Diversity, Social Goods Provision, and Performance in the Firm

Sara Fisher Ellison

Department of Economics
MIT, Cambridge, MA 02142
sellison@mit.edu

Wallace P. Mullin

Department of Economics
George Washington University
Washington, DC 20052
wpmullin@gwu.edu

Economists have studied the effect of diversity on the provision of social goods and the stock of social capital. In parallel, management scholars have studied the effect of diversity in the workplace on firm performance. We integrate these two growing literatures and explore these questions with a unique dataset. A firm provided eight years of individual-level employee survey data, which include measures of the stock of social capital, plus office-level measures of diversity and performance. We find some evidence that more gender-homogeneous offices enjoy higher levels of social goods provision but those offices do not perform any better and may actually perform worse.

1. Introduction

Legal and societal shifts in 20th Century America laid the groundwork for increased diversity in many settings. Schools ceased to be racially segregated. Barriers to female and minority employment diminished, leading to more diverse workplaces. Greater mobility resulted in neighborhoods fragmented in various dimensions. The consequent social benefits of this increased diversity, though difficult to quantify, may be quite important. With these broad social changes as a backdrop, the focus of this paper is smaller but sharper. We are interested in how diversity in a group affects the provision of social goods in the group, and then, ultimately, performance of that group. In particular, we focus attention on diversity in a market environment, that created by a firm and its workforce. Regardless of the cause of the increased workplace diversity, it is the job of the managers to encourage the greatest productivity possible from their units, maximizing profits, perhaps, or some other quantifiable outcome. It is our goal, then, to shed light on how diversity is associated with those outcomes.

Because these and related questions of diversity, social goods, and performance are of fundamental interest outside of workplace settings as well, it is not surprising that our research bears upon several strands of literature, including ones in strategic

management, organizational behavior, sociology, and economics. Economists and sociologists have focused more on the first of our two questions, how diversity in a group affects the provision of social goods. Management scholars, in contrast, have focused more on the second question, how diversity affects the performance of the group. In fact, part of the contribution we see this paper making is to serve as a sort of bridge between those literatures.

Economists' interest in the effect of diversity on the provision of social goods is ongoing. Studies have found evidence that social goods are provided at a lower level in communities exhibiting fragmentation on various dimensions. Vigdor (2004) finds that census response rates are lower in census tracts with higher ethnic fragmentation. Costa and Kahn (2003) find that desertion rates are higher in Civil War military companies with higher age and occupational fragmentation. Glaeser et al. (2000) find that trust is lower among Harvard undergraduates when race and nationality fragmentation is higher. Several studies have documented that school funding is higher in more homogenous communities (see, e.g., Goldin and Katz, 1999; Poterba, 1997; Miguel and Gugerty, 2005).[1] These results may be quite important in contexts where social goods provision is the output of interest. However, in some contexts, the social good may be an "intermediate good." In the workplace, cooperation, trust, and other social goods may be important elements of the functioning of an office, but firm owners ultimately care about an office's performance, as reflected in revenues, costs, and profits. We would also like to address this additional question, so often missing in the economics literature.

Not surprisingly, a sizeable management literature on diversity and performance exists.[2] This literature has seen a recent emphasis on the importance of "context" on the effect that diversity has on performance (see Joshi and Roh, 2009). In other words, covariates such as industry, profession, and organizational culture could play a role in determining diversity's effect. In fact, if we think of the level of social goods in a group as an element of context, our two-step analysis elucidates the role of context variables and the mechanisms by which they are important.

Finally, we draw on a paper by Rob and Zemsky (2002), which provides a formal model of social capital. Modifying their model, we elucidate how diversity could influence the accumulation of social capital and allow a separate channel for diversity to affect office performance directly. This model provides testable hypotheses, which can be evaluated in our empirical setting or by other researchers in different settings.

We have a unique data set from a firm that operates numerous small offices in the United States and abroad. They have provided us with eight years of individual-level employee survey data as well as office-level measures of diversity and performance. The survey data furnish us with several indicators of firm social capital, such the level of cooperation among employees. The data allow us to address two distinct questions. First, broadly speaking, do we find lower levels of social goods provision in more diverse offices? Such a finding has been made in the economics literature cited above, but our results provide an interesting complement to those: economists have previously focused on the effects of diversity in communities instead of workplaces,[3] and we measure diversity on two dimensions not explored in this literature, gender and tenure. We do find that higher office-level gender diversity is associated with lower employee cooperation, but that tenure diversity has little or no effect.

---

1. See also Costa and Kahn (2003) for an excellent survey of this literature.
2. Two comprehensive surveys are Williams and O'Reilly (1998) and Shore et al. (2009).
3. Costa and Kahn's study of Union Army troops is a possible exception.

In addition to our direct measures of office-level diversity, we also have employee responses to a question about whether the firm is accepting of diversity. We construct an office average of this response. Interestingly, we find that offices where the employees think the firm is accepting of diversity tend to be *more* cooperative. We could interpret these two sets of results, seemingly at odds, in the following way: employees like the idea of a diverse workplace—and may therefore provide social goods more readily in a setting that they think is supportive of diversity—but are actually more comfortable in a homogeneous setting.

Although we find these results interesting, we want to focus also on a second question: What is the effect of office diversity on performance? One can imagine a situation where diversity leads to low cooperation in a workplace, which could hinder performance. It is also possible, however, that diversity could enhance productivity in other ways. Which effect dominates becomes an empirical question, and one that we can, in principle, answer with our data. Again, we look both at the effects of diversity as well as the perception among the employees that the firm supports diversity. We find the perception the firm supports diversity has no association with revenues. Gender diversity, in contrast, is associated with higher revenues, although this latter effect diminishes when we include office-level fixed effects. In addition, both average tenure and tenure diversity in the office are important explanatory variables.

In other words, our results suggest that, consistent with the previous economics literature, employees are more cooperative in more homogenous settings. These more homogenous units, however, seem to be less productive overall.

