This activated scaling behavior is consistent with our observation of rapidly diverging exponent $z\nu$ in the vicinity of the field-induced QCP with quenched disorder. To test this paradigm, we turn to the specific analysis of SMT. According to the pair-breaking scenario of SMT, bosonic cooper pairs can be overdamped into normal single-particle excitations, which results in the quantum SMT at $T = 0$ with the clean critical exponent $\nu = ½$ (*34*). Thus, in the 2D system with SMT, the violation of the Harris criterion (with $d\nu = 1$) will most likely lead to the infinite-randomness QCP. Indeed, previous theoretical investigations revealed that the quenched disorder markedly changes the scaling behavior of SMT and results in activated scaling identical to the RTFIM (*44–47*). On this basis, we can fit the experimental results of $z\nu$ by the activated scaling law $z\nu \approx C|B - B_c^*|^{-\upsilon\psi}$ with constant C and the 2D infinite-randomness critical exponents $\nu \approx 1.2$ and $\psi \approx 0.5$ (*8*, *9*), and the good agreement between our observation and the theoretical expectation strongly supports the existence of infinite-randomness QCP (Fig. 3).

Our findings are summarized by the schematic phase diagram in Fig. 4. When approaching the infinite-randomness QCP at B_C*, the quenched disorder leads to two correlated phenomena: (i) In the zero-temperature limit, the vortex lattice deforms into a vortex glass–like phase on a length scale $L > L_P$; (ii) because of the transformation into the vortex glass–like phase, rare regions of inhomogeneous superconducting islands gradually emerge in the $B > B_2$ regime and manifest activated scaling behavior, namely, the quantum Griffiths singularity (Fig. 3). One question remains as to why this activated scaling feature has not yet been observed in SMT in previous studies. We attribute this absence to the instability of the vortex glass–like phase under thermal fluctuation. The quenched disorder will play a dominant role for length scales $L > L_P$. Thus, in the high-temperature regime, thermal fluctuations smear the inhomogeneity caused by quenched disorder, and rare regions hardly exist. Near zero temperature, the impact of quenched disorder overtakes thermal fluctuation, which results in the emergence of rare regions around the infinite-randomness QCP (Fig. 4). On the basis of these considerations, we speculate that the activated scaling feature can only be observed under extremely low temperature, which is the case in our study.

