DISCUSSION PAPER SERIES

DP16177
(v. 2022-12-16 23:11:15)

CONTRACTING IN PEER NETWORKS

Ron Kaniel and Peter DeMarzo

FINANCIAL ECONOMICS AND
ORGANIZATIONAL ECONOMICS

ISSN 0265-8003

CONTRACTING IN PEER NETWORKS

Ron Kaniel and Peter DeMarzo

Discussion Paper DP16177
First Published N/A
This Revision 21 May 2021

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

CONTRACTING IN PEER NETWORKS

Abstract

We consider multi-agent multi-firm contracting when agents benchmark their wages to their peers', using weights that vary within and across firms. When a single principal commits to a public contract, optimal contracts hedge relative wage risk without sacrificing efficiency. But compensation benchmarking undoes performance benchmarking, causing wages to load positively on peer output, and asymmetries in peer effects can be exploited to enhance profits. With multiple principals a "rat race" emerges: agents are more productive, with effort that can exceed the first-best, but higher wages reduce profits and undermine efficiency. Wage transparency and disclosure requirements exacerbate these effects.

JEL Classification: N/A

Keywords: N/A

Ron Kaniel - ron.kaniel@simon.rochester.edu
University Of Rochester and CEPR

Peter DeMarzo - pdemarzo@stanford.edu
Stanford

Contracting in Peer Networks

PETER M. DEMARZO and RON KANIEL*

This Revision: November 8, 2022

ABSTRACT

We consider multi-agent multi-firm contracting when agents benchmark their wages to those of their peers, using weights that vary within and across firms. When a single principal commits to a public contract, optimal contracts hedge relative wage risk without sacrificing efficiency. But compensation benchmarking undoes performance benchmarking, causing wages to load positively on peer output, and asymmetries in peer effects can be exploited to enhance profits. With multiple principals a "rat race" emerges: agents are more productive, with effort that can exceed the first-best, but higher wages reduce profits and undermine efficiency. Wage transparency and disclosure requirements exacerbate these effects.

* Peter DeMarzo (pdemarzo@stanford.edu) is at Stanford University; NBER. Ron Kaniel (ron.kaniel@simon.rochester.edu) is at the University of Rochester; FISF, Fudan; Reichman University; CEPR. The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC Grant Agreement no. [312842]. We are grateful to Shai Bernstein, Gian Luca Clementi, Florian Ederer, Willie Fuchs, Denis Gromb, Adriano Rampini, Jeff Zwiebel, and seminar and conference participants at Amherst, Baruch College, Boston University, Cheng Kong Graduate School of Business, Chinese University of Hong Kong, Drexel University, Emory, Federal Reserve Bank of New York, Frankfurt School of Finance and Management, Fudan, IDC Herzeliya, INSEAD, John Hopkins, London School of Economics, National University of Singapore, Pompeu Fabra and Barcelona GSE, Rutgers, Rice, Singapore Management University, Stanford, Stockholm School of Economics, Tel Aviv University, Tulane, University of Chicago, University of Gothenburg, University of Hong Kong, University of Houston, University of Lugano, University of Melbourne, University of Minnesota, University of Rome, University of Texas at Dallas, University of Toronto, Vienna University of Economics and Business, Washington University in St. Louis, Annual Meeting of the Western Finance Association, Utah Winter Finance Conference, Finance Theory Group conference, Luxembourg Asset Management Summit, and the Annual Meeting of the American Finance Association for helpful comments. Under the American Finance Association disclosure guidelines, we both have nothing to disclose. An earlier version of this paper was circulated under the title "Relative Pay for Non-Relative Performance: Keeping up with the Joneses with Optimal Contracts."

Contract and incentive theory has provided powerful insight into the optimal design of compensation schemes. Chief among them, for instance, is the idea that contracts should provide higher compensation when output suggests that the agent was more likely to have engaged in desired behavior. In particular, Holmstrom's (1992) Informativeness Principle states that any measure of performance that reveals information about the agent's effort should be included in the compensation contract. A prime example is the use of Relative Performance Evaluation (RPE), in which the agent's performance is measured relative to her peers in order to filter out common risk factors. In other words, optimal contracts should *not* "pay for luck" due to aggregate shocks, but only pay for indicators of individual performance. Yet despite this clear benefit, such performance benchmarking is observed less frequently than theory would predict.[1]

But while principals should care about relative performance, agents may also care about their relative wage. In this paper, we consider such preferences and explore their consequences. Specifically, we suppose agents have a "keeping up with the Joneses" component to their utility in which they compare their wage to a weighted average of the wages of their peers. We allow for a general network of peer relationships, where the strength of comparison may vary both within and across firms based on proximity, salience, or other factors. We characterize optimal contracts in this setting and show that many standard contracting results are overturned – agents may be *rewarded* for peer performance, and, despite seemingly weaker incentives, equilibrium effort may *exceed* the first best.

In particular, we show that when peer effects are strong enough, compensation benchmarking undoes performance benchmarking, leading to wage contracts that positively load on peer (or team) output. But although these contracts appear inefficient, we show that this effect on its own does not diminish incentives nor reduce welfare. Moreover, despite what appear to be "low power" incentives, when there are multiple principals, or if contracts are privately negotiated, a "rat race" ensues that leads to higher wages and effort that may even exceed the first best. And although rat-race effects increase productivity, wages rise even more, causing a decline in profits. Measures that increase external wage transparency (such as public disclosure requirements) are

[1] For empirical evidence of "pay for luck" in the context of CEO compensation, see for example Murphy (1985), Coughlan and Schmidt (1985), Antle and Smith (1986), Gibbons and Murphy (1990), Janakiraman, Lambert, and Larker (1992), Garen (1994), Aggarwal and Samwick (1999a,b), Murphy (1999), Frydman and Jenter (2010), and Jenter and Kanaan (2015).

likely to exacerbate these rat-race effects. Finally, if peer-effects are asymmetric, principals can exploit the asymmetry by reallocating effort to "less visible" agents. Though this reallocation inefficiently distorts production, incentive costs fall and, surprisingly, profits and welfare rise relative to when agents' preferences lack peer-effects.

Our model includes many agents who take hidden effort to produce output that is subject to both common and idiosyncratic shocks. Agents receive a compensation contract which specifies their wage as a function of their own output as well as the output of others. Agents are risk averse and have preferences that are increasing in both their own wage as well as the difference between their own wage and a weighted average benchmark of their peers' wages. We allow these weights to depend on a general peer-network structure so that, for example, agents may put more weight on close colleagues versus peers elsewhere in the firm (or in other firms). We then consider the sensitivity of the optimal contract to both the strength of agents' relative wage concerns, as well as the peer-network structure, and evaluate implications for welfare.

We begin in Section II by introducing a standard CARA-normal principal-agent model and analyzing the RPE benchmark in a multi-agent setting without peer effects. We show that, as expected, compensation in the optimal contract is based on a measure of the agent's relative performance; that is, compensation is positively related to the agent's own output and negatively related to the output of others, in proportions that depend on the correlation between agents' output. A key prediction of RPE is therefore that the incentive component of wages should be negatively correlated. Additional standard predictions are that equilibrium effort is reduced relative to the first best (to limit agents' risk) and that the optimal contract for each agent is independent of the contracts of others (and therefore, independent of how they are determined, whether they are disclosed or not, etc.).

In Section III, we consider optimal contracts with peer network effects. Because agents care about their compensation relative to their peers, the incentives for each agent will depend on both their own contract and the contracts offered to all other agents. Principals independently choose contracts for agents on their teams, which are composed of one or many agents. We assume (for now) that agents see the full set of contracts offered by their principal to their team before deciding to accept or reject their own contract.

As an initial baseline, we first consider a setting in which peer weights are symmetric (determined, for example, by a distance metric) and there is a single principal. We then demonstrate an important welfare equivalence result: the optimal contracts for a single principal will fully hedge peer effects and lead to the same productivity, profitability, and welfare as the RPE benchmark. Observed contracts, however, will have markedly different sensitivities than those predicted by RPE. In particular, to limit relative wage risk, contracts may *positively* load on peer output (contradicting RPE), implying that empirically observed deviations from RPE need not be inefficient.

Next, we consider the case with multiple principles independently setting contracts for distinct teams. In this case, there is an important "rat-race" externality across teams, as principals do not account for the impact of raising wages on their team on the welfare of agents on other teams. As a result, equilibrium wages and productivity are higher, but profitability and welfare are lower, than in the RPE or single principal benchmark. We show that the rat-race inefficiency increases with the number of principals, the strength of peer effects, and the importance of external (non-teammate) peers. In fact, the rat-race effect can be so extreme that equilibrium effort exceeds the first best, in stark contrast to standard agency models.

We then explore in greater detail the consequences of peer effects for wage correlations. We show that wage correlations are the same in both the single and multi-principal case and are therefore solely related to peer effects and independent of welfare. We show that the wage sensitivity to peer-performance increases monotonically with the strength of peer-effects, leading to positive wage correlation, consistent with empirical evidence on wage compression within firms.[2] We characterize specific cases, and show that wage contracts load more positively on "nearby" versus more distant peers. Thus, for example, if agents care more about peers within their own firm or team, we find that optimal wage contracts will always have a positive exposure to the relative performance of their team.[3]

---

[2] For example, Silva (2016) and Gartenburg and Wulf (2017) document wage convergence in multidivisional firms, which is heightened by geographic or social proximity. Shue (2013) shows similar wage convergence across executives who were former classmates. While inequity aversion is a possible explanation for wage compression (see Englmaier and Wambach (2010) and Koszegi (2014)), our analysis demonstrates that peer effects are sufficient.

[3] Ibert, Kaniel, Van Niuwerburgh, and Vestman (2018) find that fund managers elasticity of pay to fund family revenue, excluding revenue managed by the given manager, is comparable to that of manager revenue.

In Section IV, we allow for peer relationships to be asymmetric. In that case, some agents may be "more visible" than others in the sense that their wage is more salient to their peers. In that case, the principal can exploit this asymmetry and use the departure from RPE to *increase* profits. Specifically, it is optimal to reallocate effort, and therefore compensation, to less visible agents. Although this reallocation is technologically inefficient, it lowers the principal's expected cost of providing incentives sufficiently to raise total output and profits.

In Section V, we consider the consequences of alternative disclosure environments. In the standard RPE framework, wage and contract disclosure has no effect. Here, on the other hand, disclosure interacts with peer effects. For example, we show that if the principal can privately negotiate individual wage contracts, the ability to do so exacerbates the rat-race effect, raising equilibrium wages and lowering profits. Indeed, a single principal that cannot commit *not* to renegotiate may have worse outcomes than if all contracts were negotiated by independent principals.

We also consider the effect of greater wage transparency across teams or firms. Wage transparency is likely to make peer comparisons more salient, and thereby exacerbate the rat-race externalities across teams; as a result, we expect average wages to rise and their correlation to increase. We also show that when incentive contracts are disclosed externally, there is an additional incentive to increase the contract loading on external peers, but equilibrium effort is distorted downward (potentially below second best).

I. Related Literature

The key premise of our paper is that some agents have Keeping up with the Joneses (KUJ)-type preferences in which they care about their relative wage with respect to their network of peers. Our objective is to understand the implications for contracting of such preferences. Evidence supporting the presence of KUJ / relative wage-based preferences has accumulated from multiple disciplines and research designs. Early evidence was derived mostly from survey questions regarding happiness and satisfaction (for example: Luttmer 2005, Ferrer-i-Carbonell 2005).[4] More recently, researchers have utilized natural and laboratory experiments, as well as FMRI evidence supporting the underlying neuropsychological foundations, to bolster support for such preferences

[4] See Heffetz and Frank (2011) for a review.

(see for example Kuhn et.al. (2011), Card et.al. (2012), Miglietta (2014), Fliessbach et.al. (2007), and Dohmen et al. (2011)). In providing evidence for relative wage preferences the literature has considered a variety of different peer groups: co-workers (Clark and Oswald (1996)); neighbors (Luttmer (2005), and Kuhn et.al. (2011)); household members (Clark (1996)); siblings (Neumark and Postlewaite (1998)); caste members in India (Carlsson, Gupta, and Johansson-Stenman (2009)); and division managers within the same firm (Duchin, Goldberg, and Sosyura (2017)).

In the context of executive compensation, Bouwman (2013) finds that CEO pay is correlated with that of geographically-close CEOs, even if in differing industries or roles (e.g. sports stars), with stronger effects for connected CEOs (inferred from rolodex connections). After carefully controlling for a host of potentially confounding effects, she concludes that KUJ preferences is the most plausible explanation.[5] Shue (2013) demonstrates wage convergence across executives who were former HBS classmates, and shows the effect is stronger for section mates. Our modeling framework contributes to the literature by allowing for a flexible structure of peer preferences, with differential sensitivities to peers for a given agent and across agents. These heterogeneous peer effects implicitly aggregate a wide variety of factors differentially impacting the sensitivity to different agents: the size and complexity of their firms, commonality of backgrounds, geographical proximity, social interactions, person specific psychological traits, etc.

While Keeping/Catching up with the Joneses and habit formation preferences have been used in asset pricing applications starting with Abel (1990), they have received much less attention in explaining behavior in the corporate finance domain. Ederer and Patacconi (2010) introduce status considerations into a tournament setting analyzing implications for the provision of incentives. Goel and Thakor (2010) use envy-based preferences for managers to explain merger waves.[6] Dur and Glazer (2008) consider the optimal contract, with contractible effort, for an employee that is envious of his employer. Goel and Thakor (2005) consider within firm capital allocation decisions of division managers where each manager derives direct utility from wages, and in addition envies both the wages of other managers and their capital allocation as well. Their

[5] Bouwman (2013) rules out three key competing hypothesis. First, that geography may introduce commonalities in the performance-relevant characteristics of CEOs that firms in a given area emphasize in their selection of CEOs. Second, that firms may follow "leading" firms in the vicinity in setting CEO pay. Third, local labor market competition for CEOs.

[6] While closely related, envy refers to a disutility from having a lower wage than one's peer, whereas keeping up with the Joneses preferences also include a benefit from having a higher wage.

analysis focuses on induced capital distortions, ignoring the moral hazard and contracting considerations which are the focus of our analysis.

The central theme of our analysis is to underscore contracting implications of peer-dependent preferences. We contribute to the literature by showing that these preferences reverse central results from standard contract theory: Compensation can load positively rather than negatively on outputs of other agents; effort levels are too high rather than too low and can even exceed first best;[7] disclosure of contracts may reduce efficiency, instead of being neutral or better; and when peer effects are asymmetric, the principal's profits can be higher than when agents have standard preferences that do not depend on peers.

There is extensive empirical literature that has for the most part rejected the RPE hypothesis that CEO compensation should depend on relative performance, and so be negatively related to the performance of peers. Much of the evidence documents a *positive* relation, in direct contrast to the standard RPE prediction.[8] A few have noted that KUJ preferences can provide a resolution for these findings within the agency model, as these preferences lead to optimal contracts with a reduced magnitude of RPE and, when strong enough, even a positive dependence of pay on external performance measures (Fershtman, Hvide, and Weiss (2003), Itho (2004), Miglietta (2014), Bartling (2011), Liu and Sun (2016)). Bartling and von Siemens (2010) consider the impact of envy on contracts in a general moral hazard model when a principal hires two agents that are envious of each other. They argue that with risk-averse agents and without limited liability, envy can only increase the costs of providing incentives. The scope of their analysis is limited by the fact that they do not derive explicit optimal contracts. Bartling (2011) derives optimal contracts in a mean-variance framework with one principle and two agents, but emphasizes the case in which agents are inequity averse. In contrast to these papers, we consider a multi-agent, multi-principal framework within an arbitrary network of peer relationships and derive the implications for both optimal contracts and welfare. We show general conditions that allow a single principle to attain the second best, implying no loss in efficiency, and characterize the extent to which RPE

[7] The predicted link between peer effects and effort is consistent with evidence in Ghazala, and Iriberri (2010), which utilizes a natural experiment to show that adding to a students' report cards their average grade over all subjects, as well as the class average over all subjects and students led to an increase of 5% in students' grades.

[8] Examples include: Antle and Smith (1986), Barro and Barro (1990), Jensen and Murphy (1990), Janakiraman, Lambert and Lacker (1992), Hall and Liebman (1998), Joh (1999), Aggarwal and Samwick (1999a,b), Bertrand and Mullainathan (2001), and Garvey and Milbourn (2003).

predictions are reversed. Moreover, we demonstrate how the principle can exploit any asymmetry in agents' visibility within the network to raise profits and increase efficiency.

In addition to providing a more comprehensive welfare comparison, our analysis considers a market-wide equilibrium with multiple principals, employing potentially multiple agents, within a general network structure of potential peer relationships. This setting allows us to additionally analyze contracting externalities and rat-race effects across principals. Finally, we investigate the impact of the disclosure environment, both inside the firm and across firms, on contracts and efficiency. For example, we predict that compensation disclosure mandates will increase rat-race incentives and raise wages, consistent with the findings in Park, Nelson, and Huson (2001), Perry, and Zenner (2001), Schmidt (2012), Gipper (2021) and Mas (2019).

II. A Model of Peer-Contracting

A. Peers and Preferences

We consider a setting with a set $\mathcal{N}$ of $N = |\mathcal{N}| \geq 2$ identical agents. We make the standard assumption that the utility of each agent i is increasing in his own wage, w_i, and decreasing in his hidden choice of effort, a_i. We depart from the usual principal-agent framework, however, by assuming that agents also care about their wage relative to that of their peers. We model this preference by assuming agent i benchmarks his own wage against a weighted average of the wages of his peers,

$$w_{-i} \equiv \sum_{j \in \mathcal{N}} \mu_{ij} w_j \text{ with } \mu_{ij} \geq 0,\ \mu_{ii} < 1, \text{ and } \sum_{j \in \mathcal{N}} \mu_{ij} = 1. \quad (1)$$

The weights μ_{ij} capture peer network effects in which agents may care more about the wages of some peers versus others. We typically let $\mu_{ii} = 0$, so that the benchmark only includes the wages of others.[9] For simplicity, we generally assume the set of peers is irreducible (i.e. for any pair (i, j) there is a sequence connecting them with positive weights), though we will relax this assumption in some special cases.

[9] We allow $\mu_{ii} > 0$ for generality (for example if agents compare their wage to an aggregate index; see also footnote 10) but restrict $\mu_{ii} < 1$ to ensure some weight is put on peers.

Each agent i derives utility from both his absolute wage, w_i, and his relative wage, $w_i - w_{-i}$. We convert the agent's gain from his relative wage to consumption units according to

$$v_i \equiv w_i + \hat{\delta}\left(w_i - w_{-i}\right) = (1+\hat{\delta})w_i - \hat{\delta}w_{-i}, \text{ with } \hat{\delta} \in [0,\infty). \tag{2}$$

Here $\hat{\delta}$ captures the strength of agents' relative wage concerns, with $\hat{\delta} = 0$ corresponding to the standard model without peer effects ($v_i = w_i$), whereas peer effects dominate as $\hat{\delta}$ becomes large.[10] Implicit in this formulation is the standard assumption that agents care about their relative *realized* wage, rather than their ex-ante expected wage, consistent with the idea that these preferences arise, for example, from a preference for social status (which is linked to actual wealth or consumption).

Next, to avoid wealth effects and for tractability, we assume agents have CARA utility $u(c_i) = -\exp(-\lambda c_i)$, where c_i is the "adjusted consumption" level

$$c_i \equiv v_i - \Psi(a_i), \tag{3}$$

and Ψ measures the cost of effort, which we assume for simplicity is quadratic: $\Psi(a) = \frac{1}{2}ka^2$.

This specification of the agent's utility incorporates several properties that are both natural and important given our goal of comparing welfare and contract design for different levels of relative wage concerns $\hat{\delta}$. First, using CARA utility and keeping v_i *linear in wages* implies that we are not changing the concavity of the utility function. Hence, the agent's risk aversion, which determines the magnitude of the agency friction, is independent of $\hat{\delta}$. Second, because the *total weight on wages sums to 1*, the utility gain from increasing all wages equally does not depend on $\hat{\delta}$. These features are necessary for us to make meaningful welfare comparisons across settings with differing levels of peer sensitivity.[11] Our specification of peer effects in (2) is fully general given these requirements.

---

[10] While many of our results apply even if $\hat{\delta} < 0$ (altruistic agents), we do not analyze that case here. Note also that $\mu_{ii} > 0$ is equivalent to reducing the weight agent *i* places on others to $\hat{\delta}_i = \hat{\delta}\left(1 - \mu_{ii}\right)$. Thus, by adjusting μ_{ii}, our formulation allows us to consider settings in which agents have heterogeneous peer sensitivities in the range $(0, \hat{\delta})$.

[11] For example, if we had defined v_i so that the sum of the weights on all wages decreased (or increased) with $\hat{\delta}$, then the aggregate cost of compensating agents to attain a given level of utility would necessarily increase (decrease), leading to an exogenous impact on the aggregate welfare that is attainable.

