optimal incentives is *simple* if fundamental risk is large. One may wonder if there is also a simple connection between incentives, $\alpha_i^\star$, and network centrality, b_i, for any value of σ^2. Proposition 1 connects incentives to a weighted average of everyone's centrality, so it does not by itself imply that more central workers necessarily receive larger incentives. Indeed, a very central worker could in principle share large common-influences with a less-central worker. By expression (8), she could end up with a lower value of α_i than someone with less centrality but greater common influence. Our next result proves that this is never the case.

Proposition 2 (Monotonicity)**.** *Optimal incentives are a monotonic transformation of Katz-Bonacich Centrality. Formally, $b_i > b_j \iff \alpha_i^\star > \alpha_j^\star$.*

[21] Helsley and Zenou (2014) find a similar 2λ term in their welfare analysis of network interaction in cities, albeit for different reasons. In their case, the term arises from an externality in location choice.

*Figure 3: **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).*

Optimal Incentives with Negative Spillovers

To ease exposition, we have thus far discussed our results from the perspective of peer-to-peer complementarities (i.e. $\lambda > 0$). However, there are many instances where it is more natural to assume that co-workers crowd out each others' individual contribution (i.e. $\lambda < 0$).[22] Fortunately, Proposition 1 and Proposition 2 are general and characterize optimal incentive allocations for both cases. Having said this, it is not obvious what they tell us practically about how contracts differ if workers experience negative spillovers from their co-workers. In other words, since our definition of centrality $\mathbf{b} := (\mathbf{I} - \lambda \mathbf{G})^{-1}\mathbf{1}$ depends on λ, the identity of workers receiving larger incentives varies, and can look very different when λ changes signs. For instance, consider the simple network in Figure 3. There are three different types of nodes in this network and their relative size corresponds to centrality. Notice that the nodes that are most central under strategic complements are the least central under strategic substitutes. Proposition 2 therefore implies that the identity of workers receiving most incentives flips.

In recent work by Galeotti et al. (2020) on optimal interventions in networks, the authors show that when $\lambda < 0$ optimal interventions should resemble the alternating pattern reflected in the last principal component of $\mathbf{G}$. In other words, the planner likes to move the incentives of adjacent individuals in opposite directions in order to reinforce each intervention at the margin. Proposition 2 suggest that our model doesn't reflect this alternating pattern. The difference stems from the principal's cost structure in each model. In Galeotti et al. (2020) intervention costs are convex. The authors assume that any change in standalone incentives entails a quadratic cost to the principal. Formally, the cost associated to implementing vector $\boldsymbol{\alpha}$, starting from a vector $\hat{\boldsymbol{\alpha}}$, can be written as,

$$\sum_{i \in N} (\alpha_i - \hat{\alpha}_i)^2 .$$

[22]Negative spillovers in the workplace may arise from opportunities to free ride, which is documented for instance by Amodio and Martinez-Carrasco (2018) in an egg production plant in Peru. Bandiera et al. (2005) find evidence of negative spillovers in UK fruit pickers under a piece rate system. This result implies that the wage contract itself can influence the sign of the externality, an intriguing possibility that we leave for future work.

In our model, the principal faces a very different cost structure. Starting from some status-quo level of incentives $\hat{\boldsymbol{\alpha}}$, the cost to modifying worker i's incentives enters only through the increased riskiness of i's wage. Looking at equation (5) we see that the cost to the principal can be written as

$$\frac{r\sigma^2}{2}\sum_{i\in N}(\alpha_i^2 - \hat{\alpha}_i^2).$$

While it may look like a small difference, it is of great significance because lowering incentives (i.e. $\alpha_i < \hat{\alpha}_i$) can actually lower costs in our model, by decreasing the risk transferred to i. This implies that, relative to Galeotti et al. (2020), it is not profitable to amplify other interventions by exploiting the alternating sign of strategic interactions via neighboring paths.

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. $\partial\alpha_k/\partial g_{ij} \geq 0$ for all $i, j, k \in N$). Moreover, an increase in g_{ij}* strictly *increases the incentive pay of worker j and any worker k who has a common influence with j.*

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

Example 1 Continued. *Suppose worker 2's influence over worker 1 increases (i.e., g_{21} increases). Proposition 3 says that optimal performance-based compensations increase for workers 1, 3, and 5. Intuitively, an increase in g_{21} implies that worker 2's effort increases. By proposition 1, the firm provides higher incentives to worker 1 due to the increase of influence on 2. Furthermore, peer effects from worker 3 to worker 2 are now amplified due to the larger $e_2^\star$. This pushes the principal to increase $\alpha_3^\star$, which increases $e_3^\star$, to capitalize on the strengthened effects of incentive spillovers. This effect arises due to the increase of workers 1 and 3's common influence on 2. Similarly, the increase of 1's influence on 2, increases the common influence workers 1 and 5 have on 2, and the firm also provides higher incentives to worker 5.*

Fixed Payment β

The second part of Proposition 1 characterizes the fixed payment portion of the contract $\boldsymbol{\beta}^{\star}$. This unconditional payment is designed to ensure that all workers attain their reservation utility in expectation. In other words, the firm can fully wage-discriminate and extract all surplus from workers. Since workers generate positive surplus, the firm guarantees that everyone works in equilibrium. The precise expression of β_i is not very informative, however, since it depends on what we specify for workers' outside options.[23] Moreover, it responds to the network structure only through the performance-based compensation term $\boldsymbol{\alpha}$. For this reason we have focused most of our analysis so far on $\boldsymbol{\alpha}$. Having said this, our results on coarse contracts in Section 4 do rely more meaningfully on $\boldsymbol{\beta}$, and in particular on which workers the firm decides to hire when it can no longer extract all surplus.

Unknown Network Structure

So far we have assumed that the principal 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 4 where we force contracts to treat entire sections of the workforce equally. The second way to relax this assumption is to assume that the firm may simply not know the relevant peer-to-peer network.

The principal may have a model in mind as to how workers interact but might not know the realized structure of interactions. We incorporate this possibility by considering that the optimal contract now maps from a parameterized model of linking probabilities to a function of aggregate output. There are many generative models of random graphs. 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 4. *Consider a special case of the IRN model with two types in equally-sized groups. Let p represent the within-type linking probability while q represents the across-type linking prob-*

[23] In the main specification we are fixing everyone's outside option to 0, but in the appendix we solve a more general version of the model with full heterogeneity in which outside options are entirely general.

ability. Then the optimal allocation of incentives is given by

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

Notice that incentive allocations 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 4 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.[24] Notice that by comparing across different random network models we are able to vary the principal'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 of the principal may affect the optimal contract, and how this may also depend on the information workers have on the network.[25]

2.3 Organization Design: Linking Profits to Network Structure

In this section we turn to an organizational design question and ask how profits depend on the existing network structure assuming that workers receive the optimal contract from Proposition 1. First of all, notice that, for $\lambda > 0$, additional links amplify incentives and increase firm profits: the complete graph is therefore optimal.[26] If adding links is costly, it is well known from work by Belhaj et al. (2016) and Hiller (2017) that the (constrained) efficient network in games with strategic complements belongs to the class of *nested-split graphs*.[27] Unfortunately, no results currently exist that select from among this large class of networks and, even more importantly, we don't have a clear description of how network structure maps continuously to aggregate outcomes. This is important because if a firm is contemplating changes in organizational structure, a nested-split graph structure may not be feasible to implement for various reasons. But the firm might still want to know which class of networks are profit-equivalent and which networks dominate others. To tackle this we provide a direct map that links profits to the structural (i.e. spectral) properties of the network for any value of λ.

[24] 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.

[25] 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.

[26] With $\lambda < 0$ the efficient graph is the empty network.

[27] Nested split graphs is the set of all graphs with the following property: the set of neighbors of each agent is contained in the set of neighbors of each higher degree agent. The nested split graph result does not extend to cases of strategic substitute where $\lambda < 0$.

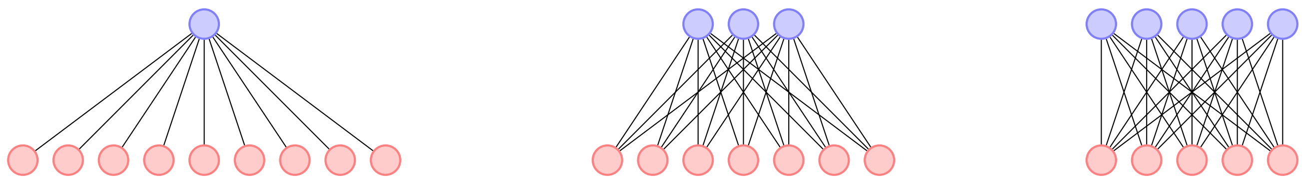


Figure 4: Complete Bipartite graphs with $N = 10$. An asymmetric bipartite graph (left panel) generates lower profits than a symmetric one (right panel).

Proposition 5 (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*:*

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

Moreover, let $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ *represent the* N *unit-eigenvectors of* $\mathbf{G}$ *associated to eigenvalues* $\mu_1 \geq \mu_2 \geq \ldots \geq \mu_n$*. Expected profits are given by:*

$$\mathbb{E}(\pi^{\star}) = \frac{1}{2}\sum_{\ell=1}^{n} \frac{(\mathbf{u}_\ell'\mathbf{1})^2}{(1+r\sigma^2)(1-\lambda\mu_\ell)^2 - (\lambda\mu_\ell)^2}. \tag{10}$$

The first part of Proposition 5 is useful because it extends a well-known result from the model of team production with multiple agents to our setting with bilateral spillovers across workers. If firms are optimizing then profits should scale one-to-one with output, no matter the organizational structure of the firm. Of course, the network structure will matter for what these profits actually look like. In fact, the second part of the proposition tells us how network structure affects profits in equilibrium by decomposing the network effects into its principal components.

Equation (10) is also useful because it can be leveraged to compare organizational structures based on their expected performance. For example, consider a firm that is debating the best way to delegate responsibilities. Assume that the firm must decide on the relative size of two divisions whose members interact, and assume for now that all relevant spillovers occur across divisions. Technically speaking, the organization must choose between all complete bipartite graphs of size N whose members are split into two divisions of size n and m, as shown in Figure 4. By leveraging spectral properties of these class of graphs,[28] we can use equation (10) in order

[28] It is well-known that the eigenvalues associated to complete bipartite graphs are $\mu_1 = \sqrt{mn} > \mu_2 = \mu_3 = \ldots = \mu_{n-1} = 0 > -\sqrt{mn} = \lambda_n$. We show in Supplementary Appendix D.1 that the unit-eigenvectors associated to the first and last eigenvalues are given by:

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

which means that $(\mathbf{u}_1'\mathbf{1})^2 + (\mathbf{u}_n'\mathbf{1})^2 = N$. Since $\sum_\ell (\mathbf{u}_\ell'\mathbf{1})^2$ must always equal N for any diagonalizable matrix,

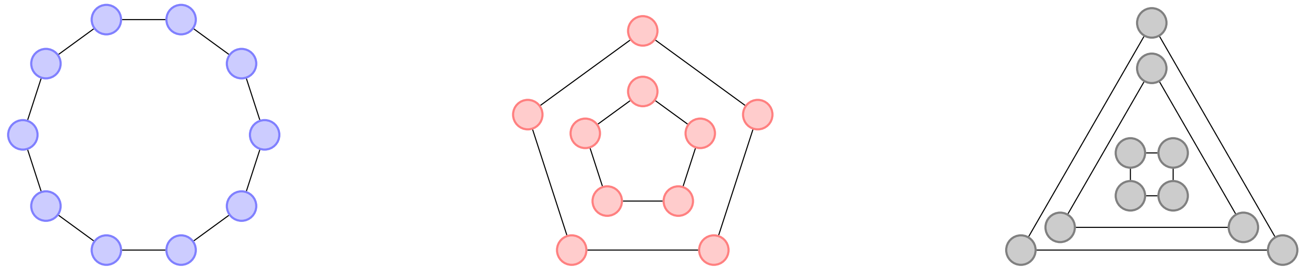


Figure 5: All 2-regular graphs with $N = 10$ give the same profits.

to write expected profits directly in terms of n and m only:

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

From this simpler expression, we can easily characterize the profit-maximizing structure among all complete bipartite graphs.

Corollary 4. *Among all complete bipartite graphs with n nodes in group A and m nodes in group B, expected profits are maximized when $n = m$.*

Imagine another application in which a fairly homogeneous organization – i.e. one in which everyone is (on average) influenced by the same number of peers – considers splitting the workforce into different divisions. Technically speaking, the CEO might want to know if her N workers should work in a single d-regular component, or should be split into separate smaller (d-regular) divisions, as shown in Figure 5. We can again leverage Proposition 5 to solve this design problem. Given well-known spectral properties for this family of graphs,[29] we can rewrite profits as follows:

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

It turns out that expected profits are only a function of the local structure of spillovers, and not on the component structure. In other words, splitting the organization into separate divisions is profit-neutral.

Corollary 5. *All $d-$regular graphs of size N generate the same expected profits.*

we know that only the first and last terms in the sum in Proposition 5 are relevant for profits.

[29] It is a well-known fact that the eigenvalues of a d-regular graph with k components of sizes $C_1, C_2, \ldots, C_k$ are $\mu_1 = \mu_2 = \ldots = \mu_k = d$. The corresponding unit-eigenvectors satisfy $\mathbf{u}_i'\mathbf{1} = \frac{C_i}{\sqrt{C_i}}$ for $i = 1, \ldots, k$, and since each $\mathbf{u}_i$ with $i \geq k+1$ is orthogonal to $\mathbf{1}$, we also know that $\mathbf{u}_i'\mathbf{1} = 0$ for $i \geq k+1$.

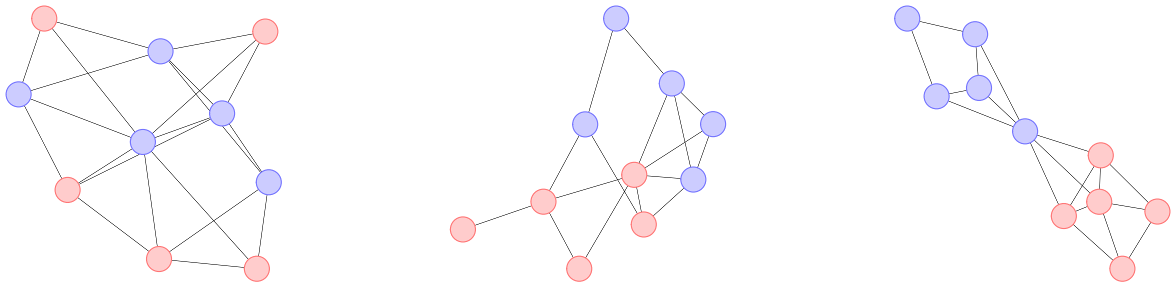


Figure 6: Planted Partition model with $n = 10$ and $p + q = 0.8$.
Panel A: $p = q = 0.4$. *Panel B: $p = 0.6, q = 0.2$.* *Panel C: $p = 0.75, q = 0.05$.*

Finally, consider how hompohily – the tendency of members of specific groups to connect disproportionately within that group – might affect firms' profits. To do this we return to the generative *planted partition random graph model* that we used for Proposition 4 where p represents the probability of connecting within a group and q the probability of connecting across groups. Notice that $p + q$ determines the average connectivity but not the level of homophily, which is determined by p/q. Decomposing profits using Proposition 5 and the well-known properties of the expected adjacency matrix $\bar{\mathbf{G}} \in \{p, q\}^{n \times n}$, we arrive at the following expression:[30]

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

It turns out that homophily (p/q) does not affect profits, controlling for average degree.

Corollary 6. *In a planted partition model with matching probabilities p and q, expected profits are only a function of $p + q$ and not of p/q.*

Comparative Statics: Investing in Workers

Firms typically invest in their workforce through various training programs that improve teamwork and/or individual skills. Should a firm train workers in order to improve their level of human capital, or should the firm instead invest on team-building exercises in order to strengthen peer complementarities? The profit decomposition result in Proposition 5 can be leveraged to answer human-resources related questions like this one in large and complex organizations. Consider an augmented version of the baseline model with the marginal cost to effort captured

[30]Notice that the only non-zero eigenvalues of $\bar{\mathbf{G}}$ are $\mu_1 = n\frac{p+q}{2}$ and $\mu_2 = n\frac{p-q}{2}$. Additionally, $\mathbf{u}_1'\mathbf{1} = n/\sqrt{n}$ and $\mathbf{u}_i'\mathbf{1} = 0$ for $i \geq 2$.

by a human-capital parameter $v \geq 1$. We can rewrite (1) as

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

Imagine that a firm can invest one dollar in order to either *decrease* v or *increase* λ.[31] Should a firm invest in lowering effort costs or increasing peer effects? That is,

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

We can invoke Proposition 5 to compute these marginal effects. Comparing the resulting expressions we characterize when it is superior to invest in team strength than human capital.

Proposition 6. *Investing uniformly in team strength (i.e. increasing λ) is superior (inferior) to investing uniformly in human capital (i.e. decreasing v) in those firms where peer networks satisfy the following condition:*

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

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.[32] 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.

Comparing across graphs, notice that human capital is a better investment if connections are scarce. The expression in Proposition 6 for instance tells us that investing in human capital is better if the network is empty (an obvious result). As we accumulate connections, however, this tendency will revert. One may wonder if, among the set of all graphs of size N, an equal number of them favor investments in human capital over investment in peer strength. The expression in Proposition 6 sheds light on this question and suggests that investing in peer strength is generally a better idea. For example, consider the class of d-regular networks. There exists a threshold $\hat{d}$ such that for all $d-$regular graphs with $d > \hat{d}$, investing in team strength is optimal. Looking at Proposition 6, however, we see that $\hat{d} \in (0, 1)$. In other words, the only regular peer structure for which investing in human capital dominates is the empty network.

A more convenient and structured way to increase network density is to increase the linking probability p of the Erdős-Rényi random graph model. As p grows from 0 (empty network)

[31] To keep notation at a minimum, we assume that both interventions have the same per-unit cost. Nothing substantial would change if these two investments in workers varied in cost.

[32] To see this recall that Assumption 1 requires $v > \lambda\mu_\ell$.

to 1 (complete network) we can identify the threshold where investing in human capital ceases to be a good idea. It turns out that the connectivity threshold coincides with the well-known threshold for a giant component. This 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.

