Firm-level determinants of wages and productivity: Management practices

Heidi L. Williams
MIT 14.662
Spring 2015

Outline

(1) Preliminaries

(2) Management and productivity:

Bertrand and Schoar (2003), Bloom and Van Reenen (2007), Bloom et al. (2013)

1 Preliminaries

Productivity - the efficiency with which firms transform inputs into outputs - is an essential concept in nearly every sub-field of economics. My goal in this lecture is to highlight some recent applied microeconomics research in this area. Most of this literature has focused on private firms, but you should think of "firms" broadly – including e.g. schools and hospitals. A broad "fact" that has motivated a great deal of productivity-related research is that there exist large and persistent differences in measured productivity levels across firms. Syverson (2011) provides an excellent recent overview; this is a fact at the heart of organizational economics, so if you are interested I would encourage you to think about sitting in on some of Bob's classes.

1.1 Conceptualizing productivity

Researchers often focus on a total factor productivity (TFP) measure like the following:

$$Y_t = A_t F(K_t, L_t, M_t) \tag{1}$$

where Y_t is output; $F(\cdot)$ is a function of observable inputs capital K_t, labor L_t, and intermediate materials M_t; and A_t is a factor-neutral shifter. Here, TFP is A_t: it captures variations in output not explained by shifts in the observable inputs that act through $F(\cdot)$. By construction, TFP is unmeasured - a residual: the variation in output that cannot be explained based on observable inputs.

While straightforward to define, a host of measurement issues arise when constructing productivity measures in practice. For labor, should one use number of employees, number of employee-hours, or other measures? How should capital be measured? Syverson (2011) discusses some current "best practices" for these types of questions, which I won't cover here.

1.2 Persistent productivity differences across firms

Analysis of firm heterogeneity has a long history in the social sciences. Bartelsman and Doms (2000) discuss how economic research on firm heterogeneity surged starting in the 1990s with the growing availability of longitudinal micro-level data sets that followed large numbers of establishments or firms over time. For example, the availability of the Longitudinal Research Database - a large panel data set of U.S. manufacturing plants developed by the U.S. Census Bureau - enabled a variety of new lines of research. Several new "facts" emerged from analyses of these datasets, one of which was the remarkable degree of heterogeneity within industries.

Syverson (2004b) provides a recent set of estimates. His Table 1 uses plant-level data from the 1977 Census of Manufactures to compute productivity distribution moments for four-digit manufacturing industries for each of four different productivity measures. His estimates imply that the plant at the 90^{th} percentile of the productivity distribution produces almost *twice* as much output with the *same measured inputs* as the 10^{th} percentile plant. Research by Hsieh and Klenow (2009) documented even larger productivity differences in China and India, with average 90-10 TFP ratios of more than 5:1.

These productivity spreads tend to be very persistent over time. Foster, Haltiwanger and Syverson (2008)'s Table 3 presents results from a regression of a producer's current TFP on its one-year-lagged TFP, which suggests autoregressive coefficients on the order of 0.8. Syverson (2011) summarizes this evidence on persistence by saying that some producers seem to have figured out their business (or at least are on their way) while others are woefully lacking.

The natural question that arises is what could be explaining these differences, and how they could persist in equilibrium. One explanation is - of course - that this "productivity dispersion" is just measurement error. That is, if we accounted properly for the differences in inputs in the production function perhaps there would be little residual dispersion in productivity. For many years, researchers "chipped away" at this measurement error concern by trying to develop better measures of input - capital, labor, materials, *etc.*. There was also a large literature that investigated how much of the residual could be accounted for by explicit measures of "intangible capital" like research and development (R&D).

This measurement error debate (which has generated a large literature) is analogous to the historical debate in the macro time series of productivity between Solow, who claimed that TFP was a large component of aggregate growth, and various critics who claim there was little role for TFP when all inputs were properly measured (*e.g.* Griliches (1996)). While difficult to rule out measurement error as an explanation, two bodies of evidence support the idea that measurement error is not the whole story:

1. Measured productivity differentials exist even within industries producing very homogenous products, such as ready mixed concrete (Foster, Haltiwanger and Syverson, 2008).

2. Measured productivity differentials are strongly correlated with firm exit and growth.

Bloom and Van Reenen (2011) summarize this literature as follows: *"In summary, there is a substantial body of evidence of persistent firm-level heterogeneity in productivity...in narrow* *industries in many countries and time periods. Differential observable inputs, heterogeneous prices, and idiosyncratic stochastic shocks are not able to adequately account for the remarkable dispersion of productivity."*

What are the potential explanations behind this dispersion in productivity? Syverson (2011) provides an excellent recent review of the literature in this area, discussing a variety of factors: managerial practice, information technology and R&D, learning-by-doing, product innovation, firm structure decisions, productivity spillovers, competition, and deregulation. I'll focus on discussing a few papers looking at the link between managerial practice and productivity as an example of labor/applied micro research in this area.

2 Management and productivity

Bloom and Van Reenen (2011) discuss how labor economics traditionally focused on the labor market rather than looking inside the "black box" of firms, but that this has dramatically changed over the last two decades. One area in which that is particularly evident is in a body of research examining the link between management and productivity.

