More versus Closer Friends: How Gender Shapes Social Networks and their Effects on Performance

Ilse Lindenlaub* Anja Prummer†

January 23, 2013

Abstract

We build a theory that connects differences between men's and women's social networks to disparities in their labor market outcomes. We first document in the AddHealth Data Set that men's and women's networks differ, which is a new stylized fact. We find that men have a larger number of friends, that is a higher degree, compared to women. On the other hand, women's friends are more likely to be friends among each other, i.e. women have a higher clustering coefficient than men. We then build a theory that shows how these differences can explain discrepancies in the job performance of men and women. We explicitly model the trade-off between network density and network span and the impact of these features on peer pressure and information acquisition, which both matter for job performance. In our model, workers repeatedly cooperate with each other to complete projects. A team member chooses his effort, which is unobservable to the other project partner, based on his network structure. A higher number of friends (sparse network) leads to more information, whereas a network with higher clustering (dense network) leads to higher effort through peer pressure. Overall, both information and peer pressure affect project completion and thus it is not a priori clear which characteristic is more advantageous. We find that women outperform men when the uncertainty about the project value is negligible. In turn, men have a better assessment of the value of the project, as they hold more information and thus perform relatively better under uncertainty. In addition, men's effort is more sensitive to the expected project reward, which can be high or low, than women's effort.

Keywords: Networks, Peer Pressure, Gender, Labor Market Outcomes
JEL Classification: D85, Z13, J16

*Contact: Department of Economics, European University Institute, Via della Piazzuola 43, 50133 Florence, Italy, ilse.lindenlaub@eui.eu

†Contact: Department of Economics, European University Institute, Via della Piazzuola 43, 50133 Florence, Italy, anja.prummer@eui.eu

1 Introduction

We document a new form of heterogeneity between men and women, concerning their social networks. In particular, we find that men have a larger number of friends, that is a higher degree, compared to women. On the other hand, women's friends are more likely to be friends among each other, i.e. women have a higher clustering coefficient than men. We then develop a theory of how this heterogeneity in social networks relates to differences in labor market outcomes of men and women. These differences are well documented. Men and women differ in terms of wages, their prevalence in leadership position and with regards to their occupational choice. Common explanations for these patterns are discrimination against women, differences in preferences and risk aversion. We offer an alternative explanation for these differences which is based on social network heterogeneity between men and women. Moreover, we add to the limited amount of network literature that evaluates the trade-off between network density and the network span. We explicitly model the impact of these network features on peer pressure and information acquisition at the work place. This allows us to gain novel insights on the impact of these network characteristics on labor market outcomes.

We show in the AddHealth Data Set that women have fewer friends than men, but their friends are more likely to be friends among each other.[1] In terms of network measures, we find that women have a lower in-, out- and overall degree, but a higher clustering coefficient. Thus, women have smaller, more closed networks, whereas men have larger, looser ego networks. According to our knowledge we are the first to document this empirical fact.

We then build a model that connects these differences in the way men and women network to their outcomes at work. In our model, workers repeatedly cooperate with each other to complete projects. Whether the project will be accomplished or not depends on the effort the partners exert, with higher effort making project completion more likely. This effort is unobservable and only the project outcome is public information. If the project is completed successfully, then the project payoff is shared between the team members. Because the output produced is split, but the costs are not, we have a team moral hazard problem, which leads the project partners to exert inefficiently low effort.

We are interested in what effort level the project partners choose and what factors influence their choice of effort. First, this choice depends on the value of the project, which can be high or low, depending on the state of the world. Individuals receive signals about the state of the world and form expectations about the value of the project. These expectations influence effort. But effort also depends on the level of peer pressure individuals face. Namely, those who are exposed to a higher level of peer pressure, will exert a higher effort.

How much information an individual has and how much peer pressure he faces, depends on his network structure. Individuals with a higher degree have more information, as they receive a higher number of signals about the state of the world, whereas a network with higher clustering, that is a more closed network, provides more peer pressure. Therefore, men, with a higher degree and lower clustering, will have an informational advantage, but less peer pressure, whereas women, with a lower degree and higher clustering, will face an informational disadvantage, but higher peer pressure.

Closed networks lead to more peer pressure, which is in line with literature that shows how closed networks can overcome free-riding problems through the creation of norms and punishment systems. We build on this idea to model peer pressure. In our setting, a failed project leads to discord between the project partners. But this discord also affects their common friends, that is

[1]We are using data of teenagers, not of men and women who are already employed. We do this as we are interested in the informal networks, not the formal ones. To our knowledge there does not exist a data set that contains information on *informal* networks at the workplace. We focus on students aged 18 and 19 as these can be considered to be adults. The network pattern we describe now holds for this age group and is either stable for the other age groups or shows a trend that works in our favor.

their disagreement spreads through their group. Common friends will also be in discord with the project partners. The idea that the disagreement spreads through the group is based on the structural balance theory. According to this theory, triads of friends are only stable as long as the relationships are balanced. This means that there is no set of three nodes such that the links among them consist of exactly two positive relationships and one negative one.[2] Agents would try to resolve this stressful situation, in which two of their friends are enemies, by taking sides. To simplify our analysis we assume that all relationships in the triad will be negative after a project failure.

When an individual undertakes a project with a partner with whom he is in current disagreement, their project will certainly fail. Therefore, the repercussions for an individual who failed a project with someone, he has many common friends with, are worse than for someone with a looser network. And thus, higher clustering leads to higher effort.

Whereas higher clustering unambiguously leads to higher effort, this is not true for information. The more information an individual has, the more precise is his belief about the state of the world. This allows for a better judgment whether high effort, in case the signals point towards the good state of the world, or low effort, if the signals make the low state more probable, should be exerted. Therefore, additional information can increase or decrease effort.

We show under which circumstances male networks lead to higher effort and wages and female networks and vice versa. We find that men outperform women when the uncertainty about the state of the world is considerable, which is particularly true when information is scarce, when signals are not completely informative and when the project reward significantly differs across the two states. They also perform better when the future is highly discounted. In addition, men's effort is more sensitive to the expected project reward, which can be high or low, than women's effort. That is, men's effort is more proportional to the rewards of the project whereas women exert similar levels of effort regardless of the expected rewards. Moreover, we find that peer pressure is beneficial for project outcomes of teams when the information flow is already large. To the contrary, if there is only little information available and the state of the world is uncertain, then the benefits of a purely information-based network outweigh the benefits of a network that also leads to peer pressure.

These predictions on project completion translate into predictions on wage dynamics. In particular, when men outperform women at the beginning of their career in terms of project completion and wages, then they continue to do so when the project reward in the future is high. Only if a low state arises in the future do women have the chance to catch up. If they do, the second-period wage advantage of women over men can never be as large as the one of men over women if the second period state is high.

Given our findings, we would expect men and women to sort into different occupations, with men choosing occupations where they face higher uncertainty while women self-select into jobs in safe environments where rewards are stable but lower. Moreover, we expect that having women in the network is particularly beneficial when information is abundant, that is at higher levels of the organizational hierarchy. Finally, our model predicts that early career wages are decisive for the future wage trajectory. This suggests that a considerable part of the wage gap is due to differences between men's and women's wages at the starting point of their careers. All of these predictions are supported by the empirical literature, which is discussed below.

The paper proceeds as follows: The remainder of the section places our findings in the context of the literature. In section 2 we show that men's and women's networks differ in the AddHealth data set. Section 3 develops the general model, which we restrict in section 4 to two periods. In section 5 we focus on the differences between men's and women's networks and how they affect effort and

[2] See for instance, Easley and Kleinberg (2010): Networks, Crowds, and Markets: Reasoning about a Highly Connected World, chapters 3 and 5 for references and summary.

wages. To gain more insights we look at a parametric example in section 6. Section 7 compares the effort chosen under the equilibrium we focus on to the effort in the static Nash equilibrium as well as to the effort a social planner would impose. Section 8 interprets our results and connects them to the stylized facts concerning differences in labor market outcomes.

1.1 Related Literature

Trade-Off Between Network Closure and Network Reach. The key feature of our paper is the trade-off between network closure, as given by higher clustering and network reach. This trade-off goes back to a debate in the sociology literature on which of these network characteristics is more important. On the one hand, Coleman (1988b) emphasizes the importance of network closure, i.e of a network structure where individuals share many common friends. This network closure facilitates cooperation and leads to a higher level of trust among individuals. An often cited example is that of diamond dealers in New York, where trade takes place in a tight network of family and religious ties. Because of this tight network, the level of trust among the traders is high, which prevents them from cheating each other. This definition of trust is similar to our concept of peer pressure. In both cases, individuals do not shirk or cheat because they fear repercussions of common friends or other members of their networks. That these concepts are indeed closely related can also be seen from Coleman (1988a). In this paper he analyzes when and how social networks can overcome public good problems. Agents have to contribute to a public good and as they are not internalizing externalities, the public good will not be provided or will be provided at a suboptimal level. Due to the lack of incentives individuals behave as free riders. However, Coleman (1988a) provides examples, which show that in many real-life situations these public goods are provided and moreover that the opposite of free rider behavior occurs, namely an excess of zeal. He claims that networks can provide additional incentives that help to overcome the free-riding problem and can even lead to this zealous behavior. He emphasizes in particular, how *closed* networks can overcome the free-riding problem through the creation of norms and punishment systems. This is what the peer pressure in our model is based on. Whereas Coleman (1988b) and Coleman (1988a) emphasize the role of network closure, Burt (1992) argues that networks with a lower closure that provide greater access to information and other resources are more beneficial. In the same spirit, Granovetter (1973) emphasizes the "strength of weak ties," namely having weak links to agents with whom one has few common friends, which is beneficial for the acquisition of information. Therefore, there are two types of social capital that derive from network structures. These two types can also be related to the concepts of bonding versus bridging social capital defined in Putnam (2000).

This trade-off has also been analyzed in Karlan et al. (2009). They consider the relative advantages of networks with high closure versus loose networks. Their paper focuses on the trust generated in a network. This trust can be used as a form of collateral, i.e. the higher the level of trust, the higher the amount of money or the value of an asset an individual can borrow through a network. Their key theoretical insight is that higher closure increases trust, but reduces access. Therefore, a higher level of trust is more important if individuals need to borrow a high-valued asset, whereas more access is more important for low-valued assets. The welfare of an individual when high-valued assets are exchanged is higher if he belongs to a more closed neighborhood, whereas in a low-value exchange environment, higher closure leads to a lower expected value.

Dixit (2003) also discusses this trade-off between sparser networks and more closed structures in a trade setting. He focuses on the role of self governance, as an alternative to official institutions, in trading relationships. Trading with more distant individuals offers higher gains, but information flows about cheating are decreasing in this distance. There is then a clear trade-off between networks that have a high closure, i.e. a local bias in trade and networks that span a larger distance. Dixit (2003) asks what determines the limits of self governance, i.e. when can communities achieve full self governance. He finds that this is only the case when these communities are sufficiently small, which shows that a minimum of trust is necessary for self-governance.

Peer Pressure There has been a growing literature on how peer pressure operates and which factors influence it and through which channels operates. Our work adds to this literature by explicitly modeling the channels of peer pressure. Kandel and Lazear (1992) model peer pressure through a simple function, where peer pressure depends on own effort, the effort of peers as well as other actions of him and his peers, which do not affect firm output directly. Their finding is that with peer pressure individuals exert higher effort, which leads to a higher profit for the firm. They then ask how firms can create peer pressure. Two possibilities of creating peer pressure are norms and mutual monitoring, where in the case of mutual monitoring the question is what the relevant group is, that is the team, the department or the entire firm. We do not address the problem of what the firm can do, but rather assume that it managed to implement a mechanism or norm in which peer pressure matters as this increases payoffs. Note that the firm has no incentive to take costs into account and therefore it might be that peer pressure leads to a decrease in welfare of the individual worker, a result which is also mentioned in Kandel and Lazear (1992).

Men and Women Have Different Labor Market Outcomes. First, men and women have different wages, i.e. even when controlling for observable characteristics, female annual earnings in OECD countries are 73%-77% of male earnings (OECD Employment Outlook (2010)), which is referred to as the gender wage gap. Looking at Canadian data, Fortin and Huberman (2002) find that a considerable component of the gender wage gap is attributable to intra-occupational wage differentials between men and women at the firm level. That is, within the same occupational group men earn more than women. In almost all European Union countries this gender wage gap is increasing across the wage distribution (Booth (2009)), which is commonly referred to as the glass ceiling effect. In a similar vein, Lalanne and Seabright (2011) show that women earn only 70% of men's salary in top executive jobs, using a data set from UK, US, German and French companies. However, these differences between men's and women's earnings are already prevalent at the beginning of the career (Gerhart (1990)). The initial salary is particularly important for the future wage trajectory of professionals (Babcock and Laschever (2003), Gerhart and Rynes (1991), Martell et al. (1996)).

But men and women do not only differ with respect to their wages. In particular, Bohnet and Saidi (2011) find gender imbalances in leadership positions, with women only holding a small fraction of leadership positions in the corporate world.[3]

Additionally, women also choose different jobs than men. In particular, women are overrepresented in categories such as teaching, health, clerical work, sales, services, and textile, which is shown in Fortin and Huberman (2002). Flabbi (2012) confirms these differences in the occupational choice between men and women: More than 70% of teachers are women. On the other hand, more than 40% of men work in physics, mathematics and engineering. Only around 10% of jobs in such professions are held by women.

Networks Differ Across Gender. There is numerous empirical evidence that men's and women's social networks differ both in private settings and at the workplace, supporting our findings from the AddHealth Data. Belle (1989) studies the social networks of children. In a literature survey,

[3]For example, at the Fortune 500 companies in 2010, 2.4% of the CEOs, 14.4% of the executive positions, and 15.7% of the board members were women.

she documents that female networks are more exclusive, intimate and self-disclosing compared to male networks both during childhood and adolescence (p.179). Moreover, boys' networks are larger and more activity-based (p.176). Eder and Hallinan (1978) study gender differences in dyadic and triadic networks of children. They find that male dyadic networks accept newcomers more often than female networks, i.e. are less exclusive. These gender differences in social networks are also evident during adulthood when women keep seeking more intimate relations compared to men, see e.g. Kürtösi (2008).

According to Kürtösi (2008) these network differences across gender carry over to the workplace despite organizational constraints stemming from the gender composition or hierarchy structure. Women at work seem to form closed empathy-based groups, which help them to overcome feelings of isolation in male dominated environments (Tattersall and Keogh (2006)). On the other hand, male networks in the workplace are activity-based and instrumental, where instrumental ties are links that are conducive to information acquisition and career advancement (Tattersall and Keogh (2006), Kürtösi (2008)), compared to those of women.

Fischer and Oliker (1983) look specifically how the number of friendships develop over the years. They find that young women have a higher number of friends than young men. But the number of friendships decreases for women over time, whereas men acquire more friends, especially at the work place, which is in line with our findings that male networks are characterized by a higher degree. Moreover, according to Marsden (1987), women have a larger kin network compared to men and a smaller non-kin network. So, even though the overall number of friends might be similar, the composition of the networks is very different.