Corollary 7. *In the Erdős-Rényi Random Graph model, investing in team strength outperforms investing in human capital if and only if* $np \geq 1$*.*

3 Modular Production

So far, we have assumed that expected output is a linear function of workers' efforts. Therefore, while workers may inhabit very different parts of the peer network, we assumed that their formal role within the organization is substitutable. This allowed us to conveniently isolate the impact of peer effects on wages and profits, but it fails to capture that a firm's production function may actually drive its organization.

Today many products are made by assembling separately-produced components, all of which are essential constituents of the final good (Garud et al., 2009; Baldwin and Clark, 2003). Partitioning firm production into separable elements can generate superior, more versatile products, as IBM uncovered when they developed the first modular computer in 1964. But failures in any one module can also affect overall production. In an extreme example popularized by Kremer (1993), the *Challenger* spacecraft exploded in 1986 because one of its many components, the O-ring, malfunctioned. In this section we consider how incentive contracts with peer complementarities are designed if firms have fragmented organizational structures.

To do this we modify our production function by incorporating *modules* and assume that final output is determined by the weakest-performing module. Formally, assume that N workers are distributed into K teams, called $k_1, k_2, \ldots, k_K$, each of which is put in charge of designing a separate module.[33] Within a team, performance is substitutable but across teams it is perfectly

[33] We will abuse notation and use K as the number of teams and the set K containing the teams $k \in K$

complementary.[34] We now write firm output as,

$$X(\mathbf{e}) = \min\left\{\sum_{i\in k_1} e_i, \sum_{i\in k_2} e_i, \ldots, \sum_{i\in k_K} e_i\right\} + \varepsilon. \quad (11)$$

We can now ask ourselves how incentive contracts should be designed with non-linear modular technologies, and how contracts depend on the network of spillovers within and across modules. Before moving on, notice that our original production function is a specific instance of (11) for the case in which all workers are contained in one single module.

In order to find the optimal contract in this setting we need to re-consider workers' equilibrium effort provision. First of all, notice that in any equilibrium, all modules must contribute the same total effort because a worker whose module contributes more than others will only gain by reducing her effort. Therefore $\sum_{i\in k} e_i = \bar{e}$ for all modules $k \in K$, and some $\bar{e} \geq 0$. The question now is which values of $\bar{e}$ constitute a Nash equilibrium? For any contract proposed by the firm, multiple equilibria may now exist. For example, $e_i = 0, \forall i \in N$ is always an equilibrium.[35] To find other equilibria, imagine that every module is contributing some arbitrary value $\bar{e}$ and consider if a worker could profitably deviate. Given technology (11), a worker can never gain by increasing effort. Decreasing effort, on the other hand, is profitable as long as the marginal cost to lowering e_i (given by α_i) is less than the marginal benefit: $e_i - \lambda\sum_{j\in N} g_{ij}e_j$. This implies that, in any equilibrium, the following no-deviation condition must also hold:

$$\alpha_i \geq e_i - \lambda\sum_{j\in N} g_{ij}e_j, \quad \forall i \in N. \quad (12)$$

Although multiple values of $\bar{e}$ may constitute an equilibrium, we focus on the *maximal equilibrium*: given a choice of $\boldsymbol{\alpha}$, workers are expected to play equilibrium value $\hat{e}(\boldsymbol{\alpha})$, such that any $\bar{e} > \hat{e}(\boldsymbol{\alpha})$ is not a Nash equilibrium.[36] Now, the firm will never offer a contract $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ with an associated maximal equilibrium, $\hat{e}(\boldsymbol{\alpha})$, that does not equate (12).[37] We can therefore conclude

[34] This specific functional form is assumed for simplicity, but since we don't assume anything on the size and composition of the different modules, it is in fact quite flexible. In a recent paper, Matouschek et al. (2023) assume that modular production is described by a network of decisions with varying degrees of pairwise coordination requirements. Our approach has the advantage that we don't need to introduce an additional "modular production network" on top of the existing peer effects network when designing optimal contracts with non-linear production.

[35] To see this notice that unilaterally raising e_i can never benefit worker i, given everyone else's equilibrium strategy and the production function assumed in equation (11).

[36] Formally, we assume that firms have rational expectations about which equilibrium will occur. We could consider alternative models under ambiguity, where firms hold varied beliefs about possible equilibria. For example, a firm might assign a uniform probability to each equilibrium associated with a contract. However, this more complex model will not alter the qualitative nature of our results, so we leave it for future research.

[37] Intuitively, for any wage offer $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ with $\hat{e}$ that does not equate (12), there exists another wage offer $(\tilde{\boldsymbol{\alpha}}, \tilde{\boldsymbol{\beta}})$ associated to the same maximal equilibrium value $\hat{e}$ that does equate (12) and generates strictly higher profits

that the following two restrictions must always hold in any optimal wage contract with modular production:

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

To proceed, we write restriction 2 above in vector notation. 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. Restriction 2 can therefore be written as,

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

where $\mathbf{1}_{\boldsymbol{K}}$ denotes a vector of ones of length K. Before we move on, we provide for clarity a brief example of matrix $\mathbf{M}$ for a small firm with two modules.

Example 2. *Imagine a firm with* 5 *workers* $i = 1, \ldots, 5$*. Suppose workers* 1*,* 2*, and* 3 *belong to module* k_1 *while workers* 4 *and* 5 *belong to module* k_2*. Then,*

$$\mathbf{M} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}, \quad \hat{e}\mathbf{1}_K = \begin{bmatrix} \hat{e} \\ \hat{e} \end{bmatrix}, \quad \mathbf{Me} = \begin{bmatrix} \sum_{i \in k_1} e_i \\ \sum_{i \in k_2} e_i \end{bmatrix} = \begin{bmatrix} \hat{e} \\ \hat{e} \end{bmatrix}.$$

We can now use matrix $\mathbf{M}$ in order to write down the principal's constrained maximization problem as optimizing profits subject to restrictions 1 and 2 above. Formally, we have

$$\max_{\boldsymbol{\alpha}} \left(\hat{e} - \frac{1}{2}\mathbf{e}'(\mathbf{I} - 2\lambda\mathbf{G})\mathbf{e} - \frac{\sigma^2 r}{2}\boldsymbol{\alpha}'\boldsymbol{\alpha} \right)$$

subject to

$$\mathbf{e} = \mathbf{C}\boldsymbol{\alpha}$$

$$\mathbf{Me} = \hat{e}\mathbf{1}_K$$

The solution to this problem describes how incentive are allocated for any network of peer effects and for any general modular structure satisfying equation (11). This is the content of our next main result (we write $\mathbf{1}$ instead of $\mathbf{1}_K$ or $\mathbf{1}_N$ from now on to ease notation).

Proposition 7 (Modular Production)**.** *The optimal allocation of incentives under modular production with module assignment* $\mathbf{M}$ *is given by:*

$$\boldsymbol{\alpha}^\star = (\mathbf{I} - \lambda\mathbf{G})\boldsymbol{\Sigma}^{-1}\frac{\mathbf{M}'\mathbf{H}^{-1}\mathbf{1}}{\mathbf{1}'\mathbf{H}^{-1}\mathbf{1}}, \tag{13}$$

because $\tilde{\boldsymbol{\alpha}} \leq \boldsymbol{\alpha}$ elemeny-by-element.

where $\boldsymbol{\Sigma}$ *is the* $N \times N$ *matrix given by* $\boldsymbol{\Sigma} := (1+\sigma^2 r)(\mathbf{I} - \lambda(\mathbf{G}+\mathbf{G}')) + \sigma^2 r(\lambda\mathbf{G})'(\lambda\mathbf{G})$*,* $\mathbf{H}$ *is the* $K \times K$ *matrix given by* $\mathbf{H} := \mathbf{M}\boldsymbol{\Sigma}^{-1}\mathbf{M}'$*.*

Proposition 7 fully characterizes optimal incentive allocations under modular production, for any assignment of workers to modules, as captured by matrix $\mathbf{M}$. Notice that the modular structure of the firm is completely independent of the peer-effects network. In other words, equation (13) is general, no matter the number of modules, their size, or the structure of connections within and across modules.

Without peer effects (i.e. $\lambda = 0$) the allocations of incentives with modular production differs in very interesting ways from the standard textbook result summarized in Corollary 1. Notice that, with $\lambda = 0$, the matrix $\boldsymbol{\Sigma}$ reduces to $(1 + r\sigma^2)\mathbf{I}$ and the allocation of incentives to worker i depends on the relative size of her module as described below.

Corollary 8 (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:*

$$\alpha_i = \frac{1}{1+r\sigma^2} \frac{\prod_{k\in K\setminus k(i)} n_k}{\sum_{r=1}^{K} \prod_{k\in K\setminus r} n_k}, \quad \forall\, i \in N. \tag{14}$$

Therefore, if modules are of equal size, $\alpha_i = \frac{1}{1+r\sigma^2}\frac{1}{K}$*. If there is only one module, then* α_i *corresponds to Corollary 1, as expected.*

Next, we consider the case with peer effects. To develop some intuition, take as a special case the situation in which everyone forms part of a single module. This case corresponds to the original production function of Section 2 and is captured by $\mathbf{M} = (1, 1, \ldots, 1)_{1\times N}$. In this case, $\mathbf{H} = \mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}$ is a scalar, and equation (13) simplifies to:

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

which, as expected, corresponds to our original result in Proposition 1.[38] At the other extreme, imagine a "weakest-link" type of technology where every worker is *essential* (i.e. each worker belongs to a separate module). In that case, $\mathbf{M} = \mathbf{I}_{N\times N}$ and every worker will exert the same effort in equilibrium. Applying Proposition 7, we now have that $\mathbf{H} = \boldsymbol{\Sigma}^{-1}$ and therefore equation (13) reduces to

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

[38] To see this, notice that $(\mathbf{I} - \lambda\mathbf{G}) = \mathbf{C}^{-1}$. Thus, we have that $\boldsymbol{\alpha}^* = \mathbf{C}^{-1}[(\mathbf{I} - \lambda(\mathbf{G}+\mathbf{G}')) + r\sigma^2(\mathbf{C}')^{-1}\mathbf{C}]^{-1}$. Next, notice we can move the $\mathbf{C}^{-1}$ from outside to inside the brackets to get $[(\mathbf{I}-\lambda(\mathbf{G}+\mathbf{G}'))\mathbf{C}+r\sigma^2(\mathbf{C}')^{-1}\mathbf{C}]^{-1}\mathbf{C}$. Finally, multiplying on the right by $(\mathbf{C}')^{-1}\mathbf{C}'$ and moving the $(\mathbf{C}')^{-1}$ inside the brackets as before, we obtain $\boldsymbol{\alpha}^* = \left[\mathbf{C}'(\mathbf{I} - \lambda(\mathbf{G}+\mathbf{G}'))\mathbf{C} + \sigma^2 r\mathbf{I}\right]^{-1}\mathbf{C}'\mathbf{1} = 1/(1+r\sigma^2)[\mathbf{I} - \lambda^2/(1+r\sigma^2)(\mathbf{G}\mathbf{C})'\mathbf{G}\mathbf{C}]^{-1}\mathbf{C}'\mathbf{1}$.

We now 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 less incoming links.

Corollary 9. *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 $\xi = \sum_{j \in N} (1-2\lambda d_j) + \sigma^2 r \left(1-\lambda d_j\right)^2$ is common across all workers.

It turns out that since modules are complementary in production, incentives don't propagate well between them. When every worker is essential, all connections are between (not within) modules, so no amplification is possible. Intuitively, all workers exert the same effort in equilibrium so paying central workers a large α_i wont 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 lowering their effort, unless variable pay is set sufficiently high.[39] 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.[40]

If modules are larger, amplification is possible within modules. Firms must then strike an optimal balance between prioritizing more outgoing paths and less incoming links. It turns out that the incentive allocation rule for any modular structure, $\mathbf{M}$, is a convex combination of the allocation rule under one module ($\mathbf{M} = \mathbf{1}'_n$) in Proposition 1 and the allocation rule under N modules ($\mathbf{M} = \mathbf{I}$) in Corollary 9. In other words,

$$\alpha_i^{\mathbf{M}} = w_i \, \alpha_i^{\mathbf{1}'_n} + (1-w_i)\, \alpha_i^{\mathbf{I}} \quad \forall i \in N,$$

The weights depend on the link structure across workers, but our simulations show that, as the network grows, they converge to a simple expression that only depends on the relative size of modules:[41]

$$w_i \underset{N\to\infty}{\longrightarrow} \frac{N - n_{k(i)} \left(\sum_k \frac{1}{n_k}\right)}{(N-1)\, n_{k(i)} \left(\sum_k \frac{1}{n_k}\right)} \quad \forall i \in N.$$

[39] 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.

[40] 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.

[41] See Supplementary Appendix E for an explicit characterization of the convex weights and the numerical results showing convergence in various different simulations.

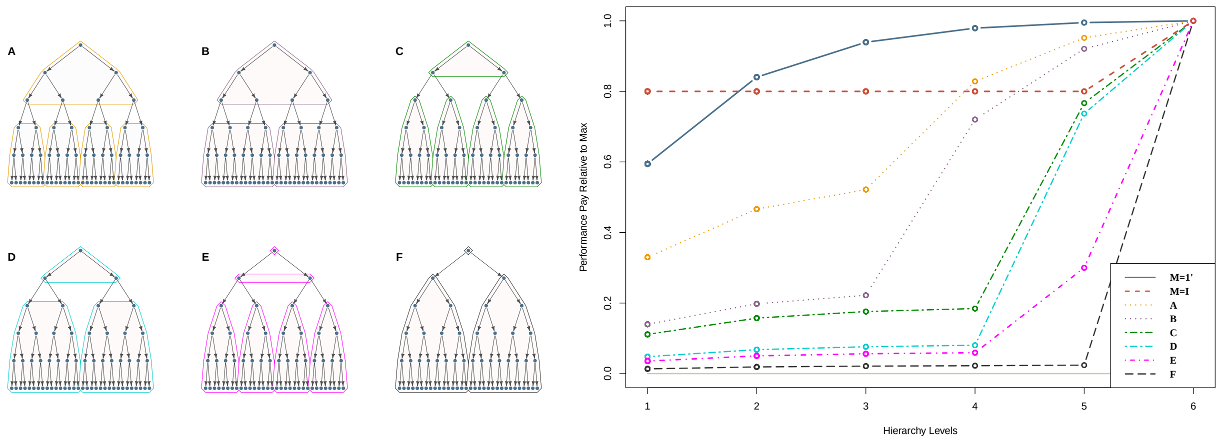


Figure 7: Performance-pay for Different Modular Configurations. Simulations are run for $\lambda = 0.2$, $\sigma^2 = 2$ *and* $r = 5$.

Before moving on, let's consider how modular structure impacts earnings profiles by revisiting the example from Figure 1. The right panel of Figure 7 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 9 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: 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.[42]

4 Wage Benchmarking

Recent hiring practices have compressed wages across firms and within occupational categories. Obloj and Zenger (2022) find a 20% decrease in the pay variance in a sample of over 100,000

[42] In Supplementary Appendix F we show that for more exotic modular configurations, earnings profiles can even be nonmonotonic.

US academics following the introduction of pay transparency tools between 1997 and 2017. Mas (2017) finds similar wage compression among city managers in California following a 2010 mandate requiring cities to disclose municipal salaries.[43] Cullen (2024) use national payroll data to study salary dispersion when firms gain access to a tool that reveals market pay benchmarks for each job title. They find that dispersion decreases 25%.[44] Finally, Hazell et al. (2022) show that 40-50% of a job's posted wages are identical across locations within a firm, even though local conditions vary substantially. How, then, should firms design wage contracts when fully discriminating across workers is ruled out?

In this section we return to the linear production function of Section 2 and instead 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. 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. This opens up the possibility that the firm may want to keep some workers from participating (i.e. keep them unemployed) because although overall output goes down, profits may actually improve.

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 $k \in K$:

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

We allow for any level of granularity in the contract because we make no restrictions on how to partition workers into categories.[45] 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.[46] 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, $\mathbf{e}^{\star} = \mathbf{C}\mathbf{T}'\hat{\boldsymbol{\alpha}}$. Before moving on, notice that if $K = N$ (i.e. if each worker is allowed to have a

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

[44] The authors find that salary benchmarks lead firms to set wages that are closer to the median salary benchmark by compressing salaries both from below and above the median. Most compression takes place among low-skilled positions (40% relative to 15%)

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

[46] Notice that matrix $\mathbf{T}$ is defined identically to the Module assignment matrix $\mathbf{M}$ in Section 3. We choose to use different notation across sections to emphasize that the K categories in this section have no relation to the K modules of Section 3. Analyzing Modular Production and Coarse Contracts simultaneously is beyond the scope of this paper, and we leave it for future research.

different job-title/contract) then $\mathbf{T} = \mathbf{I}$ and we get the original setup of Section 2.

When contracts are coarse ($K < N$) the Individual Rationality (IR) constraints look different. 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. The firm can always set β_k such that a specific worker $i \in k$ receives their reservation utility, but then the remaining workers in group k will either reject the contract or obtain positive rents from it.

In the first part of this section we begin by assuming 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.[47] Following equation (5), we can relate β_k to the effort costs of each worker $i \in k$:

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

The last term, μ_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:

$$\max_{\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\beta}}} \mathbb{E}[\pi(X, \boldsymbol{w}|\mathbf{e})] = \sum_i e_i - \sum_i w_i$$

subject to:

$$\text{CE}_i(\boldsymbol{\alpha}, \mathbf{G}, \mathbf{T}) - \mu_i = 0, \quad \forall i \in N \tag{IR}$$

$$\mu_i \geq 0, \quad \forall i \in N$$

$$\mathbf{e} = \mathbf{C}\mathbf{T}'\hat{\boldsymbol{\alpha}} \tag{IC}$$

The solution to this problem yields the optimal allocation of group-level incentives for any assignment of workers into categories, as described by matrix $\mathbf{T}$. In other words, for any level of granularity, optimal contracts are given by the following main result.