2. Social Capital, Our Setting, and the Empirical Literature on Diversity

It is useful now to define social capital and discuss ways of measuring it. It will also be useful to dive more deeply into the management literature on diversity and discuss our place there. We follow Putnam's (1995) definition: "By analogy with notions of physical and human capital—tools and training that enhance individual productivity—'social capital' refers to features of social organizations such as networks, norms, and social trust that facilitate coordination and cooperation for mutual benefit." Empirical studies of social capital, such as ours, must grapple with the question of how to measure it.[4] One possibility would be to measure the size and nature of social networks, as in Karlan et al. (2009). Such a measure would not be relevant in our setting—all employees in a particular office would know each other well—but that paper suggests that the informal exchange of favors may be an important component of social capital. In our setting, those favors would occur within a firm and could be well proxied by a measure of how cooperative an office is.

Survey responses may be the best way to gauge such phenomena.[5] Indeed, we have data from employee surveys in our firm, such as responses to the question of how cooperative the office is. Glaeser et al. (2000) use survey data as well, in conjunction with a "Trust" game, to gauge social capital.

Given the particulars of our data, there are a number of papers in the management literature that are relevant. An early paper looking at the effects of tenure diversity is

---

4. In some situations, a direct measure of the *provision* of social goods, such as whether a soldier deserts, might be available. That is an example of an act that contributes to the stock of social capital as opposed to a measure of the stock.

5. See Putnam (2000) for a detailed account of the use of survey data to measure social capital.

Zenger and Lawrence (1989). They study engineers and note that tenure homogeneity should lower communication costs, because employees hired at the same time develop a common language and shared experiences. Ancona and Caldwell (1992) also look at the effects of tenure diversity as well as functional diversity. The authors found that tenure diversity was associated with improved task work, such as setting priorities, but also detracted from team performance such as meeting budget schedules.

Like us, Pelled et al. (1999) study both gender and tenure diversity, although their performance measure is quite different from ours. They have a sample of 45 teams across three firms, and they follow Ancona and Caldwell by using a manager survey to measure team performance, in this case, emotional conflict. They found no significant effect of gender diversity on emotional conflict, even after using different measures of gender diversity as a sensitivity analysis. Company tenure diversity did have a positive effect on emotional conflict.

Recent work by members of the Diversity Research Network, a cross-discipline group of scholars studying the effects of diversity, has also explored firm or business unit performance and indicators of diversity (Kochan et al., 2003). Of particular relevance is a study by Ely (2004) that examines the effect of four types of diversity (tenure, age, sex, and race) on the performance of 486 retail bank branches. The cross-section includes demographic information of each employee, employee survey results, and branch performance measures. As a measure of team process,[6] the author created an index of survey responses to questions such as "The people I work with cooperate to get the job done" and "Employees at my work unit feel comfortable working with people from different backgrounds." One could view these questions as gauges of social capital. In contrast to our results, she finds that the diversity variables were not related to quality of team process, but quality of team process was positively related to team performance. Neither race nor sex diversity was positively related to team performance. Interestingly, the direct effect of age and tenure diversity on measured branch performance was negative. Some of these results are consistent with our findings whereas some stand in contrast. There is, though, a broad similarity between Ely's paper and ours, and we see our paper as a complement to hers.

Finally, two other papers point to related but distinct literatures: Gant et al. (2002) study social capital and productivity in a workplace. Using both survey data and direct observation, they explore the effects of different human resource practices on the social networks in firms. Hellerstein et al. (1999) study how gender composition affects productivity and wages.

3. Theory

In discussing these ideas, we find it useful to refer to a model to provide some structure and insight into the mechanisms at play. Rob and Zemsky (2002) (hereafter RZ) provide such a framework.[7] Their model shows how employees in a firm can be given incentives to contribute to social capital as well as the dynamics of how a stock of social capital can persist or deteriorate over time. Although they do not explicitly discuss diversity, their model can be interpreted and modified to incorporate various channels through which diversity can operate.

---

6. In the Williams and O'Reilly (1998) taxonomy, diversity can affect "group process" and "group performance." Although our notions of "social capital" and "performance" might not map directly onto "group process and performance" studied there, there are similarities.

7. Based on the social psychology literature, Roberge and Van Dick (2010) provide an informal theory but no formal model of when and how diversity increases group performance.

They start with a continuum of employees on the unit interval, each indexed by an idiosyncratic tendency to feel guilty, $\epsilon_j \sim \mathcal{U}[0, 1]$. Each must choose two effort levels, individual effort e_I and cooperative effort e_C. We think of the cooperative effort as contributing to the firm's social capital. The firm cannot observe the split between these two types of effort but can measure output, which is a function of both. In particular, observed output of individual j is $\hat{Q}_j = ae_I(j) + \frac{1}{2}\min(e_C(j), 1) + \frac{1}{2}\int_0^1 \min(e_C(i), 1)di$, with a a constant where $\frac{1}{2} < a < 1$. Note that an individual's cooperative effort contributes strictly less to his observed output than his individual effort. Note also that a contribution to an employee's output comes through the cooperative effort his coworkers have supplied.[8]

Driving an employee's decision about how to divide his effort is his utility function, $U_j = W_j - c(e_I(j) + e_C(j) - \bar{e})^2 - (1 - e_C(j))(r\bar{e}_C + \epsilon_j)$. Where W_j is his wage, c is a parameter governing how costly it is to him to supply effort (of either type), $\bar{e}$ is the preferred total effort level, and r is something the authors refer to as reciprocity, which we will relate below to both diversity and cooperation. We can think of the third term as representing guilt that the employee feels from not contributing the optimal amount of cooperative effort. $(1 - e_C(j))$ is the amount of shirking that he has to be guilty about and $(r\bar{e}_C + \epsilon_j)$ is the strength of his guilt. The strength of guilt is a function of reciprocity in the firm, average cooperative effort, and employee j's idiosyncratic guilt tendency.

RZ study a dynamic model where individual effort is a static best response to total collective effort in the previous period. They show that, from any initial conditions, the system converges to a steady state. In particular, the steady-state level of cooperative effort, $\bar{e}_C$, can be either zero or equal to some (positive and increasing) function of reciprocity, $f(r)$, depending on initial conditions. Reciprocity affects steady-state cooperative effort in two ways: (1) a higher r means a larger set of initial conditions from which the model converges to $\bar{e}_C = f(r)$ instead of 0, and (2) the steady state itself is a function of r. A direct implication of these results, relevant to our empirical exercise, is that the expected steady-state level of cooperative effort is increasing in reciprocity. To see this, note that $E(\bar{e}_C) = f(r)\text{Prob}(\bar{e}_{C0} > \tilde{e}_C(r))$, where $\bar{e}_{C0}$ is the initial cooperative effort and $\tilde{e}_C(r)$ is the critical value above which $\bar{e}_{C0}$ will converge to $f(r)$ in steady state. Both $f(r)$ and the probability are increasing in r, so $E(\bar{e}_C)$ is as well.