REFERENCES AND NOTES

1. B. M. McCoy, T. T. Wu, *Phys. Rev.* **176**, 631–643 (1968).
2. R. B. Griffiths, *Phys. Rev. Lett.* **23**, 17–19 (1969).
3. A. B. Harris, *J. Phys. C Solid State Phys.* **7**, 1671–1692 (1974).
4. D. S. Fisher, *Phys. Rev. Lett.* **69**, 534–537 (1992).
5. D. S. Fisher, *Phys. Rev. B* **51**, 6411–6461 (1995).
6. C. Pich, A. P. Young, H. Rieger, N. Kawashima, *Phys. Rev. Lett.* **81**, 5916–5919 (1998).
7. O. Motrunich, S. C. Mau, D. A. Huse, D. S. Fisher, *Phys. Rev. B* **61**, 1160–1172 (2000).
8. T. Vojta, A. Farquhar, J. Mast, *Phys. Rev. E* **79**, 011111 (2009).
9. I. A. Kovács, F. Iglói, *Phys. Rev. B* **82**, 054437 (2010).
10. M. C. de Andrade *et al.*, *Phys. Rev. Lett.* **81**, 5620–5623 (1998).
11. A. H. Castro Neto, G. Castilla, B. A. Jones, *Phys. Rev. Lett.* **81**, 3531–3534 (1998).
12. S. Ubaid-Kassis, T. Vojta, A. Schroeder, *Phys. Rev. Lett.* **104**, 066402 (2010).
13. D. B. Haviland, Y. Liu, A. M. Goldman, *Phys. Rev. Lett.* **62**, 2180–2183 (1989).
14. A. F. Hebard, M. A. Paalanen, *Phys. Rev. Lett.* **65**, 927–930 (1990).
15. A. Yazdani, A. Kapitulnik, *Phys. Rev. Lett.* **74**, 3037–3040 (1995).
16. A. M. Goldman, *Int. J. Mod. Phys. B* **24**, 4081–4101 (2010).
17. N. Marković, C. Christiansen, A. M. Goldman, *Phys. Rev. Lett.* **81**, 5217–5220 (1998).
18. M. P. A. Fisher, P. B. Weichman, G. Grinstein, D. S. Fisher, *Phys. Rev. B* **40**, 546–570 (1989).
19. J. Biscaras *et al.*, *Nat. Mater.* **12**, 542–548 (2013).
20. X. Shi, P. V. Lin, T. Sasagawa, V. Dobrosavljević, D. Popović, *Nat. Phys.* **10**, 437–443 (2014).
21. See supplementary materials on *Science* Online.
22. T. Nishio, M. Ono, T. Eguchi, H. Sakata, Y. Hasegawa, *Appl. Phys. Lett.* **88**, 113115 (2006).
23. R. C. Dynes, V. Narayanamurti, J. P. Garno, *Phys. Rev. Lett.* **41**, 1509–1512 (1978).
24. N. Reyren *et al.*, *Science* **317**, 1196–1199 (2007).
25. N. R. Werthamer, E. Helfand, P. C. Hohenberg, *Phys. Rev.* **147**, 295–302 (1966).
26. A. T. Bollinger *et al.*, *Nature* **472**, 458–460 (2011).
27. S. L. Sondhi, S. M. Girvin, J. P. Carini, D. Shahar, *Rev. Mod. Phys.* **69**, 315–333 (1997).
28. S. Sachdev, *Quantum Phase Transtions* (Cambridge Univ. Press, Cambridge, ed. 2, 2011).
29. M. P. A. Fisher, *Phys. Rev. Lett.* **65**, 923–926 (1990).
30. M. V. Feigel'man, A. I. Larkin, *Chem. Phys.* **235**, 107–114 (1998).
31. M. V. Feigel'man, A. I. Larkin, M. A. Skvortsov, *Phys. Rev. Lett.* **86**, 1869–1872 (2001).
32. B. Spivak, A. Zyuzin, M. Hruska, *Phys. Rev. B* **64**, 132502 (2001).
33. A. Kapitulnik, N. Mason, S. A. Kivelson, S. Chakravarty, *Phys. Rev. B* **63**, 125322 (2001).
34. S. Sachdev, P. Werner, M. Troyer, *Phys. Rev. Lett.* **92**, 237003 (2004).
35. B. Spivak, P. Oreto, S. A. Kivelson, *Phys. Rev. B* **77**, 214523 (2008).
36. D. S. Fisher, M. P. A. Fisher, D. A. Huse, *Phys. Rev. B* **43**, 130–159 (1991).
37. G. Blatter, M. V. Feigel'man, V. B. Geshkenbein, A. I. Larkin, V. M. Vinokur, *Rev. Mod. Phys.* **66**, 1125 (1994).
38. B. Rosenstein, D. Li, *Rev. Mod. Phys.* **82**, 109–168 (2010).
39. A. A. Abrikosov, L. P. Gor'kov, *Sov. Phys. JETP* **12**, 1243–1253 (1961).
40. K. Maki, *Prog. Theor. Phys.* **40**, 193–200 (1968).
41. V. M. Galitski, A. I. Larkin, *Phys. Rev. Lett.* **87**, 087001 (2001).
42. T. Vojta, *J. Phys. A* **39**, R143–R205 (2006).
43. T. Vojta, J. A. Hoyos, *Phys. Rev. Lett.* **112**, 075702 (2014).
44. J. A. Hoyos, C. Kotabage, T. Vojta, *Phys. Rev. Lett.* **99**, 230601 (2007).
45. T. Vojta, C. Kotabage, J. A. Hoyos, *Phys. Rev. B* **79**, 024401 (2009).
46. A. Del Maestro, B. Rosenow, M. Müller, S. Sachdev, *Phys. Rev. Lett.* **101**, 035701 (2008).
47. A. Del Maestro, B. Rosenow, J. A. Hoyos, T. Vojta, *Phys. Rev. Lett.* **105**, 145702 (2010).

ACKNOWLEDGMENTS

Supported by National Basic Research Program of China grants 2013CB934600, 2012CB921300, and 2015CB921102; National Natural Science Foundation of China grants 11222434, 11174007, and 11534001; and the Research Fund for the Doctoral Program of Higher Education (RFDP) of China.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/350/6260/542/suppl/DC1
Materials and Methods
Figs. S1 to S8
References (*48*, *49*)

17 January 2015; accepted 21 September 2015
Published online 15 October 2015
10.1126/science.aaa7154

ECONOMICS

Peer effects on worker output in the laboratory generalize to the field

Daniel Herbst[1] and Alexandre Mas[1,2,3]*

We compare estimates of peer effects on worker output in laboratory experiments and field studies from naturally occurring environments. The mean study-level estimate of a change in a worker's productivity in response to an increase in a co-worker's productivity (γ) is $\hat{\gamma} = 0.12$ (SE = 0.03, $n_{studies} = 34$), with a between-study standard deviation $\tau = 0.16$. The mean estimated $\hat{\gamma}$-values are close between laboratory and field studies ($\hat{\gamma}_{lab} - \hat{\gamma}_{field} = 0.04$, $P = 0.55$, $n_{lab} = 11$, $n_{field} = 23$), as are estimates of between-study variance τ^2 ($\hat{\tau}^2_{lab} - \hat{\tau}^2_{field} = -0.003$, $P = 0.89$). The small mean difference between laboratory and field estimates holds even after controlling for sample characteristics such as incentive schemes and work complexity ($\hat{\gamma}_{lab} - \hat{\gamma}_{field} = 0.03$, $P = 0.62$, $n_{samples} = 46$). Laboratory experiments generalize quantitatively in that they provide an accurate description of the mean and variance of productivity spillovers.