Proposition 8 (Wage Benchmarking)**.** *The optimal allocation of incentives under wage bench-*

[47] 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, $\bar{\psi}_k = \psi_{\bar{i}(k)}$

marking with group assignment $\mathbf{T}$ *is given by:*

$$\hat{\boldsymbol{\alpha}}^{\star} = \frac{1}{1+r\sigma^2}\left[\mathbf{T}\left(\mathbf{I} - \frac{\lambda^2}{1+r\sigma^2}(\mathbf{GC})'\mathbf{GC}\right)\mathbf{T}'\right]^{-1}\mathbf{TC}'\mathbf{1}. \tag{15}$$

The first thing to notice from Proposition 8 is that the shape of the optimal incentive allocation rule is very similar to our baseline result in Proposition 1. In fact, when $\mathbf{T} = \mathbf{I}$ equation (15) is identical to equation (6), as expected. Secondly, in the absence of peer effects, coarse contracts should also coincide with baseline, because everyone receives the same salary as shown in Corollary 1. We can confirm this here: if $\lambda = 0$, then $\hat{\boldsymbol{\alpha}}^{\star} = ((1+r\sigma^2)\mathbf{TT}')^{-1}\mathbf{T1}$, and $\alpha_i^* = 1/(1+r\sigma^2)$, as expected.[48]

Equation (15) differs from equation (6) in a very natural way. Recall the discussion after Proposition 1 on assigning incentives along an individual-specific weighted average of Katz-Bonacich centralities: $\boldsymbol{\alpha}^{\star} = \mathbf{WC}'\mathbf{1}$. Notice that incentives are now allocated to entire groups following the same logic, except that what matters now is the group-level (i.e. aggregate) Bonacich centrality measure, $\mathbf{TC}'\mathbf{1}$.[49] The weighting matrix is also a familiar transformation of the common-influences matrix introduced in equation (9). To see this, rewrite the weighting matrix in equation (15) as follows:[50]

$$\left[\mathbf{T}\left(\mathbf{I} - \frac{\lambda^2}{1+r\sigma^2}(\mathbf{GC})'\mathbf{GC}\right)\mathbf{T}'\right]^{-1} = (\mathbf{TT}')^{-1}\sum_{k=0}^{\infty}\left(\frac{\lambda^2}{1+r\sigma^2}\right)^k\left((\mathbf{TT}')^{-1}\left(\mathbf{GCT}'\right)'\mathbf{GCT}'\right)^k.$$

which looks almost like equation (9), except that we are normalizing by group size (via $\mathbf{TT}'$) and, more importantly, we are taking powers of $(\mathbf{GCT}')'\,\mathbf{GCT}'$ as opposed to powers of $(\mathbf{GC})'\mathbf{GC}$. But $\mathbf{CT}'$ simply aggregates $\mathbf{C}$ at the group level.[51] Therefore, incentives under coarse contracts are allocated following the same logic as before, but now aggregating centralities for each job category. Following this discussion, then, it is not surprising that we can extend the ordering of incentives in Proposition 2 to the case of coarse contracts as follows:

Proposition 9 (Monotonicity)**.** *Optimal incentives with wage benchmarking are a monotonic*

[48]Notice that $\mathbf{TT}'$ is a $K \times K$ diagonal matrix with the size of each group along the diagonal. Letting $|\mathbf{K}| = (|k_1|, |k_2|, \ldots, |k_K|)$ then $\mathbf{TT}' = \text{diag}(|\mathbf{K}|)$.

[49]Notice that $\mathbf{TC}'\mathbf{1}$ is a $K \times 1$ vector of group-level centralities, which simply sums the Bonacich centrality of all group members.

[50]To see this, use the well-known fact that $(\mathbf{A}+\mathbf{BD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}\left(\mathbf{I}+\mathbf{DA}^{-1}\mathbf{B}\right)^{-1}\mathbf{DA}^{-1}$. We can write the weighting matrix in the desired form by applying the following replacements: $\mathbf{A} = \mathbf{TT}'$, $\mathbf{B} = \frac{\lambda^2}{1+r\sigma^2}\left(\mathbf{GCT}'\right)'\mathbf{GCT}'$, and $\mathbf{D} = \mathbf{I}$. Substituting, we get $\left(\mathbf{I} + (\mathbf{TT}')^{-1}\mathbf{A}\sum_{k=0}^{\infty}\left((\mathbf{TT}')^{-1}\mathbf{A}\right)^k\right)(\mathbf{TT}')^{-1} = \left(\mathbf{I} + \sum_{k=1}^{\infty}\left((\mathbf{TT}')^{-1}\mathbf{A}\right)^k\right)(\mathbf{TT}')^{-1} = (\mathbf{TT}')^{-1}\sum_{k=0}^{\infty}\left((\mathbf{TT}')^{-1}\mathbf{A}\right)^k$, which gives the expression in the main text once we substitute in for $\mathbf{A}$.

[51]The (i,k) element of $\mathbf{CT}'$ measures the discounted sum of all paths from members of group k to worker i. Summing the elements of column k gives $\bar{b}_k := \sum_{i\in k} b_i$, the k-th group aggregate centrality.

transformation of the group's average Katz-Bonacich Centrality. Formally, for any two groups k and h assigned by $\mathbf{T}$*:* $\bar{b}_k/n_k \geq \bar{b}_h/n_h \iff \alpha^\star_k \geq \alpha^\star_h$.

4.1 Loss in Surplus due to Wage Benchmarking

At the beginning of the previous section we argued that, when contracts are coarse, the firm shares overall surplus with workers in the form of centrality rents. This is not the end of the story, however. Coarse contracts also 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 $S = \mathbb{E}(\pi) = \frac{1}{2}\sum_i e_i$, where the second equality comes from Proposition 5. With coarse contracts, surplus does not reduce to profits anymore, but the following lemma conveniently extends the first half of proposition 5 and allows us to conclude that, in general for any group assignment, $\mathbf{T}$, it is still the case that,

$$S = \frac{1}{2}\sum_i e_i. \tag{16}$$

Lemma 2 (Equilibrium Profits 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:*

$$\mathbb{E}(\pi) = \frac{1}{2}\sum_i e_i - \sum_i \mu_i$$

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

Using Lemma 2 and the fact that $\mathrm{CE}_i = \mu_i$ for all workers, we obtain equation (16). Define $\mathbf{e_T}$ 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 $\mathbf{T}$ and group assignment $\widetilde{\mathbf{T}}$ is given by:

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

Figure 8: ***Panel A:*** *Within-group variance is zero. No surplus loss.* ***Panel B:*** *Within-group variance is 0.53. Surplus loss, following Proposition 10, is about 0.85.*

In order to fix ideas, 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. We show that the efficiency loss can be measured very easily by computing within-group dispersion in centrality.

Proposition 10 (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. In such a scenario, offering the same contract to everyone in group k is optimal and coarse contracts generate no welfare loss. As a consequence, *regular networks* – where everyone is identical – generate zero surplus loss. More importantly, notice that Proposition 10 allows centralities to differ arbitrarily across groups. As long as within-group variability is zero, coarse contracts are efficient. In other words, it is the interaction between the peer structure $\mathbf{G}$ and group assignment $\mathbf{T}$ that actually matter. Figure 8, for instance, shows the same network and two different assignments $\mathbf{T}$, with very different consequences for $\Delta S_{\mathbf{I}-\mathbf{T}}$.

Proposition 10 provides a limit result for small λ and/or large σ^2. While this may seem restrictive, recall that the peer effects parameter λ is bounded above by the spectral radius of the network. In dense graphs λ must be very small. One may still wonder if the sum of within-group variances determines surplus loss more generally. In Figure 9 we show simulation results across many Erdős-Rényi random graphs. The figure shows that $\Delta S_{\mathbf{I}-\mathbf{T}}$ is always an increasing (albeit non deterministic) function of $\sum_k n_k Var(\mathbf{b}_k)$ and, as stated in Proposition 10, becomes an exact linear relationship as σ^2 grows. Therefore, although proportionality is only true in the limit, the efficiency cost associated to coarse contracts can still be summarized by this simple statistic.

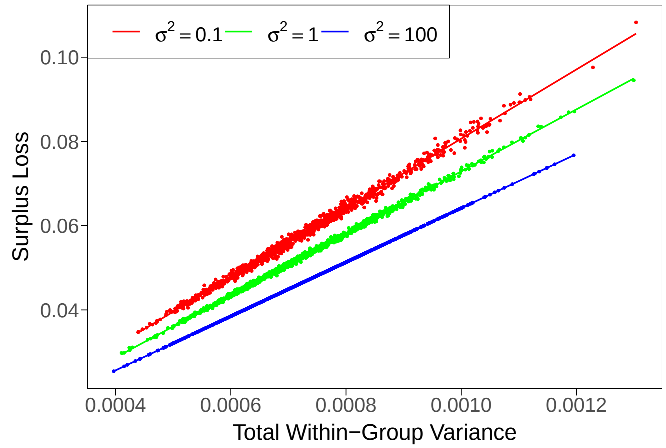


Figure 9: The loss of surplus $\Delta S_{\mathbf{I-T}}$ is increasing in within-group variance in centrality. The Relationship is linear and becomes deterministic as σ^2 increases.

4.2 Optimal Group Composition: A Theory of Unemployment

In this section we relax the assumption that the firm must write contracts that are accepted by all worker. In fact, the firm might sometimes obtain greater profits if "high-cost" workers in group k don't work. Recall that worker i extracts rents $\mu_i = \bar{\psi}_k - \psi_i$ that grow in the distance between her effort cost and that of the highest-cost member of her group, $\bar{i}(k)$. Shortening cost differences within each group will therefore decrease the rents the firm must pay to all remaining workers. Although the workforce (and hence total output) might be smaller, the firm might still gain by reducing remaining workers' rents sufficiently. The question then is which workers optimally remain employed and which don't, as a function of the existing peer network $\mathbf{G}$ and group assignment $\mathbf{T}$.[52]

The first thing to notice is that when a firm reduces workers' rents by decreasing β_k the first workers to drop out will be those with highest effort costs. The following lemma establishes that these are also the workers with lowest Bonacich centrality

Lemma 3. *Within each group, effort costs are decreasing in workers' Bonacich centrality. Formally, if $b_i > b_j$ then $\psi_i^\star < \psi_j^\star$, for $i, j \in k$.*

We can therefore describe the composition of group k simply by focusing on the distribution, F_k, of workers' centrality. By Lemma 3, for each choice of β_k there is a corresponding value $q(\beta_k)$ such that workers with $b_i < q(\beta_k)$ don't accept the offer. So far, we assumed that β_k was exactly large enough such that everyone worked. Let $\overline{\beta}_k$ represent this value: $q(\overline{\beta}_k) = b_{\bar{i}(k)}$.

[52]One could consider an alternative where a firm adjusts the composition of a group not by keeping workers out of the firm, but by re-assigning them to a different group (i.e. by adjusting $\mathbf{T}$). We don't currently consider this variant.

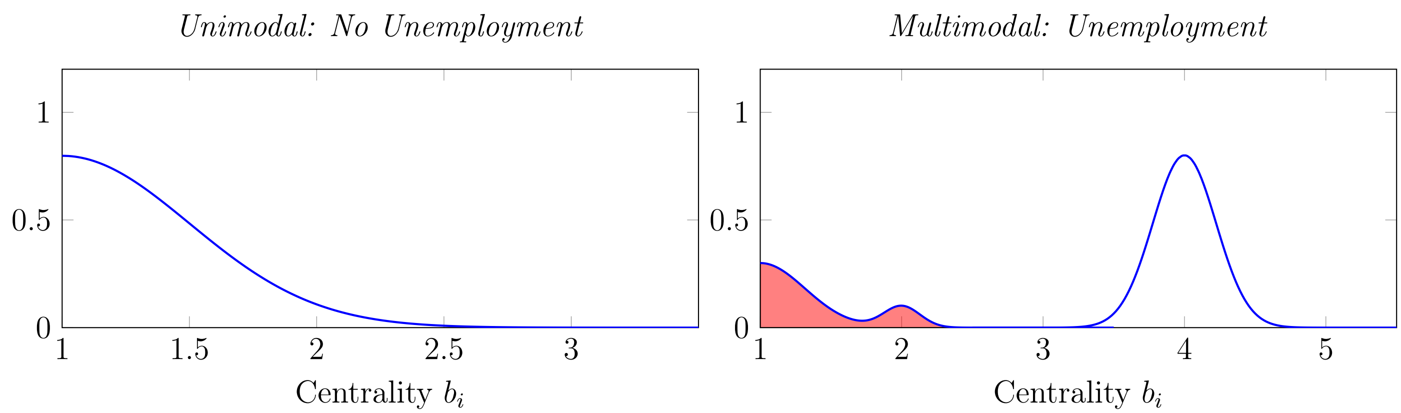


Figure 10: The shape of the distribution of centralities determines the optimal fixed-payment β_k and therefore the share of workers that remain unemployed.

The firm will now select a value of $\beta_k \leq \overline{\beta}_k$ which balances the costs (lost output) and benefits (lower centrality rents) from removing workers.

To illustrate how F_k affects the choice of β_k, imagine a group with a distribution of b_i's shown in the left panel of Figure 10. The distribution is tight and, as we will see, the cost to lowering β_k is large. On the other hand, the distribution in Panel B is spread out, with a large group of workers being much more central than others. In this case closing the centrality gap is profitable because the loss in output from losing the left-tail is more than compensated from the large centrality gap that is closed for the remaining workers. Intuitively, the firm can lower β considerably at the expense of losing only a few poorly-connected workers.

To formalize this intuition, consider a group k and assume there are no connections with other groups. This simplifies notation considerably because $e_i = (\mathbf{C}\boldsymbol{\alpha})_i$ reduces to $e_i = \alpha_k b_i$ for any worker $i \in k$.[53] Let F_k be the empirical CDF of Bonacich centrality in group k. Starting from $\overline{\beta}_k$, the firm can increase profits by choosing a lower β_k if and only if:

$$(\overline{\beta}_k - \beta_k)\left(1 - \int_{r=1}^{q(\beta_k)} dF_k(r)\right) > \alpha_k \int_{r=1}^{q(\beta_k)} r\, dF_k(r). \tag{17}$$

The right-hand-side of the inequality represents the loss in output from all workers with $b_i < q(\beta_k)$, who no longer work. The left-hand-side represents the additional rents $(\overline{\beta}_k - \beta_k)$ that the firm can now extract from all remaining workers. To compute threshold $q(\beta_k)$, we must find the centrality of the indifferent worker in group k – i.e. the worker with $\text{CE}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{G}, \mathbf{T}) = 0$.[54]

[53] The results don't change qualitatively if we allows for any arbitrary link structure across groups, but the expressions get a lot messier.

[54] There may be no worker in group k with this precise centrality but this is inconsequential because everyone who is more central will accept the contract and everyone who is less central will not accept it.

From equation (2) and the fact that $e_i = \alpha_k b_i$ we have:

$$\beta_k = z + \frac{1}{2}\alpha_k^2 b_i \left(b_i - 2\lambda \sum_{j \in N} g_{ij} b_j \right) = z + \frac{1}{2}\alpha_k^2 \left(1 - (1 - b_i)^2\right),$$

where $z = \frac{r\sigma^2}{2}\alpha_k^2 - \alpha_k \sum_i e_i$ is constant across all $i \in k$, and the second equality uses the definition of Bonacich centrality ($b_i := 1 + \sum_j g_{ij} b_j$). With this we can conclude that

$$q(\beta_k) = \sqrt{\frac{\alpha_k^2 - 2(\beta_k - z)}{\alpha_k^2}}. \tag{18}$$

Equations (17) and (18) allow us to determine exactly which fraction of workers remains unemployed in each group k, as a function of peer network $\mathbf{G}$, group assignment $\mathbf{T}$, and centrality distribution F_k.

Proposition 11 (Group Composition and Unemployment)**.** *As long as,*

$$\mathbb{E}\left[b | b < q(\beta_k)\right] \frac{\Pr\left(b < q(\beta_k)\right)}{\Pr(b > q(\beta_k))} < \frac{\overline{\beta}_k - \beta_k}{\alpha_k} \tag{19}$$

for some value of $\beta_k < \overline{\beta}_k$, then $q(\beta_k) > b_{\bar{i}(k)}$ and workers in group k with $b_i < q(\beta)$ remain unemployed. The equilibrium level of unemployment is determined by the value of $\beta_k \leq \overline{\beta}_k$ that equates (19).

This result is quite intuitive. Notice that $\mathbb{E}[b|b < q(\beta)])$ is the average centrality of unemployed workers, while $\Pr\left(b < q(\beta)\right) / \Pr(b > q(\beta))$ is the relative proportion of unemployed to employed workers given a $q(\beta)$. If a firm can increase $q(\beta)$ while keeping the left side of inequality (19) down, this means there are many workers above $q(\beta)$ for which the firm can extract more surplus, and relatively few workers below (and/or with low average centrality) that the firm can afford to lose. Consider an extreme example where everyone has the same centrality. From our previous discussion, coarse contracts don't cost the firm any rents, and no one should be unemployed. In this scenario, for any $\beta < \overline{\beta}$, the left hand side of inequality (19) diverges because $\Pr(b > q(\beta)) = 0$. Conversely, imagine a firm where a few workers have low centrality, and the rest has a much higher centrality. As long as $q(\beta)$ is less than the higher centrality, the left hand side of (19) remains small and thus can be equated for $\beta < \overline{\beta}$, leading to unemployment. The diagram in Figure 10 depicts this argument graphically for less extreme forms of the distribution.