Summing up, the evidence suggests that women form empathy-based, closed groups whereas men's networks are larger, sparser, instrumental and based on activities. These differences are prevalent across all age groups and in both private and work settings. Translating this into the network terminology of our theoretical approach, women's networks are characterized by a higher clustering and lower degree compared to male networks.

Networks and Labor Market Outcomes There is a strand of literature that studies networks and labor market outcomes. The key difference to our work is that most of this literature does not focus on the informal network at the workplace, but on the overall informal network. Moreover, we have found no other paper that analyzes the trade-off between span and density of informal networks at work and its impact on labor market performance.

Arrow and Borzekowski (2004) find that differences in networks can indeed explain wage differences. They focus on informal networks used to find new jobs and show that a difference in degree can generate a significant wage inequality. Particularly, they argue that wage differences between black and white workers are due to the fact that the black workers have a lower degree, i.e. a smaller network. Calvo-Armengol and Jackson (2004) also look at informal networks used to obtain information about job opportunities. They focus on the role of social networks in explaining unemployment and wage inequality. Specifically, they find that employment is positively correlated across time and agents. When there are two distinct groups with different unemployment rates, then the group that has initially a higher unemployment rate will continue to have a higher unemployment rate. Calvó-Armengol and Jackson (2007) is a companion paper to Calvo-Armengol and Jackson (2004) and generalizes the model used in Calvo-Armengol and Jackson (2004). Moreover, they look at wage dynamics. They find that also wages are positively correlated among time and agents and that if a group starts out with lower wages, then this difference in wages persists over time.

Further, there is empirical evidence that an individual in a larger network, keeping everything else fixed, is more likely to be employed and that his wage is higher. This is documented for migrants in Munshi (2003) for the US and in Edin et al. (2003) for Sweden.

Conti et al. (2009) look at the effects of being popular as a senior in high school on adult wages. They use data from the Wisconsin longitudinal study, which contains information on school friendship networks. Conti et al. (2009) measure, based on this network, individuals' in and out degree. They then regress wages 35 years later on these measures of degree and find that out degree is insignificant, whereas the in degree, their measure of popularity, is indeed significantly positive. For an individual with median characteristics one additional friendship nomination is associated with a 2% wage increase 35 years later. They argue that a high in degree is a measure of social skills, more accurately a measure of how good someone is in building positive personal and social relationships and in adjusting to a certain environment and situation. The authors also check whether the in degree can be considered to be innate personality trait by including measures of personality traits in the wage regression. They find that even when including these traits, the in degree is still highly significant. Therefore, the in degree captures something different from personality traits.

2 Network Properties of Men and Women

We use data from the AddHealth data set to calculate the networking properties of males and females.[4] The AddHealth data set contains data on students in grades 7-12 from a nationally representative sample of roughly 140 schools in the U.S. in years 1994-95. Every student attending the sampled schools on the interview day is asked to compile a brief questionnaire (in-school data) containing questions on respondents' demographic and behavioral characteristics, education, family background and friendship. The AddHealth website describes surveys and data in details. This sample contains information on 90,118 students.

Our main reason for using this dataset of *students* is to circumvent the problem that at the firm level networks can be shaped by the environment or that individuals with certain types of networks will sort into occupations or firms where their types of networks are more beneficial. At school level there is no such selection bias and therefore this data set gives less biased insights into the networks boys and girls tend to form.

Further, it is well documented that individuals are more likely to name others as their friends if these have a higher social status, see Marsden (2005). We believe that this is less of a problem at school compared to the workplace as at the workplace social status is connected to a higher position in the hierarchy and therefore formal power. To claim a link between individuals denotes friendship or having a good relationship instead of trying to have a connection with someone superior is therefore problematic. [5]

Last, we are interested in these network characteristics of men and women as exogenous types, comparable to different ability or skill types commonly used in the literature, where this network type is stable overtime. Burt (2011) provides evidence for the existence of different network types from a multi-role game in a virtual world. He finds that people build a similar type of network, e.g. either a closed network or a network rich in access to structural holes. [6]

[4] For a detailed description of the AddHealth data set, see `http://www.cpc.unc.edu/projects/addhealth`.

[5] It could be argued that popular children will also be named as friends by individuals who would just like to be associated with them. It is our belief that the misreporting is less of a problem with children as there are fewer ulterior motives at play.

[6] About a third of network variance is consistent with individuals across roles, but the correlation coefficient between the network formed and the network type is above 0.5.

2.1 Friendship Networks

The friendship information in the AddHealth data set is based upon actual friends nominations. Students were asked to name up to 5 male and female friends. Students named friends both from the school they attend as well as friends from outside the school. Some of the friends, who do not attend the same school attend a sister school[7] and can still be identified. The other friends cannot be identified and are dropped subsequently from the sample.[8] The friendship network constructed from this is a directed network. For this network, we calculate both the directed and undirected clustering coefficients as well as the in-, out- and overall degree.[9]

The clustering coefficient is calculated as the ratio of the number of links between a node's neighbors to the total possible number of links between the node's neighbors, both for the directed and undirected network. The in-degree denotes how often an individual was named, the out-degree gives how many friends this individuals named and the degree is then the sum of in- and out-degree.

We show the means of these network measures together with descriptive statistics in Table 1.

Table 1: Summary Statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Cl. Coeff.	0.124	0.195	0	1	108121
Cl. Coeff. (dir.)	0.093	0.153	0	1	108121
In Degree	3.755	3.522	0	37	108121
Out Degree	3.753	3.671	0	10	108121
Degree	7.508	5.966	0	44	108121
Female	0.492	0.5	0	1	108121
Age	15.003	1.712	10	19	84792
School Size	1198.034	681.887	29	2982	106982

2.2 Estimation Results

We then estimate whether gender has a significant influence on degree as well as on the clustering coefficients. We standardize all of our measures in order to improve the interpretability of our results. Further, we normalize the age by subtracting 16. In all our regressions, we also control for school, which serves to capture location effects as well as time differences from when the data was collected. Note that we are not interested in determining which other factors influence these network characteristics, as is done e.g. in Conti et al. (2009). The purpose of this section is only to show that men's and women's networks differ. Our results are given in Table 2.

We find that girls have a significantly higher clustering coefficient, independently of how the clustering coefficient is measured. Both younger and older girls have a higher clustering coefficient, i.e. this characteristic does not change as students grow older. Girls also have a higher in and out degree as well as overall degree. But older girls have a lower absolute degree, out degree and in degree than younger girls, i.e. unlike with the clustering coefficients this property changes as girls mature. However, the degree does not change much for boys as they grow older. When just taking into consideration the oldest students, i.e. those aged 18 and 19, which are the students we are most interested in as we are interested in the network properties of men and women, girls have a lower in, out and overall degree.

[7] A sister school is a school in the same community. So, if in a community there is a high school and a middle school, then the high school is the sister school of the middle school and the middle school is the sister school of the high school.

[8] Overall, less than 10% of the observations are dropped.

[9] For the undirected clustering coefficient we assume that a link exists if at least one of the individuals named the other one as a friend.

Table 2: Results All Ages

	Cl.Coeff Dir.	Cl.Coeff. Dir.	Cl. Coeff.	Cl.Coeff	In Degree	In Degree	Out Degree	Out Degree	Degree	Degree
Female	0.110***	0.0729***	0.0922***	0.0612***	0.107***	0.146***	0.186***	0.193***	0.178***	0.205***
	(0.00574)	(0.0116)	(0.00570)	(0.0116)	(0.00679)	(0.0137)	(0.00605)	(0.0122)	(0.00630)	(0.0127)
Age-16	0.00440	-0.000773	0.00456*	-0.000780	-0.0173***	0.00629	-0.0500***	-0.0375***	-0.0410***	-0.0193***
	(0.00231)	(0.00279)	(0.00230)	(0.00277)	(0.00273)	(0.00329)	(0.00244)	(0.00294)	(0.00254)	(0.00306)
Size/1000	0.0310	0.0235	-0.0170	-0.0224	0.0420	0.0329	0.0339	0.0240	0.0457	0.0342
	(0.0781)	(0.0782)	(0.0776)	(0.0777)	(0.0924)	(0.0924)	(0.0822)	(0.0823)	(0.0857)	(0.0857)
Female* Age 16-17		0.0668***		0.0594***		-0.102***		-0.0443***		-0.0878***
		(0.0105)		(0.0105)		(0.0124)		(0.0111)		(0.0115)
Female*Age 18-19		-0.0360		-0.0126		-0.356***		-0.214***		-0.342***
		(0.0219)		(0.0218)		(0.0259)		(0.0231)		(0.0240)
Female*Size/1000		0.0122		0.00858		0.0148		0.0169		0.0192*
		(0.00861)		(0.00856)		(0.0102)		(0.00907)		(0.00944)
Constant	0.0767	0.0842	0.186	0.190	-0.242	-0.210	0.165	0.189	-0.0410	-0.00819
	(0.158)	(0.158)	(0.157)	(0.157)	(0.187)	(0.187)	(0.166)	(0.166)	(0.173)	(0.173)
Additional Control: School Identifier										
Observations	84792	84792	84792	84792	84792	84792	84792	84792	84792	84792
R^2	0.198	0.199	0.205	0.205	0.132	0.134	0.170	0.170	0.205	0.207

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

2.3 How do male and female networks differ?

Our conclusion from this is that the clustering coefficient is a relatively stable property over time. Young girls seem to form closed, tight networks and so do older girls. Moreover, girls form tighter networks than boys. Unlike the clustering coefficient, the degree depends a lot more on age. Specifically, younger girls have a higher degree, independently of how this is measured, than boys. For older girls, this is no longer true. That the number of friendships changes over time has also been documented in Fischer and Oliker (1983). Their findings are in line with what we find, namely that the number of friends changes over the life cycle in a different manner, particularly at the workplace.

We use part of the table from Fischer and Oliker (1983), p. 127, to document this. In their sample Fischer and Oliker (1983) have men and women who are all employed. Still, the number of friendships with co-workers differs greatly between them.

For individuals under 36, who are unmarried and do not have children, the difference between men and women in the number of friends at the work place is small: men have on average 2.8 friends, women 2.5. But this difference increases, when men and women under 36 and married, with or without children, are compared. Without children, men have on average one more friend than women, with children they even have two more friends at the workplace. Therefore, this pattern that older girls have a lower degree than older boys does not suddenly reverse, but is also documented for man at women at the workplace across all age groups.

	Under 36, unmarried, no children	Under 36, married, no children	Under 36, married, children	36-64, married, no children	36-64, married, children
Men	2.8	3.4	3.1	2.3	2.0
Women	2.5	2.1	0.9	1.4	1.4
Men (N)	113	50	70	63	54
Women(N)	76	51	98	67	60

Table 3: Friendships with Coworkers

Taking our estimation results together with the evidence in the literature, we feel confident to assume that women in the workplace have more dense networks, with fewer friends whereas men have a greater number of friends and less tight networks.

3 Model

We consider an undirected network, g of M workers. This network describes the informal relationships of workers within the division of the firm. A link between i and j, denoted by $g_{ij} = 1$, indicates a friendship or personal relation between coworkers.

These coworkers have to complete projects together. The projects are assigned to them by the manager of their division, who picks two of his employees. The manager chooses the project team out of all possible project teams at random. The interpretation of this is that different projects require different skills. Agents are heterogeneous in skills, which is not modeled explicitly. Depending on the skill requirements of a given project, which is a random event in every period, the manager chooses two agents from the division i.e. from the network, whose skills most closely correspond to the skill requirement of the project.

The two project partners aim at completing their task successfully. Whether they are effective depends on their relationship. If they are friends, i.e. if there is a link between two chosen project partners, the project will be completed more likely, which reflects that agents work better together if they get along well. To simplify, we set the payoff of projects between unlinked agents to zero.

But even if there is a direct link between the project partners the project can still fail. Whether the project is successful or not depends on the effort the partners exert.

This effort, e, is unobservable and only the outcome of the project is public information. If the project was completed successfully, then the payoff is $2v$. If the project failed, the payoff is zero. The payoff of a successful project is shared equally between the project partners. Because the output produced is split, but the costs are not, we have a team moral hazard problem which leads the project partners to exert lower effort compared to the Pareto efficient effort.[10]

The question is now what effort level the project partners choose. Their choice depends on the magnitude of v, which can be high or low, depending on the state of the world. It also depends on dynamic considerations, which will be outlined in the following section.

3.1 Stage Game

We consider a dynamic game where in each period t, the stages of the game are as follows:
1. The project is randomly assigned to two individuals in the network. Then, the probability of being chosen and being partnered with a friend is

$$z_i = \frac{2d_i}{M(M-1)} = \frac{d_i}{a},$$

where $a \equiv \frac{M(M-1)}{2}$. We are interested in this probability as only when being partnered with a friend, the project can be completed successfully. The probability of being selected is proportional to the degree of an individual, i.e. individuals with a higher degree will be selected more often. [11]
2. The state of the world θ, and with it the payoff of the project for each partner, v, is realized, where

$$\theta = \begin{cases} \theta_h & \text{with probability} \quad q \\ \theta_l & \text{with probability} \quad 1-q, \end{cases}$$

If state θ_h (θ_l) is realized, the payoff is v_h (v_l) with $v_h > v_l$.
3. Each agent receives a signal x_i about the state of the world. If $x_i = 1$, then this indicates that the state is high. If $\theta = \theta_h$, each signal $x_i \sim \text{Bernoulli}(p)$ and if $\theta = \theta_l$, this signal is $x_i \sim \text{Bernoulli}(1-p)$. If the true state is high (low), then it is more likely to receive a high (low) signal, which is captured by $p > \frac{1}{2}$.
4. Each agent knows his own signal and learns that of his direct friends.[12] Additionally, each agent receives signals indirectly from other workers in his component or through official channels within the firm. Then, the number of signals agent i receives is

$$n_i = n_{1i} + n_{2i},$$

where $n_{1i} = d_i + 1$, where d_i denotes the degree of agent i. We interpret n_{1i} as internal information (from direct friends) and n_{2i} as external information (not from direct friends). We assume $n_{2i} \geq$

[10] See Holmstrom (1982) for moral hazard problems in teams.

[11] This is in line with Aral et al. (2011), who study project performance in a recruiting firm. They find that peripheral nodes, i.e. nodes that are not well connected, do fewer projects per unit of time than central nodes.

[12] We assume that friends forward signals truthfully. We want to abstain from modeling truthful information transmission, but we think of the underlying process as follows: suppose each worker learns with a certain probability whether the signals he receives from his friends are truthful. If he finds out he was given wrong information, then he will cut the friendship forever or can punish the colleague, who lied, through some other mechanism. In this case, no worker has an incentive to lie.