Although empirical research in this area is "new," the idea that management might be an important determinant of productivity is definitely not new. In 1887 (the second year of the *QJE*'s existence!), economist and then-President of MIT Francis Walker published a paper in the *QJE* entitled "The Source of Business Profits" in which he conjectured that variation in managerial ability is the source of differences in profits across businesses (Walker, 1887): *"It is on the account of the wide range among the employers of labor, in the matter of ability to meet these exacting conditions of business success, that we have the phenomenon in every community and in every trade, in whatever state of the market, of some employers realizing no profits at all, while others are making fair profits; others again, large profits; others, still, colossal profits. Side by side, in the same business, with equal command of capital, with equal opportunities, one man is gradually sinking a fortune, while another is doubling or trebling his accumulations."* As Syverson (2011) notes, only now - nearly 130 years later! - do we finally have the data required to generate empirical evidence on this hypothesis: *"...perhaps no potential driver of productivity differences has seen a higher ratio of speculation to actual empirical study."* Along a similar line, Bloom and Van Reenen (2007) note that while the popular press and business schools have long stressed the importance of good management, empirical economists had relatively little to say about management practices until a few very recent papers.

An important contribution in this area was the paper by Bertrand and Schoar (2003), which as we will discuss below essentially asked the question: do managers matter? They argued the data suggested a resounding answer of "yes": performance differences can be explained in part by the identity of the managers. However, this leaves open the question of what the managers do or know that affects performance: as we will discuss, Bertrand and Schoar have some information on manager characteristics, but were limited by data constraints in how deep they were able to dig into how particular CEO practices and philosophies are tied to firm performance. As we will discuss, subsequent research by Nick Bloom, John Van Reenen, and collaborators has focused on collecting new micro-datasets measuring various aspects of managerial input.[1]

2.1 Bertrand and Schoar (2003)

Bertrand and Schoar (2003) ask the question: how much do individual managers matter for firm behavior and economic performance? Despite being the focus of very little empirical research, they motivate their analysis by noting that popular perception in the business press and among managers themselves is that CEOs and other top executives are key factors in the determination of corporate practices: managers are often perceived as having their own "styles" when making investment, financing, and other strategic decisions.

The key idea of their paper is to construct a manager-firm matched panel data set which tracks individual top managers as they move across firms over time: conditioning on firm fixed effects and other variables, they ask how much of the unexplained variation in firm practices can be attributed to manager fixed effects. The focus on movers is what allows them to separate manager fixed effects from firm fixed effects: persistent differences across firms might be related to an omitted variable that is also correlated with manager fixed effects.

While not measuring productivity specifically, their results suggest that manager fixed effects are empirically important determinants of a wide range of corporate variables such as firms' returns on assets. They also tie back estimated differences in "managerial style" to a limited set of observable managerial characteristics - birth cohort and MBA graduation.

2.1.1 Data

The data they use are the Forbes 800 files from 1969-1999 (providing information on the CEOs of the 800 largest US firms) and the Execucomp data from 1992-1999 (allowing them to track the names of the top five highest paid executives in 1500 publicly traded US firms). They restrict their sample to the subset of firms for which at least one specific top executive can be observed in at least one other firm (and at each firm for at least three years). The resulting dataset consists of around 600 firms and slightly over 500 managers. For this sample of firms, they use COMPUSTAT and SDC data to construct a series of annual accounting variables.

Table 2 summarizes the transitions observed in their dataset:

[1] Much of the data collected by Bloom, Van Reenen, and colleagues is available here: `http://worldmanagementsurvey.org/`.

TABLE II
EXECUTIVE TRANSITIONS BETWEEN POSITIONS AND INDUSTRIES

to:	CEO	CFO	Other
from:			
CEO	117	4	52
	63%	75%	69%
CFO	7	58	30
	71%	71%	57%
Other	106	0	145
	60%		42%

a. This table summarizes executives' transitions across positions and industries in the manager-firm matched panel data set (as described in subsection III.A and Table I). All transitions are across firms. The first entry in each cell reports the number of transitions from the row position to the column position. The second line in each cell reports the fraction of the transitions in that cell that are between different two-digit industries.

b. "Other" refers to any job title other than CEO or CFO.

2.1.2 Estimation strategy

For each dependent variable of interest, Bertrand and Schoar estimate regressions like the following:

$$y_{it} = \alpha_t + \gamma_i + \beta X_{it} + \lambda_{\text{CEO}} + \lambda_{\text{CFO}} + \lambda_{\text{Others}} + \varepsilon_{it} \qquad (2)$$

where y_{it} is the firm-year outcome of interest, α_t are year fixed effects, γ_i are firm fixed effects, X_{it} is a vector of time-varying firm controls, and ε_{it} is an error term. The remaining variables are manager fixed effects: λ_{CEO} are fixed effects for the group of managers who are CEOs in the last position where they are observed, λ_{CFO} are fixed effects for the group of managers who are CFOs in the last position where they are observed, and λ_{Others} are fixed effects for the group of managers who are neither CEOs nor CFOs in the last position where they are observed.