While the previous discussion (and Figure 10) is entirely qualitative, we can nonetheless provide solid quantitative results by explicitly solving for β in (19) given a peer structure $\mathbf{G}$, an assignment $\mathbf{T}$, and parameters (r, σ^2, λ). For instance, Figure 11 shows the distribution of

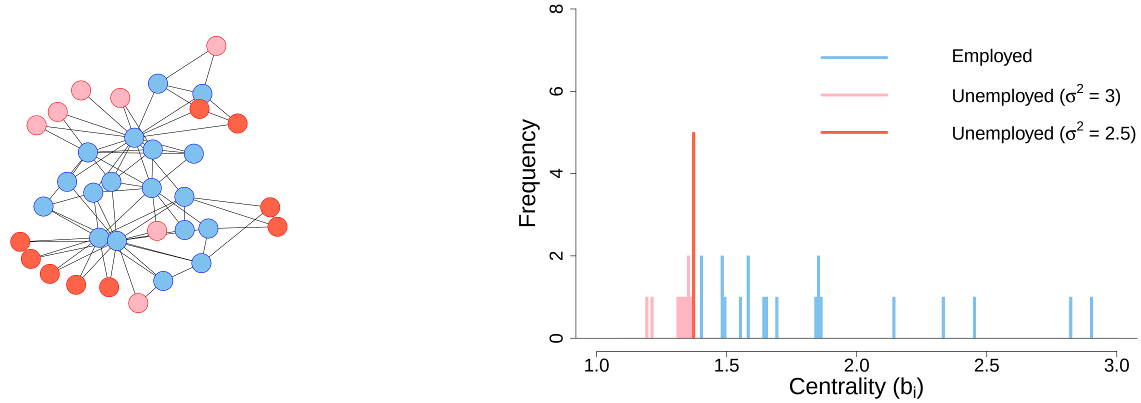


Figure 11: Endogenous unemployment in the Karate Club Network. The unemployed are always the least central. Simulations are run for $\lambda = 0.07$ *and* $r = 1$.

centrality for the very popular "karate club network" due to Zachary (1977).[55] The red and pink bars show the workers that remain unemployed for two different specifications of parameters.

A Note on Second-Order Effects of Unemployment: This subsection focused on the direct, first-order effects of not hiring certain workers, specifically the direct loss in output and increased rents from condensing groups' centralities. To maintain tractability, we omitted second-order effects, such as how a modified network shifts everyone's equilibrium contribution to total output. Using an envelope-theorem type argument, small changes in link strength negligibly affect workers' best replies, so the optimal allocation of incentives, α, should also vary negligibly. Having said this, the workers first removed are those with the lowest centrality, who minimally impact network structure. However, the complete removal of workers doesn't represent infinitesimal changes in link strength, and a thorough analysis would require re-optimizing α after any removal. While this may lead to a slightly different (and more complex) condition than (19), the qualitative features of Proposition 11 still hold. In Supplementary Appendix F, we show numerically (using the karate club network) that incorporating second-order effects slightly lowers the level of unemployment but does not change any of our results above.

5 Concluding Remarks

This paper investigates the optimal design of wage contracts in the presence of peer effects within a firm, emphasizing how network structure influences worker productivity and firm profitability.

[55] This network is very popular in the field of community detection in computer science and graph theory. We use it in this example because it is small enough and has a sufficiently diverse distribution of centrality to illustrate our point.

Our findings suggest that firms can strategically use network-based incentives to enhance overall productivity and that optimal incentive design must account for how co-worker externalities relate with the production technology and granularity of the contract.

Our framework provides a rationale for observed trends in the steepening earnings profile within firms that doesn't rely on unrealistic 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 explain why variable-pay increases as you move up the organizational hierarchy.

An important next step takes this model to the data. Using detailed data on co-worker networks within firms, we can test if performance-based compensation tracks worker centrality as the theory suggests, and we can compute untapped productivity bargains that firms could exploit with more sophisticated contracts. 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.

Future research could extend this framework to consider multiple competing firms, which would influence both firm strategies and worker behavior. Specifically, incorporating models of competition among firms as explored by Chade and Eeckhout (2023) could provide deeper insights into how firms vie for highly central and productive workers and the resulting equilibrium market structures. 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).

Moreover, considering the implications of common ownership and anti-competitive practices highlighted by Garud et al. (2009) recent work could be highly beneficial. This extension would examine how managerial incentives and firm productivity are affected when multiple firms share ownership, potentially leading to coordinated strategies that might alter the competitive landscape.

References

Akerlof, G. A. and J. L. Yellen (1986): *Efficiency wage models of the labor market*, Cambridge University Press.

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.

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., 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.

Helsley, R. W. and Y. Zenou (2014): "Social networks and interactions in cities," *Journal of Economic Theory*, 426–466.

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.

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*.

König, M. D., C. J. Tessone, and Y. Zenou (2014): "Nestedness in networks: A theoretical model and some applications," *Theoretical Economics*, 9, 695–752.

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

Lieber, E. M. and W. Skimmyhorn (2018): "Peer effects in financial decision-making," *Journal of Public Economics*, 163, 37–59.

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 (2023): "Organizing Modular Production," .

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.

ROSEN, S. (1986): "Prizes and Incentives in Elimination Tournaments." *American Economic Review*, 76.

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

SHAPIRO, C. AND J. E. STIGLITZ (1984): "Equilibrium unemployment as a worker discipline device," *The American Economic Review*, 74, 433–444.

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.

ZACHARY, W. W. (1977): "An information flow model for conflict and fission in small groups," *Journal of Anthropological Research*, 33, 452–473.

A Proofs

A.1 Proof of Lemma 1

The principal solves the following problem:

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

subject to

$$\mathrm{CE}_i(\mathbf{e}) = U_i, \quad \forall i \in N \tag{PC}$$

To simplify notation, we normalize outside options to zero for everyone, i.e. $U_i = 0$ for all $i \in N$. Notice that under enforceable (and contractible) effort, there is no incentive to pay the agents above their reservation utility and thus the participation constraint (PC) binds. This allows us to rewrite the above problem as:

$$\max_{\mathbf{e},\boldsymbol{\alpha},\boldsymbol{\beta}} E[\pi] = \sum_i e_i - \sum_i \left[\alpha_i \sum_i e_i + \beta_i\right]$$

Since the PC binds, we have:

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

Thus, the problem reduces to:

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

Notice the utility of the principal is globally decreasing in α_i^2, which implies that the optimal contract involves $\alpha_i^2 = 0 \quad \forall i$. Therefore, the problem further reduces to:

$$\max_{\mathbf{e}} E[\pi] = \sum_i e_i - \sum_i \left[\frac{1}{2}e_i^2 - \lambda e_i \sum_{j \in N} g_{ij} e_j\right]$$

This can be expressed in vector form as:

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

The first-order conditions with respect to $\mathbf{e}$ imply:

$$\mathbf{1} - \mathbf{e} + \lambda(\mathbf{G} + \mathbf{G}')\mathbf{e} = 0$$

Solving for $\mathbf{e}$, we get:

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

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

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

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

In the case of $\mathbf{G} = \mathbf{G}'$ and thus $\mathbf{G} + \mathbf{G}' = 2\mathbf{G}$, we get:

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

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

■

A.2 Proof of Proposition 1

Notice that the principal's problem can be written in matrix form as:

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

subject to

$$\mathbf{e}^\star = \mathbf{C}\boldsymbol{\alpha} \tag{IC}$$

With $\mathbf{C} \equiv (I - \lambda\mathbf{G})^{-1}$ depending exclusively on given parameters. Using the (IC), we can rewrite the problem as:

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

The first-order conditions with respect to $\boldsymbol{\alpha}$ imply:

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

where the third equality follows from $\mathbf{GC} = \mathbf{G} + \lambda\mathbf{G}^2 + \lambda^2\mathbf{G}^3 + ... = \mathbf{CG} = \frac{1}{\lambda}(\mathbf{C}-\mathbf{I})$ and the fifth equality uses the fact that $(\mathbf{GC})' = (\mathbf{CG})' = \frac{1}{\lambda}(\mathbf{C}'-\mathbf{I})$. Finally, we have that:

$$\beta_i(\boldsymbol{\alpha}, \mathbf{e}) = \frac{1}{2}e_i^2 - \lambda e_i \sum_{j\in N} g_{ij}e_k + \alpha_i^2\frac{r\sigma^2}{2} - \alpha_i \sum_k e_k$$

and thus:

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

where $\circ$ denotes the Hadamard, element-wise, product.

As $\mathbf{e}^\star = \mathbf{C}\boldsymbol{\alpha}^\star$, the above can be re-expressed as

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

■

A.3 Proof of Corollary 1

The proof follows immediately from Proposition 1.

A.4 Proof of Corollary 2

From Proposition 1, we know that with a symmetric peer-network $\mathbf{G}$ and when $\sigma^2 = 0$, optimal incentives are given by

$$\boldsymbol{\alpha}^\star = \left[\mathbf{C}'(\mathbf{I}-2\lambda\mathbf{G})\mathbf{C}\right]^{-1}\mathbf{C}'\mathbf{1}.$$

With some linear algebra, one observes that

$$
\begin{aligned}
\left[\mathbf{C}'(\mathbf{I}-2\lambda\mathbf{G})\mathbf{C}\right]^{-1}\mathbf{C}'\mathbf{1} &= \left[(\mathbf{I}-2\lambda\mathbf{G})\mathbf{C}\right]^{-1}\mathbf{1} \\
&= \mathbf{C}^{-1}\left[(\mathbf{I}-2\lambda\mathbf{G})\right]^{-1}\mathbf{1} \\
&= (\mathbf{I}-\lambda\mathbf{G})\left[\mathbf{I}-2\lambda\mathbf{G}\right]^{-1}\mathbf{1} \\
&= (\mathbf{I}-\lambda\mathbf{G})\sum_{k=0}^{\infty}(2\lambda\mathbf{G})^{k}\mathbf{1} \\
&= (\mathbf{I}-\lambda\mathbf{G})\left(\mathbf{I}+2\lambda\mathbf{G}+(2\lambda\mathbf{G})^{2}+\cdots\right)\mathbf{1} \\
&= \left(\mathbf{I}+\lambda\mathbf{G}+2(\lambda\mathbf{G})^{2}+4(\lambda\mathbf{G})^{3}+\cdots\right)\mathbf{1} \\
&= \frac{1}{2}\left(\mathbf{I}+\mathbf{I}+2\lambda\mathbf{G}+(2\lambda\mathbf{G})^{2}+(2\lambda\mathbf{G})^{3}+\cdots\right)\mathbf{1} \\
&= \frac{1}{2}\left(\mathbf{I}+(\mathbf{I}-2\lambda\mathbf{G})^{-1}\right)\mathbf{1}.
\end{aligned}
$$

Thus, when there is no fundamental risk, we have that

$$
\boldsymbol{\alpha}^{\star} = \frac{1}{2}(\mathbf{1}+\mathbf{b}(2\lambda)). \tag{20}
$$

■

A.5 Proof of Corollary 3

The proof follows immediately from Proposition 1.

A.6 Proof of Proposition 2

If $(\mathbf{C1})_i > (\mathbf{C1})_j$ then,

$$
\begin{aligned}
&(\mathbf{C1})_i - 1 > (\mathbf{C1})_j - 1 \\
&\Rightarrow ((\mathbf{C}-\mathbf{I})\mathbf{1})_i > ((\mathbf{C}-\mathbf{I})\mathbf{1})_j \\
&\Rightarrow ((\lambda\mathbf{G})\mathbf{C1})_i > ((\lambda\mathbf{G})\mathbf{C1})_j
\end{aligned}
$$

Thus, $\lambda\mathbf{G}$ is an order preserving transformation of $\mathbf{C1} = \mathbf{b}$. Also $(\mathbf{I}+\lambda\mathbf{G})\mathbf{b} = \mathbf{b}+\lambda\mathbf{G}\mathbf{b}$ preserves it as it is the addition of two order-preserving transformations of $\mathbf{b}$. We can keep iterating, such that:

$$\mathbf{b} + \lambda\mathbf{G}(\mathbf{b} + \lambda\mathbf{G}(\mathbf{b} + \lambda\mathbf{G}(\cdots))) = (\mathbf{I} + \lambda\mathbf{G} + (\lambda\mathbf{G})^2 + \cdots)\mathbf{b}$$
$$= \mathbf{Cb}$$

Thus, $\mathbf{Cb} = \hat{\mathbf{b}}$ preserves the order of $\mathbf{b}$. Combining both order-preserving transformations above, we obtain $\lambda\mathbf{GCb} = (\mathbf{C} - \mathbf{I})\mathbf{b}$, which is also an order-preserving transformation of $\mathbf{b}$. Finally, notice that $\mathbf{C} = \mathbf{I} + \lambda\mathbf{GC}$, thus using $\mathbf{Cb} = (\mathbf{I} + \lambda\mathbf{GC})\mathbf{b} = \mathbf{b} + \lambda\mathbf{GCb}$ and replacing the second $\mathbf{b}$ by its order-preserved transformation $\lambda\mathbf{GCb}$, we get $\mathbf{b} + \lambda\mathbf{GC}(\lambda\mathbf{GCb}) = (\mathbf{I} + (\lambda\mathbf{GC})\lambda\mathbf{GC})\mathbf{b}$, which by transitivity also preserves the original order of $\mathbf{b}$. Now, iterating as before, we can obtain:

$$\left(\mathbf{I} + [(\lambda\mathbf{GC})\lambda\mathbf{GC}] + [(\lambda\mathbf{GC})\lambda\mathbf{GC}]^2 + \cdots\right)\mathbf{b} = (\mathbf{I} - [(\lambda\mathbf{GC})\lambda\mathbf{GC}])^{-1}\mathbf{b},$$

which once again preserves the order of $\mathbf{b}$, and naturally means that adding a positive scalar γ also preserves such order:

$$\gamma\left(\mathbf{I} + [\gamma(\lambda\mathbf{GC})\lambda\mathbf{GC}] + [\gamma(\lambda\mathbf{GC})\lambda\mathbf{GC}]^2 + \cdots\right)\mathbf{b} = \gamma(\mathbf{I} - [\gamma(\lambda\mathbf{GC})\lambda\mathbf{GC}])^{-1}\mathbf{b}$$

Let $\gamma = \frac{1}{1+r\sigma^2}$, then:

$$\gamma(\mathbf{I} - [\gamma(\lambda\mathbf{GC})\lambda\mathbf{GC}])^{-1}\mathbf{b} = \frac{1}{1+r\sigma^2}\left(\mathbf{I} - \frac{\lambda^2}{1+r\sigma^2}(\mathbf{GC})\mathbf{GC}\right)^{-1}\mathbf{b} = \mathbf{Wb} = \boldsymbol{\alpha}$$

■

A.7 Proof of Proposition 3

Recall that the optimal performance-based compensation is given by

$$\boldsymbol{\alpha}^\star = \frac{1}{1+r\sigma^2}\left[\mathbf{I} - \frac{\lambda^2}{1+r\sigma^2}(\mathbf{GC})'\mathbf{GC}\right]^{-1}\mathbf{C}'\mathbf{1} = \frac{1}{1+r\sigma^2}\mathbf{WC}'\mathbf{1}.$$

Using the chain rule, we can write the derivative with respect to g_{ij} as the sum of two terms:

$$\frac{\partial\boldsymbol{\alpha}^\star}{\partial g_{ij}} = \frac{1}{1+r\sigma^2}\left[\frac{\partial\mathbf{W}}{\partial g_{ij}}\cdot\mathbf{C}'\mathbf{1} + \mathbf{W}\cdot\frac{\partial\mathbf{C}'\mathbf{1}}{\partial g_{ij}}\right]. \tag{21}$$

Using the rule for the derivative of the inverse of a matrix we have that

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

Next, notice that the derivatives of $\mathbf{G}$ and $\mathbf{C}$ are $\mathbf{E_{ij}}$ and $\lambda\mathbf{CE_{ij}C}$, 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 (21) as

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

which is non-negative because all matrices are non-negative. Next, focus on the second term in brackets of (21). The 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 (21) is

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

Thus, we have established that (21) is non-negative. For the second part of the proposition, consider the derivative of $\alpha_k^\star$ with respect to the link g_{ij}:

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

Let $W(j) \subseteq N$ be the set of workers that share common influence with worker j. First, consider a worker $k \notin W(j)$. Notice that for all $r \in W(j)$, it must be that $\omega_{kr} = 0$. Next, if k has no influence on any worker then $\alpha_k^\star$ is given by Corollary 1, in which case the increase in g_{ij} has no effect on k's incentives. If k has influence on some other worker $s \notin W(j)$, then it must be that k 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 k's component and so $\delta\omega_{ks}/\delta g_{ij}$ must be equal to zero. This implies that (21) is equal to zero for all $k \notin W(j)$.

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

$$\frac{\partial \omega_{kj}}{\partial g_{ij}} \cdot b_j + \omega_{kj} \cdot \frac{\partial b_j}{\partial g_{ij}} > 0,$$

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 ω_{kj}, i.e, $\partial \omega_{kj} / \partial g_{ij} > 0$. In other words, there is at least one positive term in the summation above. Lastly, recall that $w_{jj} \geq 1$. Moreover, notice that in the summation above when $k = j$ and $r = j$, the term $w_{jj}(\partial b_j / \partial g_{ij})$ must be positive. Thus, (21) is strictly positive for all $k \in W(j) \cup \{j\}$.