We can modify the model to allow diversity effects on social goods provision. If a firm or office is more homogeneous, its level of reciprocity could be higher because employees are more willing to contribute to a social good in a community where others are similar to them, for example. The model predicts $E(\bar{e}_C)$ is monotone increasing in r. If r is, as described above, a decreasing function of diversity, D, then $E(\bar{e}_C)$ is a decreasing function of D. To look for such evidence, we adopt a simple linear specification:

$$\bar{e}_{Ci} = \alpha_0 + \alpha_1 D_i + \gamma_i,$$

where i indexes different firms or offices and D_i is a measure of diversity.

A second potential channel for diversity effects is directly through observed output: we could modify $\hat{Q}$ to include an extra term $h(D)$.[9] Combining this

8. Here, we have normalized to 1 the amount of cooperation that is optimal from the firm's perspective; therefore, any additional cooperative effort above 1 will not contribute to output.

9. Although, the RZ model is hard-wired to ensure that cooperative effort never exceeds the firm's optimal level, one could imagine that employees could have a social preference for cooperation that is different from the firm's preference. In that case, the guilt term in employee utility might continue to be present even if the optimal amount of cooperation from the point of view of the firm had been achieved, and equilibria could exist where cooperative effort is oversupplied.

TABLE I.
SUMMARY STATISTICS

Variable	Obs.	Mean	Std. Dev.	Min	Max
At the employee level					
Satisfaction	1,707	3.943	0.990	1	5
Morale	1,683	3.592	1.017	1	5
Cooperate	1,541	4.038	1.036	1	5
Male	1,648	0.329	0.470	0	1
TenureYears	1,665	2.570	2.087	0.25	7
At the office-year level					
Unemploy	269	4.77	1.84	1.4	12.2
Number	269	4.94	3.12	2	19
AvgSatisfaction	269	4.06	0.58	2	5
AvgDPerception	269	4.73	0.36	3	5
AvgMorale	269	3.74	0.66	1	5
AvgCooperate	248	4.14	0.64	2	5
AvgGender	269	0.29	0.25	0	1
AvgTYears	269	2.32	1.14	0.25	6.25
GendDiversity	269	0.58	0.41	0	1
TenureDiversity	269	0.11	0.11	0	1
Revenues in thousands	269	3,794	3,660	3	23,900

modification with the results on steady-state cooperative effort gives the following: $E(\bar{Q}) = a\bar{e} + (1-a)f(r)\text{Prob}(\bar{e}_{C0} > \tilde{e}_C(r)) + h(D)$.

Here expected output is a function of a constant, reciprocity (which in turn could be a function of diversity), and an additive function of diversity. So, effects could occur through two channels and potentially cancel each other: diversity could have a negative effect on output if reciprocity is decreasing in diversity but a positive effect if diversity is valuable to the firm's productivity. Because theory is agnostic on the directions or magnitudes of these two channels, the net effect is an empirical question. We adopt the following simple linear specification:

$$\bar{Q}_i = \beta_0 + \beta_1 D_i + \eta_i.$$

Finally, two interesting but less central implications come out of the model. The fact that multiple steady states of the model can occur with high r implies (1) the possibility of a bimodal distribution of output in high r offices and (2) more output persistence in high r offices. We will revisit these implications in the results section.

4. DATA

Our data were provided by a professional services firm that operates over sixty offices in the United States and abroad. The offices range in size from just a few employees to nearly 100 at their headquarters. They administered anonymous employee satisfaction surveys approximately annually from 1995 to 2002. Table I contains summary statistics on the variables we created with these data, which we describe below.

From the survey responses, we can identify the office, gender, and tenure of the individual employees, enabling us to create office-level measures of diversity in those two dimensions. For gender, we calculated the standard deviation of a dummy variable for male for each office and scaled it linearly to fall into [0, 1], where 0 indicates an

all-male or all-female office and 1 is an office evenly divided. This variable is called *GendDiversity*. In our data, the minimum value is 0 and the maximum is 1. Note that this firm employs more women than men, and that we have both male-dominated and female-dominated offices among our observations where *GendDiversity* is near 0.

For tenure diversity, we calculated the standard deviation of tenure for each office, and then divided by the number of employees in the office. Finally, we scaled the expression linearly so that the measure takes on values of 0 for offices where everyone has worked for the firm the same amount of time and positive values for offices with some variance in the amount of time the employees have worked there, 1 being an upper bound in our data set. Note that the scaling for both of these diversity variables is arbitrary, but choosing a scale on the unit interval aided interpretation.

We use four additional questions from the survey to construct our data set. Their full text is in the Data Appendix (Supporting Information). They asked employees how accepting their workplace is of differences, how satisfied they are, how cooperative the office is, and how high morale is. We use these responses, averaged at the office level, to construct *AvgDPerception*, *AvgSatisfaction*, *AvgCooperate*, and *AvgMorale*. We also create the employee-level variables based on the last three, *Satisfaction*, *Cooperate*, and *Morale*. We view *Cooperate* as the most literal measure of social capital, or social goods provision, because employees are effectively asked to characterize the extent of the norm of reciprocity within the office. We would equate this quantity with the "cooperative effort" from the RZ model. As alternative indicators of social capital, we employ *Morale* and *Satisfaction*, which might also capture elements of social goods provision. Those three variables are coded so that higher reported satisfaction has a higher numerical value, with a maximum of 5 and a minimum of 1.[10]

We do not interpret *AvgDPerception* as a measure of social capital. It could serve as a proxy for diversity in dimensions we do not measure, such as lifestyle or ethnicity. Alternatively, one could interpret it literally, as a perception of how diversity is accepted at a particular office, which could be at odds with diversity in the office.

The variables based on survey responses are positively but not perfectly correlated. (The pairwise correlations among *Satisfaction*, *Morale*, and *Cooperate* range from 0.36 to 0.60.) The survey, then, captures a more nuanced situation than employees being uniformly "happy" or "unhappy," with that attitude pervading all responses.