The authors emphasize that their goal is not to estimate a causal effect of managers on firm practices, but rather is to assess whether there is any evidence that firm policies systematically change with the identity of the top managers.

2.1.3 Results

Tables 3 and 4 report F-tests and adjusted R^2's from the estimation for different sets of corporate policy variables. The first row reports the fit of a benchmark specification that includes only firm fixed effects, year fixed effects, and time-varying firm-level controls. The next two rows report the change in adjusted R^2 when they consecutively add CEO fixed effects and (subsequently) fixed effects for all three groups of executives (CEOs, CFOs, and other top positions). The second and third rows also report F-statistics from tests of joint significance of the different sets of manager fixed effects. The big-picture take away of the results is that manager-specific effects appear to matter both economically and statistically for the policy decisions of firms, and for firm outcomes.

TABLE III

EXECUTIVE EFFECTS ON INVESTMENT AND FINANCIAL POLICIES

Panel A: Investment policy

	F-tests on fixed effects for				
	CEOs	*CFOs*	*Other executives*	*N*	*Adjusted R^2*
Investment				6631	.91
Investment	16.74 (<.0001, 198)			6631	.94
Investment	19.39 (<.0001, 192)	53.48 (<.0001, 55)	8.45 (<.0001, 200)	6631	.96
Inv to Q sensitivity				6631	.95
Inv to Q sensitivity	17.87 (<.0001, 223)			6631	.97
Inv to Q sensitivity	5.33 (<.0001, 221)	9.40 (<.0001, 58)	20.29 (<.0001, 208)	6631	.98
Inv to CF sensitivity				6631	.97
Inv to CF sensitivity	2.00 (<.0001, 205)			6631	.98
Inv to CF sensitivity	0.94 (.7276, 194)	1.29 (.0760, 55)	1.28 (.0058, 199)	6631	.98
N of acquisitions				6593	.25
N of acquisitions	2.01 (<.0001, 204)			6593	.28
N of acquisitions	1.68 (<.0001, 199)	1.74 (.0006, 55)	4.08 (<.0001, 203)	6593	.36

Panel B: Financial policy

	F-tests on fixed effects for				
	CEOs	*CFOs*	*Other executives*	*N*	*Adjusted R^2*
Leverage				6563	.39
Leverage	0.99 (.5294, 203)			6563	.39
Leverage	0.86 (.9190, 199)	1.43 (.0225, 54)	1.21 (.0230, 203)	6563	.41
Interest coverage				6278	.31
Interest coverage	0.56 (.99, 193)			6278	.31
Interest coverage	0.35 (.99, 192)	13.85 (<.0001, 50)	2.61 (<.0001, 192)	6278	.41
Cash holdings				6592	.77
Cash holdings	2.52 (<.0001, 204)			6592	.78
Cash holdings	2.48 (<.0001, 201)	3.68 (<.0001, 54)	2.53 (<.0001, 202)	6592	.80
Dividends/earnings				6580	.65
Dividends/earnings	5.78 (<.0001, 203)			6580	.71
Dividends/earnings	4.95 (<.0001, 199)	1.07 (.3368, 54)	1.74 (<.0001, 203)	6580	.72

a. Sample is the manager-firm matched panel data set as described in subsection III.A and Table I. Details on the definition and construction of the variables reported in the table are available in the Data Appendix.

b. Reported in the table are the results from fixed effects panel regressions, where standard errors are clustered at the firm level. For each dependent variable (as reported in column 1), the fixed effects included are row 1: firm and year fixed effects; row 2: firm, year, and CEO fixed effects; row 3: firm, year, CEO, CFO, and other executives fixed effects. Included in the "Investment to Q" and "Investment to cash flow" regressions are interactions of these fixed effects with lagged Tobin's Q and cash flow, respectively. Also the "Investment," "Investment to Q," and "Investment to cash flow" regressions include lagged logarithm of total assets, lagged Tobin's Q, and cash flow. The "Number of Acquisitions" regressions include lagged logarithm of total assets and return on assets. Each regression in Panel B contains return on assets, cash flow, and the lagged logarithm of total assets.

c. Reported are the F-tests for the joint significance of the CEO fixed effects (column 2), CFO fixed effects (column 3), and other executives fixed effects (column 4). For each F-test we report the value of the F-statistic, the p-value, and the number of constraints. For the "Investment to Q" and "Investment to Cash Flow" regressions, the F-tests are for the joint significance of the interactions between the manager fixed effects and Tobin's Q and cash flow, respectively. Column 5 reports the number of observations, and column 6 the adjusted R^2s for each regression.