■

A.8 Proof of Proposition 4

We start from the following expression for firm profits in the planted partition model, which is proved below:

$$\pi = \frac{1}{2} \frac{n}{(1+\sigma^2 r)\left(1 - \frac{\lambda(p+q)}{2} n\right)^2 - \left(\frac{\lambda(p+q)}{2} n\right)^2}$$

We can therefore write individual effort as follows

$$e = \frac{2\pi}{n} = \frac{1}{(1+\sigma^2 r)\left(1 - \frac{\lambda(p+q)}{2} n\right)^2 - \left(\frac{\lambda(p+q)}{2} n\right)^2} 1$$

By a simple mean-field approximation, we assume that effort is a function of the expected adjacency $\bar{G}$, rather than the realized network: $e = (1 - \lambda \bar{G})^{-1} \alpha$. We can therefore write

$$(1 - \lambda \bar{G})^{-1} \alpha = \frac{1}{(1+\sigma^2 r)\left(1 - \frac{\lambda(p+q)}{2} n\right)^2 - \left(\frac{\lambda(p+q)}{2} n\right)^2} \mathbf{1}$$

$$\alpha = \frac{1}{(1+\sigma^2 r)\left(1 - \frac{\lambda(p+q)}{2} n\right)^2 - \left(\frac{\lambda(p+q)}{2} n\right)^2} (1 - \lambda \bar{G}) \mathbf{1}$$

$$\alpha = \frac{1}{(1+\sigma^2 r)\left(1 - \frac{\lambda(p+q)}{2} n\right)^2 - \left(\frac{\lambda(p+q)}{2} n\right)^2} (11 - \lambda \bar{G} \mathbf{1})$$

Notice that: $\bar{G}\mathbf{1} = \left(\frac{n(p+q)}{2}\right)\mathbf{1}$, then:

$$\alpha = \frac{1}{(1+\sigma^2 r)\left(1-\frac{\lambda(p+q)}{2}n\right)^2 - \left(\frac{\lambda(p+q)}{2}n\right)^2}\left(1-\lambda\frac{n(p+q)}{2}\right)1$$
$$\alpha = \frac{\left(1-\lambda\frac{p+q}{2}n\right)}{(1+\sigma^2 r)\left(1-\lambda\frac{p+q}{2}n\right)^2 - \left(\lambda\frac{p+q}{2}n\right)^2}1$$

■

A.9 Proof of Proposition 5

We first show that a firm's profits are maximized at one-half of equilibrium output. Recall that, for a symmetric $\mathbf{G}$ the optimal provision of incentives is given by

$$\boldsymbol{\alpha}^\star = \left[\mathbf{C}(\mathbf{I}-2\lambda\mathbf{G})\mathbf{C}+\sigma^2\mathbf{r}\mathbf{I}\right]^{-1}\mathbf{C}\mathbf{1},$$

and the expected profits are

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

Thus,

$$\begin{aligned} E[\pi^{*'}] &= \boldsymbol{\alpha}^{*'}\mathbf{C}\mathbf{1} - \frac{1}{2}\boldsymbol{\alpha}^{*'}\left[\mathbf{C}(\mathbf{I}-2\lambda\mathbf{G})\mathbf{C}+\sigma^2\mathbf{r}\mathbf{I}\right]\underbrace{\left[\mathbf{C}(\mathbf{I}-2\lambda\mathbf{G})\mathbf{C}+\sigma^2\mathbf{r}\mathbf{I}\right]^{-1}\mathbf{C}\mathbf{1}}_{=\boldsymbol{\alpha}^\star} \\ &= \boldsymbol{\alpha}^{*'}\mathbf{C}\mathbf{1} - \frac{1}{2}\boldsymbol{\alpha}^{*'}\mathbf{C}\mathbf{1} \\ &= \frac{1}{2}\boldsymbol{\alpha}^{*'}\mathbf{C}\mathbf{1} \\ &= \frac{1}{2}\mathbf{e}^{*'}\mathbf{1}. \end{aligned}$$

And therefore, $E[\pi^\star] = \frac{1}{2}\mathbf{X}^\star$.

We now prove the second part of the proposition. Following Proposition 1, we can write the worker's equilibrium condition as

$$\mathbf{e}^\star = \mathbf{C}\boldsymbol{\alpha} = \mathbf{C}(2\mathbf{I}-\mathbf{C}+\mathbf{C}^{-1}r\sigma^2)^{-1}\mathbf{1}$$

Rewriting, we have that

$$\left(2\mathbf{C}^{-1} - \mathbf{I} + r\sigma^2\mathbf{C}^{-2}\right)\mathbf{e}^\star = \mathbf{1} \implies \left(\mathbf{I} - 2\lambda\mathbf{G} + r\sigma^2(\mathbf{I} - \lambda\mathbf{G})^2\right)\mathbf{e}^\star = \mathbf{1}$$

Assuming that $\mathbf{G}$ is symmetric, we have $\mathbf{G} = \mathbf{U\Sigma U}'$ where $\mathbf{U} = (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n)$ is a matrix of unit-eigenvectors of $\mathbf{G}$, while $\mathbf{\Sigma} = \text{diag}(\mu_1, \mu_2, \ldots, \mu_n)$ is a diagonal matrix of corresponding eigenvalues. We can therefore write

$$\left(\mathbf{I} - 2\lambda\mathbf{U\Sigma U}' + r\sigma^2(\mathbf{I} - \lambda\mathbf{U\Sigma U}')^2\right)\mathbf{e}^\star = \mathbf{1}$$

multiplying both sides on the left by $\mathbf{U}'$ and factoring we have

$$\left((1 + r\sigma^2)(\mathbf{I} - 2\lambda\mathbf{\Sigma}) + r\sigma^2\lambda^2\mathbf{\Sigma}^2\right)\mathbf{U}'\mathbf{e}^\star = \mathbf{U}'\mathbf{1}$$

which means that we can write

$$\mathbf{e}^\star = \mathbf{U}\left((1 + r\sigma^2)(\mathbf{I} - 2\lambda\mathbf{\Sigma}) + r\sigma^2\lambda^2\mathbf{\Sigma}^2\right)^{-1}\mathbf{U}'\mathbf{1}$$

where, conveniently,

$$\left((1 + r\sigma^2)(\mathbf{I} - 2\lambda\mathbf{\Sigma}) + r\sigma^2\lambda^2\mathbf{\Sigma}^2\right)^{-1} = \\ \text{diag}\left(\frac{1}{(1 + r\sigma^2)(1 - 2\lambda\mu_1) + r\sigma^2(\lambda\mu_1)^2}, \ldots, \frac{1}{(1 + r\sigma^2)(1 - 2\lambda\mu_n) + r\sigma^2(\lambda\mu_n)^2}\right)$$

and $\mathbf{U}'\mathbf{1} = (\mathbf{u}_1'\mathbf{1}, \mathbf{u}_2'\mathbf{1}, \ldots, \mathbf{u}_n'\mathbf{1})'$. This means that we can write worker i's equilibrium effort as a function of the graph's spectral properties. Letting $u_{\ell,i}$ represent the ith element of vector $\mathbf{u}_\ell$, we have

$$e_i^\star = \sum_\ell \frac{u_{\ell,i}\left(\sum_i u_{\ell,i}\right)}{(1 + r\sigma^2)(1 - 2\lambda\mu_\ell) + r\sigma^2(\lambda\mu_\ell)^2}$$

Finally, combining this expression with our result stating that $\mathbb{E}(\pi^\star) = \frac{1}{2}\sum_i e_i^\star$ gives our final result. ■

A.10 Proof of Corollary 4

The proof follows immediately from the spectral properties of complete bipartite graphs (see footnote 30 and Supplementary Appendix D.1) and Proposition 5.

A.11 Proof of Corollary 5

The proof follows immediately from the spectral properties of regular graphs (see footnote 31 and Supplementary Appendix D.2) and Proposition 5.

A.12 Proof of Corollary 6

The proof follows immediately from the spectral properties of planted partition models (see footnote 32) and Proposition 5.

A.13 Proof of Proposition 6

We can use the general solution in Appendix B for $V = vI$, $\Lambda = G$, and $\Theta = I$ such that we have:

$$\tilde{\mathbf{C}} = \left[\mathbf{I} - \frac{\lambda}{v}\mathbf{G}\right]^{-1} \frac{1}{v}$$

$$\tilde{\mathbf{C}} = \frac{1}{v}\left[\mathbf{I} - \tilde{\lambda}\mathbf{G}\right]^{-1} = \frac{1}{v}\mathbf{C}_{\tilde{\lambda}}$$

$$\boldsymbol{\alpha} = \left(\tilde{\mathbf{C}}(v\mathbf{I} - 2\lambda\mathbf{G})\tilde{\mathbf{C}} + \sigma^2 r\mathbf{I}\right)^{-1} \tilde{\mathbf{C}}\mathbf{1}$$

$$\boldsymbol{\alpha} = \left(\frac{1}{v}\mathbf{C}_{\tilde{\lambda}} v(\mathbf{I} - 2\tilde{\lambda}\mathbf{G})\frac{1}{v}\mathbf{C}_{\tilde{\lambda}} + \sigma^2 r\mathbf{I}\right)^{-1} \frac{1}{v}\mathbf{C}_{\tilde{\lambda}}\mathbf{1}$$

$$\boldsymbol{\alpha} = \left(\mathbf{C}_{\tilde{\lambda}}(\mathbf{I} - 2\tilde{\lambda}\mathbf{G})\mathbf{C}_{\tilde{\lambda}} + v\sigma^2 r\mathbf{I}\right)^{-1} \mathbf{C}_{\tilde{\lambda}}\mathbf{1}$$

$$\mathbf{e} = \tilde{\mathbf{C}}\boldsymbol{\alpha} = \frac{1}{v}\mathbf{C}_{\tilde{\lambda}}\left(\mathbf{C}_{\tilde{\lambda}}(\mathbf{I} - 2\tilde{\lambda}\mathbf{G})\mathbf{C}_{\tilde{\lambda}} + v\sigma^2 r\mathbf{I}\right)^{-1} \mathbf{C}_{\tilde{\lambda}}\mathbf{1}$$

$$\pi = \frac{1}{2}\mathbf{1}'\mathbf{e} = \frac{1}{2}\mathbf{1}'\frac{1}{v}\mathbf{C}_{\tilde{\lambda}}\left(\mathbf{C}_{\tilde{\lambda}}(\mathbf{I} - 2\tilde{\lambda}\mathbf{G})\mathbf{C}_{\tilde{\lambda}} + v\sigma^2 r\mathbf{I}\right)^{-1} \mathbf{C}_{\tilde{\lambda}}\mathbf{1}$$

Using the diagonalization of Proposition 5, we have that:

$$\pi = \frac{1}{2v}\sum_{\ell} \frac{(\mathbf{1}'\mathbf{u}_\ell)^2}{(1 + v\sigma^2 r)(1 - \tilde{\lambda}\mu_\ell)^2 - \left(\tilde{\lambda}\mu_\ell\right)^2}$$

$$\pi = \frac{1}{2v}\sum_{\ell} \frac{(\mathbf{1}'\mathbf{u}_\ell)^2}{(1 + v\sigma^2 r)(1 - \frac{\lambda}{v}\mu_\ell)^2 - \left(\frac{\lambda}{v}\mu_\ell\right)^2}$$

To compute the derivatives, it is useful to take: $(\mathbf{1}'\mathbf{u}_\ell)^2/2 = a$, $\sigma^2 r = b$, $\mu_\ell = c$, $v = x$, $\lambda = y$, and thus, for each ℓ:

$$\frac{1}{2v}\frac{(\mathbf{1}'\mathbf{u}_\ell)^2}{(1+v\sigma^2 r)(1-\frac{\lambda}{v}\mu_\ell)^2-\left(\frac{\lambda}{v}\mu_\ell\right)^2}=\frac{a}{x\left[(1+xb)(1-\frac{y}{x}c)^2-\left(\frac{y}{x}c\right)^2\right]}$$

Then,

$$\frac{\partial}{\partial x}\left(\frac{a}{x\left[(1+xb)(1-\frac{y}{x}c)^2-\left(\frac{y}{x}c\right)^2\right]}\right)=-\frac{a\left(2b(x-cy)+1\right)}{\left(-2bcxy+cy(bcy-2)+bx^2+x\right)^2}$$

$$\frac{\partial}{\partial y}\left(\frac{a}{x\left[(1+xb)\left(1-\left(\frac{yc}{x}\right)^2\right)-\left(\left(\frac{yc}{x}\right)^2\right)\right]}\right)=\frac{2ac\left(b(x-cy)+1\right)}{\left(-2bcxy+cy(bcy-2)+bx^2+x\right)^2}$$

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

$$\begin{aligned}
2ac(b(x-cy)+1)&>a(2b(x-cy)+1)\\
2c(b(x-cy)+1)&>2b(x-cy)+1\\
2cb(x-cy)+2c&>2b(x-cy)+1\\
2cb(x-cy)-2b(x-cy)+2c&>1\\
2cb(x-cy)-2b(x-cy)+2c-2&>-1\\
2b(x-cy)(c-1)+2(c-1)&>-1\\
(b(x-cy)+1)(c-1)&>-\frac{1}{2}\\
(\sigma^2 r(v-\mu_\ell\lambda)+1)(\mu_\ell-1)&>-\frac{1}{2}\\
(1-\mu_\ell)(1+r\sigma^2(v-\mu_\ell\lambda))&<\frac{1}{2}\quad\forall\ell\text{ with }\mathbf{u}_\ell'\mathbf{1}\neq 0.
\end{aligned}$$

Thus, $(1-\mu_\ell)(1+r\sigma^2(v-\mu_\ell\lambda))<\frac{1}{2}\quad\forall\ell$ with $\mathbf{u}_\ell'\mathbf{1}\neq 0$ is a sufficient condition for $\frac{\partial\pi}{\partial\lambda}>-\frac{\partial\pi}{\partial v}$. ■

A.14 Proof of Corollary 7

Recalling the the first eigenvalue of an Erdős-Rényi graph is equal to np and that $\mathbf{u}_\ell'\mathbf{1}=0,\ \forall\ell>1$, then proof follows immediately from Proposition 6.

A.15 Proof of Proposition 7

Take the principal constrained maximization problem:

$$\max_{(\boldsymbol{\alpha})}\left(\hat{e}-\frac{1}{2}\mathbf{e}'(\mathbf{I}-2\lambda\mathbf{G})\mathbf{e}-\frac{\sigma^2 r}{2}\boldsymbol{\alpha}'\boldsymbol{\alpha}\right)$$

subject to

$$\mathbf{e}=\mathbf{C}\boldsymbol{\alpha}$$

$$\mathbf{M}\mathbf{e}=\hat{e}\mathbf{1}_{\boldsymbol{k}}$$

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

$$\mathbf{M}\mathbf{e}=\hat{e}\mathbf{1}_{\boldsymbol{k}}$$

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 $\boldsymbol{\mu}$, the Lagrangian of the above problem can be expressed as:

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

We have the following system of first order conditions:

$$\frac{\partial\mathcal{L}}{\partial\mathbf{e}}=-\frac{1}{2}\left(\breve{\boldsymbol{\Sigma}}+\breve{\boldsymbol{\Sigma}}'\right)\mathbf{e}+\mathbf{M}'\boldsymbol{\mu}=-\boldsymbol{\Sigma}\mathbf{e}+\mathbf{M}'\boldsymbol{\mu}=\mathbf{0}_{\boldsymbol{n}} \tag{22}$$

$$\frac{\partial\mathcal{L}}{\partial\boldsymbol{\mu}}=\mathbf{M}\mathbf{e}-\hat{e}\mathbf{1}_{\boldsymbol{k}}=\mathbf{0}_{\boldsymbol{k}} \tag{23}$$

$$\frac{\partial\mathcal{L}}{\partial\hat{e}}=\mathbf{1}_{\boldsymbol{k}}{}'\boldsymbol{\mu}=1 \tag{24}$$

To ease notation, take $\boldsymbol{\Sigma}=\frac{1}{2}\left(\breve{\boldsymbol{\Sigma}}+\breve{\boldsymbol{\Sigma}}'\right)=(1+\sigma^2 r)(\mathbf{I}-\lambda(\mathbf{G}+\mathbf{G}'))+\sigma^2 r(\lambda\mathbf{G})'(\lambda\mathbf{G})$.[56] From (22) we obtain

$$\mathbf{e}=\boldsymbol{\Sigma}^{-1}\mathbf{M}'\boldsymbol{\mu}, \tag{25}$$

[56]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 $\boldsymbol{\Sigma}=\breve{\boldsymbol{\Sigma}}$.

and plugging it into (23) we get

$$\mathbf{M}\boldsymbol{\Sigma}^{-1}\mathbf{M}'\boldsymbol{\mu} = \hat{e}\mathbf{1}_k \tag{26}$$

Define the $k \times k$ matrix $\mathbf{H} \equiv \mathbf{M}\boldsymbol{\Sigma}^{-1}\mathbf{M}'$. Solving for $\boldsymbol{\mu}$ we can rewrite (26) as

$$\boldsymbol{\mu} = \hat{e}\mathbf{H}^{-1}\mathbf{1}_k, \tag{27}$$

and using (24) we obtain that

$$\mathbf{1}_k'\boldsymbol{\mu} = \hat{e}\mathbf{1}_k'\mathbf{H}^{-1}\mathbf{1}_k = 1. \tag{28}$$

Next, solving for $\hat{e}$ in (28) and using it in (25) and (27) we obtain

$$\begin{aligned}
\hat{e} &= \frac{1}{\mathbf{1}_k'\mathbf{H}^{-1}\mathbf{1}_k}, \\
\boldsymbol{\mu} &= \frac{1}{\mathbf{1}_k'\mathbf{H}^{-1}\mathbf{1}_k}\mathbf{H}^{-1}\mathbf{1}_k, \\
\mathbf{e} &= \frac{1}{\mathbf{1}_k'\mathbf{H}^{-1}\mathbf{1}_k}\boldsymbol{\Sigma}^{-1}\mathbf{M}'\mathbf{H}^{-1}\mathbf{1}_k.
\end{aligned}$$

Finally, using the fact that $\boldsymbol{\alpha} = (\mathbf{I} - \lambda\mathbf{G})\mathbf{e}$, we get:

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

■

A.16 Proof of Corollary 8

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 (13), 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 $\mathbf{M}\mathbf{M}' = \text{diag}(n_1, n_2, \ldots, n_K)$ 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 8 becomes

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

■

A.17 Proof of Corollary 9

Notice from Proposition 7 that for $\mathbf{M} = \mathbf{I}$, we get $\mathbf{H} = \mathbf{M}\mathbf{\Sigma}^{-1}\mathbf{M}' = \mathbf{\Sigma}^{-1}$, and:

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

Next, notice that $\mathbf{G}\mathbf{1} = \mathbf{d}$ where $\mathbf{d} := (d_1, d_2, \ldots, d_N)'$. Therefore:

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

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

A.18 Proof of Proposition 8

We aim to maximize:

$$\max_{\hat{\alpha},\mu} \mathbb{E}[\pi(X,w|e)] = \mathbf{e}'\mathbf{1} + \lambda\mathbf{e}'\mathbf{G}\mathbf{e} - \frac{1}{2}\mathbf{e}'\mathbf{e} - \frac{1}{2}r\sigma^2(\mathbf{T}'\hat{\boldsymbol{\alpha}})'(\mathbf{T}'\hat{\boldsymbol{\alpha}}) - \mathbf{1}'\mu$$

subject to

$$\begin{aligned}
&\mu_i \geq 0 \quad \forall\, i \\
&e = (\mathbf{I}-\lambda\mathbf{G})^{-1}(\mathbf{T}'\hat{\boldsymbol{\alpha}}) \qquad \text{(IC).}
\end{aligned}$$