Table I also contains summary statistics on a measure of office performance, *Revenues*, provided from internal data by the firm. If the firm had had reliable measures of office-level profit, we might have preferred using those data, but we feel revenue provides a good alternative for a number of reasons. First, firm-reported profit is notoriously unreliable—revenue is easier to measure and not subject to arbitrary accounting conventions. Second, we view the firm as maximizing dynamic profit, and we collected data during a time of rapid growth for the firm when even economically accurate measures of static profit would have come far from matching dynamic profit. Finally, revenue is a metric often used by firms themselves to gauge office performance.

Finally, we augmented this information with economic and demographic variables for each of the cities in which an office is located. We collected annual data on city unemployment rate, *Unemploy*, from the Bureau of Labor Statistics (or comparable foreign agencies for the foreign cities). Summary statistics are included in Table I. The other economic and demographic measures, based primarily on census data, do not

---

10. The question on morale resulted in four answers of "don't know" out of the approximately 1,700 responses. These were treated as missing values.

Variable	Obs.	Mean	Std. Dev.	Min	Max
At the city level:					
CAvgAge	61	33.9	2.7	29.6	41.7
CPolitics	47	75.3	58.8	1	227
CPercMinority	64	43.6	20.5	2.0	89.5
CPercMale	64	48.8	1.2	46.5	51.4
COfficeRent	59	42.15	37.10	15.60	197.80
in annual dollars per ft^2					
CPopulation	67	1,462	1,818	81	8,008
in thousands					

vary over the course of our time period. Those are reported in Table II.[11] These variables are largely self-explanatory, but a few comments are warranted. *CPolitics* is an index of political leaning based on voting for the 2004 Presidential election (which obviously only exists for U.S. cities). Orange County had the maximum index value in our data set of 227. Detroit had the minimum at 1. Also note that for *CPercMinority*, the percentage of minority residents in a city, the definition of minority varied by country so that, for instance, whites were considered part of the minority population in Japanese cities but not in U.S. cities. Detroit, again, was at an extreme, with the maximum value in our data set of 89.5%. Nagoya, Japan, had our minimum value, 2%. Note, finally, that our office-level data set does not include the firm headquarters, only the field offices.

Our data set exhibits a number of nice features. First, the data were collected in the field, not a lab, so we are analyzing actual employees operating in authentic environments under high-powered incentives, such as career advancement. Williams and O'Reilly (1998) note that lab studies of diversity have often found much more positive effects of diversity on performance than field studies. Second, our data form a panel over eight years, allowing us to exploit a type of variation and account for factors that we would otherwise be unable to in a single cross section. Jackson and Joshi's (2004) work on diversity in social contexts notes the inference limitations created by cross section data. Third, we have offices in our data set spanning the entire set of possibilities of gender diversity and composition. In other words, we have all-male and all-female offices in addition to intermediate mixtures. This feature is quite unusual in other workplace data sets (and, in fact, other workplaces), and provides us with valuable variation for identifying our effects. Fourth, our data set largely avoids a common and serious shortcoming of field studies, that the diversity of a group cannot be randomly assigned, as in a lab experiment, but would typically be endogenously determined.[12] Our data set has the unusual feature that the diversity in each office was plausibly determined exogenously. As we explain in more detail in the results section, the firm we study was young and fast-growing during the period our data were collected. Endogenous placement of employees to target a particular gender mix was not part of a corporate staffing plan. Finally, we have an objective measure of office performance, annual office revenues, instead of having to rely on subjective measures such as manager evaluations or potentially arbitrary measures such as meeting deadlines.

11. We relied on a number of different sources to track down demographics for foreign cities and data on office rental rates. Katherine McNeill and William Wheaton were both extremely helpful in this endeavor.
12. This endogenous determination will bias econometric estimates of effects unless controlled for appropriately, a concern not widely addressed in the management literature.

TABLE III.
RESULTS OF EMPLOYEE-LEVEL REGRESSIONS

Explanatory Variables	Dep. Variable: *Cooperate*		
	(1)	(2)	(3)
GendDiversity	−0.168	−0.086	0.048
	(−1.74)	(−0.67)	(0.36)
TenureDiversity	0.409	0.013	−0.325
	(1.39)	(0.04)	(−1.07)
AvgDPerception	**0.524**	**0.535**	**0.608**
	(4.03)	(3.52)	(6.36)
TenureYears	−0.016	0.003	0.007
	(−1.24)	(0.16)	(0.52)
Year	−0.004	−0.015	−0.029
	(−0.21)	(−0.75)	(−1.78)
Male	0.002	0.093	0.014
	(0.02)	(1.32)	(0.22)
AvgGender	−0.050	−0.172	−0.336
	(−0.33)	(−0.87)	(−1.40)
Log(*CPopulation*)		−0.046	
		(−1.13)	
CPercMale		−0.029	
		(−0.72)	
CPercMinority		−0.003	
		(−0.70)	
CAvgAge		−0.008	
		(−0.32)	
COfficeRent		0.001	
		(0.92)	
Constant	1.738	3.855	1.450
	(2.70)	(1.70)	(3.06)
Observations	1,440	1,122	1,440
Office fixed effects?	No	No	Yes

Notes: Robust *t* statistics in parentheses. Coefficients in bold are significant at the 5% level.

We show correlations among our key variables in the Data Appendix (Supporting Information).

5. RESULTS

5.1 SOCIAL CAPITAL

Our theory discussion provided guidance on an empirical strategy, and, in particular, suggested two estimating equations, one where cooperative effort is a function of diversity and a second where output is a function of diversity. We now address the first of these and turn to our results on social capital within the office. We use our employee-level data and focus on explaining perceived levels of cooperation. Most particularly, we will be interested in measures of diversity as explanatory variables.