So far, we have not specified how the peer weights in (1) are determined. We will describe specific cases in applications later. Typically, we will interpret the weights as a measure of closeness or distance; e.g., agents are likely to care more about their wage in relation to peers in the same office, firm, or industry. Consistent with this interpretation, we will for now restrict the weights to be pairwise symmetric (we will relax this assumption in Section IV):

$$\mu_{ij} = \mu_{ji} \text{ for all } i, j. \tag{4}$$

Finally, purely for notational convenience and without loss of generality, we renormalize our peer sensitivity parameter to the unit interval as follows,

$$\delta \equiv \frac{\hat{\delta}}{1+\hat{\delta}} \in [0,1). \tag{5}$$

This specification will simplify many of our analytical results, and so for the remainder of the paper we will refer to δ as the agent's peer sensitivity, with $\delta = 0$ (no peer effects) or $\delta = 1$ (peer effects dominate) as corresponding to the two extremes cases.

B. Production and Wages

The production technology has additive shocks, so that the output q_i of agent i is equal to a known constant $q_0 > 0$ plus effort plus noise:[12]

$$q_i \equiv q_0 + a_i + \epsilon_i. \tag{6}$$

The random shocks ϵ_i are joint normal with mean zero and variance σ_ϵ^2, and have a pairwise correlation of $\rho \in [0,1)$.

Given the production technology and preferences, we define the first-best effort level a^{fb} to maximize net output, $a_i - \Psi(a_i)$, which is solved by $a^{fb} \equiv k^{-1}$.[13] Effort choices are hidden, however, and therefore appropriate compensation contracts are needed to motivate the agent. Even absent relative wage concerns, the correlation between output shocks implies that optimal contracts

[12] We can think of q_0 as corresponding to some component of output that is not subject to moral hazard. We assume it is large enough such that hiring an agent is strictly profitable.
[13] Here we have defined the first best in terms of productive efficiency. We will see in Section IV that with asymmetric peer effects, distortions from productive efficiency may be optimal even absent agency concerns.

will depend on both own output and peer output. We restrict attention to linear compensation contracts y_i of the form:[14]

$$w_i = y_{i0}q_0 + y_{ii}q_i + \sum_{j\in\mathcal{N}_{-i}} y_{ij}q_j = y_i \cdot q \tag{7}$$

where $\mathcal{N}_{-i} = \mathcal{N} \setminus \{i\}$ is the set of i's peers, y_{i0} determines the constant component of the wage, y_{ii} is the sensitivity of the wage to i's own output, and y_{ij} is the sensitivity of i's wage to the output of peer j.[15] Of particular interest is the contract's "relative sensitivity" to peer output, which we define by:

$$\phi_i \equiv \sum_{j\in\mathcal{N}_{-i}} \frac{y_{ij}}{y_{ii}}. \tag{8}$$

To evaluate payoffs, note that with normally distributed shocks, linear contracts, and CARA utility, agents will have mean-variance preferences. Therefore, given normally distributed consumption c, we can evaluate the agent's utility in terms of the corresponding certainty equivalent consumption level

$$u^{-1}\left(E\left[u(c)\right]\right) = E[c] - \tfrac{1}{2}\lambda Var(c). \tag{9}$$

Finally, each agent has an outside option with certainty equivalent c_0, which we can interpret as the agent choosing to leave the firm/industry (thus having a new set of peers).[16] We view this outside option as determining participation, whilst once on the job, peer comparisons determine incentives.

C. Teams

The contract for each agent is set by a risk neutral principal. Each principal manages a distinct set of agents, which we refer to as a team. Teams may correspond to different firms, or

[14] Goukasian and Wan (2010) demonstrate optimality of linear contracts with one principal and two peers in a continuous time context as in Holmstrom and Milgrom (1987). Similar results are likely to hold in our more general framework, but we have not pursued that here.

[15] It is not essential that q_0 in (7) is the same as the constant component in (6), but because the constant in (7) is arbitrary (as long as it is nonzero), we use the same term for notational convenience.

[16] Said another way, we interpret the outside option as "leaving" the peer set $\mathcal{N}$, thereby ruling out peer effects from within the firm altering the attractiveness of the outside option.

different divisions or groups within a firm. The set of teams is a partition of the set $\mathcal{N}$ of agents. The objective of the principal for team $I \subseteq \mathcal{N}$ is to offer a set of contracts to maximize the team's expected output net of wages paid:

$$\pi_I \equiv E\left[\sum_{i\in I}(q_i - w_i)\right]. \tag{10}$$

The principal must respect each agent's incentive compatibility constraint, which given (9) can be written as

$$a_i \in \arg\max\ E[c_i], \tag{11}$$

as well as the participation constraint,

$$E[c_i] - \tfrac{1}{2}\lambda Var(c_i) \geq c_0. \tag{12}$$

We assume c_0 is low enough (relative to q_0) so that hiring the agent is profitable for the principal.

The maximum team size is the total population; in this case one principal manages all peers (e.g. workers within a single firm). At the other extreme, each agent may have his own principal (for example, CEOs with distinct boards). In general, an agent will have peers both within and outside his team. We refer to agents on the same team as agent i as teammates or "internal" peers, given by the set I_{-i}, whereas agents on different teams are "external" peers, denoted by $\mathcal{N}_{-I}$. We denote the total number of peers and the number of people on a team by

$$n \equiv |\mathcal{N}_{-i}| = N-1 \text{ and } N_I \equiv |I| \in [1, N]. \tag{13}$$

D. Contracting

We assume each principal chooses contracts for their team taking as given the contracts offered and actions taken by all other teams. That is, the contracts proposed by the principal for team *I* are visible to all members of the team, but not to principals or agents outside the team. (We will consider other disclosure assumptions in Section V.) The timing of the contracting problem faced by team *I* is shown in Figure 1.[17]

[17] Note that this timing presumes that contract details are visible within a team, but only ex-post output and wages are visible across teams. Wages and output are likely to be more easily observed even externally due to standard reporting requirements as well as consumption and investment effects. That said, we explore the consequences of alternative information structures in Section V.

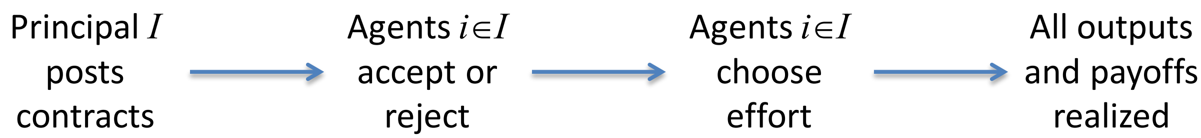

Figure 1: The Contracting Game for a Single Principal
The principal-agent problem for team I (taking all other contracts and actions as given).

Figure 2 depicts the different channels that may affect the agent's utility. The solid blue lines correspond to the standard principal-agent problem in which agents' output is uncorrelated. In that setting, the optimal wage contract depends only on the agent's own output. When there are multiple agents with correlated output shocks a new information channel is introduced, depicted by the dashed green lines. In this case, the principal can use the information in the output of others to filter out any common component of the agent's productivity shock, and relative performance evaluation (RPE) becomes optimal.

Finally, the dotted orange lines represent the new channels introduced when peer effects cause agents to also care about their relative wage. Each agent is now exposed to both the risk and the level of their wage differential with their peers. This interaction introduces two new aspects to the optimal contract. First, the principal may adjust the contract to hedge the agent against peer wage risk. Second, the principal may distort the agent's effort in order to affect the level of the average peer wage.

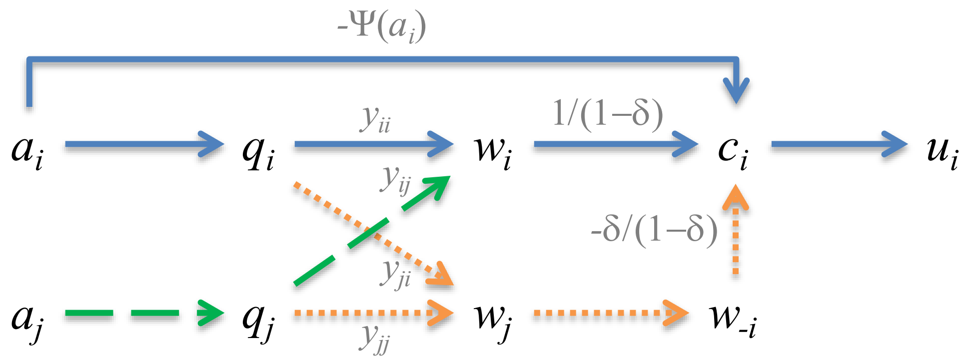

Figure 2: Alternative Channels by which Effort and Output Impact Utility
Standard principal-agent model is shown in solid blue. The information channel for relative performance evaluation is shown in dashed green. The dotted orange channel represents relative wage concerns.

E. The RPE Benchmark

As an important benchmark, before we consider peer effects we first solve for the optimal contract in the benchmark case without peer effects and obtain the standard prediction that optimal contracts should utilize relative performance evaluation (RPE): each agent's wage should have a negative sensitivity to the output of his peers.

Consider the standard setting with $\delta = 0$ so that peer effects are absent. In that case, the contracting problem for agent i does not depend on the wage contracts of other agents. The only interdependence between agents arises from the positive correlation between agents' shocks. As a result, it is optimal for the principal to reduce the agent's risk by "filtering out" the common component of output when assessing the agent's performance.

In particular, the contract can reward the agent for his performance relative to a benchmark of output based on the performance of his peers. The efficient benchmark will minimize the agent's exposure to the common risk factor. This benchmark is the multiple $\overline{\theta}$ of the average peer performance $q_{-i} = \frac{1}{n}\sum_{j \in \mathcal{N}_{-i}} q_j$ that minimizes the residual variance:

$$\overline{\theta} \equiv \arg\min_{\theta} Var(q_i - \theta q_{-i}) = \frac{Cov(q_i, q_{-i})}{Var(q_{-i})} = \frac{\rho}{\rho + \frac{1}{n}(1-\rho)} \in [0,1). \tag{14}$$

Given this benchmark, the residual uncertainty is given by

$$\bar{\sigma}^2 \equiv Var(q_i - \bar{\theta}\, q_{-i}) = \left(1 - \rho\bar{\theta}\right)\sigma_\epsilon^2 > 0\,. \tag{15}$$

The volatility $\bar{\sigma}$ of the optimal signal available to the principal determines the magnitude of the agency problem, as it is costly for the agent to bear this risk. Together the parameters $(\bar{\theta}, \bar{\sigma})$ summarize the *information* externalities in our model.

LEMMA 1 (RELATIVE PERFORMANCE EVALUATION). *Absent relative-wage concerns* $(\delta = 0)$*, the optimal contract for each agent maximizes expected output net of the cost of effort and risk-bearing, subject to incentive compatibility:*

$$\max_{y_{i\cdot}} \; a_i - \Psi(a_i) - \tfrac{1}{2}\lambda Var(c_i) \quad s.t. \;\; a_i = y_{ii}/k\,. \tag{16}$$

Optimal contracts and actions are given by

$$y_{ii}^* = \frac{1}{1 + k\lambda\bar{\sigma}^2}\,,\; y_{ij}^* = -\tfrac{1}{n}\bar{\theta} y_{ii}^*, \text{ and } a_i^* = y_{ii}^* / k = y_{ii}^* a^{fb}\,. \tag{17}$$

The constant term y_{i0}^* *is set so that the expected wage is* $E\left[w_i\right] = c_0 + \tfrac{1}{2}a_i^*$*, and the expected profit per agent is* $\pi_i^* = q_0 - c_0 + \tfrac{1}{2}a_i^*$.

PROOF: See Appendix. ■

The intuition for this result is as follows. Equation (16) states that the optimal contract maximizes the expected output net of the cost of effort and risk-bearing. The wage parameter y_{ii} determines the agent's incentives for effort, and so the (IC) constraint (11) implies that effort solves

$$\max_{a_i} \; y_{ii} a_i - \Psi(a_i)\,, \tag{18}$$

which is equivalent to $a_i = y_{ii} / k$. Finally, the sensitivity to peer output y_{ij} is used to minimize the agent's exposure to systematic risk, and $y_{ii} < 1$ reduces his exposure to project specific risk.

LEMMA 1 provides the standard contracting result that effort is attenuated relative to the first best ($a_i^* = y_{ii}^* a^{fb} < a^{fb}$) because of the cost of imposing risk on the agent. Effort decreases with the cost of effort (k), the agent's risk aversion (λ), and the residual risk ($\bar{\sigma}$). It also predicts that the relative sensitivity of the agent's wage to peer output versus his own output will correspond to the optimal benchmark in (14):

$$\phi_i^* \equiv \sum_{j \in \mathcal{N}_{-i}} \frac{y_{ij}^*}{y_{ii}^*} = -\overline{\theta}\,. \tag{19}$$

Equation (19) forms the basis for standard tests of RPE in the empirical literature, which generally conclude that compensation tends to be much less negatively correlated with peer performance than is predicted by an optimal contracting framework. Indeed, many studies often find the opposite sign – pay is *positively* related to aggregate performance. A key goal of this paper is to understand how the optimal performance benchmark changes when agents have relative wage concerns, and how this change affects wages, productivity, and profits.

Remark. In the subsequent analysis, it will sometimes be useful to consider comparative statics in the size of the peer population, n, while keeping the fundamentals of the agency problem unchanged. To do so, we can vary n while holding fixed $(\overline{\theta}, \overline{\sigma})$, which determine the information externalities, and let (ρ, σ_ϵ) adjust with n according to (14) and (15). As shown in **Lemma 1**, in the RPE setting the optimal actions a_i, relative contract sensitivities ϕ_i, and profits π_I (and thus total welfare), are all independent of n given $(\overline{\theta}, \overline{\sigma})$.

III. Peer-Contracting Equilibrium

Now we consider the case in which agents have peer-dependent preferences ($\delta > 0$). Due to relative wage concerns, the agent is exposed to the output of all other agents both directly through his own contract as well as indirectly through his concern for other's wages. As illustrated in Figure 2, and using (1), (2), (5), and (7), given a set of contracts with parameters y_{ij}, agent i's total exposure to output q_j is given by[18]

$$\beta_{ij} \equiv \frac{\partial c_i}{\partial q_j} = y_{ij} + \hat{\delta}\left(y_{ij} - \sum\nolimits_k \mu_{ik} y_{kj}\right) = \frac{y_{ij} - \delta \sum\nolimits_k \mu_{ik} y_{kj}}{1-\delta}. \tag{20}$$

[18] We write Σ_k as a shorthand for the summation is over all agents $k \in \mathcal{N}$.

As equation (20) shows, agent i has both a direct exposure (through his own contract) and an indirect exposure (through his concern for others' wages) to j's output, and hence there is now an interaction between the contracts offered to different agents. For example, the (IC) constraint (11) becomes

$$\max_{a_i} \beta_{ii} a_i - \Psi(a_i) \;\Rightarrow\; a_i = \beta_{ii} / k \,. \tag{21}$$

Therefore, the agent's effort choice depends upon the contracts given to all agents. Similarly, the agent's risk is given by

$$Var(c_i) = \left[(1-\rho)\sum_j \beta_{ij}^2 + \rho\left(\sum_j \beta_{ij}\right)^2 \right] \sigma_\epsilon^2 , \tag{22}$$

which again is a function of all contracts offered.

In the remainder of this section, we evaluate the consequences of this interdependency on equilibrium contracts. As we will see, we can restate the contracting problem for the principal in terms of the exposures β in place of y, and then solve for the contract parameters that induce these exposures. First, we evaluate the consequences for efficiency, and show that a single principal can design contracts that hedge peer effects without sacrificing efficiency. When there are multiple principals, however, competition between them leads to a rat-race effect in which agents work harder than in the RPE benchmark – and potentially harder than first best -- but firm profits and welfare are lower. We also look at the optimal wage sensitivities and correlations. There we show that, independent of welfare, peer effects lead to wage compression and, when sufficiently strong, cause wages to load positively on peer output.

A. Efficiency: The Single Team Case

It is useful to begin by considering the case with a single team, or principal ($I = \mathcal{N}$). In that case the principal has full control of all contracts, and there is no issue of coordination. We can also view the single team case as the outcome a social planner would achieve given the same information constraints.

The following result demonstrates that we can restate the contracting problem for the principal in terms of the implied exposures β in place of direct contract sensitivities y. More strikingly, despite the presence of peer effects, real outcomes and welfare are identical to the RPE setting.

PROPOSITION I (SINGLE PRINCIPAL: WELFARE EQUIVALENCE). *Suppose agents have relative wage concerns* $(\delta \geq 0)$ *and there is a single principal* $(I = \mathcal{N})$. *Then the principal will choose exposures* β *to solve (16) as in the RPE case, and hence effort, expected output and profits are equal to the RPE outcome and independent of* δ. *The optimal contracts* y^S *are a transformation of the RPE contracts* y^* *given by*

$$y^S = \Delta^{-1} y^*, \tag{23}$$

where $y^S, y^* \in \mathcal{R}^{N\times(N+1)}$ *are matrices with each row* i *representing the contract for agent* i, *and with* $\Delta \equiv \frac{1}{1-\delta}\left[\mathbf{I} - \delta\mu\right]$ *where* $\mathbf{I}$ *is the* $N \times N$ *identity matrix and* μ *is the matrix of weights* μ_{ij}.

PROOF: See Appendix. ■

The key intuition behind **PROPOSITION I** is that a single principal can effectively undo the peer effects in preferences via the wage contracts that are offered. Specifically, note that we can rewrite (20) in matrix form as $\beta = \Delta y$, and hence we can invert this relationship in (23) to obtain the RPE exposures. As a result, effort and efficiency are unaffected by the strength δ of peer effects. Observed contracts and wages, however, will be affected by δ according to (23). Indeed, in this case the principal offers a wage that combines the RPE wage with a hedge position that insures against peer effects:

COROLLARY A (OPTIMAL WAGES). *Let* w^* *be the wages paid with RPE contracts and let* w^S *be the optimal wages with a single team. Then the optimal single team contract hedges the agents' peer exposures:*

$$w_i^S = w_i^* - \delta(w_i^* - w_{-i}^S) = (1-\delta)w_i^* + \delta w_{-i}^S, \tag{24}$$

and $\sum_i w_i^S = \sum_i w_{-i}^S = \sum_i w_i^*$.

PROOF: See Appendix. ■

As the result shows, because the hedge positions aggregate to zero ($\sum_i w_i^* - w_{-i}^S = 0$), they can be implemented by the single principal at no cost.[19] The result implies that we can decompose optimal contracts into the RPE contract plus a hedging contract. For a concrete example, we calculate below the optimal contract parameters in the case of equal weights.

EXAMPLE: EQUAL WEIGHTS. *Suppose* $\mu_{ij} = 1/n$ *for all* $i \neq j$. *Then*

$$y_{ii}^S = y_{ii}^* \left[1 - \delta \left(\frac{n + \bar{\theta}}{n + \delta} \right) \right], \text{ and } y_{ij}^S = y_{ii}^* \left(\frac{\delta - \bar{\theta}}{n + \delta} \right) = y_{ij}^* + y_{ii}^* \left(\frac{\delta}{n} \right) \left(\frac{n + \bar{\theta}}{n + \delta} \right). \tag{25}$$

PROOF: See Appendix. ■

We will explore further the qualitative impacts of peer effects on optimal contracts in Section C, but first we consider the efficiency implications of having multiple teams.

B. Inefficiency: Multi-Team Rat Race

When there is more than one principal or team $(I \neq \mathcal{N})$, each principal must choose the contract for agents on her team while taking the contracts and actions of agents on other teams as given. In this case, two important externalities arise relative to the single principal case.

First, if the principal raises the incentives, and thereby the expected wage, of an agent on her team, she must also raise the wages of others on her team to compensate for the higher expected level and volatility of the peer benchmark. But if her agents also have external peers on other teams, her impact on the peer benchmark is dampened, lowering the total cost of compensation for the principal. As a result, a principal whose agents put weight on external peers is willing to increase incentives relative to the single principal case.

Second, by changing incentives and therefore the expected output of her own team, principal I can influence the realized wages of agents on other teams whose contracts put weight on the output of team I. Then, by acting to lower the expected wage of other teams, the principal

[19]While we do not pursue it here, we can also solve for w^S iteratively as the fixed point of the mapping defined by (24).

can raise the relative wage of agents on her own team. This negative externality provides an additional incentive to distort contracts.[20]

These externalities imply that we should expect lower aggregate welfare in the multi-team setting. To evaluate their quantitative effect, we first consider the optimal contract for an individual team taking the contracts and action choices of other teams as given. To simplify the analysis and notation, we impose an additional restriction that all agents on team I put the same total weight μ_I on the wages of their teammates, and thus put total weight $1-\mu_I$ on non-teammates:

$$\sum_{k\in I}\mu_{ik}=\mu_I \text{ for all } i\in I. \tag{26}$$

LEMMA 2 (TEAM CONTRACT). *Given the contracts and actions of other teams, the optimal multi-team contract y_i^M for agent i on team I has rescaled exposures $\beta_{ij}^M=(1+\alpha_i)y_{ij}^*$ and effort $a_i^M=(1+\alpha_i)a_i^*$ relative to the RPE outcome, where the scaling factor is given by*

$$\alpha_i\equiv\frac{\delta}{1-\delta}\left(1-\mu_I-\sum_{\hat{i}\in I,j\notin I}\mu_{\hat{i}j}y_{ji}\right). \tag{27}$$

PROOF: See Appendix. ■

The proof of **LEMMA 2** relies on the ability of the principal for team I to implement any set of exposures $\left(\beta_{ij}\right)_{i\in I}$ for her team members with an appropriate set $\left(y_{ij}\right)_{i\in I}$ of contracts that will depend on the contracts chosen by all other teams. As a result, it is optimal for the principal to choose the relative exposures for agent i, β_{ij}/β_{ik}, to minimize risk as in the RPE case.