Replacing $\mathbf{e} = (\mathbf{I} - \lambda\mathbf{G})^{-1}(\mathbf{T}'\hat{\boldsymbol{\alpha}}) = \mathbf{C}\mathbf{T}'\hat{\boldsymbol{\alpha}}$ eliminates the IC, reducing the problem to:

$$\max_{\hat{\alpha},\mu} \mathbb{E}[\pi] = (\mathbf{C}\mathbf{T}'\hat{\boldsymbol{\alpha}})'\mathbf{1} + \lambda(\mathbf{C}\mathbf{T}'\hat{\boldsymbol{\alpha}})'\mathbf{G}(\mathbf{C}\mathbf{T}'\hat{\boldsymbol{\alpha}}) - \frac{1}{2}(\mathbf{C}\mathbf{T}'\hat{\boldsymbol{\alpha}})'(\mathbf{C}\mathbf{T}'\hat{\boldsymbol{\alpha}}) - \frac{1}{2}r\sigma^2(\mathbf{T}'\hat{\boldsymbol{\alpha}})'(\mathbf{T}'\hat{\boldsymbol{\alpha}}) - \mathbf{1}'\mu$$

subject to

$$\mu_i \geq 0 \quad \forall\, i$$

Taking derivatives with respect to $\hat{\boldsymbol{\alpha}}$, we find the first-order conditions (FOCs) as follows:

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

■

A.19 Proof of Proposition 9

Consider two groups k and h such that $\overline{\mathbf{b}}_k = \bar{b}_k/n_k \geq \bar{b}_h/n_h = \overline{\mathbf{b}}_h$. Then we can write:

$$((\mathbf{T}\mathbf{T}')^{-1}\mathbf{T}\mathbf{b})_k > ((\mathbf{T}\mathbf{T}')^{-1}\mathbf{T}\mathbf{b})_h.$$

We know from the proof of proposition 2 that $\lambda\mathbf{G}$, as well as $\mathbf{C}\lambda\mathbf{G}$, preserve the order of the vector of Bonacich centralities $\mathbf{b}$. Moreover, it is true that $\lambda\mathbf{G}\mathbf{b} = \lambda\mathbf{G}\mathbf{C}\mathbf{1} = (\mathbf{C} - \mathbf{I})\mathbf{1} = \mathbf{b} - \mathbf{1}$. Thus, we have that

$$\begin{aligned}
&((\mathbf{T}\mathbf{T}')^{-1}\mathbf{T}\lambda\mathbf{G}\mathbf{b})_k > ((\mathbf{T}\mathbf{T}')^{-1}\mathbf{T}\lambda\mathbf{G}\mathbf{b})_h \Leftrightarrow ((\mathbf{T}\mathbf{T}')^{-1}\mathbf{T}\mathbf{b}_k - 1 > ((\mathbf{T}\mathbf{T}')^{-1}\mathbf{T}\mathbf{b}_h - 1 \\
&((\mathbf{T}\mathbf{T}')^{-1}\mathbf{T}(\lambda\mathbf{G})^x)_k > ((\mathbf{T}\mathbf{T}')^{-1}\mathbf{T}(\lambda\mathbf{G})^x)_h \\
&((\mathbf{T}\mathbf{T}')^{-1}\mathbf{T}\mathbf{C}(\lambda\mathbf{G}))_k > ((\mathbf{T}\mathbf{T}')^{-1}\mathbf{T}\mathbf{C}(\lambda\mathbf{G}))_h
\end{aligned}$$

Thus, $(\mathbf{T}\mathbf{T}')^{-1}\mathbf{T}\mathbf{C}\lambda\mathbf{G}\mathbf{b}$ preserves the order of $(\mathbf{T}\mathbf{T}')^{-1}\mathbf{T}\mathbf{b} = \overline{\mathbf{b}}$, and naturally, higher order of $(\mathbf{C}\lambda\mathbf{G})^x$ as well.

Next, recall that

$$\begin{aligned}
\hat{\boldsymbol{\alpha}} &= \frac{1}{1 + r\sigma^2}\left(\mathbf{I} + \gamma(\mathbf{T}\mathbf{T}')^{-1}(\mathbf{C}\lambda\mathbf{G}\mathbf{T}')'(\mathbf{C}\lambda\mathbf{G}\mathbf{T}') + (\gamma(\mathbf{T}\mathbf{T}')^{-1}(\mathbf{C}\lambda\mathbf{G}\mathbf{T}')'(\mathbf{C}\lambda\mathbf{G}\mathbf{T}'))^2 + \cdots\right)(\mathbf{T}\mathbf{T}')^{-1}\mathbf{T}\mathbf{B} \\
&= \frac{1}{1 + r\sigma^2}\left(\mathbf{I} + \gamma(\mathbf{T}\mathbf{T}')^{-1}(\mathbf{C}\lambda\mathbf{G}\mathbf{T}')'(\mathbf{C}\lambda\mathbf{G}\mathbf{T}') + (\gamma(\mathbf{T}\mathbf{T}')^{-1}(\mathbf{C}\lambda\mathbf{G}\mathbf{T}')'(\mathbf{C}\lambda\mathbf{G}\mathbf{T}'))^2 + \cdots\right)\overline{\mathbf{b}}.
\end{aligned}$$

We now show that $(\gamma(\mathbf{TT}')^{-1}(\mathbf{C}\lambda\mathbf{GT}')'(\mathbf{C}\lambda\mathbf{GT}'))\overline{\mathbf{b}}$ preserves the order of $\overline{\mathbf{b}}$. Notice that

$$(\mathbf{TT}')^{-1}(\mathbf{C}\lambda\mathbf{GT}')'(\mathbf{C}\lambda\mathbf{GT}')\overline{\mathbf{b}} = (\mathbf{TT}')^{-1}\mathbf{TC}\lambda\mathbf{GC}\lambda\mathbf{GT}'(\mathbf{TT}')^{-1}\mathbf{Tb},$$

and most importantly, notice that both $\mathbf{C}\lambda\mathbf{GC}\lambda\mathbf{G}$ and $\mathbf{T}'(\mathbf{TT}')^{-1}\mathbf{T}$ are symmetric matrices. Then,

$$(\mathbf{TT}')^{-1}\mathbf{TT}'(\mathbf{TT}')^{-1}\mathbf{TC}\lambda\mathbf{GC}\lambda\mathbf{Gb} = (\mathbf{TT}')^{-1}\mathbf{TC}\lambda\mathbf{GC}\lambda\mathbf{Gb} = (\mathbf{TT}')^{-1}\mathbf{T}(\mathbf{C}\lambda\mathbf{G})^2\mathbf{b},$$

which we have established above that has the same order as $(\mathbf{TT}')^{-1}\mathbf{Tb}$. Thus, we have established that $(\mathbf{TT}')^{-1}(\mathbf{C}\lambda\mathbf{GT}')'(\mathbf{C}\lambda\mathbf{GT}')(\mathbf{TT}')^{-1}\mathbf{Tb}$ also preserves the order of $(\mathbf{TT}')^{-1}\mathbf{Tb}$. Iterating we obtain that $[(\mathbf{TT}')^{-1}(\mathbf{C}\lambda\mathbf{GT}')'(\mathbf{C}\lambda\mathbf{GT}')]^x(\mathbf{TT}')^{-1}\mathbf{Tb}$ for $x \geq 2$, and thus also

$$\left(\mathbf{I} + ((\mathbf{TT}')^{-1}(\mathbf{C}\lambda\mathbf{GT}')'(\mathbf{C}\lambda\mathbf{GT}')) + ((\mathbf{TT}')^{-1}(\mathbf{C}\lambda\mathbf{GT}')'(\mathbf{C}\lambda\mathbf{GT}'))^2 + \cdots\right)\bar{\mathbf{B}}$$

preserve the order of $(\mathbf{TT}')^{-1}\mathbf{Tb}$.

Hence,

$$\begin{aligned}&\gamma\left(\mathbf{I} + \gamma(\mathbf{TT}')^{-1}(\mathbf{C}\lambda\mathbf{GT}')'(\mathbf{C}\lambda\mathbf{GT}') + (\gamma(\mathbf{TT}')^{-1}(\mathbf{C}\lambda\mathbf{GT}')'(\mathbf{C}\lambda\mathbf{GT}'))^2 + \cdots\right)(\mathbf{TT}')^{-1}\mathbf{Tb}\\ &= \frac{1}{1+r\sigma^2}\left(\mathbf{I} - \frac{\lambda^2}{1+r\sigma^2}(\mathbf{GC})\mathbf{GC}\right)^{-1}(\mathbf{TT}')^{-1}\mathbf{Tb} = \mathbf{W}(\mathbf{TT}')^{-1}\mathbf{Tb} = \hat{\boldsymbol{\alpha}}.\end{aligned}$$

and we have established that $\hat{\boldsymbol{\alpha}}$ is a monotonic transformation of the group's average Bonacich centrality. ■

A.20 Proof of Lemma 2

We know that:

$$\begin{aligned}\boldsymbol{\alpha}^\star &= \mathbf{T}'\hat{\boldsymbol{\alpha}} = \mathbf{T}'(\sigma^2 r\mathbf{TT}' + \mathbf{TC}'(\mathbf{I} - 2\lambda\mathbf{G})\mathbf{CT}')^{-1}\mathbf{TC}'\mathbf{1}\\ \mathbf{e}^\star &= \mathbf{C}\boldsymbol{\alpha}^\star = \mathbf{CT}'(\sigma^2 r\mathbf{TT}' + \mathbf{TC}'(\mathbf{I} - 2\lambda\mathbf{G})\mathbf{CT}')^{-1}\mathbf{TC}'\mathbf{1}\\ X^\star &= \mathbf{1}'\mathbf{e}^\star = \mathbf{1}'\mathbf{CT}'(\sigma^2 r\mathbf{TT}' + \mathbf{TC}'(\mathbf{I} - 2\lambda\mathbf{G})\mathbf{CT}')^{-1}\mathbf{TC}'\mathbf{1}.\end{aligned}$$

Knowing that $w = \boldsymbol{\alpha}X + \boldsymbol{\beta}$, we get:

$$\begin{aligned}\pi^\star &= \mathbf{1}'\mathbf{e}^\star - \mathbf{1}'\mathbf{w}\\ &= (\mathbf{1} - \boldsymbol{\alpha})'\mathbf{e}^\star - \mathbf{1}'\boldsymbol{\beta}.\end{aligned}$$

From the maximization problem, we know that $\boldsymbol{\beta}_i$ is defined by the agent with the highest effort cost $\bar{\psi}_k = \max_{i\in k}\{\psi_i\}$:

$$\boldsymbol{\beta}_k = \alpha_k^2 \left(\frac{r\sigma^2}{2}\right) - \boldsymbol{\alpha}_k X + \bar{\psi}_k, \forall i \in k$$

We define μ_i as the rents perceived by i due to the impossibility of fully extracting rents, thus:

$$\boldsymbol{\beta}_k = \alpha_k^2 \left(\frac{r\sigma^2}{2}\right) - \boldsymbol{\alpha}_k X + \underbrace{(\bar{\psi}_k - \psi_i)}_{=\mu_i} + \psi_i, \forall i \in k$$

Which in turn allows us to define profits as:

$$\begin{aligned}
\pi^\star &= \mathbf{1}'\mathbf{e} - \boldsymbol{\alpha}'\mathbf{e} - \frac{r\sigma^2}{2}\boldsymbol{\alpha}'\boldsymbol{\alpha} + \boldsymbol{\alpha}'\mathbf{e} - \mathbf{1}'\boldsymbol{\mu} - \mathbf{1}'\boldsymbol{\psi} \\
&= \mathbf{1}'\mathbf{e} - \frac{r\sigma^2}{2}\boldsymbol{\alpha}'\boldsymbol{\alpha} - \mathbf{1}'\boldsymbol{\psi} - \mathbf{1}'\boldsymbol{\mu} \\
&= \mathbf{1}'\mathbf{e} - \frac{r\sigma^2}{2}\boldsymbol{\alpha}'\boldsymbol{\alpha} - \underbrace{\left(\frac{1}{2}\mathbf{e}'\mathbf{e} - \mathbf{e}'\lambda\mathbf{G}\mathbf{e}\right)}_{=\mathbf{1}'\boldsymbol{\psi}} - \mathbf{1}'\boldsymbol{\mu} \\
&= \mathbf{1}'\mathbf{e} - \frac{r\sigma^2}{2}\boldsymbol{\alpha}'\boldsymbol{\alpha} - \frac{1}{2}\mathbf{e}'\left(\mathbf{I} - 2\lambda\mathbf{G}\right)\mathbf{e} - \mathbf{1}'\boldsymbol{\mu} \\
&= \mathbf{1}'\mathbf{e} - \frac{r\sigma^2}{2}\boldsymbol{\alpha}'\boldsymbol{\alpha} - \frac{1}{2}\boldsymbol{\alpha}'\mathbf{C}'\left(\mathbf{I} - 2\lambda\mathbf{G}\right)\mathbf{C}\boldsymbol{\alpha} - \mathbf{1}'\boldsymbol{\mu} \\
&= \mathbf{1}'\mathbf{e} - \frac{1}{2}\boldsymbol{\alpha}'\left(r\sigma^2\mathbf{I} + \mathbf{C}'\left(\mathbf{I} - 2\lambda\mathbf{G}\right)\mathbf{C}\right)\boldsymbol{\alpha} - \mathbf{1}'\boldsymbol{\mu} \\
&= \mathbf{1}'\mathbf{e} - \frac{1}{2}(\mathbf{T}'\hat{\boldsymbol{\alpha}})'\left(r\sigma^2\mathbf{I} + \mathbf{C}'\left(\mathbf{I} - 2\lambda\mathbf{G}\right)\mathbf{C}\right)(\mathbf{T}'\hat{\boldsymbol{\alpha}}) - \mathbf{1}'\boldsymbol{\mu} \\
&= \mathbf{1}'\mathbf{e} - \frac{1}{2}\hat{\boldsymbol{\alpha}}'\left(r\sigma^2\mathbf{T}\mathbf{T}' + \mathbf{T}\mathbf{C}'\left(\mathbf{I} - 2\lambda\mathbf{G}\right)\mathbf{C}\mathbf{T}'\right)\hat{\boldsymbol{\alpha}} - \mathbf{1}'\boldsymbol{\mu}
\end{aligned}$$

And by replacing with the optimal $\hat{\boldsymbol{\alpha}}^\star$, we get:

$$\begin{aligned}
\pi^\star &= \mathbf{1}'\mathbf{e} - \frac{1}{2}\hat{\boldsymbol{\alpha}}'\left(r\sigma^2\mathbf{T}\mathbf{T}' + \mathbf{T}\mathbf{C}'\left(\mathbf{I} - 2\lambda\mathbf{G}\right)\mathbf{C}\mathbf{T}'\right)\underbrace{(\sigma^2 r\mathbf{T}\mathbf{T}' + \mathbf{T}\mathbf{C}'(\mathbf{I} - 2\lambda\mathbf{G})\mathbf{C}\mathbf{T}')^{-1}\mathbf{T}\mathbf{C}'\mathbf{1}}_{=\hat{\boldsymbol{\alpha}}} - \mathbf{1}'\boldsymbol{\mu} \\
&= \mathbf{1}'\mathbf{e} - \frac{1}{2}\hat{\boldsymbol{\alpha}}'\underbrace{\left(r\sigma^2\mathbf{T}\mathbf{T}' + \mathbf{T}\mathbf{C}'\left(\mathbf{I} - 2\lambda\mathbf{G}\right)\mathbf{C}\mathbf{T}'\right)(\sigma^2 r\mathbf{T}\mathbf{T}' + \mathbf{T}\mathbf{C}'(\mathbf{I} - 2\lambda\mathbf{G})\mathbf{C}\mathbf{T}')^{-1}}_{=\mathbf{I}}\mathbf{T}\mathbf{C}'\mathbf{1} - \mathbf{1}'\boldsymbol{\mu} \\
&= \mathbf{1}'\mathbf{e} - \frac{1}{2}\hat{\boldsymbol{\alpha}}'\mathbf{T}\mathbf{C}'\mathbf{1} - \mathbf{1}'\boldsymbol{\mu}
\end{aligned}$$

And therefore, using $\hat{\boldsymbol{\alpha}}'\mathbf{T}\mathbf{C}' = \boldsymbol{\alpha}'\mathbf{C}' = \mathbf{e}'$ and the fact that $\mathbf{1}'\mathbf{e} = \mathbf{e}'\mathbf{1}$, we get:

$$\begin{aligned}
\pi^\star &= \mathbf{1}'\mathbf{e} - \frac{1}{2}\hat{\boldsymbol{\alpha}}'\mathbf{T}\mathbf{C}'\mathbf{1} - \mathbf{1}'\boldsymbol{\mu} \\
&= \mathbf{1}'\mathbf{e} - \frac{1}{2}\boldsymbol{\alpha}'\mathbf{C}'\mathbf{1} - \mathbf{1}'\boldsymbol{\mu} \\
&= \mathbf{1}'\mathbf{e} - \frac{1}{2}\mathbf{e}'\mathbf{1} - \mathbf{1}'\boldsymbol{\mu} \\
&= \frac{1}{2}\mathbf{1}'\mathbf{e} - \mathbf{1}'\boldsymbol{\mu}
\end{aligned}$$

And thus, $\pi^\star = \frac{1}{2}\mathbf{e}'\mathbf{1} - \mathbf{1}'\boldsymbol{\mu}$. ■

A.21 Proof of Proposition 10

Following proposition 1, we can compactly write the optimal efforts for granular contracts (G) as

$$\mathbf{e}^G = \frac{1}{1+r\sigma^2}\mathbf{C}\boldsymbol{\alpha}^G = \frac{1}{1+r\sigma^2}\mathbf{C}\mathbf{W}\mathbf{C}'\mathbf{1};$$

and, by proposition 8, we can do the analogous for coarse contracts (C) as

$$\mathbf{e}^C = \frac{1}{1+r\sigma^2}\mathbf{C}\boldsymbol{\alpha}^C = \frac{1}{1+r\sigma^2}\mathbf{C}\mathbf{T}'\hat{\boldsymbol{\alpha}}^C = \frac{1}{1+r\sigma^2}\mathbf{C}\mathbf{T}'(\mathbf{T}\mathbf{W}^{-1}\mathbf{T}')^{-1}\mathbf{T}\mathbf{C}'\mathbf{1},$$

where $W = [\mathbf{I} - \lambda^2/(1+r\sigma^2)(\mathbf{C}\mathbf{G})'\mathbf{C}\mathbf{G}]^{-1}$.