TABLE IV

EXECUTIVE EFFECTS ON ORGANIZATIONAL STRATEGY AND PERFORMANCE

	Panel A: Organizational strategy *F*-tests on fixed effects for				
	CEOs	*CFOs*	*Other executives*	*N*	*Adjusted* R^2
N of diversifying acquis.				6593	.22
N of diversifying acquis.	2.06 (<.0001, 204)			6593	.25
N of diversifying acquis.	1.23 (.0163, 202)	1.74 (.0007, 53)	3.97 (<.0001, 202)	6593	.33
R&D				4283	.78
R&D	1.86 (<.0001, 145)			4283	.79
R&D	2.27 (<.0001, 143)	3.60 (<.0001, 45)	4.46 (<.0001, 143)	4283	.83
Advertising				2584	.79
Advertising	2.88 (<.0001, 95)			2584	.81
Advertising	4.03 (<.0001, 95)	0.84 (.6665, 21)	6.10 (<.0001, 80)	2584	.84
SG&A				2397	.46
SG&A	33.55 (<.0001, 123)			2397	.83
SG&A	13.80 (<.0001, 118)	0.82 (.7934, 42)	0.77 (.9777, 146)	2397	.83

TABLE IV

(CONTINUED)

	Panel B: Performance *F*-tests on fixed effects for				
	CEOs	*CFOs*	*Other executives*	*N*	*Adjusted* R^2
Return on assets				6593	.72
Return on assets	2.04 (<.0001, 217)			6593	.74
Return on assets	2.46 (<.0001, 201)	3.39 (<.0001, 54)	4.46 (<.0001, 202)	6593	.77
Operating return on assets				5135	.34
Operating return on assets	2.61 (<.0001, 217)			5135	.39
Operating return on assets	1.60 (<.0001, 216)	0.66 (.9788, 58)	1.01 (.4536, 217)	5135	.39

a. Sample is the manager-firm matched panel data set as described in subsection III.A and Table I. Details on the definition and construction of the variables reported in the table are available in the Data Appendix.

b. Reported in the table are the results from fixed effects panel regressions, where standard errors are clustered at the firm level. For each dependent variable (as reported in column 1) the fixed effects included are row 1: firm and year fixed effects; row 2: firm, year, and CEO fixed effects; row 3: firm, year, CEO, CFO, and other executives fixed effects.

c. Also included in the "N of diversifying acquisitions," "R&D," "advertising," and "SG&A" regressions are the logarithm of total assets, return on assets, and cash flow. The "N of diversifying acquisitions" regressions also include a dummy variable for whether the firm undertook any acquisition in that year. Also included in the "Return on assets" and "Operating return on assets" regressions is the logarithm of total assets.

d. Reported in the table are *F*-tests for the joint significance of the CEO fixed effects (column 2), CFO fixed effects (column 3), and other executives fixed effects (column 4). For each *F*-test we report the value of the *F*-statistic and, in parentheses, the *p*-value and number of constraints. Also reported are the number of observations (column 5) and adjusted R^2s (column 6) for each regression.

To assess how big the observed distributions between managers are, Table 6 reports the size distribution of the manager fixed effects. The main take-away is that the variation is economically large. For example, row 1 of Table 6 shows that the difference between a manager at the 25^{th} percentile of investment level and one at the 75^{th} percentile is 0.20. This can be benchmarked against the average ratio of capital expenditures to assets in this sample, which is 0.30.

TABLE VI
SIZE DISTRIBUTION OF MANAGER FIXED EFFECTS

	Median	Standard deviation	25th percentile	75th percentile
Investment	0.00	2.80	−0.09	0.11
Inv to Q sensitivity	−0.02	0.66	−0.16	0.12
Inv to CF sensitivity	0.04	1.01	−0.17	0.28
N of acquisitions	−0.04	1.50	−0.54	0.41
Leverage	0.01	0.22	−0.05	0.09
Interest coverage	0.00	860.0	−56.0	51.7
Cash holdings	0.00	0.06	−0.03	0.02
Dividends/earnings	−0.01	0.59	−0.13	0.11
N of diversifying acquis.	−0.04	1.05	−0.28	0.21
R&D	0.00	0.04	−0.10	0.02
SG&A	0.00	0.66	−0.09	0.09
Advertising	0.00	0.04	−0.01	0.01
Return on assets	0.00	0.07	−0.03	0.03
Operating return on assets	0.00	0.08	−0.02	0.03

a. The fixed effects used in this table are retrieved from the regressions reported in Tables III and IV (row 3).
b. Column 1 reports the median fixed effect for each policy variable. Column 2 reports the standard deviation of the fixed effects. Columns 3 and 4 report the fixed effects at the twenty-fifth percentile and seventy-fifth percentile of the distribution, respectively.
c. Each fixed effect is weighted by the inverse of its standard error to account for estimation error.

The authors emphasize that the presence of managerial fixed effects does not tell us much about which specific managerial traits, characteristics, or practices might influence their decision-making. They examine the role of two managerial characteristics: MBA graduation and birth cohort/age. While limited (by data constraints), this exploratory analysis sheds some light on relevant factors.

2.2 Bloom and Van Reenen (2007)

Bloom and Van Reenen (2007) present the results of a new survey instrument which measures management practices at a sample of 732 medium-sized manufacturing firms in the US, the UK, France, and Germany. Their survey covers operations, monitoring, targets, and incentives. The survey was done via phone and targeted at plant managers in median-sized manufacturing firms. The response rate was relatively high - around 54% - and the paper contains a subset of the very detailed descriptions which are available that document the painstaking amount of work that went into trying to collect a meaningful dataset. For example, as a quality check the authors re-surveyed firms to interview different managers in different plants in the same firms using different interviewers, and found a strong correlation between these two independently collected measures.