The effort choice, however, is distorted relative to the RPE or single principal setting due to the cross-team externalities described previously. This distortion is given by the factor α_i defined in (27). Note that if there are no peer effects ($\delta=0$) or no *external* peers ($\mu_I=1$, and therefore $\mu_{ij}=0$ for $j\notin I$), then $\alpha_i=0$ and we achieve the same efficiency as the RPE

[20] In this case, the direction of the distortion is ambiguous, as it depends on the sign of the parameters of external team contracts (to be determined in equilibrium in *PROPOSITION II*).

benchmark. But if $\delta > 0$ and $\mu_I < 1$, the principal's effective cost of compensation is reduced, which raises the optimal effort choice. Second, if $y_{ji} \neq 0$, the principal can change agent j's expected wage by distorting agent i's effort. Then, by lowering j's expected wage, the principal lowers the cost of compensating any agent $\hat{i} \in I$ with $\mu_{\hat{i}j} > 0$. The definition of α_i aggregates these effects.

While **Lemma 2** reveals the potential for effort to be distorted when there are multiple teams, the direction of the distortion is still unclear as it depends upon the equilibrium contracts used by all teams. That is, the above result provides the optimal response for one team given the contracts of others. In equilibrium, all teams choose optimal responses to each other. We evaluate this equilibrium next and demonstrate that, in fact, the distortion leads to higher effort than the second-best setting, and, for δ close to 1, it will even exceed the first-best effort level.

For tractability, we make additional symmetry assumptions to compute the equilibrium. First, each team has the same size:

$$N_I = N_{\hat{I}} \text{ for all teams } I, \hat{I}. \tag{28}$$

Second, agents place the same weight on peers who are external to their team:

$$\mu_{ij} = \mu_e = \frac{1-\mu_I}{N-N_I}, \text{ for all } i \in I \text{ and } j \notin I. \tag{29}$$

Note that we allow the internal weights within teams to be different, though the total internal weight μ_I remains constant. These conditions are sufficient to imply that the equilibrium effort distortions will be symmetric (even though contracts need not be). Our next result characterizes this outcome and demonstrates a rat-race effect in which the effort of each agent is distorted upward, reducing overall efficiency.

Proposition II (Symmetric Team Equilibrium: Rat Race). *In the multi-team equilibrium, optimal contracts* $\left\{y_{ij}^M\right\}_{j>0}$ *and actions* a_i^M *are scaled versions of the single-team contract,*

$$y_{ij}^M = (1+\alpha) y_{ij}^S \text{ and } a_i^M = (1+\alpha) a_i^*, \tag{30}$$

where the scaling factor for all teams is

$$1+\alpha = \frac{1}{1-\delta\left(\frac{1-\mu_I}{1-\delta\mu_I}\right)\left(1-N_I y_e^S\right)} \geq 1. \tag{31}$$

The contract weights y^S *follow from (23), with the weight on external peers given by*

$$y_e^S = \frac{\delta\mu_e\left(1-\overline{\theta}\right)-\frac{1}{n}\overline{\theta}\left(1-\delta\right)}{1-\delta+N\delta\mu_e} y_{ii}^* < \tfrac{1}{N}. \tag{32}$$

If $\delta > 0$ *and* $\mu_I < 1$, *then* $\alpha > 0$. *Expected wages and profits per agent are given by*

$$E\left[w_i^M\right] = c_0 + \tfrac{1}{2}a_i^*\left(1+\alpha\right)^2 \text{ and } \pi_i^M = \pi_i^* - \tfrac{1}{2}a_i^*\alpha^2. \tag{33}$$

Proof: See Appendix. ■

Proposition II establishes that when principals compete with each other, they neglect the peer externality on other teams and thus provide inefficiently high-powered incentives, with contract terms and effort scaled by the factor $(1+\alpha)$. Agent's work harder, but the increase in the expected wage that is required more than offsets the increase in productivity; hence expected profits are lower than in the RPE or single-team benchmark. The next result demonstrates that these distortions increase with the strength of external peer-effects – due to a higher number or higher weight on external peers – and can cause effort to exceed even the first-best level.

Proposition III (Rat-race Monotonicity). *For a given population* N *with multiple teams and positive external weights* $\mu_e > 0$, *the rat-race distortion* α *increases with the number of teams* N/N_I, *the weight on external peers* μ_e, *and the strength* δ *of peer effects. As* $\delta \to 1$ *and peer effects dominate, equilibrium effort increases and ultimately exceeds the first best:*

$$a^M = (1+\alpha)a^* \to \left(\frac{N/N_I}{1-\overline{\theta}}\right)a^{fb}. \tag{34}$$

Proof: See Appendix. ■

The following result provides a simple characterization of the magnitude of the rat-race effect when the number of teams grows large, and shows that the strength of the effect increases with both δ and the total weight $1-\mu_I$ put on external peers:

Corollary B (Many Teams). *If we increase the number of teams while holding (N_I, μ_I) fixed, then as $N \to \infty$ the rat-race distortion converges to*

$$\alpha \to \frac{\delta}{1-\delta}\left(1-\mu_I\right). \tag{35}$$

When N is large, α increases with N if and only if $\dfrac{\delta-\bar{\theta}}{\delta\left(1-\bar{\theta}\right)} > \mu_I$.

Proof: See Appendix. ■

Finally, in the extreme case of single-agent teams, we have the following characterization, (which, by **Proposition III**, provides an upper bound on the rat-race effect for any given population size):

Example: Single-Agent Teams. *If $N_I = 1$ and $\mu_e = \frac{1}{n}$, the rat-race factor is*

$$1+\alpha = \frac{1}{1-\delta+\delta y_e^S} = \frac{1}{1-\delta+\delta\left(\dfrac{\delta-\bar{\theta}}{n+\delta}\right)y_{ii}^*}. \tag{36}$$

While our prior results have focused on symmetrically sized teams, we can also apply **Lemma 2** to evaluate equilibria when teams are asymmetric. The following result considers the case with two teams of differing size. Because each principal does not take into account the effect of its wage contracts on the other team, the rat-race externality is largest for the smaller team. As a result, agents on the smaller team will work harder, but despite having higher productivity, its per capita expected profits are lower.

PROPOSITION IV (ASYMMETRIC TEAM SIZE). *Suppose there are two teams I and J of size $1 < N_I < N_J$, with symmetric external weights $\mu_{ij} = \mu_e < 1/N$ for $i \in I$ and $j \in J$. Then for $\delta \in (0, 0.8)$, the rat-race effect is larger for the smaller team, $\alpha_i > |\alpha_j|$, and its expected profits are lower.*

PROOF: See Appendix. ■

The intuition for this result is that the rat-race externality is strongest when the size of the team is small. If the size of the team is very large (close to N) then the rat-race externality is weak. Moreover, if $\delta > 0.5$ and y_e^S is also large, it is possible that $\alpha_J < 0$; in that case the principal of the large team reduces effort so as not to indirectly benefit external agents and raise her own cost of compensation. Finally, for δ very close to 1, the rat-race distortion can become so large that at least one team becomes unprofitable (see (33)).[21]

C. Peer Sensitivity and Wage Compression

The prior results have examined the welfare consequences of peer effects. In this section we consider their effects on observed contracts and wages. A key insight behind **PROPOSITION I** and **PROPOSITION II** is that each principal can effectively undo the peer effects in preferences via the wage contracts that are offered. By hedging the peer preferences via the contract, the principal can provide a relative exposure β_i to agent i that exposes him to the same level of risk $\bar{\sigma}$ per incremental unit of output as in the RPE contract. While the level of output is impacted by the rat-race effect, optimal risk sharing is maintained.

But while the implied total exposures match the RPE outcome, the *observed* contracts y_i will be affected by the strength δ of relative wage concerns. As δ increases and peer effects become more dominant, contracts will be distorted more and more in order to hedge agents against relative wage shocks. In this section we explore the consequences of these distortions and their implications for wage sensitivity and correlation.

[21] In that case, we may expect principals to exit and their agents to join existing teams, creating an endogenous restriction on the number and size of teams that can participate in an industry before the rat-race effect makes it unprofitable. While we do not pursue it here, such dynamics could provide insight into equilibrium market structure.

Recall from **Proposition II** that when there are multiple teams the only change to the optimal contract loadings is a rescaling by the rat-race factor $(1+\alpha)$. Hence, *relative* contract sensitivities are unchanged. This result allows us to analyze the impact of peer-effects on optimal contracts in both settings together:

Proposition V (Contract Sensitivities). *In the single team setting of* **Proposition I**, *contract loadings* y^S *are compressed relative to the RPE setting:*

$$y_{ii}^* \geq y_{ii}^S \geq y_{ij}^S \geq y_{ij}^* = -\tfrac{1}{n}\overline{\theta} y_{ii}^*, \tag{37}$$

where the inequalities are strict if $\delta > 0$. *With either single or multiple teams (as in* **Proposition II**), *agent* i *'s wage sensitivity to peer output is the same, and depends on* Δ_{ii}^{-1} *according to*

$$\phi_i = \phi_i^M = \phi_i^S \equiv \sum_{j \in \mathcal{N}_{-i}} \frac{y_{ij}^S}{y_{ii}^S} = \frac{1-\overline{\theta}}{\left(1+\tfrac{1}{n}\overline{\theta}\right)\Delta_{ii}^{-1} - \tfrac{1}{n}\overline{\theta}} - 1. \tag{38}$$

When there are no peer effects ($\delta = 0$*),* $\phi_i = -\overline{\theta}$. *But as* δ *increases,* ϕ_i *strictly increases, and as* $\delta \to 1$,

$$\frac{y_{ij}^S}{y_{ii}^S} \to 1 \text{ for all } j\text{, and therefore } \phi_i \to n. \tag{39}$$

Proof: See Appendix. ■

Proposition V demonstrates that peer effects compress the contract loadings and increase the relative importance of peer output in the wage determination. Equation (38) shows that the overall peer sensitivity is characterized by the diagonal term Δ_{ii}^{-1} of the inverted peer-weighting matrix. Most strikingly, as $\delta \to 1$ and peer effects dominate, the principal will optimally adjust contracts to have equal weight on each agent's output. In other words, agents will be compensated based on aggregate output only, and wages will become perfectly correlated ($w_i^M = w_{-i}^M = w_j^M$).[22]

[22] In the single principal case, this result is immediate from **Corollary** A. In the case of multiple teams, wages are rescaled due to rat race effects.

By doing so, "relative wage risk" is eliminated. Below we illustrate these results in the case of equal weights:

Example: Equal Weights. *Suppose* $\mu_{ij} = 1/n$ *for all* $i \neq j$. *Then*

$$\phi_i = \frac{n(\delta - \overline{\theta})}{n(1-\delta) + \delta(1-\overline{\theta})}. \tag{40}$$

Figure 3 shows the impact of peer effects on wage volatility and correlation in the case with uniform weights. Note that as δ increases, wage volatility decreases and wage correlation increases and converges to 1.[23]

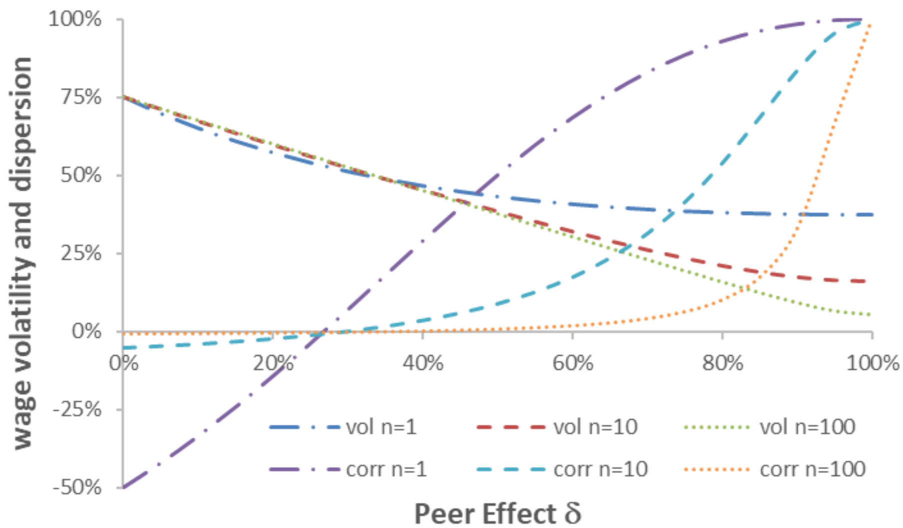

Figure 3: Wage Volatility and Dispersion

Wage volatility (expressed relative to $(1+\alpha)\overline{\sigma}$) declines, and wage correlation increases (converging to 1), with the strength δ of peer effects. (Uniform weights, $\overline{\theta} = 50\%$, $y_{ii}^* = 75\%$).

In our example with equal weights, wage contracts will load *positively* on peer output when $\delta > \overline{\theta}$, in stark contrast to the RPE prediction. The next result demonstrates that this simple

[23] Englmaier and Wambach (2010) demonstrate that wage compression results if agents are inequity averse, and thus suffer a disutility if their payoff exceeds that of others. Here, even though agents gain from being ahead of their peers, it is still optimal to reduce wage variation. Note also that wage correlation becomes positive at an even lower level of δ than is required for $\phi > 0$ due to the positive correlation in output shocks across agents.

condition holds more generally if the set of peers is large and the weights are diffuse: the hedging effect dominates the RPE effect once δ exceeds $\overline{\theta}$.

COROLLARY C (DIFFUSE WEIGHTS). *Suppose the weights on individual agents become diffuse as the population becomes large:* $\|\mu\|_{max} \to 0$ *as* $N \to \infty$. *Then*

$$\phi_i \to \frac{\delta - \overline{\theta}}{1-\delta} \text{ as } N \to \infty. \tag{41}$$

PROOF: See Appendix. ■

It is important to emphasize that despite this significant wage compression, in the single-team case effort levels are maintained and the principal's expected profit is unchanged. Hence, even if contract sensitivities appear inefficient and "low powered" relative to the RPE prediction, such contracts may be optimal and efficient with peer-dependent preferences. The inefficiencies that do arise in this framework are not due to failures of RPE, but are the result of rat-race externalities across teams which instead create *excessive* incentives and wage levels.

D. Proximity Effects

Thus far we have only analyzed the overall peer sensitivity of contracts given by ϕ_i. In this section we utilize the general network structure of preferences embodied in μ to analyze the cross-sectional variation in this sensitivity and how it depends on the "closeness" of peers. We first consider the distance to external peers, then consider team effects, and lastly look at indirect peer effects.

Recall that in our multi-team framework, the weight agents put on external peers – that is, peers on other teams – is given by μ_e. We may interpret μ_e as a measure of the distance between teams. For example, agents in different divisions of a firm may be more sensitive to the wages of peers in other divisions if they are co-located than if they are in different locations. We can consider how the wage sensitivity to external peers varies with this distance.

In particular, consider the comparative static in which we change the distance between teams as measured by μ_e. Because the peer weights must sum to one, as we change μ_e we must rescale the internal team weights correspondingly. Specifically, consider a set of weights μ^0 in which teams have the equal size N_I and agents put no weight on external peers ($\mu_e^0 = 0$). Then we can construct new weights μ with external weights $\mu_e > 0$ by rescaling internal team weights as follows:

$$\mu_{ij} = \mu_{ij}^0 \left(1 - \mu_e \left(N - N_I\right)\right) \text{ for all } i, j \in I . \tag{42}$$

We can then interpret changing μ_e as changing the distance between teams, and make the following comparison of resulting contracts.

COROLLARY D (DISTANCE EFFECTS). *Consider team weights μ defined by* (42). *As μ_e declines and teams become more distant, the contract sensitivity to external peers with a single or multiple principals (y_e^S or y_e^M) also declines.*

PROOF: See Appendix. ■

Next consider internal team effects. It is natural to assume that the wages of peers on the same team may be more salient than that of peers on different teams. To analyze the implications of this proximity effect, consider a setting of uniform internal weights and external weights. Specifically, in addition to (28) and (29), we assume agents put the same weight on each teammate, and this weight exceeds the weight put on external peers:

$$\mu_{i\hat{i}} = \mu_t \geq \mu_e \text{ for all } \hat{i} \in I_{-i} . \tag{43}$$

The following result demonstrates that in this case, optimal contracts can be written as a linear function of just three output measures: own output, team output, and aggregate output. Moreover, the weight on team output will *always* be non-negative.

PROPOSITION VI (TEAM EFFECTS). *Suppose agents have uniform internal and external weights with $\mu_t \geq 1/n \geq \mu_e$. Let $\bar{q}_I \equiv |I|^{-1} \sum_{i \in I} q_i$ be the average output of agents on team I. Then equilibrium contracts can be written in terms of own output, team output, and aggregate ouput:*

$$w_i = Y_0 + Y_{Own}\, q_i + Y_{Team}\, \bar{q}_I + Y_{Agg}\, \bar{q}_{\mathcal{N}} , \text{ with } Y_{Team} \geq 0 . \tag{44}$$

The sensitivity to team output is strictly positive unless $\mu_e = \mu_t = 1/n$ *or* $\delta = 0$.

PROOF: See Appendix for a complete characterization and additional comparative statics. ■

PROPOSITION VI shows that when agents care more about the wages of their teammates, then we should *always* see a positive loading on team output. The intuition for this result is that because the weight on aggregate output controls for both RPE effects and the relative wealth risk from external peers, the weight on team output only reflects the hedging of the incremental relative wealth risk coming from teammates.

Another notion of distance we can consider is when peers are connected only indirectly. Specifically, suppose agents i and k do not consider themselves peers $(\mu_{ik} = \mu_{ki} = 0)$. But suppose i and j are peers, as are j and k $(\mu_{ij}\mu_{jk} > 0)$. Because the optimal contract for agent j will hedge against agent k's wage, agent i will be indirectly exposed to k's wage through his concern for agent j's. Thus, we should expect contracts to reflect these peer relationship "chains."[24]

To illustrate this effect, we consider next a case in which agents are arranged in a circle and care only about their nearest neighbors; that is $\mu_{ij} = \frac{1}{2}$ if and only if $|i-j| \in \{1, n\}$. In this case, although peer relationships are strictly local, all agents are connected through a peer chain, with the distance between them determining the length of the chain. The next result demonstrates that, due to this indirect peer effect, the optimal contract sensitivity to every agent is distorted above the RPE case, with the distortion decreasing with distance to that agent. Again, the contract is distorted even for non-neighboring agents is due to the chain of influence that arises through the network of peer relationships. Finally, as long as peer-effects are non-zero, the weight on all agents within any finite neighborhood will be positive if the population is sufficiently large.

[24] Technically, this chain effect can be seen from (23) by the fact that $\Delta^{-1} = (1-\delta)\left[\mathbf{I} + \delta\mu + \delta^2\mu^2 + \ldots\right]$, with the powers of μ capturing chains of the corresponding length.

PROPOSITION VII (PEER CHAINS). *Consider a circular network of peers in which each agent puts equal weight on their two nearest neighbors. Then the contract sensitivity* y_{ij}^S *is strictly decreasing function of the distance between* i *and* j*, and exceeds the RPE benchmark even for the most distant agents. Moreover, for any fixed pair* (i, j) *and any* $\delta > 0$*,* $y_{ij}^S > 0$ *for* n *sufficiently large.*

PROOF: See Appendix for a complete characterization and additional comparative statics. ■

IV. Asymmetric Visibility and Efficiency Gains

Our prior results demonstrate that profits either stay the same (with a single principal) or decline (with multiple principals) with the strength peer effects when peer relationships are symmetric. In this section we relax the symmetry requirement and show that when some peers are "more visible" than others, the principal can exploit this asymmetry and increase profits.