Thus, the difference in outputs can be written as

$$\begin{aligned}
\mathbf{1}'(\mathbf{e}^G - \mathbf{e}^C) &= \frac{1}{1+r\sigma^2}\mathbf{1}'(\mathbf{C}\mathbf{W}\mathbf{C}'\mathbf{1} - \mathbf{C}\mathbf{T}'(\mathbf{T}\mathbf{W}^{-1}\mathbf{T}')^{-1}\mathbf{T}\mathbf{C}'\mathbf{1}) \\
&= \frac{1}{1+r\sigma^2}\mathbf{1}'\mathbf{C}[\mathbf{W} - \mathbf{T}'(\mathbf{T}\mathbf{W}^{-1}\mathbf{T}')^{-1}\mathbf{T}]\mathbf{C}'\mathbf{1} \\
&= \frac{1}{1+r\sigma^2}\mathbf{b}'[\mathbf{W} - \mathbf{T}'(\mathbf{T}\mathbf{W}^{-1}\mathbf{T}')^{-1}\mathbf{T}]\mathbf{b}.
\end{aligned} \tag{29}$$

where $\mathbf{b} = \mathbf{C}'\mathbf{1}$ is the vector of outward Bonacich centralities.

We want to show that there exists a γ such that:

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

where $\mathrm{Var}(\mathbf{b}_k)$ is the variance within group k and $\sum_k$ sums over all k groups[57]. The average

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

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, notice that the matrix $(\mathbf{TT}')^{-1}$ is a $k\times k$ diagonal matrix with (k,k) element equal to $1/n_k$. Leveraging this fact, we can write the expression above in matrix form as

$$\begin{aligned}\frac{1}{n}\sum_k n_k \mathrm{Var}(\mathbf{b}_k) &= \frac{1}{n}\left[\mathbf{b}'\mathbf{b} - (\mathbf{Tb})')(\mathbf{TT}')^{-1}\mathbf{Tb}\right] \\ &= \frac{1}{n}\left(\mathbf{b}'\left[\mathbf{I}-\mathbf{T}'(\mathbf{TT}')^{-1}\mathbf{T}\right]\mathbf{b}\right)\end{aligned} \tag{30}$$

Putting together (29) and (30) we have

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

Finally, notice that when $\lambda^2/(r\sigma^2)\to 0$ it is true that $\mathbf{W}\to\mathbf{I}$ as well as $(\mathbf{TW}^{-1}\mathbf{T}')^{-1}\to(\mathbf{TT}')^{-1}$. Thus, in the limit, as $\lambda^2/(r\sigma^2)\to 0$, we have that (29) and (30) are the same when $\gamma=\frac{n}{1+r\sigma^2}$. That is,

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

■

A.22 Proof of Lemma 3

In coarse contracts we know that, for each group k:

$$\hat{\beta}_k = \frac{1}{2}r\sigma^2\alpha_k^2 - \alpha_k X + \underbrace{\bar{\psi}_k - \psi_i}_{\mu_i} + \psi_i.$$

Notice that the first two terms above are common terms for all members of group k and can be ignored when comparing $\hat{\beta}_k$ for two members of the same group. Moreover, for $\bar{\psi}_k=\psi_i$, we have that $\hat{\beta}_k=\psi_i$, i.e., $\hat{\beta}_k$.

Next, observe that

$$\psi_i = \frac{1}{2}e_i^2 - \lambda \sum_i e_i e_j g_{ij} = \frac{1}{2}e_i(e_i - 2\lambda \sum_i e_j g_{ij}).$$

Using the first order condition agent i, we can rewrite this expression as

$$\begin{aligned}
\psi_i &= \frac{1}{2}(\alpha_k + \lambda \sum_{j \in N} e_j g_{ij})(\alpha_k + \lambda \sum_{j \in N} e_j g_{ij} - 2\lambda \sum_{j \in N} e_j g_{ij}) \\
&= \frac{1}{2}(\alpha_k + \lambda \sum_{j \in N} e_j g_{ij})(\alpha_k - \lambda \sum_{j \in N} e_j g_{ij}) \\
&= \frac{1}{2}(\alpha_k^2 - (\lambda \sum_{j \in N} e_j g_{ij})^2).
\end{aligned}$$

Thus, for two agents i, j in group k, if

$$\begin{aligned}
\psi_i > \psi_j &\Leftrightarrow \frac{1}{2}(\alpha_k^2 - (\lambda \sum_l e_l g_{il})^2) > \frac{1}{2}(\alpha_k^2 - (\lambda \sum_l e_l g_{jl})^2) \\
&\Leftrightarrow -(\lambda \sum_l e_l g_{il})^2) > -(\lambda \sum_l e_l g_{jl})^2) \\
&\Leftrightarrow (\lambda \sum_l e_l g_{il})^2) < (\lambda \sum_l e_l g_{jl})^2) \\
&\Leftrightarrow \sum_l e_l g_{il} < \sum_l e_l g_{jl} \\
&\Leftrightarrow (\mathbf{Ge})_i < (\mathbf{Ge})_j \\
&\Leftrightarrow (\mathbf{GC}\boldsymbol{\alpha})_i < (\mathbf{GC}\boldsymbol{\alpha})_j \\
&\Leftrightarrow (\lambda\mathbf{GC}\boldsymbol{\alpha})_i < (\lambda\mathbf{GC}\boldsymbol{\alpha})_j \\
&\Leftrightarrow ((\mathbf{C} - \mathbf{I})\boldsymbol{\alpha})_i < ((\mathbf{C} - \mathbf{I})\boldsymbol{\alpha})_j \\
&\Leftrightarrow (\mathbf{C}\boldsymbol{\alpha})_i < (\mathbf{C}\boldsymbol{\alpha})_j \\
&\Leftrightarrow (\mathbf{C}\dddot{\mathbf{W}}\mathbf{CTb})_i < (\mathbf{C}\dddot{\mathbf{W}}\mathbf{CTb})_j \\
&\Leftrightarrow \mathbf{b}_i < \mathbf{b}_j
\end{aligned}$$

by the previous proof. ■

A.23 Proof of Proposition 11

We drop the subscript k in the following proof to ease notation. We can rewrite equation (17) as

$$\frac{1-\int_{r=1}^{q(\beta)} dF_k(r)}{\int_{r=1}^{q(\beta)} r dF_k(r)} > \frac{\alpha}{\overline{\beta}-\beta}$$

We therefore have that

$$\frac{1}{\int_{r=1}^{q(\beta)} r dF_k(r)} - \frac{\int_{r=1}^{q(\beta)} dF_k(r)}{\int_{r=1}^{q(\beta)} r dF_k(r)} = \frac{1}{\int_{r=1}^{q(\beta)} r dF_k(r)} - \frac{1}{\mathbb{E}\left[b \mid b < q(\beta)\right]} > \frac{\alpha}{\overline{\beta}-\beta}$$

$$\implies \frac{\mathbb{E}\left[b \mid b < q(\beta)\right]}{\int_{r=1}^{q(\beta)} r dF_k(r)} - 1 > \frac{\alpha}{\overline{\beta}-\beta}\mathbb{E}\left[b \mid b < q(\beta)\right] \implies \frac{1}{\int_{r=1}^{q(\beta)} dF_k(r)} - 1 > \frac{\alpha}{\overline{\beta}-\beta}\mathbb{E}\left[b \mid b < q(\beta)\right]$$

$$\implies \frac{1-\int_{r=1}^{q(\beta)} dF_k(r)}{\int_{r=1}^{q(\beta)} dF_k(r)} > \frac{\alpha}{\overline{\beta}-\beta}\mathbb{E}\left[b \mid b < q(\beta)\right] \implies \frac{\Pr(b > q(\beta))}{\Pr(b < q(\beta))} > \frac{\alpha}{\overline{\beta}-\beta}\mathbb{E}\left[b \mid b < q(\beta)\right]$$

■

B Heterogeneous Agents and Coarse Instruments

Consider the case with agents heterogeneous in their productivity per unit of effort (θ_i), relative cost of effort (v_i), risk aversion (r_i), and reservation utility ($-\exp[-r_i(U_i)]$). Moreover, we allow for peer-effects that are not necessarily bilateral nor homogeneous, and captured by matrix $\mathbf{\Lambda}$, where $\mathbf{\Lambda}_{ij}$ is the effort reduction for i, given its interaction with j. We allow for asymmetries in $\mathbf{\Lambda}$, and thus $\mathbf{\Lambda}_{ij} \neq \mathbf{\Lambda}_{ji}$ is allowed. Under these conditions total output is given by:

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

Where θ_i represents the productivity per unit of effort of agent i. Again, the Principal focuses on the case of linear wage contracts of the form $w_i(X) = \beta_i + \alpha_i X$, as described above. In this case, the cost of effort of agent i is given by:

$$\psi_i(\mathbf{e}, \mathbf{\Lambda}) = \frac{1}{2} v_i e_i^2 - \lambda e_i \sum_{j \in N} \mathbf{\Lambda}_{ij} e_j$$

Where v_i represents the relative cost per unit of effort of agent i. As before, the certain equivalent of the utility function for each agent, which takes into account the agent's wage, cost of effort, and risk preferences is given by:

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

Following the non-cooperative game described, the optimal level of effort chosen by agent i is given by the first order conditions with respect to e_i, in this case:

$$e_i^\star = \frac{\theta_i}{v_i} \alpha_i + \frac{\lambda}{v_i} \sum_{j \in N} \mathbf{\Lambda}_{ij} e_j$$

As before, the equilibrium in this case is a Nash equilibrium. To ease computations we define $\mathbf{V} = \mathrm{diag}(v)$, $\mathbf{\Theta} = \mathrm{diag}(\theta)$, $\mathbf{R} = \mathrm{diag}(r)$. Therefore, the vector of best responses is:

$$\mathbf{e} = \mathbf{V}^{-1} \mathbf{\Theta} \alpha + \lambda \mathbf{V}^{-1} \mathbf{\Lambda} \mathbf{e}$$

In this case, any Nash equilibrium effort profile $e^\star$ of the game satisfies:

$$\mathbf{e}^\star = \left[\mathbf{I} - \lambda \mathbf{V}^{-1} \mathbf{\Lambda}\right]^{-1} \mathbf{V}^{-1} \mathbf{\Theta} \boldsymbol{\alpha}$$

The above equilibrium is well defined whenever the spectral radius of $\lambda \mathbf{V}^{-1}\boldsymbol{\Lambda}$ is less than 1 and all eigenvalues of G are distinct. Under this setup, the Principal solves:

$$\max_{\alpha,\beta} \mathbb{E}[\pi(X,w)|e] = \sum_i^n \theta_i e_i - \sum_i^n w_i$$

$$\text{subject to}$$

$$\text{CE}_i(\mathbf{e},\mathbf{G}) \geq U_i, \forall i \qquad \text{(PC)}$$

$$\mathbf{e}^\star = \left[\mathbf{I} - \lambda \mathbf{V}^{-1}\boldsymbol{\Lambda}\right]^{-1} \mathbf{V}^{-1}\boldsymbol{\Theta}\boldsymbol{\alpha} \qquad \text{(IC)}$$

As in Proposition 1 , the PC is binding. Thus the Principal's expected profits:

$$\begin{aligned}\max_{\alpha,\beta} \mathbb{E}[\pi(X,w)|e] &= \sum_i^n \theta_i e_i - \sum_i^n w_i \\ &= \left(1 - \sum_i^n \alpha_i\right) \sum_i^n \theta_i e_i - \sum_i^n \beta_i,\end{aligned}$$

become:

$$\max_{\alpha} \mathbb{E}[\pi(X,w)|e] = \sum_i^n \left\{\theta_i e_i - U_i - \frac{1}{2} v_i e_i^2 - \frac{\sigma^2}{2}\alpha_i^2 r_i + \lambda e_i \sum_{j\in N} \boldsymbol{\Lambda}_{ij} e_j\right\}$$

Which obviating the $\sum_k^n U_k$ constant terms, can be expressed in matrix form as:

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

$$\text{subject to}$$

$$\mathbf{e}^\star = \left[\mathbf{I} - \lambda \mathbf{V}^{-1}\boldsymbol{\Lambda}\right]^{-1} \mathbf{V}^{-1}\boldsymbol{\Theta}\boldsymbol{\alpha} \qquad \text{(IC)}$$

Taking $\tilde{\mathbf{C}} \equiv [\mathbf{I} - \lambda \mathbf{V}^{-1}\boldsymbol{\Lambda}]^{-1} \mathbf{V}^{-1}\boldsymbol{\Theta}$ and replacing $\mathbf{e} = \tilde{\mathbf{C}}\boldsymbol{\alpha}$ and $\mathbf{e}' = \boldsymbol{\alpha}'\tilde{\mathbf{C}}'$, the above maximization problem becomes:

$$\max_{\boldsymbol{\alpha}} \mathbb{E}[\pi] = \{\boldsymbol{\alpha}'\tilde{\mathbf{C}}'\boldsymbol{\Theta}\mathbf{1} - \frac{1}{2}\boldsymbol{\alpha}'\tilde{\mathbf{C}}'\mathbf{V}\tilde{\mathbf{C}}\boldsymbol{\alpha} - \frac{\sigma^2}{2}\boldsymbol{\alpha}'\mathbf{R}\boldsymbol{\alpha} + \lambda\boldsymbol{\alpha}'\tilde{\mathbf{C}}'\boldsymbol{\Lambda}\tilde{\mathbf{C}}\boldsymbol{\alpha}\}$$

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

$$\begin{aligned}\mathbf{0} &= \tilde{\mathbf{C}}'\boldsymbol{\Theta}\mathbf{1} - \left[\left(\tilde{\mathbf{C}}'\left(\mathbf{V} - \lambda(\boldsymbol{\Lambda}' + \boldsymbol{\Lambda})\right)\tilde{\mathbf{C}}\right) + \sigma^2\mathbf{R}\right]\boldsymbol{\alpha}^\star \\ \Rightarrow \boldsymbol{\alpha}^\star &= \left[\left(\tilde{\mathbf{C}}'\left(\mathbf{V} - \lambda(\boldsymbol{\Lambda}' + \boldsymbol{\Lambda})\right)\tilde{\mathbf{C}}\right) + \sigma^2\mathbf{R}\right]^{-1} \tilde{\mathbf{C}}'\boldsymbol{\Theta}\mathbf{1}\end{aligned}$$

Therefore, in a network of n heterogeneous agents, fully characterized by the adjacency matrix $\mathbf{\Lambda}$, the optimal linear contract is given by:

$$\boldsymbol{\alpha}^{\star} = \left[\left(\tilde{\mathbf{C}}'\left(\mathbf{V} - \lambda(\mathbf{\Lambda}' + \mathbf{\Lambda})\right)\tilde{\mathbf{C}}\right) + \sigma^2\mathbf{R}\right]^{-1}\tilde{\mathbf{C}}'\mathbf{\Theta}\mathbf{1}$$

The optimal induced effort in this case is given by $\mathbf{e}^{\star} = \tilde{\mathbf{C}}\boldsymbol{\alpha}^{\star}$ and the vector of optimal fixed payments $\boldsymbol{\beta}^{\star}$ can be recovered using for each β_i:

$$\beta_i(\alpha_i^{\star}, e_i^{\star}) = U_i - \alpha_i^{\star}\sum_{k}^{n}\theta_k e_k^{\star} + \frac{1}{2}v_i e_i^2 - \lambda e_i^{\star}\sum_{j\in N}\mathbf{\Lambda}_{ij}e_k^{\star} + (\alpha_i^{\star})^2\,\frac{r_i\sigma^2}{2}.$$

Or, in vector form:

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

Analogous to the case without agent heterogeneity, we analyze the case when the Principal is limited to offer coarse instruments for each worker i in group k, and recalling that we define linear operator T as an $n\times k$ vector-diagonal matrix such that $T_{i,j} = 1$ if individual i belongs to type j and $T_{i,j} = 0$ otherwise, the optimal contract is given by:

$$\boldsymbol{\alpha}^{\star,Coarse} = \mathbf{T}'\hat{\boldsymbol{\alpha}} = \mathbf{T}'\left(\mathbf{T}\left(\tilde{\mathbf{C}}'\left(\mathbf{V} - \lambda\left(\mathbf{\Lambda} + \mathbf{\Lambda}'\right)\right)\tilde{\mathbf{C}} + \sigma^2\mathbf{R}\right)\mathbf{T}'\right)^{-1}\mathbf{T}\tilde{\mathbf{C}}'\mathbf{\Theta}\mathbf{1}$$

$$\boldsymbol{\beta}_k = \alpha_k^2\left(\frac{r\sigma^2}{2}\right) - \boldsymbol{\alpha}_k X + \bar{\psi}_k, \forall i \in k$$

Where $\bar{\psi}_k$ is defined as highest effort cost in k, $\bar{\psi}_k = \max_{i\in k}\{\psi_i\}$.

C Incentive Provision Equivalence with Joint and Individual Production

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

$$q_i = e_i + \varepsilon_i. \tag{31}$$

The random variables $(\varepsilon_i)_{i\in N}$ are assumed to be independently and normally distributed with mean zero and variance σ^2[58]. This means that worker i's compensation is now conditional on individual output rather than the join output $X(\mathbf{e})$. That is,

$$w_i = \beta_i^{IP} + \alpha_i^{IP} q_i, \tag{32}$$

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

$$CE_i^{IP}(\mathbf{e},\mathbf{G};\alpha_i^{IP},\beta_i^{IP}) = \beta_i^{IP} + \alpha_i^{IP} e_i^{IP} - \frac{1}{2}(e_i^{IP})^2 + \lambda e_i^{IP} \sum_{j\in A} g_{ij} e_j^{IP} - (\alpha_i^{IP})^2 \frac{r_i \sigma^2}{2}. \tag{33}$$

Notice that the effort the effort-provision problem of worker i is the same when maximizing (2) and (33). 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

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

while under individual production she maximizes

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

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

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

[58]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 Spectrum Properties for Section 2.3

D.1 Complete Bipartite Graphs

Consider a complete bipartite graph of size N with two partitions of size m and n. Denote by $\mathbf{0}_{m,n}$ an $m \times n$ matrix of all zeros and $\mathbf{1}_{m,n}$ an $m \times n$ matrix of all ones. The adjacency matrix of the graph, $\mathbf{G}$, and its square, $\mathbf{G}^2$ are

$$\mathbf{G} = \left(\begin{array}{c|c} \mathbf{0_{m,m}} & \mathbf{1_{m,n}} \\ \hline \mathbf{1_{n,m}} & \mathbf{0_{n,n}} \end{array}\right), \quad \mathbf{G}^2 = \left(\begin{array}{c|c} n\mathbf{1_{m,m}} & \mathbf{0_{m,n}} \\ \hline \mathbf{0_{n,m}} & m\mathbf{1_{n,n}} \end{array}\right).$$

Since the rank of $\mathbf{G}$ and $\mathbf{G}^2$ are equal to 2, we know that there are at most two non-zero eigenvalues. Moreover, we know that if μ_i is an eigenvalue of $\mathbf{G}$ then μ_i^2 is an eigenvalue of $\mathbf{G}^2$. Thus, since the trace of a matrix equal the sum of its eigenvalues, we have

$$\begin{aligned} \text{trace}(\mathbf{G}) &= 0 = \mu_1 + \mu_2 \Rightarrow \mu_1 = -\mu_2, \\ \text{trace}(\mathbf{G}^2) &= 2mn = \mu_1^2 + \mu_2^2. \end{aligned}$$

where μ_1 and μ_2 are non-zero eigenvalues of $\mathbf{G}$. Solving for these equations we find the spectrum of $\mathbf{G}$: $spec(\mathbf{G}) = \{\sqrt{nm}, 0, \ldots, 0, -\sqrt{mn}\}$.