Laboratory experiments in the social sciences are valuable for understanding human behavior in controlled environments (*1*, *2*). However, there is an active debate about the extent to which these experiments have external validity. In an influential paper, Levitt and List have questioned whether findings in laboratory studies generalize in the real world, arguing that "because the lab systematically differs from most naturally occurring environments... experiments may not always yield results that are readily generalizable" (*3*). Indeed, on the surface, laboratory experiments appear quite different than natural settings. Subjects tend to be

[1]Department of Economics, Industrial Relations Section, Firestone Library, Princeton University, Princeton, NJ 08544, USA. [2]National Bureau of Economic Research (NBER), Cambridge, MA 02138, USA. [3]Institute for the Study of Labor (IZA), 53113 Bonn, Germany.
*Corresponding author. E-mail: amas@princeton.edu

students, and the controlled setting may appear artificial in relation to actual workplaces. In spite of the importance of laboratory studies in the social sciences, and of the many articles engaging in this debate (*4–7*), to our knowledge there has not been a systematic comparison of the same economic parameter estimated in laboratory experiments and field studies using more than a small number of studies.

We exploited the relative abundance of estimates for one particular parameter—the estimated spillover effect of worker productivity on the productivity of co-workers—in order to compare estimates from laboratory experiments with those from field studies that use data from naturally occurring environments. This parameter, which we denote γ, is useful for understanding a variety of economic phenomena, including wage setting, economic growth, the social returns to human capital investment, the effects of immigration, the optimal assignment of workers to teams, and agency problems. Over the past 15 years, there have been more than 34 studies seeking to estimate γ in a diverse set of occupations (*8–41*), including fruit pickers (*11*), supermarket cashiers (*31*), physicians (*17*), sales teams (*18*), and scientists (*40*). The large number of estimates of γ in the literature presents an opportunity to directly compare the findings of laboratory experiments and field studies.

We constructed a database of γ estimates using inclusive criteria (*42*). We included a study in the database if the paper contains an estimate of the effect of the productivity of a worker or a group of workers on another worker or group of workers, and the paper interprets the estimate as a workplace-productivity peer effect. We included laboratory experiments meant to simulate workplace environments. Estimates from both published and working papers were included to mitigate possible publication bias. This yielded papers in traditional field settings with observable productivity data but also included sports studies (*25*, *26*, *41*) and studies estimating spillovers by using administrative earnings data (*12*, *19*, *35*). Several studies that estimate worker peer effects (*43–46*) do not include all of the information required to construct comparable, variance-weighted estimates and were therefore not included. After this screen, our database contains estimates from 34 papers, of which 23 are field studies and 11 are laboratory experiments (table S1). We call this the study-level database. In this database, 50% of the studies are published (Table 1).

Studies often include several estimates from different methodologies, specifications, and/or samples. If multiple methodologies were used, we selected estimates using the methodology the authors explicitly stated as preferred. If the authors did not explicitly state their preferred methodology, we selected estimates according to the following ranking: randomized controlled trial, regression discontinuity design, instrumental variables, difference-in-differences, and ordinary least squares. For a given methodology, in cases in which there was more than one specification and the study did not clearly state the preferred estimate, we used the specification with the most controls. Last, if estimates were given for multiple independent samples, we computed a single observation-weighted average of the sample estimates and combined standard errors (*42*). Estimates were coded by the authors according to this protocol and independently verified by a second coder. Details on how estimates were coded are available in the supplementary materials (*42*).

Because some papers report estimates from different samples, we assembled a second database in which we extracted multiple estimates from a study when the estimates are from distinct samples. This database, which we denote the sample-level database, contains 46 estimates from the 34 studies. For each sample, we coded whether the job required abstract reasoning, the incentive structure of the job (fixed wages, individual piece rates, or team-based compensation), whether workers were substitutes or complements in the production process, and whether workers competed over scarce inputs to worker-level production. This sample-level database complements the study-level data set by allowing us to examine how worker spillover estimates vary by study characteristic because papers sometimes report estimates for different workplace conditions. Within the sample-level database, in 48% of samples, workers had fixed wages with no individual piece rate compensation; in 48% of samples, workers were paid with individual piece rates; and in 20% of samples, there was team-based compensation. In 24% of samples, the task or occupation required abstract reasoning (Table 1).

Twenty-two studies in our study-level database relate a worker's productivity to the productivity of a co-worker or group of co-workers (a levels-levels specification). Because the dependent and independent variables of interest are in the same units, the estimated coefficient is a unitless measure that can be interpreted as an elasticity or a standardized coefficient. In seven studies, the dependent and independent variables are in natural log units. As long as mean productivity is similar between focal and peer workers, means of the dependent and independent variables are close, and the estimated coefficients have a similar interpretation as the levels-levels specification [this follows from $dln(y)/dln(x) \approx (dy/dx)(\overline{x}/\overline{y})$]. Excluding estimates from log specifications in our data set yields similar pooled estimates of $\hat{\gamma}$ as the main sample (table S2). In five studies, the dependent and independent productivity measures are in different units, in which case we standardized both variables with respect to the individual-level productivity distribution. Because of the presence of these five studies, the most precise interpretation of the mean of $\hat{\gamma}_i$ across studies (denoted $\hat{\gamma}$) is a standardized coefficient giving the standard deviation change in worker productivity from a 1 SD change (in the individual-level productivity distribution) of co-worker productivity. However, $\hat{\gamma}$ also has an