Specifically, define the "visibility" $\hat{\mu}_i$ of agent i as the collective weight put on that agent by others. Under our current pairwise symmetry assumption (4) we have:

$$\hat{\mu}_i \equiv \sum_j \mu_{ji} = 1 \text{ for all } j. \quad (45)$$

That is, each agent has equal visibility. To see the importance of this assumption, consider the following exercise. Suppose we would like to raise agent i's certainty equivalent wage by \$1, while holding all other agents' utility constant. To do so, we need to raise w_i by $(1-\delta)$.[25] But then, to compensate agent j for the direct peer effect, we need to raise w_j by $\delta\mu_{ji}(1-\delta)$ for each agent j. Of course, we must then raise w_k by $\delta^2\mu_{kj}\mu_{ji}(1-\delta)$ to compensate each agent k for the "second-order" peer effect, and so on. Let ω_i be the total cost of these adjustments. Then we have

$$\omega_i \equiv (1-\delta)\sum_j \left[I + \delta\mu + \delta^2\mu^2 + \ldots \right]_{ji} = \sum_j \Delta_{ji}^{-1}. \quad (46)$$

[25] Recall from (2) and (4) that $v_i = w_i + \hat{\delta}(w_i - w_{-i}) = w_i / (1-\delta) - \delta w_{-i} / (1-\delta)$.

This total compensation cost ω_i is equivalent to the *Katz-centrality* of agent i in the peer network.[26] As the following result shows, total compensation costs are equal for all agents if and only if all agents have equal visibility.

LEMMA 3 (VISIBILITY, CENTRALITY, AND COMPENSATION COST). *The Katz-centrality ω_i of each agent, given by (46), represents the total cost of raising agent i's utility while leaving others unchanged. For any stochastic peer matrix μ, even if non-symmetric, the average cost of compensation, $\frac{1}{N}\sum_i \omega_i$, is equal to 1. But $\omega_i = 1$ for all i if and only if agents have equal visibility so that (45) holds.*

PROOF: See Appendix. ■

Thus, equal visibility implies that the effective cost of compensation is also equalized across all agents. It is straightforward to show that our welfare results in Section A only rely on this result, and *not* on the stronger pairwise symmetry assumption in (4).[27]

Now suppose we relax (45), so that some agents are less visible than others. Then, from **LEMMA 3**, while the average compensation cost is unchanged, individual compensation costs will differ. The principal will optimally exploit this asymmetry, providing higher incentives and wages to less-central agents in order to lower compensation costs. Thus, with asymmetric visibility there will be a tension between *compensation efficiency* and *production efficiency* (which, given symmetric convex effort costs, would allocate effort equally). The optimal tradeoff is characterized in the following result.[28]

---

[26] Katz-centrality is generally defined recursively as $\omega_i = (1-\delta) + \sum_j \delta\mu_{ji}\omega_j$, where δ is referred to as the network attenuation factor and $(1-\delta)$ is the base level importance of each node in the network. Eigenvector centrality (or Bonacich-centrality) is a parameter-free version of Katz-centrality obtained by taking the limit as $\delta \to 1$. See Katz (1953) and Newman (2018).

[27] See for example the proof of **PROPOSITION I**, which only uses symmetry to establish that $\mathbf{1}'\Delta^{-1} = \mathbf{1}'\mu = \mathbf{1}'$; that is, we only require that the matrix μ is doubly stochastic.

[28] The fact that this variation in compensation costs results in higher profits follows from the standard result that profits are convex in input prices. Note also that this gain from distorting production applies even absent agency costs (e.g. with $\overline{\sigma} = 0$).

PROPOSITION VIII (ASYMMETRIC VISIBILITY AND EFFICIENCY GAINS). *Suppose that we relax the pairwise symmetry assumption (4) and agents have asymmetric visibility. If* $\delta > 0$*, then although the relative contract sensitivities satisfy (38) as before, a single principal can distort effort to achieve higher expected profits than in the RPE case (*$\delta = 0$*). In equilibrium, the principal allocates effort (and earns profits) according to the inverse of each agent's Katz-centrality:*

$$a_i^A = \omega_i^{-1} a_i^* \quad \text{and} \quad \pi_i^A = q_0 + \omega_i^{-1}\left(\tfrac{1}{2} a_i^*\right) - c_0 = \pi_i^* + \tfrac{1}{2}\left(\omega_i^{-1} - 1\right) a_i^* . \tag{47}$$

The principal's average gain per agent is thus proportional to $\left(\frac{1}{N}\sum_i \omega_i^{-1}\right) - 1 > 0$.

PROOF: See Appendix. ■

The proposition shows that with asymmetric visibility, the principal can lower the total cost of compensation by shifting incentives (and hence wages) to less-central agents. To illustrate the relationship between the preference weights (μ, δ) and the principal's expected profit, Figure 4 shows the average gain per agent for a range of randomly generated models. Here, we have measured the asymmetry in μ according to the heterogeneity in the agents' visibility $\hat{\mu}$. As the figure shows, the principal's expected profit increases with this variation (standard deviation of $\hat{\mu}_j$) as well as with the intensity δ of the peer effects. The latter effect, monotonicity with respect to δ for a given μ, appears to hold quite generally across all numerical examples.

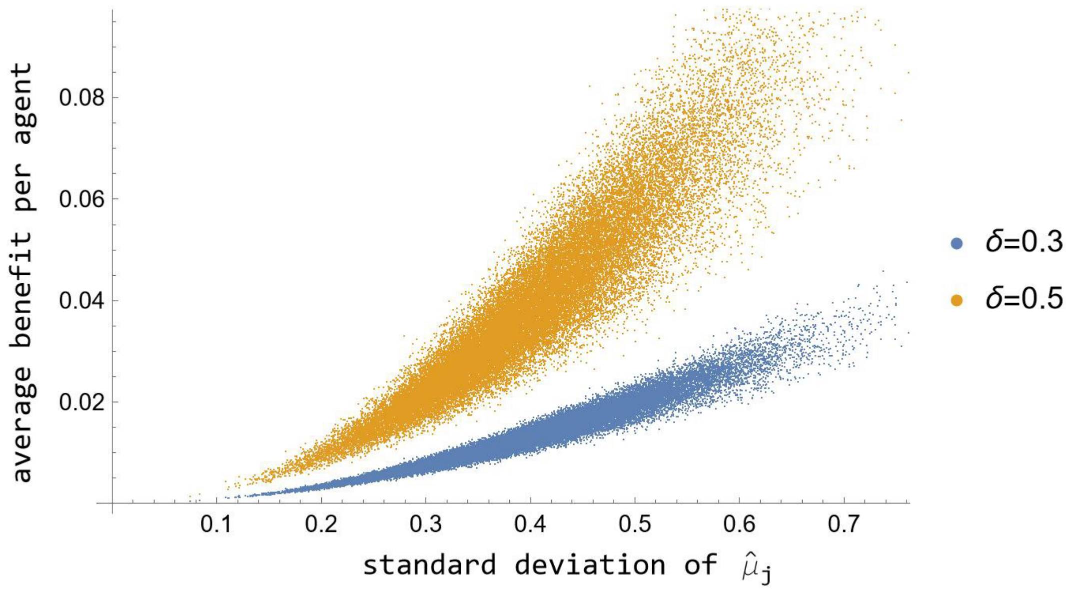

Figure 4: Benefit to Principal from Agent Heterogeneity

The average benefit per agent to the principal $\frac{1}{N}\sum_j \omega_j^{-1} - 1$ as function of standard deviation of $\hat{\mu}_j$. Simulating 40,000 random stochastic matrixes μ for $N = 10$. Blue (orange) dots are for $\delta = 0.3$ $(\delta = 0.5)$.

To further elucidate the source of increased profits to the principal, we next consider a specific natural example of asymmetry in which there are two types of agents:

- **Independent Agents:** $\mu_{jj} = 1$, $\mu_{jk} = 0$ for $k \neq j$; thus their effective δ is zero. These agents care only about their own wage, and do not care about the wages of their peers.[29]
- **Peer-Dependent Agents:** $\delta > 0$, $\mu_{mm} = 0$, $\mu_{mk} = 1/n$ for $k \neq m$. These agents benchmark their wage to the average of their peers.

Note that in this example, because peer-dependent agents care about independent agents, but not vice versa, *independent agents are more visible and central*. The following proposition shows that as a result, independent agents will be given lower equilibrium incentives and wages. We also show that the principal's profit increases with the heterogeneity of the population.

[29] Note that to make this example particularly stark, we also relax the requirement in (1) that $\mu_{jj} < 1$.

PROPOSITION IX (MIXED POPULATION). *Consider a single principal with independent agents ($j \in J$) and peer-dependent agents ($m \in \mathcal{N}_{-J}$). Then independent agents receive a standard RPE contract with $\phi_j^A = -\bar{\theta}$. Peer-dependent agents work harder than independent agents, with*

$$\frac{a_m^A}{a_j^A} = \frac{1+\frac{1}{n}\delta}{1-\delta} \equiv \psi > 1, \tag{48}$$

and furthermore,

$$\frac{1}{N}\sum_i \omega_i^{-1} = 1 + \left(\psi + \psi^{-1} - 2\right)\frac{|J|}{N}\left(1 - \frac{|J|}{N}\right), \tag{49}$$

so that the principal's profit per agent increases with an increase in population heterogeneity: that is, with an increase in δ, or shifting the proportion of independent agents closer to 50%.

PROOF: See Appendix. ■

Note that for $|J| = \frac{1}{2}N$, the principal's gain per agent decreases in N with the limit

$$\frac{1}{N}\sum_i \left(\omega_i^{-1} - 1\right) \to \frac{\delta^2}{4(1-\delta)} \text{ as } N \to \infty. \tag{50}$$

Thus, the average gain per agent does not completely dissipate even when N is very large.

The ability of the single principal to internalize the impact of peer effects across contracts allows the principal to capitalize fully on any asymmetry of visibility across agents. With multiple principals, there will be competing effects of any asymmetric visibility *within* teams,[30] which will increase profits, and the rat-race effect *across* teams, which will decrease profits.

V. Disclosure Effects

In the standard RPE contracting framework without peer effects, agent's incentives are independent of each other's contracts. As a result, the timing and visibility of contracts and wages

[30] Specifically, with equal external weights, what matters is the within-team centrality $\omega_I = \mathbf{1}'(\Delta_{II})^{-1}$, which will be equal if and only if agents have equal within team visibility $\mathbf{1}'\mu_{II}$.

are not consequential. With peer effects, however, these details will matter. We show in this section that if principals can privately negotiate contracts with individual agents on their team, then additional rat-race effects emerge within teams. In addition, the public disclosure of wage contracts across teams may increase their salience, strengthening peer effects, while also creating incentives for principals to distort their own team's contracts to affect the productivity of other teams.

A. Private Negotiation

Thus far we have assumed that each principal discloses to her team the incentive contracts to be used within the firm. This assumption implies that the principal cannot privately negotiate (or renegotiate) these contracts with individual agents. If instead individual contracts terms can be set or altered in a way that is hidden from other team members, then in equilibrium we should require that contracts be "renegotiation-proof" with respect to any principal-agent pair. That is, in equilibrium, there should be no alternative contract the principal could offer to a single agent which the agent would accept and would raise the principal's expected profit, while holding other contracts and effort as given.

When privately negotiating, the principal and agent will ignore the impact of their wage choice on the utility of other agents, as well as try to lower the wage of others through the performance benchmark, just as in the setting of multiple teams discussed earlier. Moreover, there is now an added benefit to the principal: changing the agent's effort in a way that lowers the wage of other agents within the same team contributes directly to the principal's profits.

Though there is an incentive to renegotiate, the opportunity to do so must hurt the principal ex-ante. In equilibrium, other agents within the team will anticipate the renegotiated contract and seek commensurate terms. In other words, because the renegotiation-proof contract could always be proposed in an environment with disclosure, allowing hidden renegotiation only constrains the principal. But while each principal is *individually* worse off with hidden contracting, the *equilibrium* consequence of renegotiation is less clear, as constraining contracts in this way might reduce some of the "rat-race" inefficiency that arises with multiple teams.

The following result characterizes the optimal contract when contracts are privately negotiated:

LEMMA 4 (PRIVATE NEGOTIATION CONTRACT). *Given the contracts and actions of other agents, principal I has an incentive to renegotiate privately with agent i unless agent i has exposure $\beta_i^R = (1+\alpha_i^R)y_i^*$ where the scaling factor α_i^R is given by*

$$1+\alpha_i^R \equiv \left(\frac{1-\delta\mu_{ii}}{1-\delta}\right)\left(1-\frac{\delta}{1-\delta\mu_{ii}}\sum_{j\neq i}\mu_{ij}y_{ji}-\sum_{\hat{i}\in I_{-i}}y_{\hat{i}i}\right). \tag{51}$$

PROOF: See Appendix. ■

In the case of a single-person team, private renegotiation within a team should have no effect, and thus (51) coincides with our earlier calculation in **LEMMA 2** (see (27), noting $\mu_{ii}=\mu_I$ and $I_{-i}=\varnothing$ when $N_I=1$). But with multi-person teams, new effects emerge. To understand (51) better, it is useful to consider the case with $\mu_{ii}=0$.[31] Then,

$$1+\alpha_i^R \equiv \left(\frac{1}{1-\delta}\right)\left(1-\delta\sum_{j\neq i}\mu_{ij}y_{ji}-\sum_{\hat{i}\in I_{-i}}y_{\hat{i}i}\right). \tag{52}$$

As in (27), the term $\sum_{j\neq i}\mu_{ij}y_{ji}$ captures the effect on agent i from altering other agents' expected wage, but now this term includes i's own teammates (as this own-team effect is no longer internalized by the principal). The new term $\sum_{\hat{i}\in I_{-i}}y_{\hat{i}i}$ captures the effect on the principal's expected wage bill from altering the expected wages of other agents on team I. To evaluate the equilibrium impact of renegotiation, we again consider the symmetric case, and keep $\mu_{ii}=0$ for simplicity.

PROPOSITION X (PRIVATE NEGOTIATION EQUILIBRIUM). *In the symmetric multi-team equilibrium with equal external weights and $\mu_{ii}=0$ for all i, private renegotiation exacerbates the rat-race effect, with a strictly higher scaling factor*

$$1+\alpha^R \equiv \frac{1}{1-\delta+\left(\delta-\bar{\theta}\right)y_{ii}^*-\left(N-N_I\right)y_e^S} > 1+\alpha\,, \tag{53}$$

[31] Recall that $\mu_{ii}>0$ is equivalent to adjusting δ for agent i, see footnote 10.

as long as $N_I > 1$. [32] *Moreover, distortions persist (* $\alpha^R > 0$ *) even absent peer effects (* $\delta = 0$ *).*

Proof: See Appendix. ■

Thus, private negotiation amplifies the rat-race effects that we identified in Section III. Additionally, the rat race occurs even without peer effects, since the principal has an incentive to renegotiate privately with RPE contracts (providing higher incentives to one agent lowers the expected wage paid to teammates).

Note also from (53) that private renegotiation implies a positive rat-race effect even in the case of a single principal ($N = N_I$),

$$1+\alpha^R = \frac{1}{1-\delta+\left(\delta-\bar{\theta}\right)y_{ii}^*} = \frac{1}{1-\bar{\theta}y_{ii}^* - \delta\left(1-y_{ii}^*\right)} > 1, \tag{54}$$

Therefore, with private negotiation, rat-race effects *always* appear (even without peer effects) and peer effects *always* amplify the rat race (even with a single principal).

B. Wage Transparency and Public Disclosure

Recent regulation has increased disclosure requirements regarding executive compensation. The SEC now requires compensation disclosures for the CEO, CFO, and the three additional most highly compensated officers the firm, as well as compensation peer groups.[33] Their visibility has been further enhanced by "Say on Pay" rules that require periodic shareholder approval of compensation schemes. Websites such as Glassdoor collect and provide data on salaries for a broad range of managerial positions within firms. Compensation disclosure requirements for public employees have also increased as a result of transparency and accountability measures. For example, individual faculty member salaries for all faculty at the University of California can be looked up online.

---

[32] As a benchmark, with equal weights ($\mu_{ij} = 1/n$) we can write $1+\alpha^R = \left(1-\delta+\left(N_I - 1 + \delta\right)y_e^S\right)^{-1}$, which matches (36) for single agent teams ($N_I = 1$).

[33] Jochem, Ormazabal, and Rajamani (2021) document the network structure of these peer groups and show that they have become significantly more clustered and reciprocal (i.e., symmetrically linked) over time.

In the context of our model, we interpret greater wage transparency and disclosure as leading to an increase in the salience of peer compensation.[34] As a result, individuals are likely to put higher weight on their relative versus absolute wage (an increase in δ), as well as put higher weight on more distant peers (an increase in μ_e). These effects may be due to the agent's own visibility of others' wages, or the fact that agents believe that others can more easily make such comparisons.[35]

Under this interpretation, the results of Section III allow for the following empirical implications of increased disclosure:

COROLLARY E (WAGE TRANSPARENCY). *In the symmetric team setting of* PROPOSITION II*, suppose an increase in transparency or disclosure leads to an increase in* δ*. This change will lead to an*

- Increase in rat-race effects (α): higher wages and productivity, lower profits,
- Increase in the wage contract exposure to external peers (y_e),
- Increase in the sensitivity to peer output (ϕ),
- Increase in wage correlation (with the average wage) and a decrease in wage dispersion (standard deviation as a fraction of the expected wage).

The same results apply for an increase in μ_e*, where in the case of the wage correlation and peer sensitivity, we additionally assume uniform internal weights.*

PROOF: See Appendix. ■

In the context of CEO compensation, our results suggest that SEC rules that have increased compensation disclosures and saliency should lead to higher expected compensation. Consistent with this prediction, Jochem Ormazabal, and Rajamani (2021) document a significant increase in

[34] Card et.al. (2012) provide field-based confirmation of relative pay comparisons by randomized manipulation of revelation of information on coworkers' salaries for University of California employees.

[35] For example, we can think of the weights μ_{ij} as based in part on the probability that others are able to see and compare agent *i* and *j*'s wages, and this comparison may impact agent *i*'s social status.

average pay together with a 40% decline in wage dispersion (both in aggregate and within industry-size groups) since 2006.[36]

C. External Contract Disclosure

We have assumed thus far that agents can only observe the realized *wages* of external peers after making their effort decisions. In this section we consider an additional effect that arises if *contracts* are disclosed externally across teams before effort is determined. Specifically, we modify the timing in Figure 1 so that prior to choosing effort, each agent can observe the wage contracts of agents on other teams.

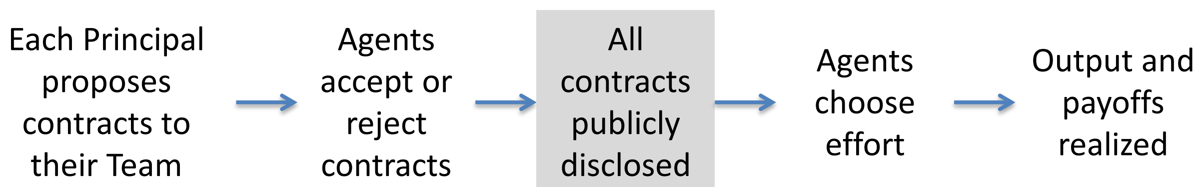

Figure 5: Contracting Game with Ex Ante External Contract Disclosure

Ex ante external contract disclosure matters because the contract of agent i on team I, once revealed, will influence the effort choice of agents on other teams ($j \notin I$). For example, an increase in y_{ij}, the loading of i's contract on j's output, will reduce agent j's effort incentives, since higher q_j would have a more positive impact on i's wage, reducing j's utility. Principal I can take advantage of this effect by offering agent i a contract with higher y_{ij}, and then, anticipating that agent j will reduce effort and earn a lower wage as a result, lower the expected wage paid to team I (given the lower expected peer benchmark).[37]

This intuition has several important implications. First, it suggests that each principal will no longer choose the implied sensitivities to minimize the agent's residual risk. Instead, the

[36] Further evidence that increased disclosure of executive compensation increases executive pay is provided in Park, Nelson, and Huson (2001), Perry, and Zenner (2001), Schmidt (2012), Gipper (2021) and Mas (2019). Gipper (2021) also shows a reverse effect for the subset of firms for whom the JOBS Act of 2012 rolled back some disclosure requirements.

[37] In the prior equilibrium, agents' incentives were determined by the *anticipated* contracts used by other teams. Public disclosure gives the principal the ability to change external expectations.

principal will raise the sensitivity to external peers in order to distort their effort downward. But, because the agent is now exposed to additional external output risk, the contract will expose him to less output risk from his own team. Finally, because these distortions raise the cost of providing incentives, the optimal level of effort declines.

The following result confirms these effects. For simplicity, we focus on the case with uniform internal and external weights, where those weights may differ from each other. We show that in the optimal contract the agent's effort (determined by β_{ii}^{D}) is distorted downward, as is the agent's implied exposure to the output of his teammates ($\beta_{i\hat{i}}^{D}$), whereas the agent's implied exposure to the output of agent's on other teams (β_{ij}^{D}) is distorted upward, relative to the setting without public disclosure.