2.2.1 Validation of the survey data

Their paper has two parts: validation of the survey, and analysis of the distribution of management practices. The thrust of the validation exercise is to match the survey data with information on firm accounts and stock market values to investigate the association between the measures of managerial practices and firm performance. These analyses uncover that better managerial practices are significantly associated with higher productivity, profitability, Tobin's Q,[2] sales growth rates, and firm survival rates. The goal is not to identify a causal relationship between the management practice measures and firm performance, but rather to check that the scores are not just "cheap talk" but rather area correlated with quantitative measures of firm performance from independent data sources.

2.2.2 Analysis of the distribution of management practices

Figure 1 shows the distribution of the average management scores per firm in all eighteen practices, plotted by country in raw form.

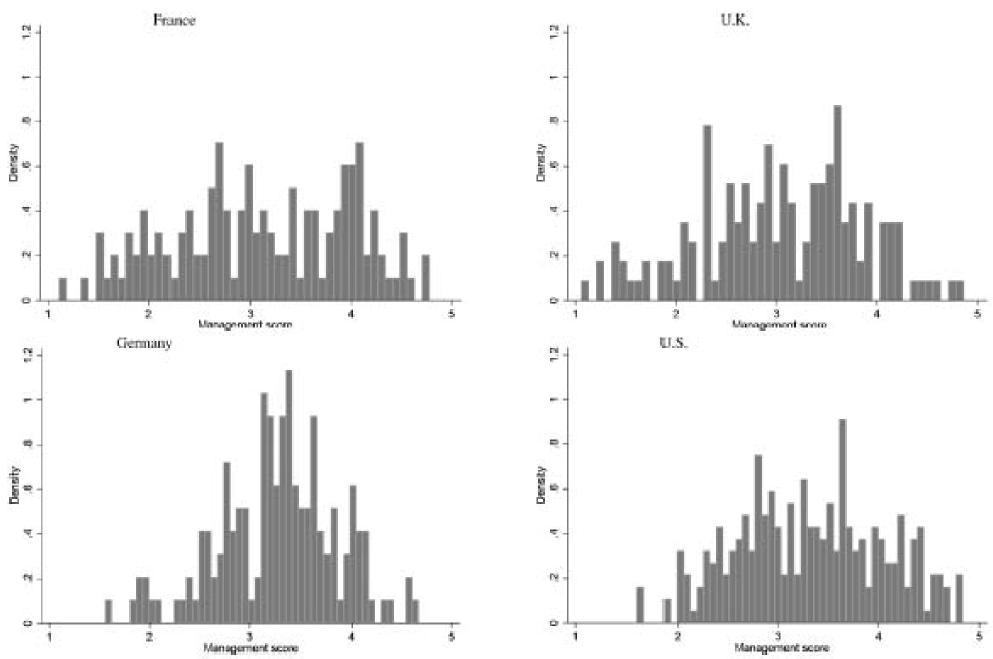

FIGURE I
Distribution of Management Scores by Country
Notes: These are the distributions of the raw management scores (simple averages across all 18 practices for each firm). 1 indicates worst practice, 5 indicates best practice. There are 135 French observations, 156 German observations, 151 UK observations, and 290 U.S. observations.

There is a huge amount of heterogeneity within each country, with firms spread across most of the distribution. About 2% of the overall variation is across countries, 42% is across countries by 3-digit industries, and 56% is within country and industry. The country-level ranking by average score is US, Germany, France, and then the UK. Bloom and Van Reenen note that the UK-US

[2] Tobin's Q is the ratio of the stock market value to net assets valued at replacement cost.

gap appears persistent over time, as evidenced by the Marshall Plan productivity mission of 1947: *"efficient management was the most significant factor in the American advantage [over the United Kingdom]."*

2.2.3 Product market competition

Table 3 investigates the relationship between product market competition and management scores. They use three measures of competition:

1. Degree of import penetration: the share of total imports relative to domestic production at the country-industry level over the years 1995-1999 (to avoid contemporaneous feedback)

2. Lerner index: $\frac{1}{\text{profit}-\text{sales}}$, calculated as the average across all firms excluding the own firm observation at the country industry level over the years 1995-1999 (to avoid contemporaneous feedback)

3. Survey measure for "no competitors".

Table 3 shows that all three measures give similar estimates: better management scores are positively and statistically significantly correlated with greater competition. The magnitude of the effects are large: for example, increasing the number of competitors from "few" to "many" (column 6) is associated with a management z-score increase of 0.140. The authors emphasize that although they condition on "stuff," they lack an instrumental variable for competition. However, they note that endogeneity bias will likely underestimate the importance of product market competition for management: for example, an exogenous quality increase in management will likely increase profitability and hence lower the Lerner index.