Next, we find the eigenvectors associate with $\mu_1 = \sqrt{mn}$. That is, we are looking for a non-zero vector, $\mathbf{v}_1' = (\mathbf{v}_{1,m}, \mathbf{v}_{1,n})$, such that

$$\mathbf{G}v_1 = \sqrt{mn}v_1 \Rightarrow \begin{pmatrix} \mathbf{1_{m,n}}\mathbf{v}_{1,n} \\ \mathbf{1_{n,m}}\mathbf{v}_{1,m} \end{pmatrix} = \sqrt{nm} \begin{pmatrix} \mathbf{v}_{1,m} \\ \mathbf{v}_{1,n} \end{pmatrix}.$$

Notice that it must be that all entries of $\mathbf{v}_{1,m}$ are identical and all entries of $\mathbf{v}_{1,n}$ are identical. Thus, we can write these vectors as $\mathbf{v}_{1,m} = \alpha \mathbf{1}_m$ and $\mathbf{v}_{1,n} = \beta \mathbf{1}_n$. Therefore, we have the following system of two equations:

$$\begin{pmatrix} \mathbf{1_{m,n}}\mathbf{v}_{1,n} \\ \mathbf{1_{n,m}}\mathbf{v}_{1,m} \end{pmatrix} = \sqrt{nm} \begin{pmatrix} \alpha\mathbf{1}_m \\ \beta\mathbf{1}_n \end{pmatrix} \Rightarrow \begin{cases} n\beta\mathbf{1}_m & = \alpha\sqrt{mn}\mathbf{1}_m \\ m\alpha\mathbf{1}_n & = \beta\sqrt{mn}\mathbf{1}_n \end{cases} \Rightarrow \begin{cases} \beta = \frac{\alpha}{n}\sqrt{mn} \\ \beta = \frac{m\alpha}{\sqrt{mn}} \end{cases} \Rightarrow \alpha = \alpha.$$

For $\alpha = 1$, the eigenvector associated with $\mu_1 = \sqrt{mn}$ is $\mathbf{v}_1' = (\mathbf{1}_m, \sqrt{m}/\sqrt{n}\mathbf{1}_n)$. Finally, dividing by its norm, i.e., $||\mathbf{v}_1|| = \sqrt{2m}$, we obtain the unit-eigenvector:

$$\mathbf{u}_1 = \begin{pmatrix} \frac{1}{\sqrt{2m}}\mathbf{1}_m \\ \frac{1}{\sqrt{2n}}\mathbf{1}_n \end{pmatrix}.$$

Similar steps lead to the unit-eigenvector associated with the last eigenvalue $\mu_n = -\sqrt{mn}$:

$$\mathbf{u}_n = \begin{pmatrix} \frac{1}{\sqrt{2m}}\mathbf{1}_m \\ -\frac{1}{\sqrt{2n}}\mathbf{1}_n \end{pmatrix}.$$

D.2 Regular Graphs

Consider a d-regular graph of size N, i.e., each node has exactly d neighbors. Since each row of $\mathbf{G}$ has exactly d ones, $\mathbf{G1}$ returns a vector with each entry equal to d. Therefore, it is easy to see that

$$\mathbf{G1} = d\mathbf{1} \Rightarrow d \text{ is an eigenvalue of } \mathbf{G}.$$

If $\mathbf{G}$ is a d-regular graph with K components of size $C_1, \ldots, C_K$ then the eigenvalues of $\mathbf{G}$ are the collection of eigenvalues of each connected component. Therefore, since component is in itself a d-regular graph, it follows that $\mathbf{G}$ has K eigenvalues equal to d. Thus, the eigenvector associated with eigenvalue μ_i for $i \in \{1, \ldots, k\}$ is given by

$$\mathbf{v}_i = \begin{pmatrix} \mathbf{1}_{C_i} \\ \mathbf{0}_{N-C_i} \end{pmatrix}.$$

Finally, when $\mathbf{G}$ is symmetric it can be diagonalized using an orthonormal basis of eigenvectors, and so, all eigenvectors associated with an eigenvalue that is not equal to d must be orthogonal to the vector $\mathbf{1}_n$. The space of vectors orthogonal to $\mathbf{1}_n$ is such that for any vector in the space it must be that the sum of its entries adds up to zero.

E Convex Weights in Modular Production

As mentioned in the text, the allocation rule under any modular structure can be written as the convex combination of the single module case $\mathbf{1}'$ and the essential workers case $\mathbf{M} = \mathbf{I}$.

Corollary 10. *The incentive allocation rule for any modular structure* $\mathbf{M}$ *is a convex combination of the allocation rule under one module (*$\mathbf{M} = \mathbf{1}'_n$*) in Proposition 1 and the allocation rule under* N *modules (*$\mathbf{M} = \mathbf{I}$*) in Corollary 9. Formally,*

$$\alpha_i^{\mathbf{M}} = w_i\, \alpha_i^{\mathbf{1}'_n} + (1 - w_i)\, \alpha_i^{\mathbf{I}} \quad \forall i \in N,$$

where the vector of weights, $\mathbf{w} = (w_1, \ldots, w_N)'$, *is given by:*

$$\mathbf{w} = \left[\mathbf{I} - \frac{\mathbf{\Sigma}}{\mathbf{1}'\mathbf{\Sigma}\mathbf{1}}\right]^{-1} \left[\frac{\mathbf{M}'\mathbf{H}^{-1}\mathbf{1}}{\mathbf{1}'\mathbf{H}^{-1}\mathbf{1}} - \frac{\mathbf{\Sigma}}{\mathbf{1}'\mathbf{\Sigma}\mathbf{1}}\mathbf{1}\right]. \tag{34}$$

E.1 Proof of Corollary 10

Let $\mathbf{Q}_{N\times 1} := \frac{\mathbf{M}'\mathbf{H}^{-1}\mathbf{1}}{\mathbf{1}'\mathbf{H}^{-1}\mathbf{1}}$ and recall that, under Proposition 7, the optimal allocation of incentives is given by:

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

Consider the convex combination, where the diagonal matrix $\mathbf{W}$ gives the weight of the optimal allocation of incentives under 1 module, i.e., $\mathbf{M} = (1, 1, \ldots, 1)_{1\times N}$, which boils down to the optimal allocation given by Proposition 1, and $\mathbf{I} - \mathbf{W}$ the weight of the optimal allocation of incentives under N modules, i.e., $\mathbf{M} = \mathbf{I}$, which is given by Corollary 9:

$$\begin{aligned}
&\mathbf{W}(\mathbf{I} - \lambda\mathbf{G})\mathbf{\Sigma}^{-1}\mathbf{1}_N + (\mathbf{I} - \mathbf{W})\frac{1}{\mathbf{1}_K'\mathbf{\Sigma}\mathbf{1}_K}(\mathbf{I} - \lambda\mathbf{G})\mathbf{1}_N \\
&= (\mathbf{I} - \lambda\mathbf{G})\mathbf{\Sigma}^{-1}\left[\mathbf{W}\mathbf{1}_N + (\mathbf{I} - \mathbf{W})\frac{1}{\mathbf{1}_K'\mathbf{\Sigma}\mathbf{1}_K}\mathbf{\Sigma}\mathbf{1}_N\right] \\
&= (\mathbf{I} - \lambda\mathbf{G})\mathbf{\Sigma}^{-1}\left[\mathbf{W} + (\mathbf{I} - \mathbf{W})\frac{1}{\mathbf{1}_K'\mathbf{\Sigma}\mathbf{1}_K}\mathbf{\Sigma}\right]\mathbf{1}_N \\
&= (\mathbf{I} - \lambda\mathbf{G})\mathbf{\Sigma}^{-1}\left[\frac{1}{\mathbf{1}_K'\mathbf{\Sigma}\mathbf{1}_K}\mathbf{\Sigma} + \mathbf{W}\left(\mathbf{I} - \frac{1}{\mathbf{1}_K'\mathbf{\Sigma}\mathbf{1}_K}\mathbf{\Sigma}\right)\right]\mathbf{1}_N.
\end{aligned}$$

Thus, we want to find a $\mathbf{w} = \mathbf{W}\mathbf{1}_N$, such that

$$\begin{aligned}
&\mathbf{Q} = \left[\frac{1}{\mathbf{1}_K'\mathbf{\Sigma}\mathbf{1}_K}\mathbf{\Sigma}\mathbf{1}_N + \mathbf{W}\left(\mathbf{I} - \frac{1}{\mathbf{1}_K'\mathbf{\Sigma}\mathbf{1}_K}\mathbf{\Sigma}\right)\mathbf{1}_N\right] \\
&\Leftrightarrow \left[\mathbf{Q} - \frac{1}{\mathbf{1}_K'\mathbf{\Sigma}\mathbf{1}_K}\mathbf{\Sigma}\mathbf{1}_N\right] = \left[\mathbf{W}\left(\mathbf{I} - \frac{1}{\mathbf{1}_K'\mathbf{\Sigma}\mathbf{1}_K}\mathbf{\Sigma}\right)\mathbf{1}_N\right].
\end{aligned}$$

Finally, pre-multiplying by $\left(\mathbf{I} - \frac{1}{\mathbf{1}_K'\mathbf{\Sigma}\mathbf{1}_K}\mathbf{\Sigma}\right)^{-1}$ on both sides of the equation above we obtain

$$\left(\mathbf{I} - \frac{1}{\mathbf{1}_K'\mathbf{\Sigma}\mathbf{1}_K}\mathbf{\Sigma}\right)^{-1}\left[\mathbf{Q} - \frac{\mathbf{\Sigma}}{\mathbf{1}_K'\mathbf{\Sigma}\mathbf{1}_K}\mathbf{1}_N\right] = \mathbf{W}\mathbf{1}_N = \mathbf{w}.$$

Notice that if there is a single production module then we have $\mathbf{M} = (1, 1, \ldots, 1)_{1\times N}$ and $\mathbf{Q} = \mathbf{1}_N$ which implies that $\mathbf{W} = \mathbf{I}$ giving the optimal provision of incentives under 1 module. Moreover, if there an N production modules then we have that $\mathbf{M} = \mathbf{I}$, $\mathbf{Q} = \frac{\mathbf{\Sigma}\mathbf{1}_N}{\mathbf{1}_N'\mathbf{\Sigma}\mathbf{1}_N}$, which

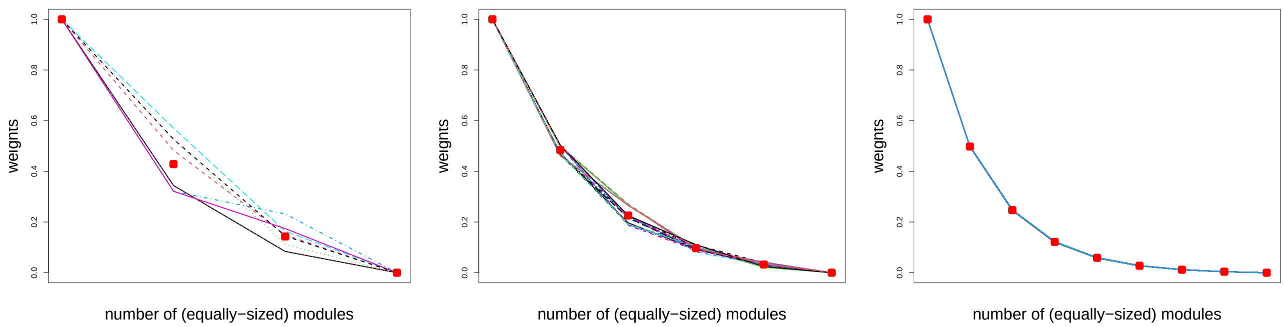


Figure 12: Each curve represents the weight w_i for each worker as we divide the workforce into more (and smaller) equally-sized modules. The red squares represent the value $\frac{\tilde{n}-1}{N-1}$ where $\tilde{n}$ is the module size. ***Panel A:*** *A random network with 8 nodes.* ***Panel B:*** *A random network with 32 nodes.* ***Panel C:*** *A random network with 256 nodes.*

implies that $\mathbf{W} = \mathbf{0}$ giving the optimal provision of incentives under N modules, as expected. ■

Figure 12 plots the weights for three simple networks with equally-sized modules. As we move to the right, the firm is split into more (and smaller) modules and the weights decrease from 1 to 0. Notice that, as per equation (34), not all workers in the same module have the same weight, but as the size of the network grows (moving across panels) the difference shrinks. We find that in the limit the weights only depend on the relative size of modules:

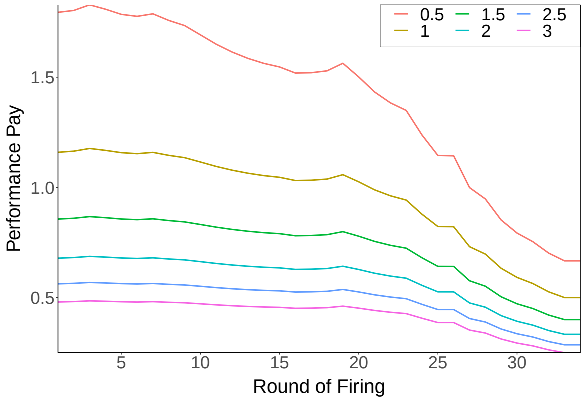
$$w_i \underset{n \to \infty}{\longrightarrow} \frac{N - n_{k(i)} \left(\sum_k \frac{1}{n_k} \right)}{(N-1)\, n_{k(i)} \left(\sum_k \frac{1}{n_k} \right)}.$$

Therefore, for a firm with equally-sized modules (i.e. $n_k = \tilde{n}$ for all $k \in K$) this implies that $w_i \rightarrow \frac{\tilde{n}-1}{N-1}$ for all $i \in N$, as in the (large) network in Panel C.

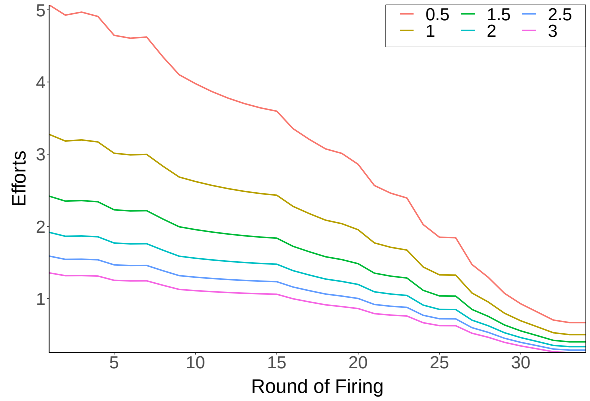
F Unemployment in the Karate Club Network

As mentioned at the end of Section 4.2, in the main body of the text we have focused on the direct, first-order effects of not hiring the least central workers. In fact, the firm may still generate unemployment even if it is possible to re-optimize contracts after the firing of the least central worker.

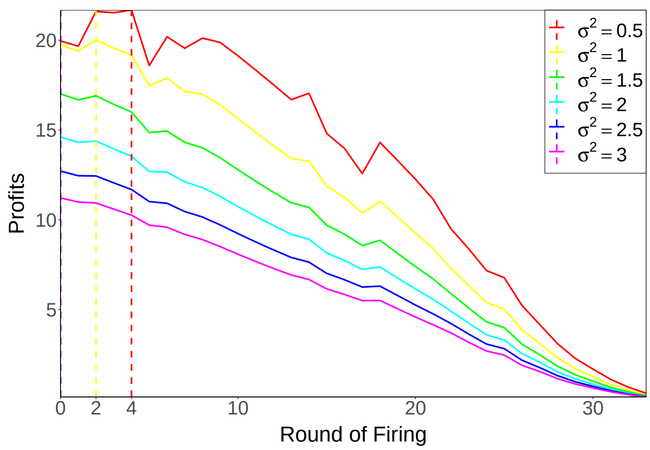
Recall that we consider the case in which all workers in the karate club network belong to the same group, i.e., $\mathbf{T} = (1, 1, \ldots, 1)_{1 \times N}$. This implies that the least central worker in the network determines the level of fix pay of all workers and allows well-connected workers to extract centrality rents from the firm.



(a) Performance Pay



(b) Equilibrium Efforts



(c) Profits after each round of firing the least central worker

Figure 13: Re-optimized performance pay, equilibrium efforts, and profits in the Karate Club Network after firing the least central workers (one at a time). Simulations are run for $\lambda = 0.07$ and $r = 1$.

Notice, first, as showed in Figures 13a and 13b for different fundamental risk levels, that firing the least central worker has very small effects on the performance-related compensation and the equilibrium levels of efforts of the remaining employed workers. Obviously, once many workers are fired, the loss in output is so large that performance pay, and efforts, decrease significantly. Lastly, Figure 13c shows the re-optimized profits of the karate club network in each round of firing the least central worker in the firm. For low levels of risk, the firm still generates unemployment even if it can re-optimize all contracts after the network has changed due to the firing of the least central worker in the firm.