Table 1. Summary statistics. Shown are unweighted summary statistics for study-level and sample-level databases. Peer-effects estimates are denoted by $\hat{\gamma}$. "*P* value" corresponds to the *P* value of the estimated $\hat{\gamma}_i$ in the study. "Sample size" gives mean and standard deviation of sample sizes across papers. "Lab experiment" is a dummy variables for whether a study was classified as a laboratory experiment. "Published" indicates whether a study was published in a peer-reviewed journal. "Fixed wage" is a dummy variable for whether compensation had a flat or hourly pay component with no individual piece rate. "Individual piece rate" and "Group piece rate" are dummy variables for whether a portion of compensation was determined by individual or group output, respectively. "Complement" is a dummy for whether authors made specific reference to complementarities between workers in their joint production function. "Perfect substitute" is a dummy for whether workers in the sample being studied generated perfectly substitutable output in the production process. "Complex job" is a dummy for whether the job or task performed required abstract reasoning. "Rival" is a dummy for whether workers in the sample competed in the production process. "Observations" reports total number of estimates in the study-level and sample-level data sets, respectively. Dashes indicate that the variable was not coded in the database.

Variable	Study		Sample	
	Mean	SD	Mean	SD
$\hat{\gamma}$	0.134	0.203	0.140	0.216
P value	0.138	0.232	0.146	0.257
Sample size	468,375	2,207,142	155,511	344,550
Lab experiment	0.324	0.475	0.326	0.474
Published	0.500	0.508	0.543	0.504
Fixed wage	—	—	0.478	0.505
Individual piece rate	—	—	0.478	0.505
Group piece rate	—	—	0.196	0.401
Complement	—	—	0.022	0.147
Perfect substitute	—	—	0.565	0.501
Complex job	—	—	0.239	0.431
Rival	—	—	0.087	0.285
Observations		34		46

approximate elasticity interpretation: A 1% increase in average co-worker productivity increases focal worker productivity by $\hat{\gamma}$ percent. Excluding the studies with standardized variables and keeping estimates that have a pure elasticity interpretation yields an almost identical average estimate of $\hat{\gamma}$, as does converting the standardized measures into elasticities when possible and computing $\hat{\gamma}$ over all studies with a natural elasticity interpretation (table S2).

The average unweighted estimate of $\hat{\gamma}$ is 0.13 (SD = 0.20) (Table 1). In the study-level database, 56% of estimates are positive and statistically significant at the 5% level, 2.9% of studies are negative and significant at the 5% level, and the remaining studies are insignificant at the 5% level (Fig. 1).

To aggregate estimates into a comprehensive summary measure, taking into account the sampling error of the estimates, we estimated a random-effects model that assumes that estimates are drawn from sampling distributions with possibly distinct means. The random-effects model assumes that each study's observed peer effects estimate is composed of three elements: the mean effect size γ, a study's divergence from the mean effect λ_i, and an error term ϵ_i around that study's estimate. For study i, the observed peer effect estimate $\hat{\gamma}_i$ is given by

$$\hat{\gamma}_i = \gamma + \lambda_i + \epsilon_i$$

To estimate the summary effect $\hat{\gamma}$, study-level estimates are weighted by inverse variance weights

$$W_i = \frac{1}{\mathrm{SE}_i^2 + \tau^2}$$

where is SE_i^2 is the squared standard error of estimate i and τ^2 is the estimated between-study