TABLE III
MANAGEMENT AND PRODUCT MARKET COMPETITION

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Estimation method	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
Dependent variable	Management z-score	Management z-score	Management z-score	Management z-score	Management z-score	Management z-score	Management z-score	Management z-score
Import penetration (5-year lagged)	0.144 (0.045)	0.166 (0.071)					0.123 (0.044)	0.180 (0.073)
Lerner index (5-year lagged)			1.516 (0.694)	1.192 (0.568)			1.204 (0.621)	1.257 (0.562)
Number of competitors					0.143 (0.051)	0.140 (0.040)	0.125 (0.043)	0.120 (0.038)
Firms	732	732	726	726	732	732	726	726
General controls	No	Yes	No	Yes	No	Yes	No	Yes

Notes. Coefficients from OLS regressions with standard errors in parentheses (robust to arbitrary heteroscedasticity and clustered by country × industry pair). Sample is a single cross section. General controls includes a full set of three-digit industry dummies, four country dummies, ln(firm size), ln(firm age), a dummy for being listed, the share of workforce with degrees, the share of workforce with MBAs, a dummy for being consolidated, and the noise controls (16 interviewer dummies, the seniority, gender, tenure, and number of countries worked in of the manager who responded, the day of the week the interview was conducted, the time of the day the interview was conducted, the duration of the interviews, and an indicator of the reliability of the information as coded by the interviewer). Import penetration = ln(import/production) in every country × industry pair with the average over 1995–1999 used. Lerner index of competition is constructed, as in Aghion et al. (2005), as the mean of (1–profit/sales) in the entire database (excluding the firm itself) for every country-industry pair (average over 1995–1999 used). Number of competitors is constructed from the response to the survey question on number of competitors, and is coded as zero for none (1% of responses), 1 for less than 5 (51% of responses), and 2 for "5 or more" (48% of responses).

The positive effect of competition on management practices could work through two mechanisms: (1) increasing management scores through greater managerial effort, or (2) increasing the exit rate of badly managed firms relative to well managed firms. Using average managerial hours as a proxy for effort, they find a statistically insignificant relationship between tougher competition and longer managerial hours. They find weak evidence that greater product market competition was associated with a reduction in the dispersion of management practices (as in Syverson (2004a) and Syverson (2004b)) - suggestive of a selection model.

2.2.4 Family firms

Family involvement in firms is commonly observed in this sample. Many firms have a family member as CEO, and of those many choose CEOs by primogeniture (succession to the eldest son). Primogeniture is much more common in France and the UK due to their Norman heritage. Table 4 presents some summary statistics.

TABLE IV
HEREDITARY FAMILY FIRM INVOLVEMENT BY COUNTRY

%	France	Germany	UK	U.S.
Family largest shareholder	30	32	31	10
(of which) Family largest shareholder and family CEO	19	11	23	7
(of which) Family largest shareholder, family CEO, and primogeniture	14	3	15	3
Founder largest shareholder	26	5	15	18
(of which) Founder largest shareholder and CEO	19	1	12	11
Number of firms	125	152	150	290

Notes. These mean values are taken from our sample of 717 firms. Family shareholding is combined across all family members. Family involvement is defined as second-generation family or beyond. Primogeniture is defined by a positive answer to the question "How *was* the management of the firm passed down: was it to the eldest son or by some other way?" Alternatives to primogeniture in frequency order are younger sons, sons-in-law, daughters, brothers, wives, and nephews. "Family largest shareholder" firms defined as those with a single family (combined across all family members, who are all second generation or beyond) as the largest shareholder; "family largest shareholder and family CEO" firms are those with additionally a family member as the CEO; "family largest shareholder, family CEO, and primogeniture" with additionally the CEO selected as the eldest male child upon succession. See Appendix II for more details on construction of the variables.

Table 5 investigates the relationship between firms' management scores and family firms. Family ownership per se is not associated with depressed firm performance, nor is family management. However, family management via primogeniture is strongly negatively and statistically significantly related to management scores. While not randomly assigned, this primogeniture correlation is robust to the inclusion of many control variables.

TABLE V
MANAGEMENT AND FAMILY FIRMS

	(1)	(2)	(3)	(4)	(5)	(6)
Estimation method	OLS	OLS	OLS	OLS	OLS	OLS
Sample	All	All	All	All	All	Family and external owners
Dependent variable	Management *z*-score	Management *z*-score	Management *z*-score	Management *z*-score	Management *z*-score	Management *z*-score
Family largest shareholder	0.005 (0.063)				0.138 (0.086)	0.137 (0.090)
Family largest shareholder and family CEO		−0.105 (0.075)			−0.010 (0.113)	−0.040 (0.114)
Family largest shareholder, family CEO, and primogeniture			−0.317 (0.096)	−0.590 (0.098)	−0.410 (0.122)	−0.379 (0.128)
Firms	732	732	732	732	732	618
Country controls	Yes	Yes	Yes	Yes	Yes	Yes
General controls	Yes	Yes	Yes	No	Yes	Yes

Notes. Coefficients from OLS regressions with standard errors in parentheses (robust to arbitrary heteroscedasticity). The sample is a single cross section. In columns (1) to (5), the complete sample is used; in column (6), founder firms are dropped. "General controls" are a full set of three-digit industry dummies, ln(firm size), ln(firm age), a dummy for being listed, share of workforce with degrees, share of workforce with MBAs, a dummy for being consolidated, and the noise controls (16 interviewer dummies, the seniority, gender, tenure, and number of countries worked in of the manager who responded, the day of the week the interview was conducted, the time of day the interview was conducted, the duration of the interviews, and an indicator of the reliability of the information as coded by the interviewer).