We estimate separate regressions for three dependent variables, *Cooperate*, *Satisfaction*, and *Morale*. The results for *Cooperate* are reported in Table III, whereas those for the other two variables are shown in Table IV. The explanatory variables consist of measures of particular interest, such as *GendDiversity*, *TenureDiversity*, and *AvgDPerception*, as well as additional control variables. Control variables at the employee level include the employee's job tenure, *TenureYears*, and a dummy variable for the gender of

TABLE IV.
ADDITIONAL RESULTS OF EMPLOYEE-LEVEL REGRESSIONS

| | Dependent Variable | | | | | |
| | *Satisfaction* | | | *Morale* | | |
Explanatory Variables	(1)	(2)	(3)	(4)	(5)	(6)
GendDiversity	−0.154	−0.128	0.020	**−0.351**	**−0.226**	−0.089
	(−1.83)	(−1.80)	(0.17)	(−3.59)	(−2.24)	(−0.77)
TenureDiversity	0.148	0.119	−0.230	0.502	0.258	0.250
	(0.74)	(0.46)	(−0.92)	(1.94)	(0.89)	(0.98)
AvgDPerception	**0.621**	**0.684**	**0.672**	**0.634**	**0.691**	**0.795**
	(7.90)	(8.89)	(7.98)	(5.16)	(5.05)	(9.33)
TenureYears	−0.035	**−0.049**	−0.019	**−0.063**	**−0.054**	**−0.042**
	(−1.87)	(−2.84)	(−1.53)	(−5.66)	(−3.02)	(−3.33)
Year	**−0.072**	**−0.064**	**−0.082**	**−0.066**	**−0.058**	**−0.082**
	(−5.36)	(−4.62)	(−6.34)	(−3.66)	(−3.28)	(−6.25)
Male	0.021	−0.039	0.029	−0.043	−0.048	−0.035
	(0.41)	(−0.72)	(0.52)	(−0.74)	(−0.62)	(−0.63)
AvgGender	−0.102	0.077	−0.359	0.215	0.226	−0.028
	(−0.42)	(0.70)	(−1.77)	(1.05)	(1.39)	(−0.14)
Log(*CPopulation*)		0.025			−0.038	
		(0.68)			(−0.71)	
CPercMale		−0.004			−0.020	
		(−0.13)			(−0.65)	
CPercMinority		−0.002			0.000	
		(−0.86)			(−0.12)	
CAvgAge		−0.002			−0.002	
		(−0.12)			(−0.08)	
COfficeRent		0.001			**0.004**	
		(1.16)			(2.89)	
Constant	1.635	1.491	1.411	1.270	2.098	0.474
	(4.39)	(0.78)	(3.38)	(2.09)	(1.12)	(1.12)
Observations	1,579	1,233	1,579	1,558	1,216	1,558
Office fixed effects?	No	No	Yes	No	No	Yes

Notes: Robust *t* statistics in parentheses. Coefficients in bold are significant at the 5% level.

the respondent, *Male*. Other control variables are year of the response, *Year*, the fraction of males in an office, *AvgGender*, and, in some specifications, city-level measures such as the percent male, *CPercMale*, the percent minority, *CPercMinority*, the log of population, *LogCPopulation*, the average age, *CAvgAge*, and office rental rates, *COfficeRent*. We do not include office fixed effects in the first two specifications, but include them in the last.[13]

Because we view *Cooperate* as our most literal measure of social capital, Table III contains our preferred results. Columns (1)–(3) are the three specifications, a baseline, or more parsimonious one, one containing office-level control variables, and one with office fixed effects. We focus on results in column (1) because much of our identification of diversity effects comes from the cross-section, but we do present and discuss the augmented specifications as well.

Turning to specification (1), we see that higher levels of gender diversity (a more equal mix of men and women) are associated with lower levels of cooperation. This

---

13. Blanchflower and Oswald (2004) use a qualitative response model to explain determinants of survey responses due to their discrete nature. We agree with the logic of their approach, although for ease of interpretation we will report linear regression results. Our major conclusions are robust to estimating an ordered probit model.

result, represented by the estimated coefficient of -0.168 on *GendDiversity*, has a *p* value of 0.09. The magnitude suggests that moving from an office evenly split between men and women to either an all-male or all-female office, holding constant other characteristics, would increase cooperation about one-sixth of a point on a five-point scale.[14] We also see that higher levels of tenure diversity (a mix across number of years that employees in an office had worked in the firm) were associated with higher levels of cooperation, although this result is not statistically significant at traditional levels.

Another interesting result to come out of specification (1) is the importance of *AvgDPerception*. Offices where the employees, on average, believe their employer to be accepting of diversity are more cooperative. In addition to being highly significant (with a *t*-statistic of 4.03), its magnitude is also noteworthy. The estimated coefficient of 0.524 suggests that increasing office-average response to the question on diversity perception by one point increases cooperation more than a half point on the same scale. This result bears a more careful examination. Initially, it seems at odds with the first result that more gender diversity is associated with less cooperation. One can think of at least two ways to reconcile these results. First, it is possible that our measure *AvgDPerception* is a proxy for diversity in an office, but diversity on dimensions other than gender and tenure. In addition to gender differences, the question specifically mentions ethnic and lifestyle differences, dimensions on which we have no data. We believe a more likely explanation is that there is a distinction between a company that provides an environment accepting of diversity and one that has diversity. The employees seem more cooperative (and more satisfied overall, as we see below) in an environment *supportive* of but lacking in diversity.

One might be concerned that these results are affected by "common method bias," a bias resulting from the fact that our dependent variable and one explanatory variable, *AvgDPerception*, were gathered with a common method, an employee survey, and could therefore share a component attributable to the method as opposed to the underlying constructs.[15] Such a situation could result in estimated coefficients being biased upward. The raw correlation between *Cooperate* and *AvgDPerception*, 0.16, suggests the problem is unlikely to be serious, coupled with the fact that *AvgDPerception* is an office average of individual answers, not the answers themselves. Nevertheless, we also reestimated specification (1) without *AvgDPerception*. The results for the remaining variables, including our variable of primary interest, *GendDiversity*, were largely unaffected. No estimated coefficients changed sign at all or magnitude substantially, and only one 5% significance test changed—the coefficient of *GendDiversity* actually became significant at the 5% level, strengthening our main results. Nevertheless, our preferred specification is the one we present in Table III with a caveat: we can be confident in the other results, but caution may be warranted in the interpretation of the estimated coefficient of *AvgDPerception*.

The impact of *TenureYears* is negative but not statistically significant. One might imagine that those with higher tenure would be more well-integrated into the office culture, but other factors, such as boredom or job fatigue, might offset this. We included *Year*, a time trend, as a parsimonious way to absorb any possibly spurious association with time. It is not significant in specification (1).[16]

We wanted to ensure that changes in office-level gender diversity were not affecting the sample's level of *Cooperate* merely by adding more men or women to the sample. So

14. Of course one cannot vary *GendDiversity* in an office without also varying *Male* and *AvgGender*, but their estimated effects on *Cooperate* were small enough to ignore for this counterfactual.
15. See Podsakoff et al. (2003).
16. Including year fixed effects instead did not materially alter results. In addition, all but one of the year fixed effects were insignificant.

we control for the gender of the respondent with *Male* and the fraction of the office that is male with *AvgGender*. In specification (1) these effects are not significant.