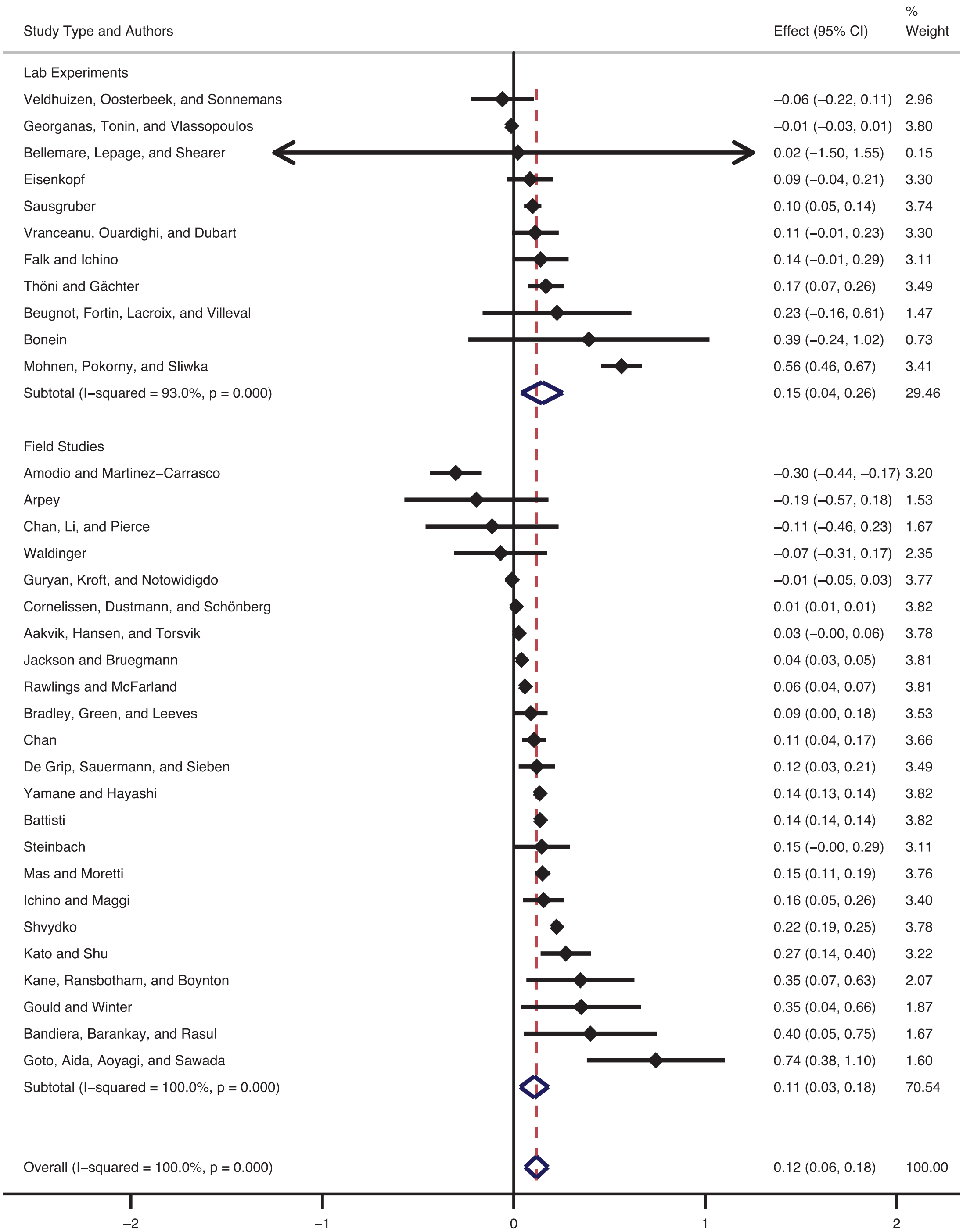

Fig. 1. All study-level estimates, summary effects by study type, and the overall summary effect. Solid black diamonds and lines show 95% confidence intervals around each study's estimated peer effect. The first two blue diamonds are centered around estimated summary peer effects by subgroup (laboratory versus field), and the dashed line and lowermost diamond are centered around the estimated overall summary effect. Widths of blue diamonds correspond to estimated 95% confidence intervals for summary effects. "% Weight" corresponds to the inverse-variance weight, where variance is computed as the reciprocal sum of each study's sample variance and the between-study variance: $W_i = (\sigma_i^2 + \tau^2)^{-1}$, normalized to sum to 100%.

Table 2. Summary estimates by paper type. Shown are the random-effects summary estimate of γ across laboratory and field studies. Standard errors are in parentheses. Observations are inverse variance–weighted, where variance is computed as the reciprocal sum of each study's sample variance and the between-study variance: $W_i = (\sigma_i^2 + \tau^2)^{-1}$. Standard error of the summary estimate is computed as the square root of the reciprocal sum of W_i. Final column reports P values for a comparison of means z test for equality in summary effect and τ^2 terms between laboratory and field studies. P value for τ^2 difference was computed using bootstrapped standard errors.

	Overall	Laboratory	Field	P value
Summary effect	0.120	0.148	0.107	0.545
	(0.031)	(0.055)	(0.038)	
τ^2	0.025	0.023	0.026	0.890
Observations	34	11	23	

Table 3. Laboratory-field comparisons controlling for sample characteristics. Shown are meta-regression coefficients for the sample-level data set with standard errors in parentheses. *$P < 0.10$, **$P < 0.05$, ***$P < 0.01$. Dependent variable is study's reported worker peer effect estimate γ_i for sample i. "Laboratory experiment" is an indicator for whether or not an estimate was taken from a laboratory study. "Group piece rate" is a dummy variables for whether a portion of compensation was determined by group output. "Fixed wage" is a dummy variable for whether compensation had a flat or hourly pay component with no individual piece rate. "Published" indicates whether a study was published in a peer-reviewed journal. "Complex job" is a dummy for whether the job or task performed required abstract reasoning. "Perfect substitute" is a dummy for whether workers in the sample being studied generated perfectly substitutable output in the production process. "Complement" is a dummy for whether authors made specific reference to complementarities between workers in their joint production function. "Rival" is a dummy for whether workers in the sample competed in the production process. "GPR*Lab," "FW*Lab," and "PS*Lab" are interaction terms for "Laboratory experiment" with "Group piece rate," "Fixed wage," and "Perfect substitute," respectively. Estimates were obtained via variance-weighted least squares regression, where weights and standard errors are consistent with a random-effects meta-analysis model: Inverse-variance weights are computed as the reciprocal sum of each study's sample variance and the between-study variance: $W_i = (\sigma_i^2 + \tau^2)^{-1}$.