Bloom and Van Reenen note that a substantial fraction of firms have surprisingly bad management practices: scores of two or less imply only basic shop-floor management, very limited monitoring of processes or people, ineffective and inappropriate targets, and poor incentives and firing mechanisms. Most of the cross-country average differences are explained by the left tail. They use their data to quantify that competition and primogeniture account for over half of the tail of badly managed firms. Figure 2 plots the management histogram for all firms reporting low competition and/or primogeniture family succession, relative to the remaining high-competition and no-primogeniture succession. The tail of badly managed firms is substantially larger in the low-competition and primogeniture sample.

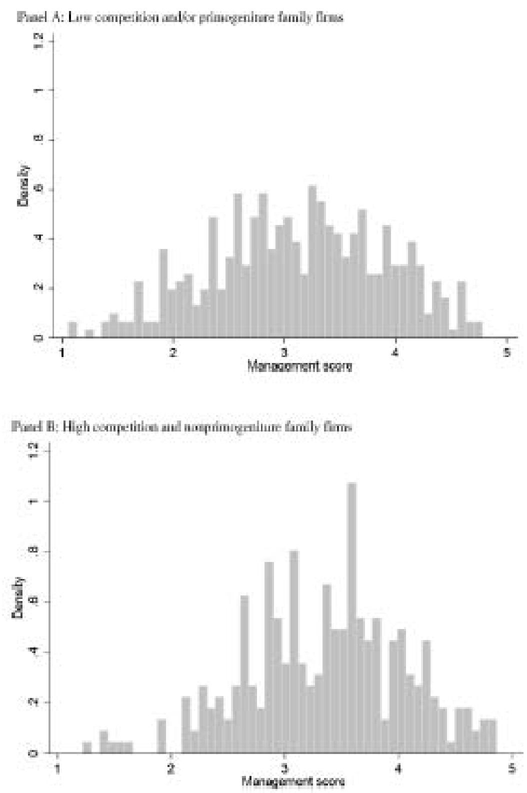

FIGURE II
The Distribution of Management Scores Split by Production Market Competition and Family Firms

Notes: Panel A shows average management scores for the 414 firms which (i) report facing "few" or "no" competitors, and/or (ii) have a family (second generation or more) as the largest shareholder with a family CEO chosen by primogeniture. Split by country is France (95), Germany (101), UK (84) and the U.S. (134). Overall 9.7% of the sample score two or less. 1 indicates worst practice, 5 indicates best practice. Panel B shows average management scores for the 308 firms which report facing "many" competitors and do not have a family (second generation or more) as the largest shareholder with a family CEO chosen by primogeniture. Split by country is France (34), Germany (51), UK (67) and the U.S. (156). Overall 2.9% of the sample score two or less. 1 indicates worst practice, 5 indicates best practice.

Bloom and Van Reenen are very clear in this paper that they are not able to isolate a causal effect of management scores and firm outcomes, but this is something they have pursued in subsequent research.

2.3 Bloom, Eifert, Mahajan, McKenzie and Roberts (2013)

Bloom, Eifert, Mahajan, McKenzie and Roberts (2013) investigate whether differences in management practices can explain differences in productivity across firms by carrying out a field experiment on large Indian textile firms. The treatment is free consulting on management practices provided by an international consulting firm.

2.3.1 Experimental design

1. Empirical setting. The authors argue that Indian firms are broadly representative of firms in emerging economies in terms of (poor) management practices as measured by the management practice scores. The authors focus on the textile industry because it is the largest manufacturing industry in India, and more specifically they focus on large woven cotton fabric firms located near Mumbai. These firms purchase yarn from upstream spinning firms, weave cotton yarn into cotton fabric for suits, shirts and home furnishings, and send their fabric to downstream dyeing and processing firm. The authors choose large firms because they argue that having complicated multi-plant structures means that systematic management practices are likely of value to them.

2. Sample selection. Out of 66 potential subject firms, only 17 firms selected to be in the experiment; these are referred to as *project firms* in the paper. While there is no statistically significant difference between the project firms and other non-project firms based on pre-intervention observables, the selection into the experimental sample could of course be driven by unobservables. The authors argue that since policy efforts to offer management training will also rely on firms volunteering to participate, the estimate for this sample is relevant for the types of firms that take advantage of help when it is offered. The resulting sample consists of 28 plants across 17 firms.

3. Pre-intervention conditions. These firms are all family owned and managed by male family members. The authors document that disorganized production practices lead to frequent quality defects, which require extensive checking and mending processes which employed 19% of the factory manpower on average.

4. Randomization. The authors randomized at the firm level - six firms as control firms, and eleven firms as treatment firms. Due to funding and capacity constraints, the experiment was conducted in two rounds. In the first round, a random plant from four randomly chosen treatment firms received treatment, and the rest received treatment in second round. Eight plants were "non experimental plants" due to a lack of historical performance data and did not directly receive any consulting services.