$m - n_{1i}$, which implies that each individual observes at least as many signals as there are individuals in his component, which is of size m. Note that $\forall i, j \in \{1, \ldots, m\}$, $n_i = n_j \equiv n$, i.e. every worker, who is part of the same component, observes the same signals. The signal realization in a period is denoted by x^t.[13] A sufficient statistic for the signal realizations is the number of high signals, $y^t \equiv \sum_{i=1}^{n} x_i^t$. The number of low signals is then simply $n - y^t$.
5. Based on the signal realization y^t, the agents calculate the posterior probability of the state of the world via Bayesian updating. This posterior probability is given by:

$$Pr(\theta_h|y^t) = \frac{Pr(y^t|\theta_h)Pr(\theta_h)}{Pr(\theta_h)Pr(y^t|\theta_h) + Pr(\theta_l)Pr(y^t|\theta_l)}$$

6. The two agents in a team choose their effort $e_i(h^t, y^t)$ simultaneously. The effort depends on the history, which is denoted by h^t and the signals about the true state of the world. The histories are public histories in the sense that they condition on whether the relationship with the current project partner is intact or not, which depends on the last period's project outcome and on who participated in the project team, but not on the private actions of the agents. Exertion of effort leads to cost $c(e_i(h^t, y^t))$.
7. The outcome of the project is realized. With probability $f(e_i, e_j) \in [0, 1)$ the project is successful, with probability $1 - f(e_i, e_j)$ it fails. The outcome of the project can be observed by the entire network.[14]

3.2 Static Decision Problem

Before going to the dynamic case, where friendship history matters, we are interested in the solution of the static problem. More precisely we ask what are the pure strategy Nash equilibria in the one period problem, where histories are irrelevant. This one period problem is given by

$$f(e_i, e_j)(Pr(\theta_h|y)v_h + (1 - Pr(\theta_h|y))v_l) - c(e_i) \tag{1}$$

We write here e_i instead of $e_i(y)$ as y only matters through its influence on the expected first period payoff $Pr(\theta_h|y)v_h + (1 - Pr(\theta_h|y))v_l$ and does not influence e_i directly. We only write $e_i(y)$ if we want to emphasize that e_i varies with the observed signals.
We maintain throughout our entire analysis the following assumptions regarding $f(\cdot, \cdot)$ and $c(\cdot)$:

Assumption 1. *The effort function $f(\cdot, \cdot)$ has the following properties:*

1. *Symmetry: e_i and e_j enter $f(e_i, e_j)$ symmetrically.*
2. $f_1(e_i, e_j) = f_2(e_i, e_j) > 0$, $f_{11}(e_i, e_j) = f_{22}(e_i, e_j) < 0$.
3. *Strict Supermodularity: $f_{12}(e_i, e_j) = f_{21}(e_i, e_j) > 0$.*
4. $f(e_i, 0) = 0$, $f(0, e_j) = 0$.
5. *Non-Increasing Returns to Scale: $f_{11}(e_i, e_j) + f_{12}(e_i, e_j) = f_{22}(e_i, e_j) + f_{21}(e_i, e_j) \leq 0$.*

The properties of the cost function $c(.)$ are

1. $c'(.) > 0$, $c''(.) > 0$.
2. $c(0) = 0$.
3. *Symmetry: $c_i(.) = c_j(.) = c(.)$.*

[13] As only individuals who are direct friends can complete a project successfully and direct friends are necessarily in the same component, the signal realizations observed by them are the same and therefore do not have to be indexed by i.
[14] We could alternatively assume that the project outcome can only be observed by the project partners and their common friends.

The assumption of supermodularity implies that e_i and e_j are complements. We focus on complements as with substitutes a manager would not have an incentive to involve two workers in a project, but rather, he would assign the task to one individual, therefore eliminating the team moral hazard problem.

Given these assumptions, we can derive the first order condition of (1), which is given by

$$f_1(e_i, e_j)(Pr(\theta_h|y)v_h + (1 - Pr(\theta_h|y))v_l) - c'(e_i) = 0, \tag{2}$$

and equivalently for j. Then, we can determine the pure strategy Nash equilibria of the game.

Proposition 1. *Nash Equilibria One-Shot Game*
There exist exactly two pure strategy Nash equilibria, which are symmetric,

$$e_i^*(y) = e_j^*(y) = 0 \quad \forall y,$$
$$e_i^*(y) = e_j^*(y) = e^*(y) > 0 \quad \forall y.$$

The proofs for for this and all other propositions as well as lemmas can be found in the Appendix. The one-period problem has two pure strategy Nash equilibria, which are symmetric. The fact that efforts are symmetric is driven by the symmetry of $f(\cdot,\cdot)$ and the fact that both players have the same cost function. Further, in one of the equilibria, the effort exerted is zero *independently* of the signals. In the other one, the effort is strictly positive, but this effort depends on the signals received. There exists exactly *one* Nash equilibrium with strictly positive effort for each y. This is due to the supermodularity of $f(\cdot,\cdot)$ and the convexity of the cost function, as well as the fact that increasing returns to scale are precluded by our assumptions.

Based on this we continue to discuss the dynamic game, the game where each worker takes into account that his effort choice today has an influence on his friendship histories in the future.

3.3 Dynamic Decision Problem

As we have restricted the effort to depend only on public histories, we can write the payoff function, the project partners maximize, recursively. As is common we indicate variables in $t+1$ with a *prime*.

$$V_{i,t}(h, y; e_j) = f(e_i(h), e_j)(Pr(\theta_h|y)v_h + (1 - Pr(\theta_h|y))v_l) - c(e_i(h))$$
$$+ \beta EV_{i,t+1}(h', y'; e'_k).$$

Each partner maximizes the payoff of the project today, $f(e_i(h), e_j)(Pr(\theta_h|y)v_h + (1 - Pr(\theta_h|y))v_l) - c(e_i(h))$, and the expected continuation value $\beta EV_{i,t+1}(h', y', e'_k)$. Note that in $t+1$ worker i is teamed with worker k, who can be the same or different from the team partner j in t.

The solution concept applied to the problem we defined is that of a public perfect equilibrium.

Definition 1. *A public perfect equilibrium is a profile of public strategies σ that for any public history h^t and any signal realization y^t specifies a Nash equilibrium for the repeated game i.e. $\forall h^t, \forall y^t$ $\sigma|_{h^t,y^t}$ is a Nash equilibrium.*

We focus now on the two period model as it allows us to illustrate key insights of how differences in men's and women's networks lead to differences in project completion and wages in the simplest manner possible.

4 Two Period Model

In the two period model we define the following two possible friendship histories with the current project partner:

1. A negative friendship history (denoted by $\underline{h}$), which occurs when the agent had a project failure
   - with the current teammate during the last period or
   - with a common friend of his current project partner during the last period.
2. A positive friendship history (denoted by $\overline{h}$), which occurs whenever 1. does not occur.

A failed project affects how the project is done in the next period if the project is done again by the same two partners or if it is done with one of the previous partners and a common friend, a node that is directly linked to both partners of the previous period. Otherwise, a failed project does not have any consequences. We assume a positive history in $t = 1$. As previously mentioned, a project failure affects the relationship between project partners as well as their relation to a common friend. We believe it is intuitively plausible that a project failure leads to a discord among the team partners. The intuition behind it is that when a project failed, they have to justify of why it was the case in front of their boss, maybe face a reevaluation, which is disagreeable to them and also affects their relationship. This discord then spreads also to their common friends. The idea is based on the well-established, empirically documented *structural balance theory*. As mention previously, according to this theory, triads of friends are only stable as long as the relationships are balanced. This implies that there cannot be a set of three nodes such that there are two positive and one negative relations. The workers will resolve the situation by taking sides. To simplify our analysis, we assume that all relationships in a triad will be negative after a project failure.

The histories are of importance as they influence the effort choice. We now select a strategy in which the histories can help to alleviate the team moral hazard. We assume that this kind of strategy has emerged through some cultural process, which we do not specify further.

We know from the static game that it is a Nash equilibrium to choose zero effort in every period, regardless of the signals. Because of this we can set for all y $\underline{e}_i(y) \equiv e(y, \underline{h}) = 0$. As $f(0,0) = 0$, the project, when i is teamed with someone he has a negative friendship history with, will fail for sure and the payoff is zero.

The problem then reduces to finding the optimal $\overline{e}_i(y) \equiv e(y, \overline{h}) > 0$, $\overline{e}_i'(y') \equiv e'(y', \overline{h}') > 0$. In each period the workers take the observed signals into account which depends on the observed information today. Moreover, the worker also takes the expected project value *tomorrow* into account, that is he realizes that today's project success or failure affect the next period through the friendship histories. What matters is how likely it is for the worker to enter a review phase in the second period if the project failed in the first period. This probability is given by

$$u_{ij} \equiv \frac{1 + \sum_{k,k\neq i,k\neq j} g_{ik}g_{jk}}{d_i} \equiv \frac{K_{ij}}{d_i}.$$

The term $\sum_{k,k\neq i,k\neq j} g_{ik}g_{jk}$ denotes the number of common friends of i and j and therefore K_{ij} is a proxy for $i's$ and $j's$ common friends. Hence, u_{ij} is the probability that in the next period agent i is doing a project with the previous project partner or with a common friend of him and the previous project partner, given that i is chosen for a project with a friend in the next period. The probability of having a positive friendship history in the second period is then given by $z_i(f(\overline{e}_i, \overline{e}_j) + (1 - u_{ij})(1 - f(\overline{e}_i, \overline{e}_j)))$.

We focus now on the effort that maximizes the payoff of each worker, given everyone adheres to the specified strategy. Therefore, we consider maximization problem of each worker, which is given by

$$\max_{\overline{e}_i,\overline{e}_i'} \quad f(\overline{e}_i,\overline{e}_j)(Pr(\theta_h|y)v_h + (1-Pr(\theta_h|y))v_l) - c(\overline{e}_i) \tag{3}$$
$$+\beta z_i(f(\overline{e}_i,\overline{e}_j) + (1-u_{ij})(1-f(\overline{e}_i,\overline{e}_j)))E_{y'}\left(f(\overline{e}_i',\overline{e}_k')(Pr(\theta_h|y')v_h + (1-Pr(\theta_h|y'))v_l) - c(\overline{e}_i')\right).$$

We solve the problem by backward induction, starting in the second period. As j and k belong to the same set, namely the friends of i, we can replace k in the second period by j. The high effort level in the second period is then given by

$$\overline{e}_i'^* \equiv \arg\max V_{i,2} = \left(f(\overline{e}_i',\overline{e}_j')(Pr(\theta_h|y')v_h + (1-Pr(\theta_h|y'))v_l) - c(\overline{e}_i')\right), \tag{4}$$

for i and is defined analogously for j. Note that this is the same problem as the static problem. The first order condition, which is necessary and sufficient given our assumptions, is then the same as equation (2) and is only repeated as a reminder.

$$f_1(\overline{e}_i',\overline{e}_j')(Pr(\theta_h|y')v_h + (1-Pr(\theta_h|y'))v_l) - c'(\overline{e}_i') = 0 \tag{5}$$

and analogously for worker j. We denote the maximized payoff in the second period by $V_{i,2}^*$ and $V_{j,2}^*$. Then, the maximization problem of agent i in the first period reads

$$\max_{\overline{e}_i} \quad f(\overline{e}_i,\overline{e}_j)(Pr(\theta_h|y)v_h + (1-Pr(\theta_h|y))v_l) - c(\overline{e}_i) \tag{6}$$
$$+\beta z_i(f(\overline{e}_i,\overline{e}_j) + (1-u_{ij})(1-f(\overline{e}_i,\overline{e}_j)))E_{y'}V_{i,2}^*$$

and analogously for agent j. The FOCs of worker i, which holds at the equilibrium efforts, are given by:

$$f_1(\overline{e}_i,\overline{e}_j)(Pr(\theta_h|y)v_h + (1-Pr(\theta_h|y)v_l + \beta z_i u_{ij} E_{y'}V_{i,2}^*)) - c'(\overline{e}_i) = 0 \tag{7}$$

Based on these first order conditions we can characterize the equilibrium effort.

Proposition 2. *There exists exactly one Nash equilibrium after $\overline{h}$, in which efforts are strictly positive and symmetric,*

$$\overline{e}_i^*(y) = \overline{e}_j^*(y) \equiv \overline{e}^*(y) \quad \forall y,$$
$$\overline{e}_i'^*(y) = \overline{e}_j'^*(y) \equiv \overline{e}'^*(y') \quad \forall y'.$$

We already know from Proposition 1 that this is true in the second period. But efforts are also symmetric and unique in the first period as $z_i u_{ij} = \frac{K_{ij}}{a}$ is constant across project partners, which leads $\beta z_i u_{ij} E_{y'} V_i^*$ to be the same for both workers.[15] Note that the symmetry in efforts only holds in the two period model, as in the last period the network structure does not matter anymore and therefore $E_{y'}V_i^* = E_{y'}V_j^* = E_{y'}V^*$. More accurately, if this game is played for T periods, then in $T-1$, the efforts are symmetric. So far we considered general network structures, but did not take the network differences between men and women into account. From now on we concentrate on these disparities.

[15] We now drop the time subscript from the value V since there is no confusion that this is the value of the second period.

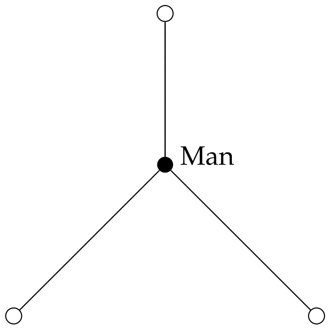


Figure 1: Male Network

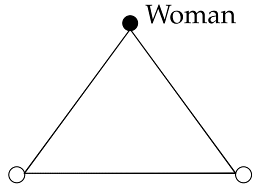


Figure 2: Female Network

5 Men vs Women: Comparative Statics

We focus on ego networks, where the ego network of a man is characterized by a higher degree and that of a woman by a higher number of common friends.

In what follows, we choose to model the ego network of a man as a star and the ego network of the woman as a ring, as these networks have the features we are looking for. To keep matters simple, we consider as our baseline cases a man who has three friends and a woman, who has two friends, who are also connected among each other. This is depicted in figures (1) and (2).[16]

These ego-networks can be part of a larger network. To make networks comparable we fix the number of nodes and links. This implies that women, due to their more clustered network will lose information compared to men.[17]

We then approximate the common friends in the man's network, given by the star, by K^S. The common friends in the woman's network are denoted by K^R, as their network is depicted as a ring.[18] Based on the previous discussion, $K^S < K^R$. Similarly, we denote the number of signals, i.e. the information gathered from the ego network as n_1^S and n_1^R. Note that n_1^S and n_1^R approximate the degree. The information obtained from outside the ego network is the same for both agents, i.e. $n_2^R = n_2^S = n_2$. Then, it also has to be that $n^R < n^S$.

We focus on *effort* as performance measure because the success probability of a project, and thus project completion, is monotonously increasing in effort. We are also interested in how degree and clustering affect *wages* in both periods. We will provide a definition for wages once we have discussed the determinants of effort in detail.

We base our analysis on equation (7) evaluated at the *equilibrium* effort. To simplify notation, we define $\pi(y) \equiv Pr(\theta_h|y)v_h + (1 - Pr(\theta_h|y)v_l$. Equation (7) then becomes

$$f_1(\bar{e}, \bar{e})(\pi(y) + \beta zuE_{y'}V^*) - c'(\bar{e}) = 0, \tag{8}$$

where we drop the stars for notational convenience. We also drop the indices as we know that effort levels are symmetric at equilibrium and zu is the same across project partners.

Then, $\bar{e}$ depends on $\pi(y)$, zu and $E_{y'}V_2^*$. The question is how $\bar{e}$ changes with $\pi(y)$, zu and $E_{y'}V^*$.