	(1)	(2)	(3)	(4)
Laboratory experiment	0.052	0.016	0.029	–0.050
	(0.059)	(0.062)	(0.059)	(0.102)
Group piece rate		0.149**	0.162**	0.174*
		(0.073)	(0.070)	(0.097)
Fixed wage			0.099*	0.046
			(0.049)	(0.059)
Published		0.054	0.044	0.026
		(0.055)	(0.053)	(0.056)
Complex job		–0.087	–0.085	–0.090
		(0.067)	(0.064)	(0.065)
Perfect substitute		–0.106*	–0.108**	–0.113*
		(0.053)	(0.050)	(0.059)
Complement		–0.135	–0.082	–0.122
		(0.171)	(0.164)	(0.177)
Rival		–0.188**	–0.155*	–0.182*
		(0.089)	(0.086)	(0.092)
GPR*Lab				0.002
				(0.146)
FW*Lab				0.181
				(0.114)
PS*Lab				–0.023
				(0.130)
Constant	0.095***	0.152***	0.096*	0.143**
	(0.032)	(0.052)	(0.056)	(0.064)
Observations	46	46	46	46

variance of estimates. The between-study variance is estimated with the empirical Bayes method (*47*), which has been shown to be more accurate than alternatives when the heterogeneity of estimates is large (*48*). The standard error of $\hat{\gamma}$ is given by

$$\mathrm{SE}(\hat{\gamma}) = \frac{1}{\sqrt{\sum_i W_i}}$$

In the random-effects model, the majority of papers receive comparable weight, with a small handful of papers with imprecise estimates receiving little weight (Fig. 1). The average value of $\hat{\gamma}_i$ in the pooled sample is $\hat{\gamma} = 0.12$ (SE = 0.03, n = 34 studies) with an estimated between-study variance $\tau^2 = 0.025$ (Table 2). The I^2 statistic, which describes the percentage of total variation across studies that is due to heterogeneity rather than chance (*42*), is almost 100% for the pooled sample. The average value of $\hat{\gamma}_i$ for laboratory and field samples respectively is 0.148 (SE = 0.055, n_{lab} = 11 studies) and 0.107 (SE = 0.038, n_{field} = 23 studies). The random-effect average pooled estimates for both the laboratory and field estimates are statistically distinguishable from zero at conventional levels. We cannot reject equality of the means of γ in the laboratory and field samples (P = 0.55). The between-study variances are also similar in the laboratory and field samples: The between-study variability of estimates, τ^2, is estimated as 0.023 in the laboratory sample and 0.026 in the field sample, and we cannot reject equality of these parameters using bootstrapped standard errors (P = 0.89).

Studies differ in compensation schemes, production processes, and task complexity. These features of the studies may influence the extent of peer effects. The small mean difference between estimates in the laboratory and field samples is stable after controlling for study and workplace characteristics in a regression framework. We used the sample-level data set to estimate the following equation using a modified variance-weighted least-squares regression:

$$\hat{\gamma}_i = \alpha_i + \beta \mathrm{lab}_i + \delta X_i + \epsilon_i$$

where lab_i is an indicator for whether the estimate was taken from a laboratory experiment, X_i is a vector of variables controlling for study and workplace characteristics, ϵ_i is the error term, and each observation is weighted by its inverse variance W_i (Table 3). Without controls, the estimated laboratory experiment indicator coefficient ($\hat{\beta}$) is similar in magnitude to the value found in the study-level data set ($\hat{\beta}$ = 0.05, P = 0.38, n = 46 samples). The estimates attenuated when we added workplace and paper characteristics, including variables for worker compensation incentives (group piece-rate indicator and fixed-wage indicator), publication status, an indicator for whether the occupation requires abstract reasoning, and an indicator for whether workers in the study are rivals ($\hat{\beta}$ = 0.029, P = 0.62, n = 46 samples).

The coefficients on the workplace characteristics in the regression setting are worth additional discussion. As has been long noted in the literature on organizations (*49*), workplaces with team-based production and fixed wages are susceptible to the free-rider problem because effort is a public good when there is imperfect monitoring. Assuming workers only have self-interested motives, introducing a more productive co-worker in such a workplace will lead other workers to be less productive because their colleague is taking on a greater share of the workload. However, the free-rider problem can be overcome with mutual monitoring or the threat of social sanctions. In this case, the presence of a more productive co-worker might lead other workers to increase their effort because of these considerations. The coefficients on the group piece-rate and fixed-wage coefficients are both positive and statistically significant, suggesting that positive co-worker productivity spillovers are particularly important in these settings. This finding is evidence that positive co-worker peer effects, possibly because of the threat of social sanctions, help mitigate the free-rider problem. The negative and significant coefficient on substitutes indicates that production processes in which workers are perfect substitutes have less pronounced peer-productivity spillovers. We tested whether these conclusions differ between laboratory and field studies. Estimating a model with interactions of the laboratory indicator with indicators for group piece-rate, fixed wages, and perfect substitutability (variables for which there is sufficient overlap between laboratory and field studies), we found coefficients on interaction terms for group piece rate and substitutability that are close to zero and insignificant, whereas the coefficient on the fixed-wage and laboratory interaction is positive but not statistically significant because of a large standard error (Table 3, column 4). The estimates on the interaction terms imply that laboratory and field studies yield similar relationships between workplace characteristics and the magnitude of peer effects for two of the three characteristics we examined (and inconclusive for the third), with the caveat that these estimates are somewhat imprecise.