In interpreting these results, it is important to note that specification (1) does not contain office fixed effects. One might think that a hypothetical San Francisco office differs systematically from a hypothetical Sheboygan office, and these differences should be controlled for in the estimation. It is also the case, however, that our identification of certain effects might be coming primarily from the cross-section, an identification that would be wiped out with the inclusion of fixed effects. In particular, we have, on average, four years of data[17] for each office, a length of time when most offices would not have experienced significant turnover, which suggests that much of our identification of the diversity effects would come off of the cross-section. In order both to control for some city (office) characteristics and to preserve some identification off of the cross-section, we include specification (2), adding a rich set of city-level covariates. None of these we include are statistically significant in this regression, but the other results are affected. In particular, the coefficient on *GendDiversity* is cut in half and is no longer even marginally significant. Note, though, that we lose a relatively large fraction of our observations when we include the extra covariates due to missing observations.

Finally, we include a fixed effects model, specification (3). The results are consistent with our concern about being able to identify effects off of time series variation alone. In particular, *GendDiversity* is not significant. Notably, though, the coefficient on *AvgDPerception* increases somewhat and becomes more significant with the inclusion of the office fixed effects. This finding is less surprising given that *AvgDPerception* could be driven in part by firm-wide policy changes over time and, therefore, have its effect identified more by the time series.

Recall that although *Cooperate* was our preferred measure of social goods provision, we have alternative measures, *Satisfaction* and *Morale*. *Satisfaction* seems less well-suited as a proxy for social goods provision because it is possible that the sources of employee satisfaction could be entirely individual in nature. *Morale*, however, has a more cooperative, group-based, connotation. The results for *Satisfaction* and *Morale*, found in Table IV, are similar to those for *Cooperate*, but stronger statistically. Higher levels of *AvgDPerception* are associated with significantly higher levels of *Satisfaction* and *Morale*. But higher levels of gender diversity, *GendDiversity*, are associated with lower indicators of well-being, and this association seems more persistent and significant than in the first set of regressions. The coefficient on *GendDiversity* is marginally significant in specifications (1) and (2) and solidly so in specifications (4) and (5). *TenureYears*, insignificant in specification (1), is negative and significant in the *Morale* regressions, and marginally so in the *Satisfaction* regressions. The estimated magnitudes for both of those effects are small, however. The control for *Year* also becomes significant in all six specifications. The same broad patterns in the results emerge, though. These proxies for workplace social capital are not perfectly correlated, so the estimated relationships reflect three similar but distinct patterns.

We take the following broad lessons from the results at this stage: gender diversity is associated with lower levels of social capital (at least marginally), whereas the perception at the office level that the firm supports diversity is associated with higher levels of social capital. This latter result is present even after controlling for a fixed geographic effect.

---

17. We have data for the maximum eight years for about a quarter of our offices. Quite a few offices were either opened or closed during the eight-year period, and for others, data are missing for a year or two in the middle of the period.

We find it interesting that most other explanatory variables were not particularly close to being significant—we would not have been surprised to find significantly different answers to these survey questions between men and women respondents, for instance, or in male-dominated versus female-dominated offices. Those differences were largely absent, though.

These results are suggestive of patterns where diversity can have important effects in the workplace. We offer them, however, with a caveat. One might be concerned about the potentially endogenous placement of employees or management in offices. In particular, a firm might hire employees to achieve a certain gender mix, for instance, and could possibly focus that hiring in offices with lower morale or cooperation. Or a firm might move a manager from an office with high levels of cooperation to one with low levels to improve the latter office, mechanically changing the composition of those offices. Although we cannot dismiss a concern such as this out of hand, we would argue that this concern is not likely to be so important in our particular setting. The firm we study was quite young at the time and experiencing rapid growth. Although now it is a much more well-established and mature firm, in the late 1990s, it was run by a set of college friends who largely hired additional friends of theirs to start offices in cities where they were interested in moving. The firm was run on a shoestring, and expenditures like corporate consultants to advise the firm on corporate culture and diversity in hiring would not have been in the budget. Furthermore, even though employee surveys were conducted, a top executive in the firm claimed that no analysis other than ours had ever been performed on those data. Hiring and staffing were not random, of course, but elements of those processes that could lead to troublesome endogeneity for us were likely to have been absent. In addition, the fact that we see all-male, all-female, and mixed offices in the data also suggests that the firm was not interested in targeting a certain gender mix.

5.2 Performance

Although we care about these indicators of employee satisfaction as proxies for social capital, they remain intermediate inputs. A firm's ultimate aim is to generate revenues and profits. So in Table V, we look at the association between office-level attributes and the log of office-level *Revenues*. Of course in interpreting these and other results, we are careful about inferring causality where correlation is established.

In assessing revenue effects, we add *AvgTYears*, office-average tenure, *Year*, a linear year effect, and *Unemploy*, the unemployment rate in the office's closest metropolitan area, among our control variables. We also include *YearsOpen*, a variable equal to the year of the observation minus the year of that office's first observation, which controls for smaller revenues that offices would generate before they became established in a city.

Table V presents results from four specifications, our base specification with a smaller set of city covariates, our augmented specification with the full set of city co-variates (but a smaller number of observations due to missing values), a specification including number of employees in the office, and a specification with office fixed effects. Looking across all specifications, the estimated effect of *AvgDPerception* is not significantly different from zero. Recall that this perception of the firm's acceptance of diversity was a significant variable in the employee-level regressions. But, interestingly, it does not appear to be associated with a revenue payoff. Note that such a result emphasizes the distinction between "intermediate goods" such as firm social capital and the ultimate outcome of interest for a firm.