PROPOSITION XI (EXTERNAL CONTRACT DISCLOSURE). *Suppose teams have the same size* N_I *and agents have uniform internal and external peer weights:* $\mu_{ii}=0$, $\mu_{i\hat{i}}=\mu_{\iota}$ *for* $\hat{i}\in I_{-i}$ *and* $\mu_{ij}=\mu_{e}$ *for* $j\notin I$. *Then the equilibrium with external contract disclosure has uniform internal and external loadings:* y_{ii}^{D}, $y_{i\hat{i}}=y_{\iota}^{D}$ *and* $y_{ij}=y_{e}^{D}$ *with exposures that are distorted from the multi-team equilibrium:*

$$\begin{aligned} &\beta_{ii}^{D}<\left(1+\alpha_{I}\right)y_{ii}^{*} && \text{(lower effort incentives),} \\ &\beta_{i\hat{i}}^{D}<-\tfrac{1}{n}\bar{\theta}\beta_{ii}^{D} && \text{(lower internal exposure), and} \\ &\beta_{ij}^{D}>-\tfrac{1}{n}\bar{\theta}\beta_{ii}^{D} && \text{(higher external exposure)} \end{aligned} \tag{55}$$

where $\alpha_{I}=\dfrac{\delta}{1-\delta}\left(1-\mu_{I}\right)\left(1-N_{I}y_{e}^{D}\right)$. *Without ex ante external disclosure,* ***Error! Reference source not found.*** *holds with equality.*

PROOF: See Appendix. ■

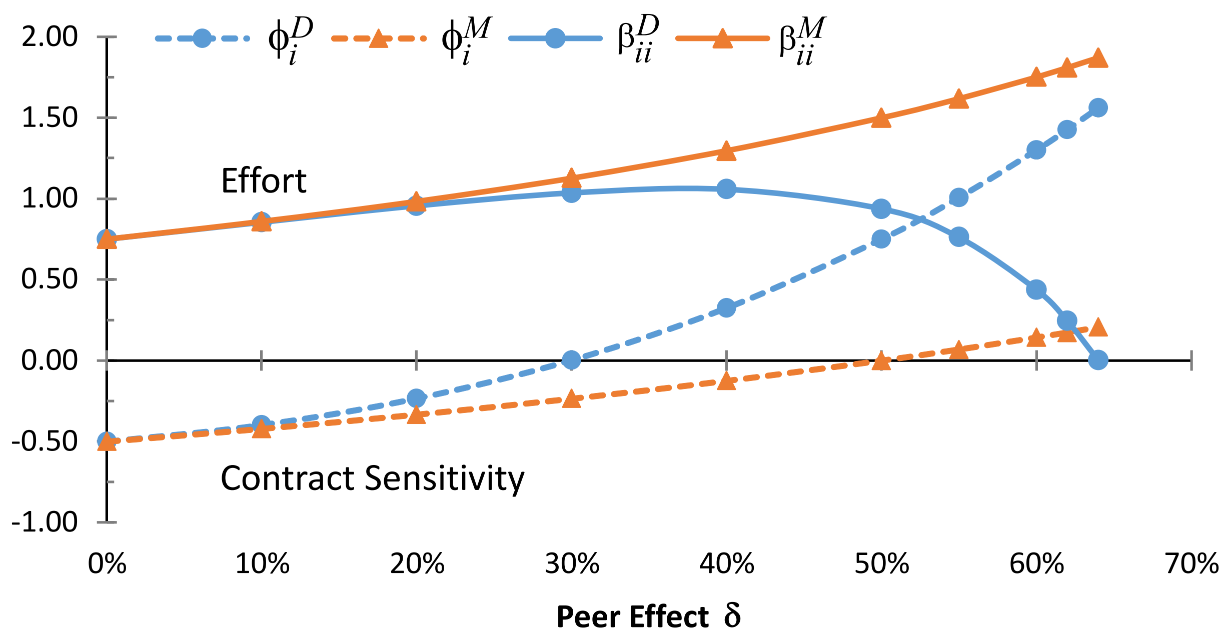

Figure 6: Effort and Contract Sensitivity with Ex Ante External Contract Disclosure
Ex ante disclosure increases contract sensitivity to external peers and decreases effort relative to case when contracts are not externally disclosed. (Parameters: $N = 2, N_I = 1, \overline{\theta} = 50\%, y_{ii}^* = 75\%$).

Figure 6 shows the effect of external disclosure on both effort and contract sensitivity. Note that the effect of public disclosure increases with strength of peer effects δ. Indeed, when δ is sufficiently high, public disclosure can cause effort to collapse well below the second-best level, in contrast the settings without disclosure where it is always above the second-best.

Thus, external contract disclosure is likely to lead to even greater departures from "RPE" in observed contracts. And although public disclosure lowers equilibrium productivity, because effort is inefficiently high due to the rat-race effect, this decrease in effort can raise profitability.

VI. Conclusion

In this paper we have extended a standard moral hazard optimal contracting framework to a setting in which agents care about both their absolute wage, as well as how their wage compares to that of their peers. We allow for a general network of peer relationships, both within and across firms. Our results overturn standard predictions from contracting models. We find that rat-race effects across teams raise equilibrium effort, and that compensation benchmarking offsets performance benchmarking, so that optimal contracts load more positively on peer output (including indirect peers) than the standard RPE model would imply. When peer effects are sufficiently strong, effort can exceed first best, while at the same time wages are driven primarily by aggregate (rather than individual) performance. Finally, principals can exploit asymmetric peer effects within their teams and raise profits, relative to when agents' preferences are devoid of peer effects, by inefficiently shifting effort to less visible agents.

We also considered the implications of different levels of disclosure and transparency. On one hand, if private renegotiation is possible, a rat-race effect emerges even within teams which raises team effort and output and lowers profits. On the other hand, external disclosure across teams is likely to increase the saliency of peer comparisons, exacerbating the rat-race effect across teams and increasing wage levels and correlations.

In addition to these broad predictions, our model makes more nuanced predictions regarding the relationship between contract sensitivities and the details of peer relationships. We show that relative to RPE contracts, peer effects imply that contracts will be less sensitive to the agent's own output and will have higher (more positive) sensitivity to peer output. The contract loading on peers will be highest for peers who are "closer" in the network, such as teammates or neighbors. With data on social networks becoming more readily available to researchers, one could envision merging such data with data on compensation to empirically evaluate these predictions.

To highlight these effects, we have simplified other aspects of the principal-agent model. For example, in practice the correlation between agents' output will also likely vary with some measure of distance. In that case the contract loadings will depend on relative distance between peers in the social network versus the production matrix.

We have also simplified the model by maintaining a standard linear CARA-normal framework, both to facilitate tractability and precise quantitative conclusions as well as to allow comparisons to existing benchmarks. Future work could analyze how results might change under a different utility specification (such as CRRA), as well as different ways to embed KUJ preferences. For example, one could assume different risk aversions with respect to absolute and relative income and/or one could assume the two components are additively separable. Future empirical or experimental work characterizing how KUJ preferences are most accurately represented could help guide researchers as to which specifications to consider. Nonetheless, we believe the forces identified here will continue to arise: optimal contracts will contain a hedging component, rat-race effects will emerge, principals may gain by distorting production toward less visible agents, and disclosure environments will matter.

In the multi-team context, we have also treated team and population size as exogenous. It would be interesting to explore the consequences of peer effects on industry structure when entry and exit is endogenous. For example, although we have assumed constant returns to scale in our model, larger teams are more efficient due to their ability to internalize rat-race effects and thereby prevent wages from escalating. Conversely, rat-race effects may be an important barrier to entry for smaller entrants to an industry.

REFERENCES

Aggarwal, Rajesh K., and Andrew A. Samwick, 1999a, The Other Side of the Trade-Off: The Impact of Risk on Executive Compensation, *Journal of Political Economy* 107, 65–105.

———, 1999b, Executive Compensation, Strategic Competition, and Relative Performance Evaluation: Theory and Evidence, *Journal of Finance* 54, 1999–2043.

Antle, Rick, and Abbie Smith, 1986, An Empirical Investigation of the Relative Performance Evaluation of Corporate Executives, *Journal of Accounting Research* 24, 1–39.

Barro, Jason R., and Robert J. Barro, 1990, Pay, Performance, and Turnover of Bank CEO's, *Journal of Labor Economics* 8, 448-81.

Bartling, Björn, 2011, Relative Performance or Team Evaluation? Optimal Contracts for Other-Regarding Agents, *Journal of Economic Behavior & Organization* 79, 183– 193.

Bartling, Björn, and Ferdinand A. von Siemens, 2010, The intensity of incentives in firms and markets: moral hazard with envious agents, *Labour Economics* 17 (2), 598–607.

Bertrand Marianne, Mullainathan Sendhil, 2001, Are CEOs rewarded for luck? The ones without principals are, *Quarterly Journal of Economics* 116(3):901-32.

Bouwman, Christa H.S, 2013, The Geography of Executive Compensation, Wharton Financial Institutions Center working paper.

Card, D., A. Mas, E. Moretti, and E. Saez, 2012, Inequality at Work: The Effect of Peer Salaries on Job Satisfaction, *American Economic Review*, 2981-3003.

Carlsson, Fredrik, Gautam Gupta, and Olof Johansson-Stenman, 2009, Keeping up with the Vaishyas? Caste and relative standing in India, Oxford Economic Papers 61, 52–73.

Clark, Andrew E., 1996, L'utilité est-elle relative? Analyse à l'aide de données sur les ménages. Economie et Prévision 121, 151-164.

Clark, Andrew E., and Andrew J. Oswald, 1996, Satisfaction and comparison income, *Journal of Public Economics* 61, 359-81.

Coughlan, Anne T., and Ronald M. Schmidt, 1985, Executive Compensations, Management Turnover, and Firm Performance and Empirical Investigation, *Journal of Accounting and Economics* 7, 43-46.

DeMarzo, Peter, Michael J. Fishman, Zhiguo He, and Neng Wang, 2012, Dynamic Agency and the *q* Theory of Investment, *Journal of Finance* 67 (6), 2295-2340.

Dohmen, Thomas, Armin Falk, Klaus Fliessbach, Uwe Sunde and Bernd Weber, 2011, Relative versus absolute income, joy of winning, and gender: Brain imaging evidence, *Journal of Public Economics* 95, 279-285.

Duchin, Ran, Amir Goldberg, and Denis Sosyura, 2017, Spillovers inside Conglomerates: Incentives and Capital, *Review of Financial Studies* 30 (5), 1696–1743.

Dur, Robert., and Amihai Glazer, 2008, Optimal contracts when a worker envies his boss, *Journal of Law, Economics, and Organization* 24 (1), 120–137.

Englmaier, Florian, and Achim Wambach, 2010, Optimal Incentive Contracts under Inequity Aversion, *Games and Economic Behavior* 69 (2), 312–28.

Ederer, Florian, and Andrea Patacconi, 2010, Interpersonal Comparison, Status and Ambition in Organizations, *Journal of Economic Behavior & Organization* 75, 348–363.

Ferrer-i-Carbonell, Ada, 2005, Income and well-being: an empirical analysis of the comparison income effect, *Journal of Public Economics* 89: 997 – 1019.

Fershtman, C., H.K. Hvide, and Y. Weiss, 2003, A behavioral explanation of the relative performance evaluation puzzle, *Annales d'Economie et de Statistique* (71–72), 349–361.

Fliessbach, Klaus, Bernd Weber, Peter Trautner, Thomas Dohmen, Uwe Sunde, Christian E. Elger, and Armin Falk, 2007, Social Comparison Affects Reward-Related Brain Activity in the Human Ventral Striatum, *Science* 318 (5854), 1305-1308.

Frydman, Carol, and Dirk Jenter, 2010, CEO Compensation, *Annual Review of Financial Economics* 2, 75-102.

Gartenberg, Claudine, and Julie Wulf, 2017, Pay Harmony? Peer Comparison and Performance Compensation in Multi-business Firms, *Organization Science* 28(1), 39–55.

Garvey, Gerald T., and Todd T. Milbourn, 2003, Incentive Compensation When Executives Can Hedge the Market: Evidence of Relative Performance Evaluation in the Cross Section, *Journal of Finance* 58, 1557–81.

Ghazala, Azmat, and Nagore Iriberri, 2010, The importance of relative performance feedback information: Evidence from a natural experiment using high school students, *Journal of Public Economics* 94, 435-452.

Gibbons, Robert, and Kevin J. Murphy, 1990, Relative Performance Evaluation for Chief Executive Officers*, Industrial and Labor Relations Review* 43(3), 30S-51S.

Gipper, Brandon, 2021, The economic effects of expanded compensation disclosures Assessing the Effects of Disclosing Management Compensation, *Journal of Accounting and Economics* 71(1), 101338.

Goel, Aanad M., and Anjan V. Thakor, 2005, Green with Envy: Implications for Corporate Investment Distortions. *Journal of Business* 78(6), 2255-2288.

Goel, Anan M., and Anjan V. Thakor, 2010, Do envious CEOs cause merger waves?, *Review of Financial Studies* 23, 487–517.

Goukasian, Levon, and Xuhu Wan, 2010, Optimal Incentive Contracts Under Relative Income Concerns, *Mathematics and Financial Economics* 4, 57-86.

Hall, Brian J., and Jeffrey B. Liebman, 1998, Are CEOs Really Paid Like Bureaucrats?, *Quarterly Journal of Economics* 3, 653-691.

Heffetz, Ori, and Robert H. Frank, 2011, Preferences for Status: Evidence and Economic Implications. In Jess Benhabib, Alberto Bisin, and Matthew Jackson (eds.), *Handbook of Social Economics* (vol. 1A, pp. 69-91), Amsterdam: Elsevier.

Ibert, Markus, Ron Kaniel, Stijn Van Nieuwerburgh, and Roine Vestman, 2018, Are Mutual Fund Managers Paid For Investment Skill?, *Review of Financial Studies* 31(2), 715–772.

Janakiraman, Surya N., Richard A. Lambert, and David F. Larcker, 1992, An Empirical Investigation of the Relative Performance Evaluation Hypothesis. *Journal of Accounting Research* 30, 53–69.

Jenter, Dirk, and Fadi Kanaan, 2015, CEO Turnover and Relative Performance Evaluation, *Journal of Finance 70(5)*, 2155-2183.

Jensen, Michael C., and Kevin J. Murphy, 1990, Performance Pay and Top-Management Incentives, *Journal of Political Economy* 98, 225–264.

Jochem, Torsten, Gaizka Ormazabal, and Anjana Rajamani, 2021, Why Have CEO Pay Levels Become Less Diverse? European Corporate Governance Institute – Finance Working Paper.

Joh, Sung Wook, 1999, Strategic Managerial Incentive Compensation in Japan: Relative Performance Evaluation and Product Market Collusion, *The Review of Economics and Statistics* 81(2), 303-313.

Katz, Leo, 1953, A New Status Index Derived from Sociometric Analysis, *Psychometrika* 18(1), 39-43.

Koszegi, B, 2014, Behavioral Contract Theory, *Journal of Economic Literature* 52(4), 1075-1118.

Liu, Qi, and Bo Sun, 2016, Relative Wealth Concerns, Executive Compensation, and Systemic Risk-Taking, *American Economic Journal: Microeconomics* (forthcoming).

Luttmer, Erzo F.P., 2005, Neighbors as negatives: Relative earnings and well-being, *Quarterly Journal of Economics* 120, 963-1002.

Mas, Alexandre, 2019, Does Disclosure affect CEO Pay Setting? Evidence from the Passage of the 1934 Securities and Exchange Act, Working paper, Princeton University.

Miglietta, Salvator, 2014, Incentives and Relative Wealth Concerns, *The Quarterly Journal of Finance* 4(4), 1450013.

Newman, Mark, 2018, *Networks*, Oxford University Press.

Neumark, David, and Andrew Postlewaite, 1998, Relative income concerns and the rise in married women's employment, *Journal of Public Economics* 70, 157–183.

Park, Yun W., Toni Nelson, and Mark R. Huson, 2001, Executive Pay and Disclosure Environment: Canadian Evidence, *Journal of Financial Research*, 347-365.

Perry, Tod, and Mark Zenner, 2002, Pay for performance? Government regulation and the structure of compensation contracts, *Journal of Financial Economics*, 435-488.

Schmidt, Cornelius, 2012, Does Transparency Increase Executive Compensation?, Working paper, University of Lausanne.

Shue, Kelly, 2013, Executive Networks and Firm Policies: Evidence from the Random Assignment of MBA Peers, *Review of Financial Studies* 26, 1401-1442.

Silva, Rui C., 2013, Internal Labor Markets and Investment in Conglomerates, Working Paper, London Business School.

Appendix

Proof of Lemma 1: When $\delta = 0$, the contracting problem for agent i is independent of the wage contracts of other agents, except to the extent that other agents' actions affect the expectation of the benchmark q_{-i}. Because $E[c_i] = \left(y_i \cdot 1\right) q_0 + y_{ii} a_i + \sum_{j \in \mathcal{N}_{-i}} y_{ij} a_j - \Psi(a_i)$, the agent will choose effort a_i to maximize $y_{ii} a_i - \Psi(a_i)$, and thus $a_i = y_{ii}/k$. The principal will choose y_{i0} so that the agent's participation constraint binds; therefore $E[c_i] - \frac{1}{2}\lambda Var(c_i) = c_0$. Because $c_i \equiv v_i - \psi(a_i)$,

$$\overbrace{E[v_i] = E[w_i]}^{\delta=0} = c_0 + \Psi(a_i) + \tfrac{1}{2}\lambda Var(c_i)\,. \tag{A1}$$

The principal therefore chooses y_{ij}/y_{ii} to minimize variance as in (14), so that the volatility of consumption is $y_{ii}\bar{\sigma}$. Then, because $a_i = y_{ii}/k$, the principal chooses y_{ii} to trade off incentives for effort and the cost of risk-bearing in order to maximize

$$\begin{aligned} \pi_i = E[q_i - w_i] &= q_0 + a_i - \Psi(a_i) - \tfrac{1}{2}\lambda Var(c_i) - c_0 \\ &= q_0 + y_{ii}/k - \tfrac{1}{2}k(y_{ii}/k)^2 - \tfrac{1}{2}\lambda y_{ii}^2\bar{\sigma}^2 - c_0, \end{aligned} \tag{A2}$$

which implies (16) and (17). The principal's expected profit is $\pi_i = q_0 + \frac{1}{2}a_i^* - c_0$. Finally, note that the constant term of the wage contract is set so that the participation constraint (A1) binds; hence, because $E[q_j] = E[q_i] = q_0 + a_i^*$ by symmetry,

$$y_{i0}^* q_0 = c_0 + \tfrac{1}{2}k(a_i^*)^2 + \tfrac{1}{2}\lambda\left(y_{ii}^*\right)^2\bar{\sigma}^2 - \sum_j y_{ij}^* E[q_j] = c_0 + \tfrac{1}{2}a_i^* - k a_i^*\left(1 - \bar{\theta}\right)\left(q_0 + a_i^*\right). \tag{A3}$$

■

Proof of Proposition I: Note first that the participation constraint for each agent must bind, as otherwise the principal could cut the fixed component of the wage, increasing profits and (due to peer effects) relaxing the participation constraint of other agents. Therefore,

$$\begin{aligned} E[v_i] &= c_0 + \psi(a_i) + \tfrac{1}{2}\lambda Var(c_i) \\ &= c_0 + \psi\left(\beta_{ii}/k\right) + \tfrac{1}{2}\lambda\left[(1-\rho)\sum_j \beta_{ij}^2 + \rho\left(\sum_j \beta_{ij}\right)^2\right]\sigma_\epsilon^2 \equiv C(\beta_i), \end{aligned} \tag{A4}$$

where the function C is the implied total cost borne by the agent given sensitivities β. Rewriting (A4) in column vector form by stacking the equation for each agent i, and using the fact that $v = \Delta w$ and Δ is invertible (it has a strictly dominant diagonal), we have

$$\Delta^{-1}E[v] = \Delta^{-1}E[\Delta w] = E[w] = \Delta^{-1}C(\beta), \tag{A5}$$

where we write $C(\beta)$ to be the column vector with row i equal to $C(\beta_i)$. Next let $\mathbf{1} \in \mathcal{R}^{N\times 1}$ denote the column vector of ones. Then the principal's objective becomes

$$\begin{aligned} E\left[\sum_{i\in\mathcal{N}} q_i - w_i\right] &= \mathbf{1}'E[q-w] = \mathbf{1}'E[q] - \mathbf{1}'\Delta^{-1}C(\beta) = \mathbf{1}'\left(E[q] - C(\beta)\right) \\ &= E\left[\sum_{i\in\mathcal{N}} q_i - C(\beta_i)\right], \end{aligned} \tag{A6}$$

where the penultimate step follows because μ is doubly stochastic ($\mu\mathbf{1} = \mathbf{1}$ from (1) and from (4) μ is symmetric, hence $\mathbf{1}'\mu = \mathbf{1}'$) and therefore Δ and Δ^{-1} are doubly stochastic (Δ is a convex combination of $\mathbf{I}$ and μ, which are both doubly stochastic, and $\mathbf{1} = \Delta^{-1}\Delta\mathbf{1} = \Delta^{-1}\mathbf{1}$), and hence $\mathbf{1}'\Delta^{-1} = \mathbf{1}'$. Because $E[q_i] = c_0 + a_i$ also depends only on β_{ii}, the solution β^* to (A6) does not depend on δ, and so the solution matches the RPE case with $\delta = 0$; i.e., $\beta^* = y^*$. The optimal contract y^S then follows from $\beta^* = \Delta y^S$. ■