[16]We do not make any assumptions about the gender of these friends. Further, we do not depict their ego-networks.

[17]This assumes that the number of connections within a network is not too large and that there will not be one large component. Further, the rest of the network has to be the same to make a comparison sensible.

[18]We drop here the subscript as $K_{ij} = K \, \forall j$.

$$\frac{\partial \bar{e}}{\partial \pi(y)} = \frac{f_1(\bar{e},\bar{e})}{c''(\bar{e}) - (f_{11}(\bar{e},\bar{e}) + f_{12}(\bar{e},\bar{e}))\,(\pi(y) + \beta zu E_{y'}V^*)}$$
$$\frac{\partial \bar{e}}{\partial zu} = \frac{f_1(\bar{e},\bar{e})\beta E_{y'}V^*}{c''(\bar{e}) - (f_{11}(\bar{e},\bar{e}) + f_{12}(\bar{e},\bar{e}))\,(\pi(y) + \beta zu E_{y'}V^*)}$$
$$\frac{\partial \bar{e}}{\partial E_{y'}V^*} = \frac{f_1(\bar{e},\bar{e})\beta zu}{c''(\bar{e}) - (f_{11}(\bar{e},\bar{e}) + f_{12}(\bar{e},\bar{e}))\,(\pi(y) + \beta zu E_{y'}V^*)}$$

By assumption, $f_{11}(\bar{e},\bar{e}) + f_{12}(\bar{e},\bar{e}) \leq 0 \; \forall \bar{e}$ and therefore, we know that the equilibrium effort increases when $\pi(y)$, zu or $E_{y'}V^*$ increase.

Next, we discuss the determinants of effort $\pi(y)$, zu and $E_{y'}V^*$ in detail and see how they compare between men and women.

1. Common Friends zu. The first determinant of effort of agent i is zu, the term that approximates peer pressure. In particular, zu defines the probability of being selected together with agent j *and* having a link with j *and* node j being the previous project partner of i or a common friend of i and the previous project partner,

$$zu = \frac{1 + \sum_{k,k\neq i,k\neq j} g_{ik}g_{jk}}{a} = \frac{K}{a} \tag{9}$$

As mentioned previously, it is equal across project partners, which leads to the same level of effort despite possibly differing degrees.

We know that zu increases in the number of common friends as $\frac{K+1}{a} > \frac{K}{a}$. When an agent has more common friends, the probability of being paired in the next period with someone who cares about this period's project outcome is higher. As $K^R > K^S$, and thus $(zu)^R > (zu)^S$ we know that the level of peer pressure is higher for the ring than for the star.

2. Expected First Period Payoff $\pi(y)$. We are interested in how $\pi(y)$ changes with information as this is one of the channels through which information influences effort, and therefore also one way in which the informational advantage of the star matters. As we can say only little about the signal *realizations* and because the value of information depends on the realized state of the world, we consider the effect of information on $E_y(\pi(y)|\theta_h)$ and $E_y(\pi(y)|\theta_l)$. We will analyze effort for the two possible states separately. This is an *ex-post* perspective, in order to take into account that, depending on the state, some signal structures are more likely to occur than others, which stems from the assumption that signals are informative. Of course, the agents in the model make their effort choices *ex-ante* only based on their signals but without knowing with certainty which state of the world has been realized. We choose this approach to highlight that differences between men's and women's performance are state-dependent.

In order to simplify our analysis, we first establish that $\pi(y)$ has the martingale property.

Lemma 1. *$\pi(y'_n)$ satisfies the martingale property:* $\pi(y'_n) = E(\pi(y'_{n+1})|y'_n)$.

Based on Lemma 1 we can establish the properties of the expected first period payoff, which will help us to make statements under which conditions men have a higher first period payoff than women. We refer to $\pi(y_n)$ as the first period payoff, when an agent receives n signals.

Lemma 2 ($E_y\left(\pi(y_n)\right)$ and $E_y\left(\pi(y_n)|\theta\right)$**)**. .

1. *Information does not matter:*

$$E_y\left(\pi(y_n)\right) = E_y\left(\pi(y)\right) \quad \forall n$$

2. *Information matters conditional on* θ

$$E_y\left(\pi(y_{n+1})|\theta_h\right) > E_y\left(\pi(y_n)|\theta_h\right) \qquad E_y\left(\pi(y_{n+1})|\theta_l\right) < E_y\left(\pi(y_n)|\theta_l\right)$$

(i) unless

(a) $v_l \to v_h$ *for* $\theta \in \{\theta_h, \theta_l\}$

(b) $p \to 1$ *or* $p \to 0.5$ *for* $\theta \in \{\theta_h, \theta_l\}$,

(c) $q \to 1$ *if* $\theta = \theta_h$, $q \to 0$ *if* $\theta = \theta_l$.

Then,

$$E_y\left(\pi(y_n)|\theta\right) = E_y\left(\pi(y)|\theta\right) \quad \forall n$$

(ii) Abundance of Information: Let $n_2 \to \infty$ *and, hence,* $n \to \infty$*. Then,*

$$\lim_{n\to\infty} E_y\left(\pi(y_n)|\theta_h\right) = v_h, \qquad \lim_{n\to\infty} E_y\left(\pi(y_n)|\theta_l\right) = v_l.$$

The unconditional expectation over $\pi(y)$ is independent of the number of signals. This follows from the martingale property of $\pi(y)$.

However, he expectation over $\pi(y)$, given that a state of the world has been realized does depends on the number of signals. It increases in the number of signals if the state of the world is high and decreases in the number of signals it the state of the world is low. The more signals are available, the more accurate the posterior belief about the state of the world (part 2). We know that the star node obtains more signals and has more accurate information than the ring which is why the star's posterior is closer to the true value than the ring. This implies that, given $\theta = \theta_h$ (high state), the star node has a higher expected payoff compared to the ring node because he receives on average more high signals than low signals and has more of them. Given $\theta = \theta_l$ (low state), the reverse is true because the ring node has fewer signals and hence less info. In her posterior she puts more weight on the high state then the star node, even though this is a mistake.

According to part (i), the expected payoff becomes *independent* of the number of overall signals n when the value in the low and the high state are the same or if the signals are either completely informative or completely uninformative. Moreover, there is no dependence on the overall number of signals if agents have a prior, q, which reflects complete certainty about the state of the world. This implies that the effect of information on the expected payoff is reinforced when the uncertainty of the underlying environment is large so that an additional signal has a large impact on the posterior belief. This happens when (a) the v_h is considerably higher than v_l, (b) when the prior about the state of the world does not reflect certainty (c) when signals are neither completely informative nor completely uninformative. Therefore, the differences between the expected first period payoff in star and ring die out when the uncertainty of the underlying environment vanishes, i.e. when the low value approaches the high value or when the prior about the state of the world reflects certainty. When there is nearly no uncertainty about the state of the world, then the additional information of the star has no bearing on the expected payoff. The same is true for completely informative or uninformative signals.

If information becomes abundant, which happens e.g. when the number of external signals that an agent receives becomes large, then, given that signals are informative, agents know with certainty that either the high or the low state has occurred. Additional information does not change this belief (part (ii)). The expected payoff converges to the high value when the state is high and to the low project value when it is low. As expected payoffs are equal to the "true" payoff because the belief of all agents about the state is accurate, there is no difference between the star and ring node.

In sum, the star node has more accurate information about the state of the world, which is why his posterior belief more closely reflects the true state compared to the ring. Differences in expected payoff between star and ring node vanish whenever the additional information available to the star node has no bearing. This can happen for four reasons: both states lead to the same project value; the prior is such that signals add no new information; the (external) information becomes abundant such that the additional internal information of the star has no impact; last, signals are either completely informative or uninformative.

3. Second Period Expected Value. The second period expected value is another determinant of effort and the second channel through which the informational advantage of the star matters.

Recall that it is given by

$$E_y'(V^*) = E_y'\left(f(\overline{e}', \overline{e}')\pi(y) - c(\overline{e}')\right), \tag{10}$$

where $\overline{e}'$ denote the second period *equilibrium* efforts. Notice, that (10) is computed by the agents within the model *before* the second period state has realized. Hence, this is an *ex-ante* expectation over the whole range of possible future signal structures y' without conditioning on the state of the world θ'. We are interested, first, in the properties of $E_{y'}(V^*)$ and, second, in how it differs between star and ring.

We can express V^* as a function of $\pi(y')$, and write $V^*(y')$ to emphasize the dependence of V^* on the signals, i.e.

$$V^*(y') = g(\pi(y')) \tag{11}$$

As $\pi(y')$ is a martingale, we know that when g is a convex function, then $g(\pi(y'))$ is a submartingale. If g is a concave function, then $g(\pi(y'))$ is a supermartingale. Last, if g is both convex and concave, then $g(\pi(y'))$ is a martingale.[19]

A submartingale has the property that $E(V_n^*) < E(V_{n+1}^*))$, whereas a for a supermartingale $E(V_n^*) > E(V_{n+1}^*))$.

Lemma 3 (Second Period Expected Value $E_{y'}(V^*)$)**.** .

1. *Information doesn't matter*

$$\frac{\partial^2 g(\pi(y'))}{\partial \pi(y')\partial \pi(y')} = 0: \qquad E_{y'}(V_n^*) = E_{y'}(V_{n+1}^*) \quad \forall n$$

2. *Information matters*

$$\frac{\partial^2 g(\pi(y'))}{\partial \pi(y')\partial \pi(y')} > 0: \qquad E_{y'}(V_n^*) < E_{y'}(V_{n+1}^*)$$
$$\frac{\partial^2 g(\pi(y'))}{\partial \pi(y')\partial \pi(y')} < 0: \qquad E_{y'}(V_n^*) > E_{y'}(V_{n+1}^*)$$

[19] See Davidson (1994) p.233 Theorem 15.5. The theorem states only the case for the submartingale, but can be extended to the case of a supermartingale. For this theorem to hold it has to be the case that $E_{y'}(V_n^*) < \infty$, which is always fulfilled as $0 \leq E_{y'}(V_n^*) < v_h \; \forall n$.

(i) unless

(a) $v_l \to v_h$

(b) $p \to 1$ *or* $p \to 0.5$ *for* $\theta \in \{\theta_h, \theta_l\}$,

(c) $q \to 1$ *or* $q \to 0$.

Then,

$$E_{y'}(V_n^*) = E_{y'}(V^*) \quad \forall n.$$

(ii) Abundance of Information: Let $n_2 \to \infty$ *and, hence,* $n \to \infty$. *Then,*

$$\lim_{n\to\infty} E_{y'}(V_n^*) = E_{y'}(V^*).$$

Part 1 states that the expected value in the second period is independent of the number of signals, if V^* is a martingale. But if V^* is a submartingale, having an additional signal leads to a higher expected value. The reverse is true if V^* is a supermartingale: then an additional signal decreases the expected value. The dependence of the second period value on information is reinforced when the difference between the payoff in the high and low state is large, when signals are not completely (un)informative and when the prior belief is bounded away from extreme values one and zero (part (i)). Under such circumstances, information matters even more. Last, if external information and hence overall information becomes abundant, an additional signal has no impact on payoff variability (part(ii)).

Based on this, the star node has the same expected value as the ring if V^* is a martingale. If V^* is a submartingale, the star has a higher second period expected value than the ring. If V^* is a supermartingale, then the ring has a higher second period expected value than the star. When the ring and the star have different second period expected values, then this difference is negligible if the payoff in the high and low state are equivalent, signals are completely informative or uniformative or if the prior is correct. The difference between ring and star also disappears whenever there is an abundance of information.

Now that we know how the three determinants of effort, zu, $\pi(y)$, and $E_{y'}(V^*)$ differ for the ring and the star, we can start to show under which circumstances the star and the ring performs better.

The overall effect of a larger ego network size (at the expense of common friends) on effort is still ambiguous. Agents with a larger ego component have more signals, which has two effects: First, it makes the expected posterior belief given a state in the first period more accurate. Second, it affects the second period expected payoff, where the agents take expectations *before* any state in the second period is realized. The second effect is ambiguous and can push or decrease first period effort depending on whether V^* is a sub- or a supermartingale. The first effect pushes effort if the realized state is high and hence if the first period signals point towards the high state. Hence, a larger ego network size encourages effort when the state of the world is high relatively to the situation where the state of the world is bad. On the other hand, men with their larger ego network have no common friends, which works as a disincentive of effort *in both states* because the threat of second period punishment after a first period project failure is weaker. Despite these ambiguities, we can draw some conclusion depending on the state of the world on period one. In order to make this state dependence clear, we write $e(\theta)$.

Proposition 3 (First Period Effort in Star versus Ring)**.** .

(i) State-Dependence: $\bar{e}^S(\theta_h) > \bar{e}^S(\theta_l)$ *and* $\bar{e}^R(\theta_h) > \bar{e}^R(\theta_l)$.

(ii) Increasing Differences: $\bar{e}^S(\theta_h) - \bar{e}^S(\theta_l) > \bar{e}^R(\theta_h) - \bar{e}^R(\theta_l)$.

(iii) Uncertainty:

(a) If $v_l \rightarrow v_h$ then $\bar{e}^R(\theta) > \bar{e}^S(\theta)\ \forall \theta$.

(b) If $q \rightarrow 1$ then $\bar{e}^R(\theta_h) > \bar{e}^S(\theta_h)$; if $q \rightarrow 0$ then $\bar{e}^R(\theta_l) > \bar{e}^S(\theta_l)$.

(iv) Signal Informativeness: If $p \rightarrow 0.5$ or $p \rightarrow 1$, then $\bar{e}^R(\theta) > \bar{e}^S(\theta)\ \forall \theta$.

(v) Abundance of Information: If $n_2 \rightarrow \infty$, then $\bar{e}^R(\theta) > \bar{e}^S(\theta) \quad \forall \theta$.

(vi) Discounting: If $\beta \rightarrow 0$, then $\bar{e}^R(\theta_l) > \bar{e}^S(\theta_l)$ and $\bar{e}^R(\theta_h) < \bar{e}^S(\theta_h)$.

It is clear that both men and women put more effort when signals point towards the high state (part (i)). Interestingly, the difference between effort in high versus low state is larger in the star compared to the ring (part (ii)). Men's effort is more sensitive to the state of the world. Our interpretation is that men choose effort more accurately: High effort when signals indicate the high state and low effort when signals point towards low state. To the contrary, women's posterior belief about the state of the world in the first period is less correct, which is why their effort across states is more balanced. They exert relatively little effort when the state is high and relatively much when the state is low.

Part (iii) states that women outperform men in terms of effort if the uncertainty of the underlying environment vanishes: (a) when values in both states are the same ($v_l \rightarrow v_h$) or (b) when they have a prior which reflects certainty about the state (for $q = 0$ when $\theta = \theta_l$ or $q = 1$ when $\theta = \theta_h$). When there is no uncertainty, men's ego networks lose their advantage, which is purely information-based. Since the effort-pushing effect of clustering does not directly depend on the uncertainty of the environment, women put more effort than men no matter which state has been realized. Stated differently, a *necessary condition* for men to put more effort than women is that there is uncertainty in the underlying environment: (a) when $v_h - v_l$ is large; (b) for $0 << q << 1$. In uncertain environments, men might reap the benefits of their ego network structure and perform relatively better than women, but this depends primarily on the second period expected value and its properties.