TABLE V.
RESULTS OF OFFICE-LEVEL REGRESSIONS

| | Dependent Variable: | | | |
| | Log(*Revenues*) | | | |
Explanatory Variables	(1)	(2)	(3)	(4)
AvgTYears	**0.302**	**0.234**	**0.201**	**0.251**
	(3.75)	(3.38)	(3.34)	(3.47)
AvgDPerception	0.036	−0.081	0.100	0.007
	(0.20)	(−0.29)	(0.37)	(0.04)
GendDiversity	**0.406**	**0.311**	0.057	−0.124
	(2.30)	(2.08)	(0.37)	(−0.68)
TenureDiversity	**−1.682**	**−1.514**	−0.641	**−1.615**
	(−3.20)	(−2.74)	(−1.06)	(−4.02)
AvgGender	0.132	0.442	**0.514**	0.180
	(0.53)	(1.98)	(2.40)	(0.54)
Year	**−0.370**	−0.174	**−0.196**	
	(−6.08)	(−1.87)	(−2.13)	
YearsOpen	**0.465**	**0.262**	**0.250**	**0.121**
	(6.82)	(2.63)	(2.55)	(3.73)
Number			**0.129**	
			(5.70)	
Unemploy	−0.041	**−0.106**	**−0.095**	−0.051
	(−1.02)	(−2.32)	(−2.20)	(−0.86)
Log(*CPopulation*)	0.052	**0.237**	**0.178**	
	(0.56)	(2.62)	(2.08)	
CPercMale		0.066	0.012	
		(1.74)	(0.31)	
CPercMinority		0.005	0.002	
		(1.32)	(0.49)	
CAvgAge		−0.046	−0.017	
		(−1.28)	(−0.47)	
COfficeRent		**0.022**	**0.021**	
		(3.77)	(4.06)	
CPolitics		**−0.004**	**−0.002**	
		(−3.77)	(−2.46)	
constant	14.005	11.147	11.932	13.811
	(13.43)	(4.05)	(4.63)	(16.88)
Observations	269	200	200	269
Office fixed effects?	No	No	No	Yes

Notes: Robust *t* statistics in parentheses. Coefficients in bold are significant at the 5% level.

Turning to additional results in Table V, higher levels of *GendDiversity* are positively and significantly associated with office revenue in our base specification (1). The estimated coefficient of 0.41 implies that going from an office that is either all male or all female to an office split equally between the sexes would be associated with a revenue gain of 41%(!). Of course, the implications for firm behavior are less clear cut, because the firm might have to make additional changes in order to change the gender composition of its workforce, but the relationship uncovered in the sample is still of interest. These results are consistent with a conclusion that the gender diversity of an office improves office performance significantly. Note that our dependent variable is Log(*Revenues*), which should mitigate concerns over outliers, like large offices, driving our results.[18] To be sure, we also symmetrically trimmed observations associated with

---

18. Recall that the headquarters office, by far the largest, is not included in this analysis.

the largest and smallest residuals and obtained qualitatively similar results.[19] These results are available upon request.

We are interested in controlling for any source of spurious correlation. We do, though, have the same concern regarding identification in the presence of office fixed effects as we had previously, so we included specification (2) as an intermediate step. We control for various city characteristics that could be correlated with both diversity in an office and revenues. The core results are robust to the inclusion of additional covariates, despite the significant drop in observations. In particular, the sign and significance of the *GendDiversity* coefficient remains unchanged.[20]

In column (4), we also report a specification with office-level fixed effects. In that specification, the estimated contribution of *GendDiversity* to office level revenue is no longer significant. As in our results for employee satisfaction, much of our identification of a genuine effect may be coming from the cross-sectional variation.

TenureDiversity is associated with a large, negative, and statistically significant revenue effect, a result that survives including the fixed effects. This result stands in contrast with those from the employee-level regressions where *TenureDiversity* has no significant effect. In other words, we find no evidence that having an office where employees were hired at different times is associated with lower levels of cooperation, but we do find evidence that it is associated with lower levels of performance. One possibility is that employees arriving at a firm at the same time might have a shared language and can communicate more effectively, a phenomenon that might not be reflected in our measures of social goods provision but would affect office performance. We should note that a new office could have difficulty in generating revenue compared to an older, more established office, and the new office could also have lower tenure diversity because a large group of employees could be hired at the same time at the opening of the office. We do, however, include *YearsOpen* and *AvgTYears* to control for that situation, both of which have positive and significant effects.

In column (3) we included a specification with number of employees as an explanatory variable, although we suggest caution in interpreting those results. Number of employees could proxy for one component of firm cost, of course, but absent wage data, it would likely be a poor proxy. Furthermore, we feel that the strength of the relationship between office revenues and office employees would mostly be arising from the mechanical need to hire additional employees as office revenues increased. The potential endogeneity induced by that phenomenon makes interpretation of column (3) results problematic and also highlights that results in the other columns should be viewed as a reduced-form estimate of patterns in the data as opposed to any causal relationship.

5.3 Additional Results

In our discussion of the RZ model, we noted two additional implications. The fact that multiple steady states of the model can occur with high *r* implies (1) the possibility of a bimodal distribution of output in high *r* offices (or at least higher dispersion)[21] and (2) more output persistence in high *r* offices. To investigate these possibilities, we took

19. Depending on the number of observations trimmed, the estimated coefficient on *GendDiversity* ranged from 0.23 to 0.41, typically retaining significance at the 5% level.
20. At the suggestion of a referee, we estimated this specification excluding data from the sole field office in the same state as the headquarters. Magnitudes and significance of the results were virtually unchanged.
21. Multiple equilibria for high *r* offices could but need not result in a bimodal distribution of output. If the equilibria were close enough together relative to the variance of any error in the system, the result could simply be higher dispersion but not bimodality.

the residuals from our base model (1) of Log(*Revenues*) from Table V. If there was, in fact, a bimodal distribution of output for high r offices, we would expect to see a bimodal distribution of residuals from that regression for those offices. The same would be true for higher dispersion. A kernel regression of the residuals from offices with *Gend Diversity* ≤ 0.5, which we interpret as high r offices, did not reveal any obvious bimodality. Residuals from high r offices did, however, exhibit higher variance, 1.37 versus 1.00, consistent with the existence of multiple steady states. Second, we regressed the residuals on lagged residuals by office as well as an interaction between lagged residuals and *Gend Diversity*. Greater output persistence for high r offices should be manifested in a negative coefficient estimate on the interaction term. We found, instead, that the estimated coefficient was positive and significant. Tests for specific empirical implications of RZ, then, leave us with mixed results. Researchers may find it useful to test these implications in other settings, however.