PROOF OF COROLLARY A: From (23), $w^* = y^* q = \Delta y^S q = \Delta w^S$ and given the definition of Δ, we have

$$(1-\delta)w^* = w^S - \delta\mu w^S = w^S - \delta w^S_{-i}, \tag{A7}$$

which is equivalent to (24). Pre-multiplying by $\mathbf{1}'$, and using $\mathbf{1}'\mu = \mathbf{1}'$, shows that the aggregate wages and the wage benchmarks are all identical. ■

PROOF OF EXAMPLE: EQUAL WEIGHTS: With equal weights,

$$\Delta = \frac{1+\frac{1}{n}\delta}{1-\delta}\mathbf{I} - \frac{\frac{1}{n}\delta}{1-\delta}\mathbf{1}\mathbf{1}', \; \Delta^{-1} = \frac{1-\delta}{1+\frac{1}{n}\delta}\mathbf{I} + \frac{\frac{1}{n}\delta}{1+\frac{1}{n}\delta}\mathbf{1}\mathbf{1}', \text{ and } y^* = y^*_{ii}\left[\left(1+\tfrac{1}{n}\bar{\theta}\right)\mathbf{I} - \tfrac{1}{n}\bar{\theta}\mathbf{1}\mathbf{1}'\right],$$

and the result then follows from (23). ■

Proof of Lemma 2: Let Δ_I be the sub-matrix of Δ formed from rows $i \in I$, and Δ_J be the matrix formed from rows $J = \mathcal{N} \setminus I$, and let Δ_{II} and Δ_{IJ} select the corresponding rows *and* columns (including column 0). Using similar notation to select the rows of β and y, because $\beta = \Delta y$ we have

$$\beta_I = \Delta_I y = \begin{bmatrix} \Delta_{II} & \Delta_{IJ} \end{bmatrix} \begin{bmatrix} y_I \\ y_J \end{bmatrix} = \Delta_{II} y_I + \Delta_{IJ} y_J , \tag{A8}$$

which we can rearrange as,

$$\Delta_{II} y_I = \beta_I - \Delta_{IJ} y_J . \tag{A9}$$

Because Δ_{II} is invertible (it has a strictly dominant diagonal), we can therefore solve for y_I in terms of β_I and y_J as follows:

$$y_I = \left(\Delta_{II}\right)^{-1} \left[\beta_I - \Delta_{IJ} y_J\right]. \tag{A10}$$

Equation (A10) shows that the principal for team I can implement any set of exposures β_I with an appropriate set of contracts y_I. Using the definitions $w = yq$ and $\beta q = \Delta yq = \Delta w = v$, together with the binding participation constraint (A4), we have

$$\begin{aligned} E\left[w_I\right] &= E\left[y_I q\right] = \left(\Delta_{II}\right)^{-1} E\left[\beta_I q - \Delta_{IJ} y_J q\right] = \left(\Delta_{II}\right)^{-1} E\left[v_I - \Delta_{IJ} y_J q\right] \\ &= \left(\Delta_{II}\right)^{-1} \left[C(\beta_I) - \Delta_{IJ} E\left[w_J\right]\right] \end{aligned} \tag{A11}$$

Because $\Delta_{IJ} \le 0$, (A11) reflects the fact that an increase in the wage of agent $j \notin I$ raises the cost of compensating agent $i \in I$. Because Δ_{II} has constant row (and column) sums equal to $\left(\frac{1-\mu_I \delta}{1-\delta}\right)$, the principal's expected profit π_I is equal to

$$E\left[\sum_{i\in I} q_i - w_i\right] = \mathbf{1}'E\left[q_I\right] - \mathbf{1}'\left(\Delta_{II}\right)^{-1}\left[C(\beta_I) - \Delta_{IJ} y_J E\left[q\right]\right]$$
$$= \mathbf{1}'E\left[q_I\right] - \left(\frac{1-\delta}{1-\mu_I\delta}\right)\left[\mathbf{1}'C(\beta_I) - \mathbf{1}'\Delta_{IJ} y_J E\left[q\right]\right]$$
$$= \left(\mathbf{1}' + \left(\frac{1-\delta}{1-\mu_I\delta}\right)\mathbf{1}'\Delta_{IJ} y_{JI}\right)E\left[q_I\right] - \left(\frac{1-\delta}{1-\mu_I\delta}\right)\left[\mathbf{1}'C(\beta_I) - \mathbf{1}'\Delta_{IJ} y_{JJ} E\left[q_J\right]\right]$$
$$= \left(\frac{1-\delta}{1-\mu_I\delta}\right)\left(\left(\left(\frac{1-\mu_I\delta}{1-\delta}\right)\mathbf{1}' + \mathbf{1}'\Delta_{IJ} y_{JI}\right)E\left[q_I\right] - \mathbf{1}'C(\beta_I) + \mathbf{1}'\Delta_{IJ} y_{JJ} E\left[q_J\right]\right)$$

Finally, removing terms that do not depend on y_I, maximizing π_I is equivalent to maximizing

$$\sum_{i\in I}\left(\left(\frac{1-\mu_I\delta}{1-\delta}\right) + \sum_{\hat{ij}} \Delta_{\hat{ij}} y_{ji}\right)a_i - C(\beta_i)$$
$$= \sum_{i\in I}\left(1 + \frac{\delta}{1-\delta}\left(1 - \mu_I - \sum_{\hat{ij}} \mu_{\hat{ij}} y_{ji}\right)\right)a_i - C(\beta_i) \qquad \text{(A12)}$$
$$= \sum_{i\in I}\left(1+\alpha_i\right)a_i - C(\beta_i)$$

Hence, principal I will choose β_i to maximize $\left(1+\alpha_i\right)a_i - C(\beta_i)$. From the incentive constraint $a_i = \beta_{ii}/k$, β_{ii} maximizes

$$(1+\alpha_i)\beta_{ii}/k - C(\beta_i). \qquad \text{(A13)}$$

As in ***LEMMA 1***, at the optimum $\beta_{ij}^M = -\frac{1}{n}\overline{\theta}\beta_{ii}^M$ to minimize variance, but the agent's sensitivity to his own output is distorted by the factor α_i, with $\beta_{ii}^M = (1+\alpha_i)y_{ii}^*$. The optimal contract for team I then follows from (A10). ■

PROOF OF PROPOSITION II: In a symmetric equilibrium, $\alpha_i = \alpha$, and ***LEMMA 2*** implies the principals choose sensitivities $\beta^M = (1+\alpha)y^*$. Because $\beta^M = \Delta y^M$,

$$y^M = \Delta^{-1}\beta^M = \Delta^{-1}(1+\alpha)y^* = (1+\alpha)y^S. \qquad \text{(A14)}$$

Conditions (28) and (29) imply that $\mu_I = 1-\mu_e(N-N_I)$ and also that $\Delta_{ij}^{-1} = \lambda_e$ is constant for agents on different teams. Because $\Delta^{-1}\Delta = \mathbf{I}$,[38]

$$\lambda_e = \frac{\delta\mu_e}{1-\delta(1-N\mu_e)} \in \left[0, \frac{1}{N}\right). \tag{A15}$$

Therefore, from $y^S = \Delta^{-1}y^*$,

$$\frac{y_e^S}{y_{ii}^*} = \left(\lambda_e - \tfrac{1}{n}\bar{\theta}(1-\lambda_e)\right) = \frac{\delta\mu_e\left(1-\bar{\theta}\right) - \frac{1}{n}\bar{\theta}(1-\delta)}{1-\delta+N\delta\mu_e} \in \left[-\tfrac{1}{n}\bar{\theta}, \lambda_e\right). \tag{A16}$$

As an aside, note that $y_e^S > 0$ iff

$$\frac{\delta}{1-\delta} > \frac{\bar{\theta}}{n\mu_e\left(1-\bar{\theta}\right)}. \tag{A17}$$

Finally, from (27),

$$\alpha = \frac{\delta}{1-\delta}\left(1-\mu_I - \sum_{\hat{i}\in I, j\notin I}\mu_e y_e^M\right) = \frac{\delta}{1-\delta}\left(1-\mu_I\right)\left(1-N_I(1+\alpha)y_e^S\right), \tag{A18}$$

and (31) follows by solving for $1+\alpha$. Because $N_I y_e^S < N_I / N < 1$, we have $\alpha > 0$. The expected wages and profit per agent follow because, by symmetry, $E\left[w_i\right] = E\left[v_i\right] = C_i$, and

$$\begin{aligned} \pi_i^M = E[q_i - w_i] &= q_0 + a_i^M - \Psi(a_i^M) - \tfrac{1}{2}\lambda Var(c_i) - c_0 \\ &= q_0 + (1+\alpha)a_i^* - \Psi((1+\alpha)a_i^*) - \tfrac{1}{2}\lambda(1+\alpha)^2 y_{ii}^{*2}\bar{\sigma}^2 - c_0 \\ &= \pi_i^* - \frac{y_{ii}^*}{2k}\alpha^2 = \pi_i^* - \tfrac{1}{2}a_i^*\alpha^2. \end{aligned} \tag{A19}$$

Again, the constant component of the wage contract is set so that the agent's participation constraint binds:

---

[38] We can write $\Delta = A - \frac{\delta}{1-\delta}\mu_e\mathbf{11'}$ and $\Delta^{-1} = B + \lambda_e\mathbf{11'}$ where both A and B are block diagonal with $AB = I$. Because $A\mathbf{1} = \Delta\mathbf{1} + \frac{\delta}{1-\delta}\mu_e\mathbf{11'1} = \mathbf{1} + \frac{\delta}{1-\delta}\mu_e N\mathbf{1}$ and $A^{-1}\mathbf{1} = B\mathbf{1} = \Delta^{-1}\mathbf{1} - \lambda_e\mathbf{11'1} = \mathbf{1} - \lambda_e N\mathbf{1}$, the value of λ_e follows from the fact that we must have $1 + \frac{\delta}{1-\delta}\mu_e N = \left(1-\lambda_e N\right)^{-1}$.

$$y_{i0}^M q_0 = c_0 + \tfrac{1}{2} a_i^* (1+\alpha)^2 - k a_i^* (1+\alpha)(1-\overline{\theta})(q_0 + a_i^*(1+\alpha)).$$

■

PROOF OF PROPOSITION III: From (31), as $\delta \to 1$, $1+\alpha \to \frac{1}{N_I y_e^S}$. For $\mu_e > 0$ (equivalently, $N_I < N$ and $\mu_I < 1$), from (A16), $y_e^S \to \frac{1}{N}(1-\overline{\theta}) y_{ii}^*$. Combining these results and using $a^* = y_{ii}^* / k$, we have

$$(1+\alpha)a^* \to \frac{N}{N_I(1-\overline{\theta})} \frac{1}{k}, \tag{A20}$$

which implies that effort will exceed the first-best level $1/k$ for δ sufficiently close to 1.

To establish monotonicity in δ, from (31) it suffices to show

$$\frac{\partial}{\partial \delta}\left[\delta\left(\frac{1-\mu_I}{1-\delta\mu_I}\right)(1-N_I y_e^S)\right] > 0. \tag{A21}$$

Plugging in for y_e^S from (A16), taking the derivative, and simplifying we can show

$$\overbrace{\left[\frac{(N-1)(1-\delta+\delta N\mu_e)^2(1-\delta\mu_I)^2}{1-\mu_I}\right]}^{A>0} \frac{\partial}{\partial \delta}\left[\delta\left(\frac{1-\mu_I}{1-\delta\mu_I}\right)(1-N_I y_e^S)\right] =$$
$$\underbrace{\left((1-\delta)(N-1+N_I y_{ii}^*\overline{\theta}) + (N-1)\delta(N-N_I y_{ii}^*(1-\overline{\theta}))\mu_e\right)(1-\delta+\delta N\mu_e)}_{B>0} \tag{A22}$$
$$-\underbrace{N_I y_{ii}^* \delta(N-1+\overline{\theta})\mu_e(1-\delta\mu_I)}_{C\geq 0}$$

Because $0 \leq \delta, y_{ii}^*, \overline{\theta}, \mu_I \leq 1$, the terms A, B, and C are positive; we thus need to show that $B - C \geq 0$. Note that a lower bound on B is obtained by setting $\overline{\theta} = 0$ and $y_{ii}^* = 1$, and an upper bound for C is obtained by setting $\overline{\theta} = y_{ii}^* = 1$. These substitutions, together with the fact that $(N-N_I)\mu_e = 1-\mu_I$, imply

$$
\begin{aligned}
B-C &\geq (N-1)\big(1-\delta+\delta(N-N_I)\mu_e\big)(1-\delta+\delta N\mu_e)-N_I\delta N\mu_e(1-\delta\mu_I) \\
&=(N-1)(1-\delta+\delta N\mu_e)(1-\delta\mu_I)-NN_I\delta\mu_e(1-\delta\mu_I)
\end{aligned}
\tag{A23}
$$

Dividing by $(1-\delta\mu_I)$ maintains the sign, and yields

$$
(N-1)(1-\delta)+N\delta(N-1-N_I)\mu_e>0, \tag{A24}
$$

where the final inequality holds given $N_I<N-1$ with more than one principal.

To show monotonicty in N_I, we show $1-\delta\left(\frac{1-\mu_I}{1-\delta\mu_I}\right)(1-N_I y_e^S)$ increases in N_I , where we of course need to keep in mind that μ_I and y_e^S depend on N_I as well. Furthermore, it suffices to focus on $N_I \leq \frac{1}{2}N$ since for $N_I=N$, $\alpha=0$. Taking the derivative with respect to N_I and multiplying by

$$
(\delta\mu_e)^{-1}\big(1-\delta+(N-N_I)\delta\mu_e\big)^2(N-1)(1-\delta+\delta N\mu_e)>0 \tag{A25}
$$

and simplifying yields

$$
\begin{aligned}
&\big[(N-1)(1-\delta)(1-\delta+\delta N\mu_e)\big] \\
&\quad + y_{ii}^*\Big((N-2N_I)(1-\delta)+(N-N_I)^2\,\delta\mu_e\Big)\Big((N-1)\delta\mu_e-\overline{\theta}\big(1-\delta+(N-1)\delta\mu_e\big)\Big)
\end{aligned}
\tag{A26}
$$

Note that the square bracket is positive. Since $N \geq 2N_I$, $\Big((N-2N_I)(1-\delta)+(N-N_I)^2\,\delta\mu_e\Big)\geq 0$ and is decreasing in N_I. Hence, if $\Big((N-1)\delta\mu_e-\theta\big(1-\delta+(N-1)\delta\mu_e\big)\Big)\geq 0$, we are done. Otherwise, a lower bound for the whole expression is obtained by setting $y_{ii}^*=\overline{\theta}=N_I=1$, which after simplification yields

$$
\begin{aligned}
&\big[(N-1)(1-\delta)(1-\delta+\delta N\mu_e)\big]-(1-\delta)\Big((N-2)(1-\delta)+(N-1)^2\,\delta\mu_e\Big) \\
&\quad=(N-1)(1-\delta)\left[\left(\frac{1}{N-1}\right)(1-\delta)+\delta\mu_e\right]>0
\end{aligned}
\tag{A27}
$$

Finally, we show $1-\delta\left(\frac{1-\mu_I}{1-\delta\mu_I}\right)\left(1-N_I y_e^S\right)$ decreases in μ_e. Taking the derivative with respect to μ_e, multiplying by

$$\left[\left(N-N_I\right)\left(1-\delta\right)\delta\right]^{-1}\left(N-1\right)\left(1-\delta+\left(N-N_I\right)\delta\mu_e\right)^2\left(1-\delta+\delta N\mu_e\right)^2>0 \tag{A28}$$

and simplifying yields

$$\begin{aligned}&-\left(N-1\right)\left(1-\delta+\delta N\mu_e\right)^2\\&+y_{ii}^* N_I\left[\begin{aligned}&\left(N-1\right)\delta\mu_e\left(2\left(1-\delta\right)+\delta\mu_e\left(2N-N_I\right)\right)\\&-\bar{\theta}\underbrace{\left(1-2\delta+\delta^2+2\left(N-1\right)\left(1-\delta\right)\delta\mu_e+\left(N_I+\left(N-2\right)N\right)\delta^2\mu_e^2\right)}_{>0}\end{aligned}\right]\end{aligned} \tag{A29}$$

Because the first term is negative, if the square bracket term is negative as well we are done. We can maximize the above by setting $\bar{\theta}=0$ and $N_I=1$ inside $[\cdot]$, and $y_{ii}^*=1$ and $N_I=N/2$ outside $[\cdot]$. Substituting and simplifying yields

$$\begin{aligned}&\left(N-1\right)\left[-\left(1-\delta+\delta N\mu_e\right)^2+\delta N\mu_e\left(1-\delta+\delta N\mu_e-\tfrac{1}{2}\delta\mu_e\right)\right]\\&<-\left(N-1\right)\left(1-\delta+\delta N\mu_e\right)\left(1-\delta\right)<0\end{aligned} \tag{A30}$$

■

PROOF OF COROLLARY B: The limiting case is immediate from (31) and $\left|y_e^S\right|\leq\frac{1}{n}$ from (A16). Also from (31), α is increasing in N iff y_e^S is decreasing. As N increases, $y_e^S\to 0$, and the highest order term in N of $\frac{\partial}{\partial N}y_e^S$ is proportional to $N^{-2}\left(\bar{\theta}\left(1-\delta\mu_I\right)-\delta\left(1-\mu_I\right)\right)<0$. Hence α increases with N for N sufficiently large if $\bar{\theta}<\delta\frac{\left(1-\mu_I\right)}{1-\delta\mu_I}$ (it can also be shown if this holds as an equality by checking the N^{-1}-order term). Otherwise, the convergence is either decreasing or hump-shaped. ■

Proof of Proposition IV: From ***Lemma 2***, because $\mu_I + N_J\mu_e = 1$ and, in equilibrium, $y_{ji} = (1+\alpha_j)y_e^S$,

$$\begin{aligned}\alpha_i &\equiv \frac{\delta}{1-\delta}\left(1-\mu_I - \sum_{\hat{i}\in I, j\notin I}\mu_{\hat{i}j}y_{ji}\right) = \frac{\delta}{1-\delta}\left(N_J\mu_e - N_I N_J \mu_e\left(1+\alpha_j\right)y_e^S\right) \\ &= \underbrace{\frac{\delta}{1-\delta}\left(N_J\mu_e - N_I N_J\mu_e y_e^S\right)}_{A_j} - \underbrace{\frac{\delta}{1-\delta}N_I N_J \mu_e y_e^S}_{B}\alpha_j\end{aligned} \tag{A31}$$

and symmetrically for α_j. Thus,

$$\begin{aligned}\alpha_i - \alpha_j &= A_i - A_j + B\left(\alpha_i - \alpha_j\right) \\ &= \frac{A_i - A_j}{1-B} = \frac{\delta\mu_e}{1-\delta}\frac{N_J - N_I}{1-B} > 0,\end{aligned} \tag{A32}$$

where the final inequality follows because $N_J > N_I$ (by assumption) and $B < 1$ as we will show below. Next note that

$$\begin{aligned}\alpha_i + \alpha_j &= A_i + A_j - B\left(\alpha_i + \alpha_j\right) \\ &= \frac{A_i + A_j}{1+B} > 0,\end{aligned} \tag{A33}$$

where the last inequality follows because A_i and A_j are both positive (which follows from $y_e^S \le 1/N$ from (A16)), and $B > -1$ as we show next. Together, (A32) and (A33) establish $\alpha_i > \left|\alpha_j\right|$, and the ranking of expected profits follows by the logic of (33).

Finally, we need to confirm $\left|B\right| < 1$. From (A16),

$$\mu_e y_e^S = \mu_e\left(\frac{\delta\mu_e}{1-\delta(1-N\mu_e)}\left(1+\tfrac{1}{n}\overline{\theta}\right) - \tfrac{1}{n}\overline{\theta}\right)y_{ii}^*, \tag{A34}$$

which is minimized by setting $\overline{\theta} = y_{ii}^* = 1$ and $\mu_e = 1/N$; hence,

$$\mu_e y_e^S \ge \frac{1}{N}\left(\frac{\delta}{N}\left(\frac{N}{N-1}\right) - \frac{1}{N}\right) = \frac{-(1-\delta)}{N\left(N-1\right)}. \tag{A35}$$

Therefore

$$B = \frac{\delta}{1-\delta} N_I N_J \mu_e y_e^S \geq -\delta \frac{N_I N_J}{N(N-1)} > -1. \tag{A36}$$

Next note that (A34) is maximized by setting $\overline{\theta} = 0, y_{ii}^* = 1,$ and $\mu_e = 1/N$; hence, $\mu_e y_e^S \leq \frac{\delta}{N^2}$.