Women outperform men in terms of effort and thus project completion if the signals are either completely informative or completely uninformative, i.e. for $p \rightarrow 0.5$ or $p \rightarrow 1$ (part (iv)). If signals are completely uninformative, then the additional information of men is not valuable. On the other hand, if they are completely informative, then already the first signal reveals the state, which is why the additional information of men does not matter either. This holds independent of the state realization. Hence, a *necessary condition* for men to have a higher project completion rate than the ring (i.e. higher effort) is that the signals are neither completely informative nor completely uninformative. Only then men's additional information is valuable.

If we make external information abundant for both men and women (part (v)), then women have a higher first period project completion rate independent of the state of the world. This is due to a saturation effect of information which takes place when the information held suffices to have a precise idea about the state. At this point, the additional (internal) information of men do not impact the posterior belief and thus not their effort choices. Effort differences between men and women are solely determined by differences in clustering, with the women having higher clustering, which drives up their effort.

Finally, as we increase discounting ($\beta \rightarrow 0$), women's ego network loses its advantage because the effort-pushing effect from clustering is only at work if the time horizon is dynamic. Women can only reap the benefits from their informal networks when the decision today matters for tomorrow.

After having discussed, effort in detail, we turn now to wages. We define first and second period wages as follows.

Definition 2 (Equilibrium Wages)**.** .
The first period wage and second period wage are respectively given by:

$$w_i(\theta) = f(\bar{e}(\theta), \bar{e}(\theta))v(\theta) \tag{12}$$
$$w_i'(\theta, \theta') = \beta z_i[f(\bar{e}(\theta), \bar{e}(\theta)) + (1 - u_{ij})(1 - f(\bar{e}(\theta), \bar{e}(\theta)))]f(\bar{e}'(\theta'), \bar{e}'(\theta'))v(\theta') \tag{13}$$

where $\theta \in \{\theta_l, \theta_h\}$ *(*$\theta' \in \{\theta_l', \theta_h'\}$*) is the realized state in the first (second) period.*

Note that $\bar{e}(\theta)$ and $\bar{e}'(\theta')$ do not carry any indices because they indicate first and second period *equilibrium* effort levels, which are symmetric across project partners. Both wages are state-dependent. They are given by the project value (conditional on the state) weighted by the probability of project success. Not only the first but also the second period wage depend on the realization of the first period state. This dynamic effect exists because the second period project success hinges on the likelihood of punishment and hence on the project success in the first period. The second period wage additionally depends on the realization of the second period state.

At this stage we cannot characterize the wages any further analytically. Moreover, Proposition 3 gives only results in the limit. But we are also interested in what happens if we move further away from the different limits. Therefore,we look at a parametric example to illustrate this.

6 Parametric Example

In this section, we solve the two-period model for a parametric example in closed form, which we use in the next section to see what happens as we move away from the limit. Specify the effort function as $f(e_i, e_j) = \sqrt{e_i e_j}$ and let the cost for effort be given by $c(e_i) = \frac{1}{2}e_i^2$. Both functional forms satisfy Assumption (1) from above for $e_i, e_j \in [0, 1)$. From equation (5) and its counterpart for j, we can derive

$$\bar{e}' = \bar{e}_i' = \bar{e}_j' = \frac{1}{2}\pi(y') \tag{14}$$

Now, the optimal choices in the first period can be worked out, taking into account that optimal choices in the second period are foreseen. From equations (7) and (**??**), we find first period effort levels:

$$\bar{e} = \bar{e}_i = \bar{e}_j = \frac{1}{2}\left(\pi(y) + \beta z u \frac{3}{8} E_{y'}\left(\pi(y')^2\right)\right) \tag{15}$$

As discussed above, we are mostly interested in men's and women's effort choices *given* a particular state has realized. The first period effort given a state is denoted by

$$\bar{e}(\theta) = \left(\frac{1}{2}\left(E_y(\pi(y)|\theta) + \beta z u \frac{3}{8} E_{y'}\left(\pi(y')^2\right)\right)\right), \tag{16}$$

for $\theta \in \{\theta_l, \theta_h.\}$ The second line in (16) follows by linearity of expectations. Similarly, for the second period effort, averaging over the second period signal structures *given* a second period state, we obtain from (14):

$$\bar{e}'(\theta') = \frac{1}{2}E_{y'}\left(\pi(y')|\theta'\right) \tag{17}$$

for $\theta' \in \{\theta_l', \theta_h'\}$. We can now state for this example how effort in both period varies depends on the n and K.

Using the definition of wages from the general model, first and second period wages in this parametric example are given by Definition (3), where we condition them on the realized state.

Definition 3 (Equilibrium Wages)**.** .
First and second period wage are respectively defined as:

$$w_i(\theta) = \sqrt{\overline{e}(\theta)^2}v(\theta) = \overbrace{\frac{1}{2}\left(E_y\left(\pi(y)|\theta\right) + \beta zu\frac{3}{8}E_{y'}\left(\pi(y')^2\right)\right)}^{\overline{e}(\theta)} v(\theta) \tag{18}$$

$$\begin{aligned} w_i'(\theta,\theta') &= \beta z_i(\sqrt{\overline{e}(\theta)^2} + (1-u_{ij})(1-\sqrt{\overline{e}(\theta)^2}))\sqrt{\overline{e}'(\theta)^2}v(\theta') \\ &= \beta z_i\left(1-u_{ij}\left(1-\underbrace{\frac{1}{2}\left(E_y\left(\pi(y)|\theta\right)+\beta zu\frac{3}{8}E_{y'}\left(\pi(y')^2\right)\right)}_{\overline{e}(\theta)}\right)\right)\underbrace{\frac{1}{2}E_{y'}\left(\pi(y')|\theta'\right)}_{\overline{e}'(\theta')}v(\theta') \end{aligned} \tag{19}$$

The following Lemma directly follows from wage equations (18) and (19).

Lemma 4 (Equilibrium Wages)**.** .

$$\frac{\partial w_i(\theta)}{\partial E_y\left(\pi(y)|\theta\right)} > 0 \quad \frac{\partial w_i(\theta)}{\partial E_{y'}\left(\pi(y')^2\right)} > 0 \quad \frac{\partial w_i(\theta)}{\partial zu} > 0$$

$$\frac{\partial w_i'(\theta,\theta')}{\partial E_y\left(\pi(y)|\theta\right)} > 0 \quad \frac{\partial w_i'(\theta,\theta')}{\partial E_{y'}\left(\pi(y')^2\right)} > 0 \quad \frac{\partial w_i'(\theta,\theta')}{\partial E_{y'}\left(\pi(y')|\theta'\right)} > 0 \quad \frac{\partial w_i'(\theta,\theta')}{\partial zu} \lesseqgtr 0 \quad \frac{\partial w_i'(\theta,\theta')}{\partial z_i} > 0$$

As it is straightforward to show this, we omit the proof. The first period wage positively depends on first period expected payoff $E_y\left(\pi(y)|\theta\right)$, which reflects the instantaneous gain, the second moment of second period payoff $E_{y'}\left(\pi(y')^2\right)$ and zu, all pushing first period effort and thus the probability of project success.

Likewise, the second period wage positively depends on $E_y\left(\pi(y)|\theta\right)$ and $E_{y'}\left(\pi(y')^2\right)$ because of their positive effect on first period effort, which makes punishment in the second period less likely. Moreover, the second period wage positively depends on the second period expected payoff $E_{y'}\left(\pi(y')|\theta'\right)$, which reflects the instantaneous gain in that period. Notice that (19) ambiguously depends on zu. There is a positive effect of zu because of its positive effect on first period effort. But there is also a negative effect of zu because first period failures are more likely to be punished by playing the low effort equilibrium, which yields zero payoff for both project partners. Finally, there is a positive effect of the probability to be selected, z_i, which positively depends on one's degree.

Having analyzed in detail the determinants of equilibrium effort and wages and how they differ between men and women, we can return to our central question. Under which circumstances do men outperform women in terms of project completion (i.e. effort) and when is the opposite true? Whether the star or the ring node exerts more first period effort depends on the model's parameters:

Our simulations confirm these results. Our baseline parameters are given in Table 4.

Table 4: Baseline Parameters

v_l	v_h	p	q	β	d_i^R	d_i^S	K^R	K^S	M
0	1	0.75	0.5	0.9	2	3	2	1	10

In figure 3, we compare men's and women's effort as a function of external information first when the high state is the true first period state (left) and then for the low state (right).[20] Under the high state of the world, men exert higher effort than women for low amounts of external information, whereas the opposite is true for high amounts of external information. We denote by n_2^* the information cutoff, i.e. the amount of external information where men's and women's effort functions intersect. The intuition behind this result is that men's additional information is valuable only when overall information is scarce. Once there is much information, the additional signal of men has no bearing and women perform better. Notice that as n_2 becomes large, the effort functions are flat. Any additional piece of information leaves beliefs and hence effort unchanged. If the first period state is low (panel 2), both men's and women's effort converges to the low value ($v_l = 0$) as information becomes abundant. However, women's effort is higher than men's for all amounts of external information (for all n_2). As opposed to the high state, there is no cutoff point here.

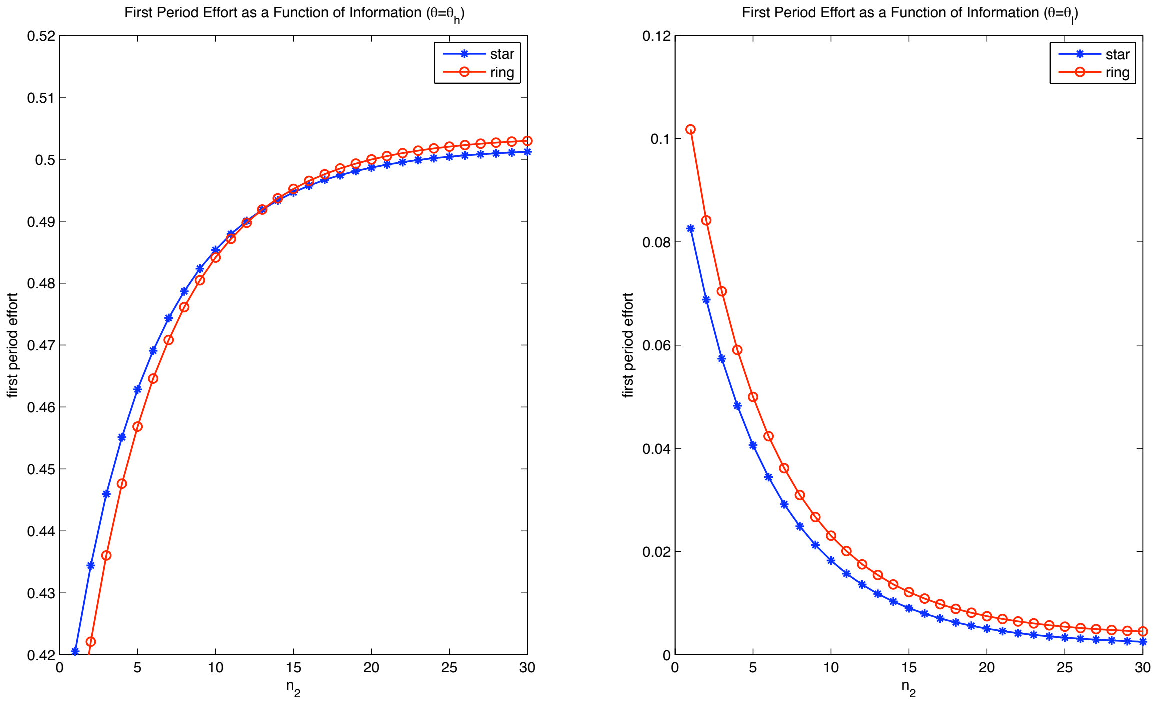
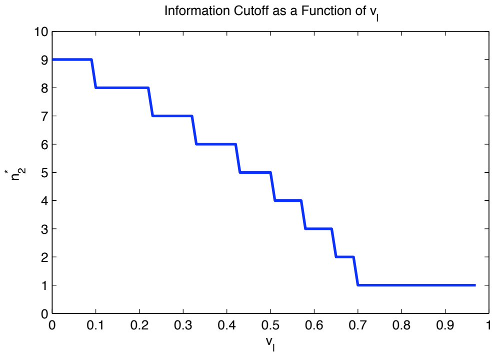


Figure 3: Effort in Star and Ring Depending on First Period State

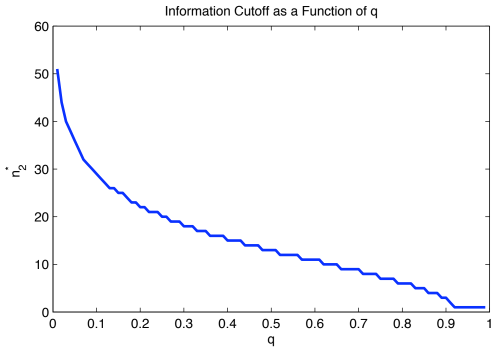
The information cutoff n_2^* indicates the amount of information below which information is more conducive to effort than peer pressure. A higher n_2^* reflects a stronger advantage of men over women. In the following simulations we show that this threshold changes with the underlying uncertainty of the environment as well as with the informativeness of the signals, confirming our results (iii) and (iv) of Proposition 3. We plot the information cutoff n_2^* as a function of the low value v_l (figure 4(a)), the probability of the high state q (figure 4(b)), and signal informativeness p (figure 4(c)). First, men's informational advantage is more pronounced when the project outcome is uncertain. This is the case when low and high project values are very different or when the probability that the high state occurs is very low.

Second, the informational advantage of men are more pronounced when signals are neither completely informative nor uninformative. When signals are completely uninformative, the informational advantage of men vanishes. In turn, when they are extremely informative, then women get a good grasp about the state even though they have less signals than men. This is why the information cutoff

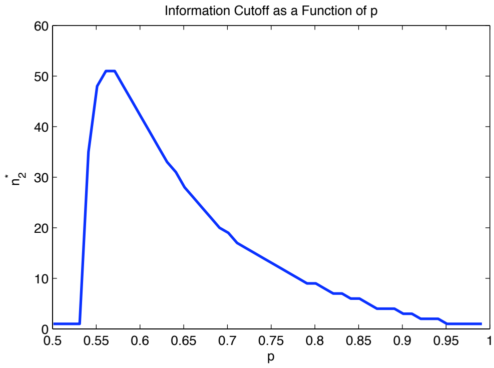
[20]Recall that we set $n_2^R = n_2^S = n_2$.



(a) Cutoff v_l



(b) Cutoff q



(c) Cutoff p

Figure 4: Cutoffs

n_2^* as a function of p is non-monotonous in p, as pictured in the graph below.
These predictions on effort in star and ring readily translate in the following results about their wages.