We conclude that for estimation of γ, laboratory studies generalize quantitatively. This is a surprising finding because even proponents of laboratory experiments have argued that laboratory experiments may only generalize qualitatively (*7*), and it suggests that laboratory experiments have more external validity than previously recognized. One caveat is that there is between-study dispersion reflecting unobserved heterogeneity in worker peer effects estimates both for field and laboratory studies, so that any individual study differs from the mean of the summary $\hat{\gamma}$ estimate. However, we also found that between-study standard deviations are comparable in the laboratory and field samples, as are the prediction intervals. Consequently, laboratory experiments provide a representative depiction of the overall distribution of γ values.

An additional question is why estimation of productivity spillovers translates well in the laboratory. The findings suggest that it is possible to simulate realistic work environments in the laboratory, particularly experiments with real-effort tasks. However, there are aspects of real workplaces that are missing in even the most complex laboratory studies. In laboratory studies, subjects are usually aware that they are being observed, and it is known that in a number of environments, there are effects of social facilitation on performance that are mediated by whether the task requires performance of learned skills or learning new skills (*50*). It is also not possible to simulate long-term employment relationships in the laboratory, a feature of many real workplaces. Our results indicate that these special features of the laboratory, among others, may not be as important for estimating productivity spillovers as the ones that are modeled.