Therefore,

$$B = \frac{\delta}{1-\delta} N_I N_J \mu_e y_e^S \leq \frac{\delta^2}{1-\delta} \frac{N_I N_J}{N^2} < \frac{\frac{1}{4}\delta^2}{1-\delta} < 1, \tag{A37}$$

where the last inequality holds for $\delta < 2(\sqrt{2}-1) \approx 0.828$. ■

PROOF OF PROPOSITION V: We can write y^* as

$$y^* = y_{ii}^* \left[\left(1 + \tfrac{1}{n}\overline{\theta}\right)\mathbf{I} - \tfrac{1}{n}\overline{\theta}\mathbf{1}\mathbf{1}' \right], \tag{A38}$$

and because $\Delta\mathbf{1} = \mathbf{1} \Rightarrow \Delta^{-1}\mathbf{1} = \mathbf{1}$, we have

$$y_{ij}^S = \left[\Delta^{-1} y^*\right]_{ij} = y_{ii}^* \left[\left(1 + \tfrac{1}{n}\overline{\theta}\right)\Delta^{-1} - \tfrac{1}{n}\overline{\theta}11' \right]_{ij} = y_{ii}^* \left(\left(1 + \tfrac{1}{n}\overline{\theta}\right)\Delta_{ij}^{-1} - \tfrac{1}{n}\overline{\theta} \right). \tag{A39}$$

Next note that $\Delta \equiv \frac{1}{1-\delta}\left[\mathbf{I} - \delta\mu\right]$ implies

$$\Delta^{-1} = (1-\delta)\mathbf{I} + \delta\mu\Delta^{-1} = (1-\delta)\left[\mathbf{I} + \delta\mu + \delta^2\mu^2 + \ldots\right], \tag{A40}$$

and therefore Δ^{-1} is a symmetric stochastic matrix and, if $\delta > 0$, strictly positive (since μ is irreducible). Equations (A39) and (A40) thus imply $y_{ij}^S / y_{ii}^* \in [-\tfrac{1}{n}\overline{\theta}, 1]$, and $y_{ii}^S > y_{ij}^S$ follows from $\Delta_{ii}^{-1} > \Delta_{ij}^{-1}$, which we establish below.

For the overall peer sensitivity, $y^S\mathbf{1} = \Delta^{-1} y^* \mathbf{1} = y_{ii}^* \left[\left(1 + \tfrac{1}{n}\overline{\theta}\right) - \tfrac{n+1}{n}\overline{\theta} \right]\mathbf{1}$ implies

$$\phi_i^S \equiv \frac{\sum_{j\in\mathcal{N}_{-i}} y_{ij}^S}{y_{ii}^S} = \frac{\sum_j y_{ij}^S}{y_{ii}^S} - 1 = \frac{(y^s 1)_i}{\left(\Delta^{-1} y^*\right)_{ii}} - 1 = \frac{\left(1+\frac{1}{n}\overline{\theta}\right) - \frac{n+1}{n}\overline{\theta}}{\left(1+\frac{1}{n}\overline{\theta}\right)\Delta_{ii}^{-1} - \frac{1}{n}\overline{\theta}} - 1$$
$$= \frac{1-\overline{\theta}}{\left(1+\frac{1}{n}\overline{\theta}\right)\Delta_{ii}^{-1} - \frac{1}{n}\overline{\theta}} - 1 \tag{A41}$$

Thus the behavior of ϕ_i^S is determined by Δ_{ii}^{-i}. Specifically, $-\overline{\theta} < \phi_i^S < n$ is equivalent to $1 > \Delta_{ii}^{-1} > 1/N$. Clearly $\Delta_{ii}^{-1} < 1$ if $\delta > 0$ (since from above, Δ^{-1} is stochastic and strictly positive).

To bound Δ_{ii}^{-1} from below, we show next that the diagonal elements of Δ^{-1} are larger than the off-diagonal elements. Suppose instead that there exists $\hat{j} \neq i$ such that $\Delta_{i\hat{j}}^{-1} = \max_j \Delta_{ij}^{-1}$. But then $\delta\left(\Delta^{-1}\mu\right)_{i\hat{j}} = \delta\sum_k \Delta_{ik}^{-1}\mu_{k\hat{j}} \leq \delta \max_k \Delta_{ik}^{-1} < \Delta_{i\hat{j}}^{-1}$, contradicting $\Delta^{-1} = \left(1-\delta\right)I + \delta\Delta^{-1}\mu$. Hence $\Delta_{ii}^{-1} = \max_j \Delta_{ij}^{-1} > 1/N$.

In the limit when $\delta = 1$, (A40) implies $\Delta^{-1} = \mu\Delta^{-1} = \mu^t\Delta^{-1}$. Because μ is ergodic and doubly stochastic, $\mu_{ij}^t \to 1/N$ as $t \to \infty$, and therefore $\Delta_{ij}^{-1} \to 1/N$. Thus, $y_{ij}^S \to y_{ii}^*(1-\overline{\theta})/N$.

Finally, from (A40),

$$\frac{\partial}{\partial\delta}\Delta^{-1} = \frac{\partial}{\partial\delta}\left[\left(1-\delta\right)\mathbf{I} + \delta\mu\Delta^{-1}\right] = -\left(\mathbf{I} - \mu\Delta^{-1}\right) + \delta\mu\frac{\partial}{\partial\delta}\Delta^{-1}$$
$$= -\left(\mathbf{I} - \delta\mu\right)^{-1}\left(\mathbf{I} - \mu\Delta^{-1}\right) \tag{A42}$$
$$= -\underbrace{\left(\frac{1}{1-\delta}\Delta^{-1}\right)}_{\substack{\text{from} \\ \Delta\equiv\frac{1}{1-\delta}[I-\delta\mu]}}\underbrace{\left(\frac{1}{\delta}\left(\mathbf{I} - \Delta^{-1}\right)\right)}_{\substack{\text{from} \\ \Delta^{-1}=(1-\delta)I+\delta\mu\Delta^{-1}}} = \frac{-1}{\delta(1-\delta)}\left(\Delta^{-1} - \Delta^{-1}\Delta^{-1}\right)$$

Because $\Delta_{ii}^{-1} = \max_j \Delta_{ij}^{-1} > \min_j \Delta_{ij}^{-1}$ for $\delta \in (0,1)$, we have

$$\left(\Delta^{-1} - \Delta^{-1}\Delta^{-1}\right)_{ii} = \Delta_{ii}^{-1} - \sum_j \left(\Delta_{ij}^{-1}\right)^2$$
$$> \Delta_{ii}^{-1} - \left(\max_j \Delta_{ij}^{-1}\right)\underbrace{\sum_j \Delta_{ij}^{-1}}_{=1} \tag{A43}$$
$$= \Delta_{ii}^{-1} - \Delta_{ii}^{-1} = 0$$

and thus $\frac{\partial}{\partial\delta}\Delta_{ii}^{-1} < 0$. ■

PROOF OF COROLLARY C: From (A40), $\Delta^{-1} = \left(1-\delta\right)\mathbf{I} + \delta\mu\Delta^{-1}$ implies $\Delta_{ii}^{-1} \in \left[1-\delta, 1-\delta+\delta\left\|\mu\right\|_{max}\right]$. Then from (38),

$$\phi_i^S = \frac{1-\overline{\theta}}{\left(1+\frac{1}{n}\overline{\theta}\right)\Delta_{ii}^{-1} - \frac{1}{n}\overline{\theta}} - 1 \rightarrow \frac{1-\overline{\theta}}{1-\delta} - 1, \tag{A44}$$

and the result follows. ■

PROOF OF COROLLARY D: From (32), $y_e^S = \dfrac{\delta\mu_e\left(1-\overline{\theta}\right) - \frac{1}{n}\overline{\theta}\left(1-\delta\right)}{1-\delta+N\delta\mu_e} y_{ii}^*$ which is increasing in μ_e.

Also, $y_e^M = (1+\alpha^M)y_e^S$, and α^M is increasing in μ_e from PROPOSITION III. ■

PROOF OF PROPOSITION VI: Let $n_I = N_I - 1$. With uniform internal and external weights, the symmetric matrix μ can be diagonalized with three distinct eigenvalues:

$$\{e_1, e_a, e_b\} = \{1,\ \overbrace{\mu_I - N_I\mu_e}^{\text{multiplicity } \frac{N}{N_I}-1},\ \overbrace{-\mu_t = -\mu_I / n_I}^{\text{multiplicity } \frac{N}{N_I}n_I}\}.$$

It is straightforward to show that Δ^{-1} has the same structure with eigenvalues

$$\left\{1,\ a = \frac{1-\delta}{1-\delta e_a},\ b = \frac{1-\delta}{1-\delta e_b}\right\},$$

and thus has corresponding external, internal, and diagonal elements:

$$\Delta_e^{-1} = \frac{1-a}{N},\ \Delta_t^{-1} = \Delta_e^{-1} + \frac{a-b}{N_I},\ \Delta_{ii}^{-1} = \Delta_t^{-1} + b. \tag{A45}$$

Equilibrium contracts are given by $y^M = \left(1+\alpha\right)y^S = \left(1+\alpha\right)\Delta^{-1}y_{ii}^*$, with external, internal, and own exposures

$$\frac{y_x^S}{y_{ii}^*} = \left(1+\frac{1}{n}\overline{\theta}\right)\Delta_x^{-1} - \frac{1}{n}\overline{\theta},\ \text{for } x \in \{e, t, ii\}. \tag{A46}$$

Note that $e_a - e_b = \mu_I - N_I\mu_e + \mu_I / n_I = N_I\left(\mu_I / n_I - \mu_e\right) = N_I\left(\mu_t - \mu_e\right) \geq 0$. Therefore, $a \geq b$ and hence $\Delta_t^{-1} \geq \Delta_e^{-1}$. Then from (A46), $y_t^S \geq y_e^S$ and this inequality is strict unless $\mu_t = \mu_e$ or $\delta = 0$. The result in (44) then follows from $Y_{Team} = \left(1+\alpha\right)N_I\left(y_t^S - y_e^S\right)$. Similar logic implies $Y_{Own} > 0$.

Because $\Delta_e^{-1} = \frac{1-a}{N} = \frac{\delta\left(1-e_a\right)}{N\left(1-\delta e_a\right)} = \frac{\delta\mu_e}{1-\delta+N\delta\mu_e}$, we have

$$y_e^S = \frac{\delta\mu_e\left(1-\overline{\theta}\right) - \frac{1}{n}\overline{\theta}\left(1-\delta\right)}{1-\delta+N\delta\mu_e} y_{ii}^* ,$$

which is increasing in μ_e. Because the rat-race factor α is also increasing in μ_e from ***Proposition III***, y_e^M is also increasing in μ_e.

Note that

$$\tfrac{\partial}{\partial\mu_e}\Delta_{ii}^{-1} = \left(\frac{1}{N_I} - \frac{1}{N}\right)a' + \left(1 - \frac{1}{N_I}\right)b' = \left(\frac{1}{N_I} - \frac{1}{N}\right)\frac{\delta\left(1-\delta\right)}{\left(1-\delta e_a\right)^2}e_a{}' + \left(\frac{n_I}{N_I}\right)\frac{\delta\left(1-\delta\right)}{\left(1-\delta e_b\right)^2}e_b{}' ,$$

and because $e_b \leq e_a \leq 1$, $e_a{}' = -N$, and $e_b{}' = (N - N_I)/n_I$,

$$\tfrac{\partial}{\partial\mu_e}\Delta_{ii}^{-1} = \delta\left(1-\delta\right)\left(\frac{N}{N_I} - 1\right)\left(\frac{1}{\left(1-\delta e_b\right)^2} - \frac{1}{\left(1-\delta e_a\right)^2}\right) \leq 0,$$

with equality if and only if $\mu_e = 1/n$. Hence (38) implies ϕ_i is increasing in μ_e.

When $\mu_e = 0$, then $\mu_I = e_a = a = 1$ and so $\Delta_e^{-1} = 0$ and $y_e = -\frac{1}{n}\overline{\theta}$. Because $\alpha = 0$ and $\Delta_t^{-1} = \frac{1-b}{N_I} = \frac{\delta\left(1-e_b\right)}{N_I\left(1-\delta e_b\right)} = \frac{\delta\left(1+\frac{1}{n_I}\right)}{N_I\left(1+\frac{1}{n_I}\delta\right)} = \frac{\delta}{n_I+\delta}$, the optimal contract has team weights

$$y_t^M = y_t^S = \left(\frac{\delta - \frac{n_I}{n}\overline{\theta}}{n_I + \delta}\right)y_{ii}^* .$$

■

Proof of Proposition VII: Let $d(i,j)$ be the distance between i and j on the circle. Then we can check that

$$\Delta_{ij}^{-1} = \sqrt{\frac{1-\delta}{1+\delta}}\left(\frac{z^{d(i,j)} + z^{N-d(i,j)}}{1-z^N}\right) \text{ where } z = \frac{1}{\delta} - \sqrt{\frac{1}{\delta^2} - 1} \in (0,1). \tag{A47}$$

Therefore, $\Delta_{ij}^{-1} > 0$ is decreasing in $d(i,j)$, and thus

$$y_{ij}^S = \sum_k \Delta_{ik}^{-1} y_{ij}^* = \sum_k \Delta_{ik}^{-1}\left(-\tfrac{1}{n}\overline{\theta} y_{ii}^* + \left(1+\tfrac{1}{n}\overline{\theta}\right) y_{ii}^* \mathbf{1}_{k=j}\right) = -\tfrac{1}{n}\overline{\theta} y_{ii}^* + \Delta_{ij}^{-1}\left(1+\tfrac{1}{n}\overline{\theta}\right) y_{ii}^*$$

is decreasing in $d(i,j)$ and exceeds the RPE benchmark $-\tfrac{1}{n}\overline{\theta} y_{ii}^*$. Note also that $y_{ij}^S > 0$ iff

$$\Delta_{ij}^{-1} > \frac{\overline{\theta}}{n+\overline{\theta}},$$

which holds for sufficiently large n since $\Delta_{ij}^{-1} \to \sqrt{\frac{1-\delta}{1+\delta}} z^{d(i,j)} > 0$ as $n \to \infty$. Moreover, if we define d_N^0 as the maximum distance for which the wage sensitivity is positive, then d_N^0 grows at rate $\ln(N)$.

For the total sensitivity, using (38) we have $\phi_i = \frac{1-\overline{\theta}}{\left(1+\frac{1}{n}\overline{\theta}\right)\Delta_{ii}^{-1} - \frac{1}{n}\overline{\theta}} - 1 \to \frac{1-\overline{\theta}}{\Delta_{ii}^{-1}} - 1$. Then since

$$\Delta_{ii}^{-1} = \sqrt{\frac{1-\delta}{1+\delta}}\left(\frac{1+z^N}{1-z^N}\right) \to \sqrt{\frac{1-\delta}{1+\delta}} \text{ as } N \to \infty,$$

we have

$$\lim_{N\to\infty} \phi_i > 0 \text{ iff } \overline{\theta} < 1 - \sqrt{\frac{1-\delta}{1+\delta}}. \tag{A48}$$

■

Proof of Lemma 3: Note first that $\mathbf{1}'\Delta = \frac{1}{1-\delta}\mathbf{1}'\left[\mathbf{I} - \delta\mu\right] = \frac{1}{1-\delta}\left[\mathbf{1}' - \delta\hat{\mu}\right]$; therefore $\hat{\mu} = \mathbf{1}'$ if and only if $\mathbf{1}'\Delta = \mathbf{1}'$. Next, $\mathbf{1}'\Delta = \mathbf{1}' \Leftrightarrow \mathbf{1}'\Delta^{-1} = \mathbf{1}'$, and so $\hat{\mu} = \mathbf{1}'$ if and only if $\omega = \mathbf{1}'$. Moreover, because $\omega\mathbf{1} = \mathbf{1}'\Delta^{-1}\mathbf{1} = \mathbf{1}'\mathbf{1} = N$, so that the average ω_i equals 1. Finally, because $v = \Delta w$, the total wage bill to provide v is equal to $\mathbf{1}'w = \mathbf{1}'\Delta^{-1}v = \omega v$. ■

Proof of Proposition VIII: Recall from (A5) that the principal maximizes

$$E\left[\sum_i q_i - w_i\right] = \mathbf{1}'E[q-w] = \mathbf{1}'E[q] - \mathbf{1}'\Delta^{-1}C(\beta) = E\left[\sum_i q_i - \omega_i C(\beta_i)\right], \tag{A49}$$

from which the optimal effort level follows. By ***Lemma 3,*** $\hat{\mu} \neq \mathbf{1}'$ implies $\omega \neq \mathbf{1}'$, and hence by Jensen's inequality, $\frac{1}{N}\sum_i \omega_i^{-1} > 1$. The principal's total expected payoff is

$$N(q_0 - c_0) + \tfrac{1}{2}a_i^*\left(\sum_j \omega_j^{-1}\right) > N\left(q_0 - c_0 + \tfrac{1}{2}a_i^*\right) = N\pi_i^*,$$

which improves upon the RPE payoff. ■

Proof of Proposition IX: In this setting, we can represent Δ in block format as $\Delta = \begin{bmatrix} \mathbf{I}_J & 0_{J,N-J} \\ \Delta_{N-J,J} & \Delta_{N-J,N-J} \end{bmatrix}$ with

$$\Delta_{N-J,J} = -\frac{1}{n}\frac{\delta}{1-\delta}\mathbf{1}_{N-J}\mathbf{1}_J', \text{ and } \Delta_{N-J,N-J} = \frac{1+\frac{1}{n}\delta}{1-\delta}\mathbf{I}_{N-J,N-J} - \frac{1}{n}\frac{\delta}{(1-\delta)}\mathbf{1}_{N-J}\mathbf{1}_{N-J}'.$$

Computing the inverse yields $\Delta^{-1} = \begin{bmatrix} \mathbf{I}_J & 0_{J,N-J} \\ \Delta^{-1}_{N-J,J} & \Delta^{-1}_{N-J,N-J} \end{bmatrix}$ with

$$\Delta^{-1}_{N-J,J} = \frac{1}{n}\frac{\delta}{1-(1-J/n)\delta}\mathbf{1}_{N-J}\mathbf{1}_J', \text{ and}$$

$$\Delta^{-1}_{N-J,N-J} = \frac{1-\delta}{1+\frac{1}{n}\delta}\mathbf{I}_{N-J,N-J} + \frac{1}{n}\frac{\delta}{\left(1+\frac{1}{n}\delta\right)}\frac{(1-\delta)}{\left(1-(1-J/n)\delta\right)}\mathbf{1}_{N-J}\mathbf{1}_{N-J}'.$$

Finally, by direct calculation, the effective cost for each agent type is

$$\omega_j = 1 + \frac{(n+1-J)\delta}{n(1-\delta)+J\delta} = \frac{n+\delta}{n(1-\delta)+J\delta} \text{ and}$$

$$\omega_m = \frac{1-\delta}{1+\frac{1}{n}\delta} + \frac{(n+1-J)\delta}{(1+\frac{1}{n}\delta)} \frac{(1-\delta)}{n(1-\delta)+J\delta} = \frac{n(1-\delta)}{n(1-\delta)+J\delta} = \frac{1-\delta}{1+\frac{1}{n}\delta}\omega_j < \omega_j .$$

Letting $\psi = \frac{1+\frac{1}{n}\delta}{1-\delta} > 1$, from ***PROPOSITION VIII*** we have $\frac{a_m^A}{a_j^A} = \frac{\omega_m^{-1}}{\omega_j^{-1}} \equiv \psi$, and

$$\begin{aligned}
\frac{1}{N}\sum_j \omega_j^{-1} &= \frac{1}{N}\left(J\omega_j^{-1} + (N-J)\omega_m^{-1}\right) = \frac{1}{N}\omega_j^{-1}\left(J + (N-J)\psi\right) \\
&= \left(\frac{n(1-\delta)+J\delta}{n+\delta}\right)\left(\frac{J}{N} + \left(1-\frac{J}{N}\right)\psi\right) \\
&= \left(\psi^{-1} + \frac{J}{N}\left(1-\psi^{-1}\right)\right)\left(\psi + \frac{J}{N}(1-\psi)\right) \\
&= 1 + \left(\psi + \psi^{-1} - 2\right)\frac{J}{N}\left(1-\frac{J}{N}\right).
\end{aligned}$$