Proposition 4 (Wages in Star versus Ring (Parametric Example))**.** .
(i) First period wage $w(\theta)$*: If* $\bar{e}^S(\theta) > (<)\bar{e}^R(\theta)$ *then* $w^S(\theta) > (<)w^R(\theta)$.
(ii) Wage Dynamics: If $w_i^S(\theta) > w_i^R(\theta)$*, then* $w_i^S(\theta, \theta') > w_i^R(\theta, \theta')$ *for* $\theta' = \theta_h$*. This wage gap persistence is weakened when*

(a) $v_l \to v_l$*,*

(b) when $p \to 1$ *or* $p \to 0.5$*,*

(c) when $q \to 1$ *or* $q \to 0$*.*

(b) If $w_i^S(\theta) > w_i^R(\theta)$*, for* $w_i'^S(\theta, \theta') < w_i'^R(\theta, \theta')$ *to happen, it must be that* $\theta' = \theta_l$*.*

Part (i) states that whether star or ring node has a higher first period wage $w_i(\theta)$ solely depends their relative project completion rate: If the star node has a higher project completion in the first period, then his wage is higher for either state. If the ring node has the higher project completion, the opposite is true. Due to the tight connection between first period project completion and first period wage, the results from Proposition 3 can be readily applied: Women earn a higher wage when the uncertainty about the underlying environment is low, when the signals are completely (un)informative or when (external) information becomes abundant. Men fare relatively better under uncertainty and high discounting of the future.

The second period wage does not only depend on the second period state but also on the first period state through first period effort, which determines the likelihood of punishment in the second period. In situations, where the man in the star already had a higher first period project completion compared to the woman in the ring and hence earned a higher first period wage (by part (i)), the wage gap persists if the state in the second period is high (part (ii)). The more uncertainty there is in the underlying environment, the larger the second period wage advantage of the man. Only if the second period state is low, the woman can revert the wage inequality. If this happens, the wage advantage of women over men can never be as large as the one of men over women if the second period state is high. This shows the strong impact of the performance at the beginning of the career on the future wage trajectory of men relative to women.

7 Comparison of Effort Levels

To show whether the effort increases in an efficient manner when using this trigger strategy, we first compare the optimal effort as chosen by a social planner to the high static Nash equilibrium effort as well as to the high effort chosen when sticking to the trigger strategy. The planner's problem is given by

$$\max_{e_i,e_j} f(e_i,e_j)2(Pr(\theta_h|y)v_h + (1-Pr(\theta_h|y))v_l) - c(e_i) - c(e_j)$$

Taking derivatives yields

$$2f_1(e_i,e_j)(Pr(\theta_h|y)v_h + (1-Pr(\theta_h|y))v_l) = c'(e_i)$$
$$2f_2(e_i,e_j)(Pr(\theta_h|y)v_h + (1-Pr(\theta_h|y))v_l) = c'(e_j)$$

There are again two solutions, one with zero effort and one with strictly positive effort. As the strictly positive effort yields a higher payoff, this is the solution we focus on. Due to the assumptions made on $f(\cdot,\cdot)$ and $c(\cdot)$, $e_i^* = e_j^* \equiv e^*$.

The effort level in the high effort static Nash equilibrium is the same as the high effort level in the second period when playing the trigger strategy. Recall that this effort is given by

$$f_1(\overline{e}_i',\overline{e}_j')(Pr(\theta_h|y)v_h + (1-Pr(\theta_h|y))v_l) - c'(\overline{e}_i') = 0$$
$$f_2(\overline{e}_i',\overline{e}_j')(Pr(\theta_h|y)v_h + (1-Pr(\theta_h|y))v_l) - c'(\overline{e}_j') = 0$$

To emphasize that this is the static high level Nash equilibrium, we define $e^{NE} \equiv \overline{e}'$

Last, the first period high effort level of the trigger strategy is given by equations,

$$f_1(\overline{e}_i,\overline{e}_j)(Pr(\theta_h|y)v_h + (1-Pr(\theta_h|y))v_l + \beta zuE_{y'}V_{i,2}^*) = c'(\overline{e}_i)$$
$$f_2(\overline{e}_i,\overline{e}_j)(Pr(\theta_h|y)v_h + (1-Pr(\theta_h|y))v_l + \beta zuE_{y'}V_{j,2}^*) = c'(\overline{e}_j)$$

Comparing these conditions yields the following proposition.

Proposition 5. *The static Nash equilibrium effort level is lower than the first period trigger strategy equilibrium effort level* $e^{NE}(y) < \overline{e}(y), \forall y$

The proof can be found in the Appendix. The trigger strategy leads to an increase in effort compared to the effort selected in the static high effort Nash equilibrium. The effort in the high static Nash equilibrium is lower than the planner's effort, as can be seen in Proposition (6).

Proposition 6. *The effort level of the static high effort Nash equilibrium is lower than the effort the planner chooses,* $\overline{e}' < e^*$*, for any signal realization.*

The proof is omitted as it works along the same lines as the proof of Proposition (5).

It is not clear however, whether the high effort exerted in the first period when playing according to the trigger strategy is higher or lower than the planner's effort. In fact, anything is possible.

Proposition 7. *The first period effort chosen when sticking to the trigger strategy can be higher, lower or the same as the effort implemented by a planner.*

We compare the following two equations:

$$2f_1(e_i, e_j)(Pr(\theta_h|y)v_h + (1 - Pr(\theta_h|y))v_l) = c'(e_i)$$
$$f_1(\bar{e}_i, \bar{e}_j)(Pr(\theta_h|y)v_h + (1 - Pr(\theta_h|y))v_l + \beta zuE_{y'}V^*_{i,2}) = c'(\bar{e}_i)$$

Whether the effort under the trigger strategy is higher, the same or lower than the planner's effort depends on the whether $\beta zuE_{y'}V^*_{i,2}$ is higher, the same or lower than $(Pr(\theta_h|y)v_h + (1 - Pr(\theta_h|y))v_l)$. The over-effort occurs when $\beta zuE_{y'}V^*_{i,2} > (Pr(\theta_h|y)v_h + (1 - Pr(\theta_h|y))v_l)$. This is more likely if the expected payoff in the first period is sufficiently low. But we know from our previous analysis that women choose higher effort than men if the number of low signals is high. Therefore, women are more likely to exert more effort than socially optimal.

Given the effort in the first period when playing the trigger strategy is at most for one signal realization the same as the effort of the planner, it is clear that in expectation the utility for an agent is higher when the planner prescribes an effort than if they choose the effort according to the trigger strategy. Whenever $e^{NE} < \bar{e} < e^*$, then the trigger strategy leads for sure to a higher first period utility than playing the Nash equilibrium. If this condition does not hold, it might be that the first period utility under the Nash equilibrium is higher than the utility given by the trigger strategy.

8 Discussion

In the comparative statics exercises we compared efforts and wages of two agents, a man and a woman, whose ego networks are star and ring respectively. The man under consideration is the central node in the star who is doing a project with one of the peripheral nodes and the woman is the node in a ring who is doing a project with another ring-node. This way of modeling gender ego networks is based on our empirical findings from the AddHealth data that women's networks are characterized by higher levels of clustering and that women have less friends. From our comparative statics exercises, we derive a series of predictions, which we here connect to the gender differences in labor market outcomes described in the introduction and literature review.

Women Outperform Men when Uncertainty About the State of the World is Negligible. Because there is a lower information flow in the ring compared to the star, women do better than men when information is not valuable, i.e. when the uncertainty about the state of the world is negligible. This is the case (a) when the signals are completely (un)informative, (b) when the variance between low and high value is small, (c) when the prior probability are extreme so that the agents are ex-ante convinced that either the low or the high state will materialize, (d) when (external) information is abundant, making internal information from the friends in the ego network futile. In turn, men have a more accurate idea about the state of the world because they hold more info, which is informative. They perform relatively better under uncertainty and when the realized state of the world is high, whereas women do relatively better in terms of project completion when the realized state is low. This suggests several reasons of why women do worse on the labor market than observationally similar men. One is that men are drawn to occupations and tasks, which yield a higher return if things work out, but a low one if not. Instead, women put *over-effort* (compared to men) into projects that yield a low return. We would expect men and women to sort into different occupations, namely men choosing occupations that are risky, because they are relatively better at dealing with this risk compared to women. Women, on the other hand, more than men seek lines of work, such as education and health care or work in public and non-profit organizations, where they reap lower market rewards (Fortin and Huberman (2002)), but these rewards are certain. Our theory offers an alternative explanation of why women choose less risky occupation than men, in contrast to the standard explantation of women being more risk averse, see Eckel and Grossman (2008) for an overview of the literature that finds women to be more risk averse than men.

Differences in First Period Wage are Decisive for the Future Wage Gap. In our model, when men outperform women at the beginning of the career (period 1), then they continue to do so in the second period in states that yield high payoff. Only if a low state realizes in the future, women have the chance to catch up. If they do, the second period wage advantage of women over men can never be as large as the one of men over women if the second period state is high. This is in line with the literature that points out the importance of the gender wage gap at early stages of the career for the future wage path (Babcock and Laschever (2003), Gerhart and Rynes (1991), Martell et al. (1996)).

Moreover, the relatively poor performance of women might stem from possibly large differences in outcomes across gender when overall information is scarce so that internal information from the ego network is crucial. This is likely to be the case at the beginning of the career. In contrast, women would outperform men when information is abundant, which is is typically the case at higher levels of the organizational hierarchy, where managers obtain information from many divisions in the company and a large informal network. Hence, women might be outperformed by men at the starting point of their career, which is why they get stuck at low levels of the hierarchy while men move up the career letter. This is in line with a finding by Blinder (1973) that a considerable part of the gender wage gap stems from different age-wage profiles. Women exhibit an almost flat age-wage profile, i.e. their wages do not rise over the life cycle, whereas male wages do.

Men Outperform Women When the Future is Highly Discounted. The benefits of women's network structure (i.e. higher clustering compared to that of men which creates peer pressure) are of dynamic nature: Women tend to put more effort into projects to prevent a failure because this would not only jeopardize the relationship with the project partner but also the one with the common friend. If there is no subsequent period or if the discount factor is low, this mechanism fails to work. This suggests that women do better in stable environments and permanent positions, where they face the same potential project partners in every period.

Women's Ego Network Structure is Particularly Beneficial when Information is Abundant. The model suggests that having women in the network is particularly beneficial in work environments where information is abundant. There, a woman who brings some closure to the network is more beneficial than an additional man who brings even more info because information has already reached some saturation level. Complementarities between information and peer pressure kick in, which is why project completion in a diversified network is higher than in a network that is fully info-based.

This might shed light on the finding by Lalanne and Seabright (2011) that having females in the network is beneficial to both male and female executives but not to agents at lower levels in the organizational hierarchy. The underlying assumption is that executives have more sources of information than workers at lower levels of the organizational hierarchy, which is why additional information is more valuable at these lower levels.

References

ARAL, S., E. BRYNJOLFSSON, AND M. W. VAN ALSTYNE (2011): "Information, Technology and Information Worker Productivity," *Information Systems Research, Forthcoming*.

ARROW, K. AND R. BORZEKOWSKI (2004): "Limited network connections and the distribution of wages," *FEDS Working Paper No. 2004-41*.

BABCOCK, L. AND S. LASCHEVER (2003): *Women Don't Ask: Negotiation and the Gender Divide.*, Princeton University Press.

BELLE, D. (1989): *Children's social networks and social supports*, vol. 136, John Wiley & Sons Inc.

BLINDER, A. (1973): "Wage Discrimination: Reduced Form and Structural Estimates," *The Journal of Human Resources*, 8, 436–455.

BOHNET, I. AND F. SAIDI (2011): "Informational Differences and Performance: The Role of Organizational Demography," *Working Paper*.

BOOTH, A. L. (2009): "Gender and Competition," *IZA Discussion Paper*.

BURT, R. (1992): *Structural holes: The social structure of competition*, Harvard Univ Pr.

——— (2011): "Structural holes in virtual worlds," Unpublished.

CALVO-ARMENGOL, A. AND M. JACKSON (2004): "The effects of social networks on employment and inequality," *The American Economic Review*, 94, 426–454.

CALVÓ-ARMENGOL, A. AND M. JACKSON (2007): "Networks in labor markets: Wage and employment dynamics and inequality," *Journal of economic theory*, 132, 27–46.

COLEMAN, J. (1988a): "Free riders and zealots: The role of social networks," *Sociological Theory*, 6, 52–57.

——— (1988b): "Social capital in the creation of human capital," *American journal of sociology*, 95–120.

CONTI, G., A. GALEOTTI, G. MUELLER, AND S. PUDNEY (2009): "Popularity," Tech. rep., Institute for Social and Economic Research.

DAVIDSON, J. (1994): *Stochastic Limit Theory: An Introduction for Econometricicans*, Oxford University Press, USA.

DIXIT, A. (2003): "Trade expansion and contract enforcement," *Journal of Political Economy*, 111, 1293–1317.

EASLEY, D. AND J. KLEINBERG (2010): *Networks, crowds, and markets*, Cambridge Univ Press.

ECKEL, C. AND P. GROSSMAN (2008): "Men, women and risk aversion: Experimental evidence," *Handbook of experimental economics results*, 1, 1061–1073.

EDER, D. AND M. HALLINAN (1978): "Sex differences in children's friendships," *American Sociological Review*, 237–250.

EDIN, P., P. FREDRIKSSON, AND O. ÅSLUND (2003): "Ethnic enclaves and the economic success of immigrants—Evidence from a natural experiment," *The Quarterly Journal of Economics*, 118, 329–357.

FISCHER, C. AND S. OLIKER (1983): "A research note on friendship, gender, and the life cycle," *Social Forces*, 62, 124–133.

FLABBI, L. (2012): "Gender Differences in Education, Career Choices and Labor Market Outcomes on a Sample of OECD Countries," *World Development Report, Gender Equality and Development*.

FORTIN, N. AND M. HUBERMAN (2002): "Occupational gender segregation and women's wages in Canada: an historical perspective," *Canadian Public Policy/Analyse de Politiques*, 11–39.

GERHART, B. (1990): "Gender Differences in Current and Starting Salaries: The Role of Performance, College Major and Job Title," *Industrial and Labor Relations Review*, 43, 418–433.

GERHART, B. AND S. RYNES (1991): "Determinants and Consequences of Salary Negotiations by Male and Female MBA Graduates," *Journal of Applied Psychology*, 76, 256–262.

GRANOVETTER, M. (1973): "The strength of weak ties," *American journal of sociology*, 1360–1380.

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

KANDEL, E. AND E. LAZEAR (1992): "Peer pressure and partnerships," *Journal of Political Economy*, 801–817.

KARLAN, D., M. MOBIUS, T. ROSENBLAT, AND A. SZEIDL (2009): "Trust and social collateral," *The Quarterly Journal of Economics*, 124, 1307–1361.

KÜRTÖSI, Z. (2008): "Differences in female and male social networks in a work setting Differences in female and male social networks in a work setting Differencesin female and male social networks in a work setting," Ph.D. thesis, Sociology Department Budapest.

LALANNE, M. AND P. SEABRIGHT (2011): "The Old Boy Network: Gender Differences in the Impact of Social Networks on Remuneration in Top Executive Jobs," Unpublished.

MARSDEN, P. (1987): "Core discussion networks of Americans," *American sociological review*, 122–131.

——— (2005): "Recent developments in network measurement," *Models and methods in social network analysis*, 8, 30.

MARTELL, R., D. LANDE, AND C. EMRICH (1996): "Male-Female Differences: A Computer Simulation," *American Psychologist*, 51, 157–158.