REFERENCES AND NOTES

1. A. Falk, E. Fehr, *Labour Econ.* **10**, 399–406 (2003).
2. G. Charness, P. Kuhn, *Handb. Labor Econ.* **4**, 229 (2010).
3. S. D. Levitt, J. A. List, *J. Econ. Perspect.* **21**, 153–174 (2007).
4. C. Camerer, in *Handbook of Experimental Economic Methodology*, G. Fréchette, A. Schotter, Eds. (Oxford Univ. Press, New York, 2015), chap. 14.
5. A. Falk, J. J. Heckman, *Science* **326**, 535–538 (2009).
6. S. D. Levitt, J. A. List, *Can. J. Econ.* **40**, 347–370 (2007).
7. J. Kessler, L. Vesterlund, The external validity of laboratory experiments: Qualitative rather than quantitative effects; working paper; available at www.pitt.edu/~vester/External-Validity.pdf.
8. A. Aakvik, F. Hansen, G. Torsvik, Dynamic peer effects in sales teams; working paper; available at https://web.archive.org/web/20150122070301/www.uib.no/sites/w3.uib.no/files/attachments/wp10.13.pdf.
9. F. Amodio, M. A. Martinez-Carrasco, Input allocation, workforce management and productivity spillovers: Evidence from personnel data; working paper; available at http://economics.yale.edu/sites/default/files/amodio_paper.pdf.
10. N. Arpey, Peer effects among weavers: Evidence from a Chinese textile firm with a relative group incentive scheme; working paper; available at http://blogs.colgate.edu/economics/files/2014/09/Arpey-2014-Peer-Effects-among-Weavers.pdf.
11. O. Bandiera, I. Barankay, I. Rasul, *Rev. Econ. Stud.* **77**, 417–458 (2010).
12. M. Battisti, High wage workers and high wage peers; Ifo Working Paper no. 168; available at http://hdl.handle.net/10419/84189.
13. C. Bellemare, P. Lepage, B. Shearer, *Labour Econ.* **17**, 276–283 (2010).
14. J. Beugnot, B. Fortin, G. Lacroix, M. C. Villeval, Social networks and peer effects at work; working paper; available at http://ssrn.com/abstract=2342248.
15. A. Bonein, Social comparison and peer effects with heterogeneous ability; working paper; available at http://crem.univ-rennes1.fr/wp/2014/201411.pdf.
16. S. Bradley, C. Green, G. Leeves, *Labour Econ.* **14**, 319–334 (2007).
17. D. C. Chan, Teamwork and moral hazard: Evidence from the emergency department; working paper; available at http://web.stanford.edu/~chand04/papers/ed_paper.pdf.
18. T. Y. Chan, J. Li, L. Pierce, *Manage. Sci.* **60**, 1965–1984 (2014).
19. T. Cornelissen, C. Dustmann, U. Schönberg, Peer effects in the workplace; CESifo Working Paper Series no. 4398; available at http://ssrn.com/abstract=2330320.
20. A. De Grip, J. Sauermann, I. Sieben, The role of peers in estimating tenure-performance profiles: Evidence from personnel data; IZA Discussion Paper no. 6164; available at http://ssrn.com/abstract=1970759.
21. G. Eisenkopf, *Econ. Educ. Rev.* **29**, 364–374 (2010).
22. A. Falk, A. Ichino, *J. Labor Econ.* **24**, 39–57 (2006).
23. S. Georganas, M. Tonin, M. Vlassopoulos, Peer pressure and productivity: The role of observing and being observed; CESifo Working Paper Series no. 4572; available at http://ssrn.com/abstract=2390559.
24. J. Goto, T. Aida, K. Aoyagi, Y. Sawada, *J. Behav. Econ. Fin.* **4**, 94 (2011).
25. E. D. Gould, E. Winter, *Rev. Econ. Stat.* **91**, 188–200 (2009).
26. J. Guryan, K. Kroft, M. J. Notowidigdo, *Am. Econ. J. Appl. Econ.* **1**, 34 (2009).
27. A. Ichino, G. Maggi, *Q. J. Econ.* **115**, 1057–1090 (2000).
28. C. K. Jackson, E. Bruegmann, *Am. Econ. J. Appl. Econ.* **1**, 85 (2009).
29. G. C. Kane, S. Ransbotham, A. Boynton, Is high performance contagious among knowledge workers?; working paper; available at http://aisel.aisnet.org/cgi/viewcontent.cgi?article=1001&context=icis2012.
30. T. Kato, P. Shu, Performance spillovers and social network in the workplace: Evidence from rural and urban weavers in a Chinese textile firm; IZA Discussion Paper no. 3340; available at http://repec.iza.org/dp3340.pdf.
31. A. Mas, E. Moretti, *Am. Econ. Rev.* **99**, 112–145 (2009).
32. A. Mohnen, K. Pokorny, D. Sliwka, *J. Labor Econ.* **26**, 693–720 (2008).
33. C. M. Rawlings, D. A. McFarland, *Soc. Sci. Res.* **40**, 1001–1017 (2011).
34. R. Sausgruber, *Exp. Econ.* **12**, 193–201 (2009).
35. T. Shvydko, Essays in labor economics: Peer effects and labor market rigidities. working paper; available at https://cdr.lib.unc.edu/indexablecontent/uuid:bdb79e81-8334-4f03-9fb5-f54a2f2b22cb.
36. D. Steinbach, Compositional peer effects in team production; working paper; available at https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbnxzdGVpbmJhY2h-kYW5ueXxneDoxNGQzY2VmYTQ4YTVlYmlw.
37. C. Thöni, S. Gächter, Understanding social interaction effects in the workplace: An experimental approach; working paper; available at www.utdallas.edu/negcent/seminars/gaechter/Gachter-Thoni.pdf.
38. R. V. Veldhuizen, H. Oosterbeek, J. Sonnemans, Peers at work: From the field to the lab; Tinbergen Institute Discussion Paper 14-051; available at http://ssrn.com/abstract=2430717.
39. R. Vranceanu, F. E. Ouardighi, D. Dubart, Coordination in teams: A real effort-task experiment with informal punishment; ESSEC Working Paper; available at https://hal-essec.archives-ouvertes.fr/hal-00857364.
40. F. Waldinger, *Rev. Econ. Stud.* **79**, 838–861 (2012).
41. S. Yamane, R. Hayashi, Peer effects of swimmers; working paper; available at www.iser.osaka-u.ac.jp/coe/dp/pdf/no.236_dp.pdf.
42. Materials and methods are available as supplementary materials on *Science* Online.
43. B. H. Hamilton, J. A. Nickerson, H. Owan, *J. Polit. Econ.* **111**, 465–497 (2003).
44. E. Lazear, K. Shaw, C. Stanton, The value of bosses; working paper; available at www.nber.org/papers/w18317.pdf.
45. P. Arcidiacono, J. Kinsler, J. Price; available at http://public.econ.duke.edu/~psarcidi/multiable.pdf.
46. S. Kaur, M. Kremer, S. Mullainathan, *Am. Econ. Rev.* **100**, 624–628 (2010).
47. C. N. Morris, *J. Am. Stat. Assoc.* **78**, 47–55 (1983).
48. K. Sidik, J. N. Jonkman, *Stat. Med.* **26**, 1964–1981 (2007).
49. E. Kandel, E. Lazear, *J. Polit. Econ.* **100**, 801 (1992).
50. R. B. Zajonc, *Science* **149**, 269–274 (1965).

ACKNOWLEDGMENTS

The authors thank W. Dobbie, F. Goncalves, A. Krueger, J. Wagner, I. Walker, and S. Zhu. The authors contributed equally to this study; order of authorship is alphabetical. All data and code are available at the Harvard Dataverse (http://dx.doi.org/10.7910/DVN/CGM3XW).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/350/6260/545/suppl/DC1
Materials and Methods
Tables S1 and S2
Reference (*51*)

16 July 2015; accepted 11 September 2015
10.1126/science.aac9555