■

PROOF OF LEMMA 4: Following the proof of ***LEMMA 2***, holding the contracts and actions of all other agents $j \in J = N_i$ fixed,

$$\begin{aligned}
E[q_i - w_i] &= E[q_i] - (\Delta_{ii})^{-1}\left[C(\beta_i) - \Delta_{iJ} y_{Ji} E[q_i] - \Delta_{iJ} y_{JJ} E[q_J]\right] \\
&= E[q_i] - \left(\frac{1-\delta}{1-\delta\mu_{ii}}\right)\left[C(\beta_i) - \Delta_{iJ} y_{Ji} E[q_i] - \Delta_{iJ} y_{JJ} E[q_J]\right] \\
&= \left(1 + \left(\frac{1-\delta}{1-\delta\mu_{ii}}\right)\Delta_{iJ} y_{Ji}\right)E[q_i] - \left(\frac{1-\delta}{1-\delta\mu_{ii}}\right)C(\beta_i) + \left(\frac{1-\delta}{1-\delta\mu_{ii}}\right)\Delta_{iJ} y_{JJ} E[q_J] \\
&= \left(1 - \frac{\delta}{1-\delta\mu_{ii}}\mu_{iJ} y_{Ji}\right)E[q_i] - \left(\frac{1-\delta}{1-\delta\mu_{ii}}\right)C(\beta_i) + \left(\frac{1-\delta}{1-\delta\mu_{ii}}\right)\Delta_{iJ} y_{JJ} E[q_J]
\end{aligned}$$

In addition, by changing $E[q_i]$, the principal also changes the expected wage to agent $\hat{i}$ by $y_{\hat{i}i}$. Therefore, letting $\hat{I} = I_i$, the principal will choose β_i to maximize

$$\left(1 - \frac{\delta}{1-\delta\mu_{ii}}\mu_{iJ} y_{Ji} - 1' y_{\hat{I}i}\right)E[q_i] - \left(\frac{1-\delta}{1-\delta\mu_{ii}}\right)C(\beta_i) = \left(\frac{1-\delta}{1-\delta\mu_{ii}}\right)\left[\left(1+\alpha_i^R\right)E[q_i] - C(\beta_i)\right],$$

for

$$1+\alpha_i^R \equiv \left(\frac{1-\delta\mu_{ii}}{1-\delta}\right)\left(1-\frac{\delta}{1-\delta\mu_{ii}}\sum_{j\neq i}\mu_{ij}y_{ji}-\sum_{\hat{i}\in I_i}y_{\hat{i}i}\right), \tag{A50}$$

which is solved by $\beta_i^R = (1+\alpha_i^R)y_i^*$. ■

Proof of Proposition X: Using (51) with symmetric α^R,

$$\begin{aligned}
1+\alpha^R &\equiv \left(\frac{1-\delta\mu_{ii}}{1-\delta}\right)\left(1-\frac{\delta}{1-\delta\mu_{ii}}\sum_{j\neq i}\mu_{ij}y_{ji}-\sum_{\hat{i}\in I_{-i}}y_{\hat{i}i}\right)\\
&=\frac{1}{1-\delta}\left(1-\delta\sum_{j\neq i}\mu_{ij}y_{ji}+y_{ii}-\sum_{\hat{i}\in I}y_{\hat{i}i}\right)=\frac{1}{1-\delta}+\Delta_{i\cdot}y_{\cdot i}-\frac{1}{1-\delta}\mathbf{1}'y_{Ii}\\
&=\frac{1}{1-\delta}+\left(1+\alpha^R\right)\left(\left(\Delta\Delta^{-1}y^*\right)_{ii}-\frac{1}{1-\delta}\sum_{\hat{i}\in I}y_{\hat{i}i}^S\right)\leftarrow \text{using } y=(1+\alpha^R)y^S \text{ and } y^S=\Delta^{-1}y^*\\
&=\frac{1}{1-\delta}+\left(1+\alpha^R\right)\left(y_{ii}^*-\frac{1}{1-\delta}\sum_{\hat{i}\in I}y_{\hat{i}i}^S\right)\\
&=\frac{1}{(1-\delta)\left(1-y_{ii}^*\right)+\sum_{\hat{i}\in I}y_{\hat{i}i}^S}\leftarrow \text{moving } \left(1+\alpha^R\right) \text{ to the left-hand side}\\
&=\frac{1}{(1-\delta)\left(1-y_{ii}^*\right)+\left(y_{ii}^*\left(1-\overline{\theta}\right)-\left(N-N_I\right)y_e^S\right)}
\end{aligned}$$

where the final step follows because

$$\sum_{\hat{i}\in I}y_{\hat{i}i}^S+\left(N-N_I\right)y_e^S=\mathbf{1}'y^S=\mathbf{1}'\Delta^{-1}y^*\overset{(\text{by symmetry of } y^*)}{=}\mathbf{1}'y^*=y^*\mathbf{1}=y_{ii}^*\left(1-\overline{\theta}\right).$$

To show that $\alpha^R>\alpha$, note from (A15) and (A16) that

$$\lambda_e=\frac{\delta\mu_e}{1-\delta(1-N\mu_e)}=\frac{\delta\left(\frac{1-\mu_I}{N-N_I}\right)}{1-\delta+\delta N\left(\frac{1-\mu_I}{N-N_I}\right)}=\frac{\delta\left(1-\mu_I\right)}{N\left(1-\delta\mu_I\right)-N_I\left(1-\delta\right)},$$

and

$$y_e^S=\frac{1}{N-1}\left(\lambda_e\left(N-1+\overline{\theta}\right)-\overline{\theta}\right)y_{ii}^*=\frac{1}{N-1}\left(\frac{\delta\left(1-\mu_I\right)\left(N-1+\overline{\theta}\right)}{N\left(1-\delta\mu_I\right)-N_I\left(1-\delta\right)}-\overline{\theta}\right)y_{ii}^*.$$

Therefore,

$$\overbrace{1-\delta\left(\frac{1-\mu_I}{1-\delta\mu_I}\right)\left(1-N_I y_e^S\right)}^{\frac{1}{1+\alpha}}-\overbrace{\left(1-\delta+\left(\delta-\overline{\theta}\right)y_{ii}^{*}-\left(N-N_I\right)y_e^S\right)}^{\frac{1}{1+\alpha^R}}$$

$$=1-\delta\left(\frac{1-\mu_I}{1-\delta\mu_I}\right)-\left(1-\delta\right)-\left(\delta-\overline{\theta}\right)y_{ii}^{*}+\left(\delta\left(\frac{1-\mu_I}{1-\delta\mu_I}\right)-1\right)N_I y_e^S+N y_e^S$$

$$=\frac{\delta\left(1-\delta\right)\mu_I}{1-\delta\mu_I}-\left(\delta-\overline{\theta}\right)y_{ii}^{*}+\left(\frac{N\left(1-\delta\mu_I\right)-N_I\left(1-\delta\right)}{1-\delta\mu_I}\right)y_e^S$$

$$=\frac{(1-\delta)\mu_I}{1-\delta\mu_I}\left(\delta-\left(\delta-\frac{\left(N_I-1\right)}{\left(N-1\right)\mu_I}\overline{\theta}\right)y_{ii}^{*}\right)$$

$$\geq 0,$$

and the inequality is strict as long as $N_I>1$ so that $\mu_I>0$. The formula for α^R with equal weights on all agents follows from (25).

Finally, for the case $\delta=0$, note from (32) that $y_e^S=-\frac{1}{n}\overline{\theta}y_{ii}^{*}$ and so for $N_I>1$,

$$1+\alpha^R=\frac{1}{1-\delta+\left(\delta-\overline{\theta}\right)y_{ii}^{*}-\left(N-N_I\right)y_e^S}=\frac{1}{1-\left(\frac{N_I-1}{N-1}\right)\overline{\theta}y_{ii}^{*}}>1.$$

■

Proof of Corollary E: The rat-race results follow from ***Proposition III***. The effects on y_e follow because, from (A16),

$$y_e^S=\frac{\delta\mu_e\left(1-\overline{\theta}\right)-\frac{1}{n}\overline{\theta}\left(1-\delta\right)}{1-\delta+N\delta\mu_e}y_{ii}^{*},$$

which is monotone in both δ and μ_e. The result on peer sensitivity follows from ***Proposition V*** and the results for μ_e follow from the characterization in the proof of ***Proposition VI***, and the fact that from (A45) and (A46),

$$\operatorname{sgn}\tfrac{\partial}{\partial\mu_e}\left(y_t^S-y_e^S\right)=\operatorname{sgn}\left(a'-b'\right)<0\ .$$

Finally, for the wage correlation and dispersion, let $\mathbf{C}_q$ be the $N \times N$ covariance matrix for the output q. Normalize $\sigma_\epsilon^2 = 1$. Then

$$\mathbf{C}_q = \rho\mathbf{11}' + (1-\rho)\mathbf{I}.$$

Consider first the single principal case. Because $w = yq = \Delta^{-1}y^*q$, and because both Δ^{-1} and y^* are symmetric, the covariance matrix for wages is given by

$$\mathbf{C}_w = y\mathbf{C}_q y = \Delta^{-1}y^*\mathbf{C}_q y^*\Delta^{-1}.$$

Because $y^*/y_{ii}^* = -\frac{1}{n}\overline{\theta}\mathbf{11}' + \left(1+\frac{1}{n}\overline{\theta}\right)\mathbf{I}$, we have

$$\mathbf{C}_w = \left(y_{ii}^*\right)^2\left(1+\tfrac{1}{n}\overline{\theta}\right)(1-\rho)\left[-\tfrac{1}{n}\overline{\theta}\mathbf{11}' + \left(1+\tfrac{1}{n}\overline{\theta}\right)\Delta^{-1}\Delta^{-1}\right].$$

Because $\Delta^{-1}\mathbf{1} = \mathbf{1}$, then $\mathbf{C}_w\mathbf{1}$ does not depend on δ. As a result, $Cov(w_i, \overline{w})$ and $Var(\overline{w})$ don't depend on δ. So the correlation of each wage with the average wage only depends on the variance of each wage. Thus, it is sufficient to show that $\left[\Delta^{-1}\Delta^{-1}\right]_{ii}$, and hence the variance of each agent's wage, declines with δ.

Because μ is symmetric, it can be diagonalized as $\mu = U\Lambda_\mu U'$ where Λ_μ is the diagonal matrix of eigenvalues of μ, which we denote by $\lambda_\mu(j)$ for $j = 1,\ldots,N$. Because μ is stochastic, $\left|\lambda_\mu(j)\right| \le 1$.

Finally, we have

$$\Delta = \frac{1}{1-\delta}\left[\mathbf{I} - \delta\mu\right] = \frac{1}{1-\delta}\left[\mathbf{I} - \delta U\Lambda_\mu U'\right] = U\left[\frac{\mathbf{I} - \delta\Lambda_\mu}{1-\delta}\right]U'$$

$$\Rightarrow \Delta^{-1}\Delta^{-1} = U\left[\frac{\mathbf{I} - \delta\Lambda_\mu}{1-\delta}\right]^{-2}U' = \sum_j\left[\frac{1-\delta\lambda_\mu(j)}{1-\delta}\right]^{-2}u(j)u(j)'$$

$$\Rightarrow \left[\Delta^{-1}\Delta^{-1}\right]_{ii} = \sum_j\left[\frac{1-\delta\lambda_\mu(j)}{1-\delta}\right]^{-2}u_i(j)^2$$

The result therefore follows from

$$\partial_\delta\left[\frac{1-\delta\lambda_\mu(j)}{1-\delta}\right]=\frac{1-\lambda_\mu(j)}{(1-\delta)^2}\geq 0\,.$$

In the case of multiple principals, wages are rescaled by the rat-race factor. This rescaling will not change the correlation of individual wages with the average wage, nor the dispersion (variance) as a fraction of the expected wage. ■

PROOF OF PROPOSITION XI: Following the steps of the proof of ***LEMMA 2*** up to (A12), but now retaining the terms that depend on q_J, we see that maximizing π_I is equivalent to maximizing

$$\left(\sum\nolimits_{i\in I}\left(\left(\frac{1-\mu_I\delta}{1-\delta}\right)+\sum\nolimits_{\hat{i}\hat{j}}\Delta_{\hat{i}\hat{j}}y_{ji}\right)E\left[q_i\right]-C(\beta_i)\right)+\mathbf{1}'\Delta_{IJ}y_{JJ}E\left[q_J\right]$$
$$=\left(\sum\nolimits_{i\in I}\left(1+\frac{\delta}{1-\delta}\left(1-\mu_I-\sum\nolimits_{\hat{i}\hat{j}}\mu_{\hat{i}\hat{j}}y_{ji}\right)\right)E\left[q_i\right]-C(\beta_i)\right)-\left(\frac{\delta}{1-\delta}\right)\sum\nolimits_{j\in J}\left(\sum\nolimits_{\hat{i}k}\mu_{\hat{i}k}y_{kj}\right)E\left[q_j\right]$$
$$=\left(\sum\nolimits_{i\in I}\left(1+\frac{\delta}{1-\delta}\left(1-\mu_I-N_I(N-N_I)\mu_e y_e\right)\right)E\left[q_i\right]-C(\beta_i)\right)-\left(\frac{\delta N_I\bar{y}}{1-\delta}\right)\sum\nolimits_{j\in J}E\left[q_j\right]$$

where $\bar{y}=\mu_e\left(y_{jj}+(N_I-1)y_t+(N-2N_I)y_e\right)$. Recall that $a_i=\beta_{ii}/k$, and hence we can equivalently maximize

$$\left(\sum\nolimits_{i\in I}(1+\alpha_I)\beta_{ii}-kC(\beta_i)\right)-\gamma_I\sum\nolimits_{j\in J}\beta_{jj}\,, \tag{A51}$$

where $\alpha_I=\dfrac{\delta}{1-\delta}\left(1-\mu_I-N_I(N-N_I)\mu_e y_e\right)=\dfrac{\delta}{1-\delta}\left(1-\mu_I\right)\left(1-N_I y_e\right)$ and $\gamma_I=\dfrac{\delta}{1-\delta}N_I\bar{y}$.

Recall that $\beta_{ij}\equiv\dfrac{\partial c_i}{\partial q_j}=\dfrac{y_{ij}-\delta\sum_k\mu_{ik}y_{kj}}{1-\delta}$ and

$$C(\beta_i)\equiv c_0+\psi\left(\beta_{ii}/k\right)+\tfrac{1}{2}\lambda\left[(1-\rho)\sum\nolimits_j\beta_{ij}^2+\rho\left(\sum\nolimits_j\beta_{ij}\right)^2\right]\sigma_\epsilon^2\,. \tag{A52}$$

Taking the first-order condition of (A51) with respect to β_{ii} we have:

$$1+\alpha_I=k\tfrac{\partial}{\partial\beta_{ii}}C(\beta_i)=\beta_{ii}+k\lambda\sigma_\epsilon^2\left[(1-\rho)\beta_{ii}+\rho\sum\nolimits_j\beta_{ij}\right] \tag{A53}$$

Next, the first-order condition with respect to $\beta_{i\hat{i}}$ yields

$$0 = k\tfrac{\partial}{\partial\beta_{i\hat{i}}}C(\beta_i) = k\lambda\sigma_\epsilon^2\left[(1-\rho)\beta_{i\hat{i}} + \rho\sum\nolimits_j \beta_{ij}\right], \tag{A54}$$

or equivalently,

$$\beta_{i\hat{i}} = -\frac{\rho}{1-\rho}\sum\nolimits_j \beta_{ij} \tag{A55}$$

Finally, the first-order condition with respect to y_{ij} implies

$$0 = k\tfrac{\partial}{\partial\beta_{ij}}C(\beta_i)\tfrac{\partial\beta_{ij}}{\partial y_{ij}} + \gamma_I\tfrac{\partial\beta_{jj}}{\partial y_{ij}} = k\lambda\sigma_\epsilon^2\left[(1-\rho)\beta_{ij} + \rho\sum\nolimits_j \beta_{ij}\right]\frac{1}{1-\delta} - \gamma_I\frac{\mu_e\delta}{1-\delta},$$

which is equivalent to

$$\beta_{ij} = -\frac{\rho}{1-\rho}\sum\nolimits_j \beta_{ij} + \frac{\mu_e\delta\gamma_I}{k\lambda\sigma_\epsilon^2(1-\rho)}. \tag{A56}$$

Summing (A55) and (A56),

$$\sum\nolimits_j \beta_{ij} = \beta_{ii} - \frac{n\rho}{1-\rho}\sum\nolimits_j \beta_{ij} + \frac{\left(N-N_I\right)\mu_e\delta\gamma_I}{k\lambda\sigma_\epsilon^2(1-\rho)},$$

and so,

$$\left(1+\frac{n\rho}{1-\rho}\right)\sum\nolimits_j \beta_{ij} = \frac{n\rho}{\overline{\theta}\left(1-\rho\right)}\sum\nolimits_j \beta_{ij} = \beta_{ii} + \frac{\left(N-N_I\right)\mu_e\delta\gamma_I}{k\lambda\sigma_\epsilon^2(1-\rho)}.$$

Therefore, letting $\nu = \tfrac{1}{n}\overline{\theta}(N-N_I)\mu_e\delta\gamma_I$, and noting that $k\lambda\sigma_\epsilon^2\left(1-\rho\right) = \dfrac{k\lambda\overline{\sigma}^2}{1+\frac{1}{n}\overline{\theta}} = \dfrac{1-y_{ii}^*}{\left(1+\frac{1}{n}\overline{\theta}\right)y_{ii}^*}$,

$$\begin{aligned}\beta_{i\hat{i}}^D &= -\frac{\rho}{1-\rho}\sum\nolimits_j \beta_{ij}^D = -\tfrac{1}{n}\overline{\theta}\left(\beta_{ii}^D + \frac{(N-N_I)\mu_e\delta\gamma_I}{k\lambda\sigma_\epsilon^2(1-\rho)}\right) = -\tfrac{1}{n}\overline{\theta}\beta_{ii}^D - \frac{\nu}{k\lambda\sigma_\epsilon^2(1-\rho)}\\ &= -\tfrac{1}{n}\overline{\theta}\beta_{ii}^D - \left(1+\tfrac{1}{n}\overline{\theta}\right)\left(\frac{y_{ii}^*}{1-y_{ii}^*}\right)\nu < -\tfrac{1}{n}\overline{\theta}\beta_{ii}^D\end{aligned} \tag{A57}$$

$$\begin{aligned}\beta_{ij}^D &= -\tfrac{1}{n}\overline{\theta}\beta_{ii}^D - \frac{\nu}{k\lambda\sigma_\epsilon^2(1-\rho)} + \frac{\mu_e\delta\gamma_I}{k\lambda\sigma_\epsilon^2(1-\rho)} = -\tfrac{1}{n}\overline{\theta}\beta_{ii}^D + \frac{\nu}{k\lambda\sigma_\epsilon^2(1-\rho)}\left(\frac{N-1}{\overline{\theta}(N-N_I)}-1\right)\\ &= -\tfrac{1}{n}\overline{\theta}\beta_{ii}^D + \left(\frac{n}{\overline{\theta}(N-N_I)}-1\right)\left(1+\tfrac{1}{n}\overline{\theta}\right)\left(\frac{y_{ii}^*}{1-y_{ii}^*}\right)\nu > -\tfrac{1}{n}\overline{\theta}\beta_{ii}^D\end{aligned} \tag{A58}$$

Finally, from (A53),

$$\begin{aligned} 1+\alpha_I &= \beta_{ii} + k\lambda\sigma_\epsilon^2\left[(1-\rho)\beta_{ii} + \rho\sum\nolimits_j \beta_{ij}\right] \\ &= \beta_{ii} + k\lambda\sigma_\epsilon^2(1-\rho)\left[\beta_{ii} + \tfrac{1}{n}\overline{\theta}\beta_{ii} + \frac{\nu}{k\lambda\sigma_\epsilon^2(1-\rho)}\right] \\ &= \left(1+k\lambda\overline{\sigma}^2\right)\beta_{ii} + \nu = y_{ii}^{*-1}\beta_{ii} + \nu \end{aligned}$$

and so,

$$\beta_{ii}^D = \left(1+\alpha_I\right)y_{ii}^* - \nu y_{ii}^* < \left(1+\alpha_I\right)y_{ii}^* \tag{A59}$$

Note that without public disclosure, $\frac{\partial\beta_{ij}}{\partial y_{ij}} = 0$, so that the last term in (A56) is zero. Hence the solution is the same as with $\nu = 0$, which matches the result of ***LEMMA 2***. Note also, in equilibrium, the sensitivities for each agent should be consistent with the contract parameters; that is $\beta = \Delta y$:

$$\begin{aligned} \left(1-\delta\right)\beta_{ii} &= y_{ii} - \delta\sum\nolimits_k \mu_{ik}y_{ki} = y_{ii} - \delta\left(\mu_I y_t + (N-N_I)\mu_e y_e\right), \\ \left(1-\delta\right)\beta_{i\hat{i}} &= y_{i\hat{i}} - \delta\sum\nolimits_k \mu_{ik}y_{k\hat{i}} = y_t - \delta\left(\mu_t y_{ii} + (\mu_I - \mu_t)y_t + (N-N_I)\mu_e y_e\right) \\ \left(1-\delta\right)\beta_{ij} &= y_{ij} - \delta\sum\nolimits_k \mu_{ik}y_{kj} = y_e - \delta\left(\mu_e y_{ii} + (N_I - 1)\mu_e y_t + (N-2N_I)\mu_e y_e + \mu_I y_e\right) \end{aligned} \tag{A60}$$

We thus have six linear equations we can solve to determine the six unknowns (β^D, y^D). ■