MUNSHI, K. (2003): "Networks in the modern economy: Mexican migrants in the US labor market," *The Quarterly Journal of Economics*, 118, 549–599.

PUTNAM, R. (2000): *Bowling Alone: America's Declining Social Capital*, Simon and Schuster.

TATTERSALL, A. AND C. KEOGH (2006): "Women, Networking and the Impact on Pay, Position and Status in the UK ICT Industry," Unpublished.

Appendix

Derivation of z_i and u_{ij}

The probability that one agent is chosen is given by

$$Pr(C) = \frac{M-1}{\frac{1}{2}M(M-1)} = \frac{2}{M},$$

and the probability that this agent i is linked to the suggested project partner j, given that he is selected by

$$Pr(g_{ij} = 1|C) = \frac{d_i}{M-1}.$$

Then, the probability of being chosen *and* being partnered with a friend is

$$z_i \equiv Pr(g_{ij} = 1 \wedge C) = Pr(g_{ij} = 1|C)Pr(C) = \frac{2d_i}{M(M-1)}.$$

$$u_{ij} \equiv Pr(\sum_{k,k\neq i,k\neq j} g_{ik}g_{jk}|g_{ij} = 1 \wedge C)$$

Proof of Proposition 1: One-Shot Game

Effort is symmetric Distinguish two cases: First, in the low-effort equilibrium (non-interior), given that one agent believes that the other one exerts zero effort, he himself also exerts zero effort, hence there exists one equilibrium in which for all y $e_i^*(y) = e_j^*(y) = 0$. Second, in the high-effort equilibrium (interior), we have $e_i^*, e_j^* \neq 0$ and the FOCs for i and j hold at the equilibrium effort levels. The first order condition for worker i is given by

$$f_1(e_i^*, e_j^*)(Pr(\theta_h|y)v_h + (1 - Pr(\theta_h|y))v_l) - c'(e_i^*) = 0. \tag{20}$$

Since the signals in a component are public so that the posterior about the state of the world is the same among project partners, we can combine their first order conditions to obtain:

$$\frac{f_1(e_i^*, e_j^*)}{f_2(e_i^*, e_j^*)} = \frac{c'(e_i^*)}{c'(e_j^*)} \tag{21}$$

Proceed by contradiction and assume that $e_j^* > e_i^*$. Due to convexity of the cost functions, the RHS of (21) is smaller than one. Due to concavity and supermodularity of the effort function, we have $f_1(e_i^*, e_j^*) > f_2(e_i^*, e_j^*)$, which is why the LHS is larger than one. A contradiction. A similar argument shows that $e_j^* < e_i^*$ also yields a contradiction. Hence, both agents exert the same effort level in the second period $e_i^* = e_j^*$.
There are two Nash equilibria in pure strategies. We know that all possible equilibria are symmetric. We also know that when one worker chooses zero effort, then the other worker will also choose zero effort. It remains to show that there is exactly one Nash equilibrium where both workers exert strictly positive effort. In order to show this, it suffices to show that the FOCs (which under symmetry become a function of one variable) have one zero under the condition that effort is strictly positive.

Imposing symmetry of the effort levels $e_i^* = e_j^* = e^*$ to obtain

$$f_1(e^*, e^*)(Pr(\theta_h|y)v_h + (1 - Pr(\theta_h|y))v_l) = c'(e^*) \tag{22}$$

where $f_1(e, e)$ is non-increasing in e: Under CRS of $f(.)$, $f_1(e, e)$ is constant in e. Under DRS, $f_1(e, e)$ is decreasing in e. Increasing returns to scale are ruled out by Assumption 1. In turn, the function $c'(e)$ is increasing in e, also due to Assumption 1, starting in the origin. Hence, the two functions have a unique intersection, implying a unique symmetric equilibrium with strictly positive effort.

Proof of Proposition 2: Nash equilibrium after $\overline{h}$

We have already excluded zero effort, that is, equilibrium effort has to be strictly positive. The symmetry and uniqueness in the second period, the last period, can be shown exactly as in the proof of Proposition 1, as there are no more dynamic considerations. It remains to show symmetry and uniqueness of the equilibrium in the first period

Both agents exert the same effort in the first period. The argument is similar to the one used for the proof of Proposition 1. In the first period, we focus on the high-effort equilibrium (interior solution) because we assume a positive friendship history. Hence, $\overline{e}_i, \overline{e}_j \neq 0$. Due to step one, $E_{y'}V_{i,2}^* = E_{y'}V_{j,2}^* \equiv E_{y'}V_2^*$. Also notice that $z_i u_{ij} = z_j u_{ji} = zu$. Taken together the two conditions imply that $Pr(\theta_h|y)v_h + (1 - Pr(\theta_h|y)v_l + \beta zu E_{y'}V_2^*$ is the same across agents. We can then combine the first-period FOCs (7) and the corresponding one for worker j to obtain:

$$\frac{f_1(\overline{e}_i, \overline{e}_j)}{f_2(\overline{e}_i, \overline{e}_j)} = \frac{c'(\overline{e}_i)}{c'(\overline{e}_j)} \tag{23}$$

By the same argument as in the proof for Proposition 1, it follows that equilibrium effort levels in the first period are symmetric across project partners, or $\overline{e}_i = \overline{e}_j$.

Uniqueness of $\overline{e} > 0$ The argument for the first period is analogous to the argument provided in the Proof of Proposition 1 and is therefore omitted.

Proof of Lemma 1: $\pi(y)$ has the martingale property

$$\pi(y_n') = Pr(\theta_h'|y_n')v_h + (1 - Pr(\theta_h'|y_n'))v_l$$

Define $\theta_n' \equiv Pr(\theta_h'|y_n')$. We know that the stochastic process $\{\theta_n'\}$ is a martingale as

$$E(\theta_{n+1}'|y_n') = E(E(\theta'|y_{n+1}')|y_n') = E(\theta'|y_n') = \theta_n',$$

where the second equality follows from the *tower property* of conditional expectations. Then,

$$\begin{aligned} E(\pi(y_{n+1}')|y_n') &= E(\theta_{n+1}'v_h + (1 - \theta_{n+1}')v_l|y_n') = E(\theta_{n+1}'v_h|y_n') + E((1 - \theta_{n+1}')v_l|y_n') \\ &= v_h E(\theta_{n+1}'|y_n) + v_l(1 - E(\theta_{n+1}'|y_n')) = \theta_n'v_h + (1 - \theta_n')v_l \\ &= \pi(y_n') \end{aligned}$$

Proof of Lemma 2: Properties of $E_y(\pi(y_n))$ and $E_y(\pi(y_n)|\theta)$

1. The number of signals do not matter for $E_y(\pi(y))$ due to the martingale property of $\pi(y)$.

$$E(\pi(y_{n+1})) = E(E(\pi(y_{n+1})|y_n)) = E(\pi(y_n))$$

2. We note that the posterior is given by

$$Pr(\theta_h|y) = \frac{Pr(y|\theta_h)Pr(\theta_h)}{Pr(\theta_h)Pr(y|\theta_h) + Pr(\theta_l)Pr(y|\theta_l)} = \frac{qp^y(1-p)^{n-y}}{qp^y(1-p)^{n-y} + (1-q)p^{n-y}(1-p)^y} \tag{24}$$

$$Pr(\theta_l|y) = \frac{Pr(y|\theta_l)Pr(\theta_l)}{Pr(\theta_h)Pr(y|\theta_h) + Pr(\theta_l)Pr(y|\theta_l)} = \frac{(1-q)p^{n-y}(1-p)^y}{qp^y(1-p)^{n-y} + (1-q)p^{n-y}(1-p)^y} = 1 - Pr(\theta_h|y). \tag{25}$$

What matters is the difference in the number of high and low signals. To see this we first set $y = n-y$. In this case, $Pr(\theta_h|y) = q$ and $Pr(\theta_l|y) = 1-q$, i.e. no updating. Moreover, $\frac{\partial Pr(\theta_h|y)}{\partial y} > 0$ and $\frac{\partial Pr(\theta_l|y)}{\partial (n-y)} > 0$, which implies that high (low) signals lead to the belief that the high (low) state more likely.
This is the basis for what follows from now on. The proof of

$$E_y\left(\pi(y_{n+1})|\theta_h\right) > E_y\left(\pi(y_n)|\theta_h\right) \qquad E_y\left(\pi(y_{n+1})|\theta_l\right) < E_y\left(\pi(y_n)|\theta_l\right)$$

will be left to part (ii).
(i) Additional signals do not matter.
(a) For $v_l \to v_h$,

$$\begin{aligned}\lim_{v_l \to v_h} E_y\left(\pi(y)|\theta_h\right) &= \sum_{i=0}^{n} \frac{(n)!}{i!(n-i)!}\left(p^{n-i}(1-p)^i\right) v_h \\ &= (p+1-p)^{n+1} v_h = v_h\end{aligned}$$

which is independent of n. Similarly for $E_y\left(\pi(y)|\theta_l\right)$. Moreover,

$$\begin{aligned}\lim_{v_h \to v_l} E_y\left(\pi(y)|\theta_h\right) &= \sum_{i=0}^{n} \frac{(n)!}{i!(n-i)!}\left(p^{n-i}(1-p)^i\right) v_l \\ &= (p+1-p)^{n+1} v_l = v_l\end{aligned}$$

which is independent of n. Similarly for $E_y\left(\pi(y)|\theta_l\right)$. Hence, $\lim_{v_l \to v_h} \frac{E_y(\pi(y)|\theta)}{\partial n} = \lim_{v_h \to v_l} \frac{E_y(\pi(y)|\theta)}{\partial n} = 0$.
(b) Assume $p \to 0.5$. Then,

$$\begin{aligned}\lim_{p \to 0.5} E_y\left(\pi(y)|\theta_h\right) &= \sum_{i=0}^{n} \frac{(n)!}{i!(n-i)!}\left(p^{n-i}(1-p)^i\right)\left(\frac{0.5^n}{0.5^n}\frac{qv_h + (1-q)v_l}{q + (1-q)}\right) \\ &= \frac{qv_h + (1-q)v_l}{q + (1-q)}\end{aligned}$$

and similarly for $E_y\left(\pi(y)|\theta_l\right)$.
Assume $p \to 1$. Then,

$$\begin{aligned}\lim_{p \to 1} E_y\left(\pi(y)|\theta_h\right) &= \lim_{p \to 1} \sum_{y=0}^{n} \frac{(n)!}{y!(n-y)!}\left(p^y(1-p)^{n-y}\right)\left(\frac{qp^y(1-p)^{n-y}v_h + (1-q)p^{n-y}(1-p)^y v_l}{qp^y(1-p)^{n-y} + (1-q)p^{n-y}(1-p)^y}\right) \\ &= \lim_{p \to 1} \frac{(n)!}{n!(n-n)!}\left(p^n(1-p)^{n-n}\right)\left(\frac{qp^n(1-p)^{n-n}v_h + (1-q)p^{n-n}(1-p)^n v_l}{qp^n(1-p)^{n-n} + (1-q)p^{n-n}(1-p)^n}\right) \\ &= \lim_{p \to 1} p^n \left(\frac{qp^n v_h + (1-q)(1-p)^n v_l}{qp^n + (1-q)(1-p)^n}\right) \\ &= v_h,\end{aligned}$$

and analogue for $\theta = \theta_l$

(c) Assume $q \to 1$. Then,

$$\lim_{q\to 1} E_y\left(\pi(y)|\theta_h\right) = \sum_{i=0}^{n} \frac{(n)!}{i!(n-i)!}\left(p^{n-i}(1-p)^i\right) v_h$$
$$= (p+1-p)^n v_h = v_h$$

which is independent of n. Similarly for $q \to 0$ and $E_y\left(\pi(y)|\theta_l\right)$.

(ii) Note that $y \sim$ Binomial$(np, np(1-p))$ if $\theta = \theta_h$ and $y \sim$ Binomial$(n(1-p), np(1-p))$ if $\theta = \theta_l$. Then, $\lim_{n\to\infty}(y-(n-y)) = \infty$ if $\theta = \theta_h$ and $\lim_{n\to\infty}(y-(n-y)) = -\infty$ if $\theta = \theta_l$. To see this note that $y-(n-y) = 2y-n$. By the central limit theorem, as $n \to \infty$,

$$\text{if} \quad \theta = \theta_h \quad y \xrightarrow{p} np \quad \Rightarrow \lim_{n\to\infty}(2np-n) = \infty$$
$$\text{if} \quad \theta = \theta_l \quad y \xrightarrow{p} n(1-p) \quad \Rightarrow \lim_{n\to\infty}(2n(1-p)-n) = -\infty.$$

Then, $\lim_{n\to\infty} Pr(\theta_h|y) = 1$ if $\theta = \theta_h$ and $\lim_{n\to\infty} Pr(\theta_h|y) = 0$ if $\theta = \theta_l$ as

$$\lim_{n\to\infty} Pr(\theta_h|y) = \lim_{n\to\infty} \frac{1}{1+\frac{1-q}{q}\left(\frac{1-p}{p}\right)^{2y-n}}$$

Next, we can show that $Pr(\theta_h|y)$ is increasing in n if $\theta = \theta_h$ and decreasing in n if $\theta = \theta_l$. We show this in detail for $\theta = \theta_h$ and leave the proof of $\theta = \theta_l$ to the reader. If $\theta = \theta_h$, $Pr(\theta_h|y) = \frac{1}{1+\frac{1-q}{q}\left(\frac{1-p}{p}\right)^{2y-n}} \equiv \frac{1}{1+\frac{1-q}{q}\left(\frac{1-p}{p}\right)^{\hat{y}}}$, where $\hat{y}$ denotes the difference between the number of high and low signals. To simplify notation, we set $\frac{1-p}{p} \equiv \tilde{p}$ and $\frac{1-q}{q} \equiv \tilde{q}$

Given we are in the high state, an additional high signal arrives with probability p and a low signal with probability $1-p$. Then, for $Pr(\theta_h|y)$ is increasing in n if $\theta = \theta_h$ it has to hold that

$$\frac{1}{1+\tilde{q}\tilde{p}^{\hat{y}}} < \frac{p}{1+\tilde{q}\tilde{p}^{\hat{y}+1}} + \frac{1-p}{1+\tilde{q}\tilde{p}^{\hat{y}-1}}$$
$$\Leftrightarrow \quad (1+\tilde{q}\tilde{p}^{\hat{y}+1})(1+\tilde{q}\tilde{p}^{\hat{y}-1}) < p(1+\tilde{q}\tilde{p}^{\hat{y}})(1+\tilde{q}\tilde{p}^{\hat{y}-1}) + (1-p)(1+\tilde{q}\tilde{p}^{\hat{y}})(1+\tilde{q}\tilde{p}^{\hat{y}+1})$$
$$\Leftrightarrow \quad \tilde{p}^{\hat{y}+1} + \tilde{p}^{\hat{y}-1} + \tilde{q}\tilde{p}^{2\hat{y}} < p(\tilde{p}^{\hat{y}} + \tilde{p}^{\hat{y}-1} + \tilde{q}\tilde{p}^{2\hat{y}-1}) + (1-p)(\tilde{p}^{\hat{y}}\tilde{p}^{\hat{y}+1} + \tilde{q}\tilde{p}^{2\hat{y}+1})$$
$$\Leftrightarrow \quad p\tilde{p}^{\hat{y}+1} + (1-p)\tilde{p}^{\hat{y}-1} + \tilde{q}\tilde{p}^{2\hat{y}} < \tilde{p}^{\hat{y}} + \tilde{q}(p\tilde{p}^{2\hat{y}-1} + (1-p)\tilde{p}^{2\hat{y}+1})$$
$$\Leftrightarrow \quad p\tilde{p}^2 + (1-p) - \tilde{p} < \tilde{q}\tilde{p}^{\hat{y}}(p + (1-p)\tilde{p}^2 - \tilde{p})$$