6. Conclusion

The managers of firms face the challenge of assembling a workforce and a culture that will succeed in the task at hand. The results of this paper shed light on how actual and perceived diversity is associated with indicators of firm social capital and measures of ultimate office performance, revenues.

We find that the perception that a firm is supportive of diversity in an office is positively associated with indications of the level of cooperation in that office. Other proxies for social capital or corporate culture, such as employee morale and satisfaction, were also higher in offices in which this perception was higher. Nevertheless, the presence of actual gender diversity was a significant factor in *reducing* these same measures of social capital. At the same time, tenure diversity had little measurable association with proxies for social capital.

In our second set of results, we investigate the determinants of office-level revenues. We find that the perception that the firm accepts diversity has no estimated payoff in this dimension. Interestingly, gender diversity is associated with a positive contribution to revenues, although this effect is diminished once office-level fixed effects are included. In contrast, tenure diversity is associated with lower revenues.

We note an interesting contrast between our results on tenure diversity and gender diversity. Tenure diversity had little association with our measure of social capital but a strong negative association with performance. Gender diversity seems to affect the functioning of the firm in quite a different way. Our first set of results suggests that gender diversity could have detrimental impacts on the formation of firm social capital, but the revenue results suggest that whatever impacts it had were outweighed or at least canceled by the direct contribution of gender diversity in the office.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web site:

Data Appendix
Table S1: Correlations of Variables from Employee-level Regressions ($n = 1414$)
Table S2: Correlations of Variables from Office-level Regressions ($n = 269$)

References

Ancona, D. and D. Caldwell, 1992, "Demography and Design: Predictors of New Product Team Performance," *Organization Science*, 3, 321–341.

Blanchflower, D. and A. Oswald, 2004, "Well-Being Over Time in Britain and the USA," *Journal of Public Economics*, 88, 1359–1386.

Costa, D. and M. Kahn, 2003, "Cowards and Heroes: Group Loyalty in the American Civil War," *Quarterly Journal of Economics*, 118, 519–548.

Costa, D. and M. Kahn, 2003. "Civic Engagement and Community Heterogeneity: An Economist's Perspective," *Perspectives on Politics*, 1, 103–112.

Ely, R., 2004, "A Field Study of Group Diversity, Participation in Diversity Education Programs, and Performance," *Journal of Organizational Behavior*, 25, 755–780.

Gant, J., C. Ichniowski, and K. Shaw, 2002, "Social Capital and Organizational Change in High-Involvement and Traditional Work Organizations," *Journal of Economics and Management Strategy*, 11, 289–328.

Glaeser, E., D. Laibson, J. Scheinkman, and C. Soutter, 2000, "Measuring Trust," *Quarterly Journal of Economics*, 115, 811–846.

Goldin, C. and L. Katz, 1999, "Human Capital and Social Capital: The Rise of Secondary Schooling in America, 1910 to 1940," *Journal of Interdisciplinary History*, 29, 683–723.

Hellerstein, J.K., D. Neumark, and K. Troske, 1999, "Wages, Productivity, and Worker Characteristics: Evidence from Plant-Level Production Functions and Wage Equations," *Journal of Labor Economics*, 17, 409–446.

Jackson, S.E. and A. Joshi, 2004, "Diversity in Social Context: A Multi-attribute Multilevel Analysis of Team Diversity and Sales Performance," *Journal of Organizational Behavior*, 25, 675–702.

Joshi, A. and H. Roh, 2009, "The Role of Context in Work Team Diversity Research: A Meta-Analytic Review," *Academic of Management Journal*, 3, 599–627.

Karlan, D., M. Mobius, T. Rosenblat, and A. Szeidl, 2009, "Trust and Social Collateral," *Quarterly Journal of Economics*, 124, 1307–1361.

Kochan, T., K. Berukova, R. Ely, S. Jackson, A. Joshi, K. Jehn, J. Leonard, D. Levine, and D. Thoma, 2003, "The Effects of Diversity on Business Performance: Report of the Diversity Research Network," *Human Resource Management*, 3, 3–21.

Miguel, E. and M. Gugerty, 2005, "Ethnic Diversity, Social Sanctions, and Public Goods in Kenya," *Journal of Public Economics*, 89, 2325–2368.

Pelled, L., K. Eisenhardt, and K. Xin, 1999, "Exploring the Black Box: An Analysis of Work Group Diversity, Conflict and Performance." *Administrative Science Quarterly*, 44, 1–28.

Podsakoff, P., S. MacKenzie, J. Lee, and N. Podsakoff, 2003, "Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies," *Journal of Applied Psychology*, 88, 879–903.

Portes, A. and E. Vickstrom, 2011, "Diversity, Social Capital, and Cohesion," *Annual Review of Sociology*, 37, 461–479.

Poterba, J., 1997, "Demographic Structure and the Political Economy of Public Education," *Journal of Policy Analysis and Management*, 16, 48–66.

Putnam, R., 1995, "Bowling Alone: America's Declining Social Capital," *Journal of Democracy*, 6, 65–78.

Putnam, R., 2000, *Bowling Alone*, New York: Simon and Schuster.

Rob, R. and P. Zemsky, 2002, "Social Capital, Corporate Culture, and Incentive Intensity," *RAND Journal of Economics*, 33, 243–57.

Roberge, M. and R. Van Dick, 2010, "Recognizing the Benefits of Diversity: When and How Does Diversity Increase Group Performance?" *Human Resource Management Review*, 20, 295–308.

Shore, L., B. Chung-Herrera, M. Dean, K. Ehrhart, D. Jung, A. Randel, and G. Singh, 2009, "Diversity in Organizations: Where Are We Now and Where Are We Going?" *Human Resource Management Review*, 19, 117–133.

Vigdor, J., 2004, "Community Composition and Collective Action: Analyzing Initial Mail Response to the 2000 Census," *Review of Economics and Statistics*, 86, 303–312.

Williams, K. and C. O'Reilly, 1998, "Demography and Diversity in Organizations: A Review of 40 Years of Research," *Research in Organizational Behavior*, 20, 77–140.

Zenger, T. and B. Lawrence, 1989, "Organizational Demography: The Differential Effects of Age and Tenure Distributions on Technical Communication," *Academy of Management Journal*, 32, 77–140.