Note that $p\tilde{p}^2 + (1-p) - \tilde{p} = 0$. Then,

$$0 < \tilde{q}\tilde{p}^{\hat{y}}(p + (1-p)\tilde{p}^2 - \tilde{p}),$$

which holds for $p > \frac{1}{2}$. Then we can apply the Monotone Convergence Theorem, which implies that $\lim_{n\to\infty} E_y(Pr(\theta_h|y)v_h) = E_y(\lim_{n\to\infty} Pr(\theta_h|y)v_h)$. From this it follows that $\lim_{n\to\infty} E_y\left(\pi(y)|\theta_h\right) = v_h$ and $\lim_{n\to\infty} E_y\left(\pi(y)|\theta_l\right) = v_l$. From the monotone convergence, it also follows that

$$E_y\left(\pi(y_{n+1})|\theta_h\right) > E_y\left(\pi(y_n)|\theta_h\right) \qquad E_y\left(\pi(y_{n+1})|\theta_l\right) < E_y\left(\pi(y_n)|\theta_l\right).$$

Proof of Lemma 3: $E_{y'}(V*)$

1. The number of signals does not matter if the second period expected value is a martingale. This follows from the definition of a martingale.
2. These properties follow from the definitions of sub- and supermartingales, respectively.
(i) (a) $v_l \to v_h$

$$\lim_{v_l \to v_h} E_{y'}(V_n^*) = \lim_{v_l \to v_h} \sum_{y=0}^{n} \frac{n!}{(y)!(n-y)!}(qp^y(1-p)^{n-y} + (1-q)p^{n-y}(1-p)^y) \\ \times (f(\overline{e}'(y), \overline{e}'(y))\pi(y) - c(\overline{e}'(y)))$$

As the other terms are constant in v_l, all that matter is

$$\begin{aligned} \lim_{v_l \to v_h} (f(\overline{e}'(y), \overline{e}'(y))\pi(y) - c(\overline{e}'(y))) &= \lim_{v_l \to v_h} f(\overline{e}'(y), \overline{e}'(y)) \lim_{v_l \to v_h} \pi(y) - \lim_{v_l \to v_h} c(\overline{e}'(y)) \\ &= \lim_{v_l \to v_h} f(\overline{e}'(y), \overline{e}'(y)) v_h - \lim_{v_l \to v_h} c(\overline{e}'(y)) \end{aligned}$$

Note that

$$\lim_{\pi(y) \to v_h} \overline{e}'(\pi(y)) = \overline{e}'_{v_h},$$

i.e. the effort converges against some constant as $\pi(y) \to v_h$ and

$$\lim_{\overline{e}' \to \overline{e}'_{v_h}} f(\overline{e}', \overline{e}') = b,$$

with b being another constant. As $f(.,.)$ is continuous, i.e. $f(\overline{e}'_{v_h}, \overline{e}'_{v_h}) = b$, we know that

$$\lim_{\pi(y) \to v_h} f(\overline{e}', \overline{e}') = b.$$

The argument is the same for $c(.)$ and is therefore omitted. Then, we can write

$$\lim_{v_l \to v_h} (f(\overline{e}'(y), \overline{e}'(y))\pi(y) - c(\overline{e}'(y))) = b_{v_l},$$

which is constant and thus independent of n. Therefore,

$$\lim_{v_l \to v_h} E_{y'}(V_n^*) = b_{v_l}$$

(b) $p \to 1$ or $p \to 0.5$ for $\theta \in \{\theta_h, \theta_l\}$,

$$\lim_{p \to 1} E_{y'}(V_n^*) = \lim_{p \to 1} \sum_{y=0}^{n} \frac{n!}{(y)!(n-y)!}(qp^y(1-p)^{n-y} + (1-q)p^{n-y}(1-p)^y) \\ \times (f(\overline{e}'(y), \overline{e}'(y))\pi(y) - c(\overline{e}'(y)))$$

$$\lim_{p\to 1}\pi(y) = v_h \quad \text{if} \quad n-2y<0$$
$$\lim_{p\to 1}\pi(y) = qv_h + (1-q)v_l \quad \text{if} \quad n-2y=0$$
$$\lim_{p\to 1}\pi(y) = v_l \quad \text{if} \quad n-2y>0$$

As $\pi(y)$ converges to some constant, so does $(f(\overline{e}'(y),\overline{e}'(y))\pi(y) - c(\overline{e}'(y)))$. We denote by $V^*(v_h)$ $(V^*(v_l))$ $[V^*(v)]$ the limit when $\pi(y)$ converges to v_h (v_l) $[qv_h+(1-q)v_l]$.
Note further that if $n-2y<0$

$$\begin{aligned}\lim_{p\to 1}(qp^y(1-p)^{n-y} + (1-q)p^{n-y}(1-p)^y) &= \lim_{p\to 1} qp^y(1-p)^{n-y}\left(1+\frac{1-q}{q}\left(\frac{p}{1-p}\right)^{n-2y}\right)\\ &= \lim_{p\to 1} qp^y(1-p)^{n-y}\end{aligned}$$

Then we know that

$$\lim_{p\to 1} = \begin{cases} q & \text{if} \quad y=n \\ 0 & \text{otherwise}\end{cases}$$

If $n-2y>0$

$$\begin{aligned}\lim_{p\to 1}(qp^y(1-p)^{n-y} + (1-q)p^{n-y}(1-p)^y) &= \lim_{p\to 1}(1-q)p^{n-y}(1-p)^y(\frac{q}{1-q}\left(\frac{1-p}{p}\right)^{n-2y}+1)\\ &= \lim_{p\to 1}(1-q)p^{n-y}(1-p)^y.\end{aligned}$$

It follows that

$$\lim_{p\to 1} = \begin{cases} 1-q & \text{if} \quad y=0 \\ 0 & \text{otherwise}\end{cases}$$

Last, if $n-2y=0$

$$\lim_{p\to 1}(qp^y(1-p)^{n-y} + (1-q)p^{n-y}(1-p)^y) = \lim_{p\to 1} p^y(1-p)^{n-y} = 0,$$

as $y,n>0$ From this it then follows that

$$\lim_{p\to 1} E_{y'}(V_n^*) = qV^*(v_h) + (1-q)V^*(v_l),$$

which is independent of n.
Next, let $\lim_{p\to .5}$

$$\lim_{p\to .5} E_{y'}(V_n^*) = \lim_{p\to 1}\sum_{y=1}^{n}\frac{n!}{(y-1)!(n+1-y)!}(qp^{n+1-y}(1-p)^{y-1} + (1-q)p^{y-1}(1-p)^{n+1-y})$$
$$\times\,(f(\overline{e}'(y),\overline{e}'(y))\pi(y) - c(\overline{e}'(y)))$$

Then again, $\lim_{p\to .5} E_{y'}(V_n^*)\pi(y) = qv_h + (1-q)v_l$, that is the limit is constant. Therefore, by the same logic as in (i)

$$\lim_{p \to .5} E_{y'}(V_n^*) = b_p,$$

where b_p is a constant.
(c) $q \to 1$

$$\lim_{q \to 1} E_{y'}(V_n^*) = \lim_{q \to 1} \sum_{y=1}^{n} \frac{n!}{(y-1)!(n+1-y)!} (qp^{n+1-y}(1-p)^{y-1} + (1-q)p^{y-1}(1-p)^{n+1-y}) \\ \times (f(\overline{e}'(y), \overline{e}'(y))\pi(y) - c(\overline{e}'(y)))$$

Then,

$$\lim_{q \to 1} (qp^{n+1-y}(1-p)^{y-1} + (1-q)p^{y-1}(1-p)^{n+1-y}) = p^{n+1-y}(1-p)^{y-1},$$
$$\lim_{q \to 1} \pi(y) = v_h.$$

$\lim_{q \to 1} E_{y'}(V_n^*)$ is then again a constant and independent of n.
$q \to 0$

$$\lim_{q \to 0} (qp^{n+1-y}(1-p)^{y-1} + (1-q)p^{y-1}(1-p)^{n+1-y}) = p^{y-1}(1-p)^{n+1-y},$$
$$\lim_{q \to 0} \pi(y) = v_l.$$

$\lim_{q \to 0} E_{y'}(V_n^*)$ is then again a constant and independent of n
(ii) Abundance of Information: We want to show that

$$\lim_{n \to \infty} E_{y'}(V_n^*) = E_{y'}(V^*).$$

Note that if $\{V_n^*\}$ is a submartingale sequence and $\sup_n E(V_n^*) \leq M < \infty$, then $V_n^* \to V^*$ almost surely.[21] We know that $\sup_n E(V_n^*) \leq M < \infty$, as $E(V_n^*) < v_h$. Then we can apply the monotone convergence theorem, from which is follows that $\lim_{n \to \infty} E_{y'}(V_n^*) = E_{y'}(V^*)$.

Proof of Proposition 3: effort of star and ring

Assume $n_1^S > n_1^R$ (and hence $n^S > n^R$) and $K^S < K^R$.
(i) Follows from the assumption $n_1^S > n_1^R$ and Lemma 2.
(ii) $\overline{e}^S(\theta_h) - \overline{e}^S(\theta_l) = E_y\left(\pi(y^S)|\theta_h\right) - E\left(\pi(y^S)|\theta_l\right) > E_y\left(\pi(y^R)|\theta_h\right) - E\left(\pi(y^R)|\theta_l\right) = \overline{e}^R(\theta_h) - \overline{e}^R(\theta_l)$ where the inequality follows from the assumption $n_1^S > n_1^R$ and Lemma 2.
(iii) .
(a) For $v_l \to v_h$, Lemmas 2 and 3 imply that $\overline{e}^R(\theta) - \overline{e}^S(\theta) = (zu)^R - (zu)^S > 0$, which is positive $\forall\theta$.
(b) For $q \to 0$, $q \to 1$, Lemmas 2 and 3 imply that $\overline{e}^R(\theta) - \overline{e}^S(\theta) = (zu)^R - (zu)^S > 0$, which is positive $\forall\theta$.
(iv) For $p \to 0.5$, $p \to 1$, Lemmas 2 and 3 imply that $\overline{e}^R(\theta) - \overline{e}^S(\theta) = (zu)^R - (zu)^S > 0$, which is positive $\forall\theta$.
(v) Lemmas 2 and 3 imply that $\overline{e}^R(\theta) - \overline{e}^S(\theta) = (zu)^R - (zu)^S > 0$, which is positive $\forall\theta$.

[21] See Davidson (1994) p.235 Theorem 15.7.

(vi) If $\beta \to 0$, then $\bar{e}^S(\theta) = E_y\left(\pi(y^S)|\theta\right)$ and $\bar{e}^R(\theta) = E_y\left(\pi(y^R)|\theta\right)$. The result then follows from Lemma 2.

Proof of Proposition 4: wages of star and ring

Assume $n_1^S > n_1^R$ and (hence $n^S > n^R$) and $K^S < K^R$. Define $w(\theta)$ and $w'(\theta, \theta')$ as in Definition 3.
(i) Follows from Definition 3.
(ii) . Suppose $w_i^S(\theta) > w_i^R(\theta)$, which implies $\bar{e}^S(\theta) > \bar{e}^R(\theta)$ by part (i). $w_i(\theta, \theta')$ is given by:

$$w_i'(\theta, \theta') = \beta z_i \underbrace{(\bar{e}(\theta) + (1 - u_{ij})(1 - \bar{e}(\theta)))}_{\equiv X} \bar{e}'(\theta) v(\theta')$$

Notice that $u_{ij}^S = \frac{1}{3}$ and $u_{ij}^R = 1$. Hence,

$$X^S = \frac{2}{3} + \frac{1}{3}\bar{e}^S(\theta)$$
$$X^R = \bar{e}^R(\theta).$$

It follows that $X^S > X^R$ for $\bar{e}(\theta)^S > \bar{e}(\theta)^R$. Moreover, for $\theta' = \theta_h$, it holds that $\bar{e}'(\theta')^S > \bar{e}'(\theta'^R)$ (by Proposition **??**). Also, $z_i^S > z_i^R$ due to $d_i^S > d_i^R$. This together implies that $w_i'^S(\theta, \theta') > w_i'^R(\theta, \theta')$.
The stated dependence on the parameters follows from Propositions 3 and **??**.

Proof of Proposition 5: $\bar{e} > e^{NE}$

We know that for a given signal realization e^{NE} is given by

$$f_1(e^{NE}, e^{NE})(Pr(\theta_h|y)v_h + (1 - Pr(\theta_h|y))v_l) = c'(e^{NE})$$

and $\bar{e}$

$$f_1(\bar{e}, \bar{e})(Pr(\theta_h|y)v_h + (1 - Pr(\theta_h|y))v_l + \beta z_i u_{ij} E_{y'} V_{i,2}^*) = c'(\bar{e})$$

Dividing the two equations,

$$\frac{f_1(e^{NE}, e^{NE})(Pr(\theta_h|y)v_h + (1 - Pr(\theta_h|y))v_l)}{f_1(\bar{e}, \bar{e})(Pr(\theta_h|y)v_h + (1 - Pr(\theta_h|y))v_l + \beta z_i u_{ij} E_{y'} V_{i,2}^*)} = \frac{c'(e^{NE})}{c'(\bar{e})}$$
$$\frac{(Pr(\theta_h|y)v_h + (1 - Pr(\theta_h|y))v_l)}{(Pr(\theta_h|y)v_h + (1 - Pr(\theta_h|y))v_l + \beta z_i u_{ij} E_{y'} V_{i,2}^*)} = \frac{f_1(\bar{e}, \bar{e})c'(e^{NE})}{f_1(e^{NE}, e^{NE})c'(\bar{e})} \tag{26}$$

As $\beta z_i u_{ij} E_{y'} V_{i,2}^* > 0$, we know that

$$\frac{(Pr(\theta_h|y)v_h + (1 - Pr(\theta_h|y))v_l)}{(Pr(\theta_h|y)v_h + (1 - Pr(\theta_h|y))v_l + \beta z_i u_{ij} E_{y'} V_{i,2}^*)} < 1$$
$$\Rightarrow \frac{f_1(\bar{e}, \bar{e})c'(e^{NE})}{f_1(e^{NE}, e^{NE})c'(\bar{e})} < 1 \tag{27}$$

Given the assumptions on $f(\cdot, \cdot)$ and $c(\cdot)$, for (27) to hold, it has to be the case that $e^{NE} < \bar{e}$. As this holds for any signal realization, we know that $e^{NE} < \bar{e}, \forall y$.