Table 1. Table 1 presents summary statistics for the firms in the sample. The treatment and control firms are not statistically different across any of the observable characteristics.

TABLE I

THE FIELD EXPERIMENT SAMPLE

	All				Treatment	Control	Diff
	Mean	Median	Min	Max	Mean	Mean	p-value
Number of plants	28	n/a	n/a	n/a	19	9	n/a
Number of experimental plants	20	n/a	n/a	n/a	14	6	n/a
Number of firms	17	n/a	n/a	n/a	11	6	n/a
Plants per firm	1.65	2	1	4	1.73	1.5	0.393
Employees per firm	273	250	70	500	291	236	0.454
Employees, experimental plants	134	132	60	250	144	114	0.161
Hierarchical levels	4.4	4	3	7	4.4	4.4	0.935
Annual sales ($m) per firm	7.45	6	1.4	15.6	7.06	8.37	0.598
Current assets ($m) per firm	8.50	5.21	1.89	29.33	8.83	7.96	0.837
Daily mtrs, experimental plants	5,560	5,130	2,260	13,000	5,757	5,091	0.602
BVR management score	2.60	2.61	1.89	3.28	2.50	2.75	0.203
Management adoption rates	0.262	0.257	0.079	0.553	0.255	0.288	0.575
Age, experimental plant (years)	19.4	16.5	2	46	20.5	16.8	0.662
Quality defects index	5.24	3.89	0.61	16.4	4.47	7.02	0.395
Inventory (1,000 kilograms)	61.1	72.8	7.4	117.0	61.4	60.2	0.945
Output (picks, million)	23.3	25.4	6.9	32.1	22.1	25.8	0.271
Productivity (in logs)	2.90	2.90	2.12	3.59	2.91	2.86	0.869

5. Treatment. The management intervention ran from August 2008 to November 2011. The authors hired consultants in the Indian office of a large international consulting firm (at a pro bono rate) to provide management consulting. The intervention was carried out in three phases. In the first "diagnosis" phase, the consultants diagnosed opportunities for improvement in a set of 38 operational management practices during the first month. These practices form a set of precisely defined binary indicators that are used to measure changes in management practices as a result of the consulting intervention. The second "implementation" phase is the treatment - four months of intensive support for the implementation of recommendations from the diagnosis phase. The control firms received no support during this phase. During the third "measurement" phase, the consultants collected performance and management data from all treatment and control firms. The control firms were provided with 273 consultant hours on average whereas the treatment firms were provided with 781 consultant hours on average.

Figure 5. Figure 5 plots the average management practice adoption for the 14 treatment plants, the 6 control plants, the 8 non experimental plants, and 96 non-project firms surveyed in 2011. All plants started off with low baseline adoption rates, with poor initial adoption rates, but increased by the end of implementation phase, and persisted onward. The firms appeared to adopt the practices that were the easiest to implement and/or had the largest perceived short-run pay-offs.

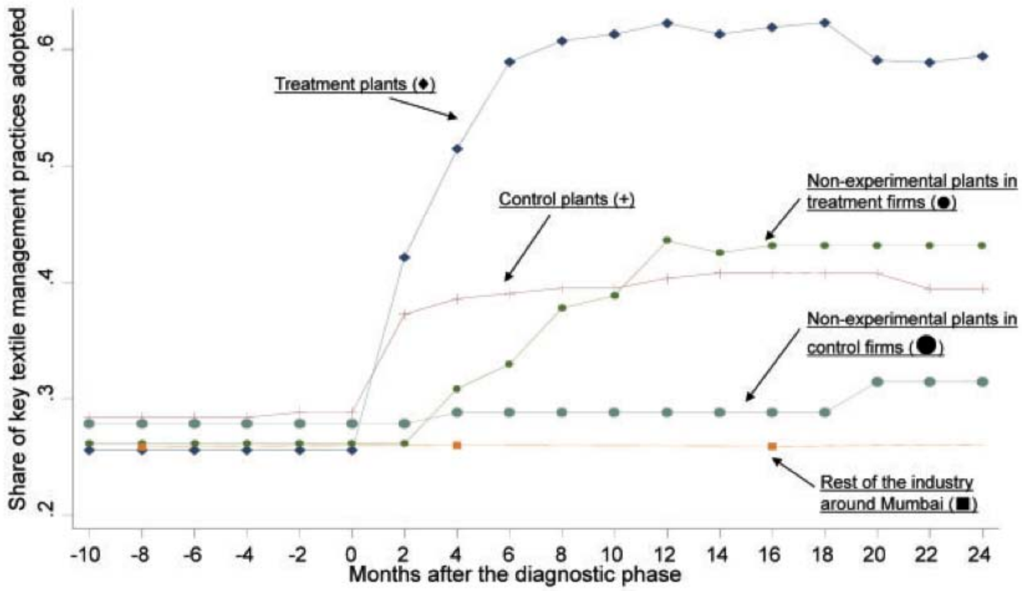

FIGURE V

The Adoption of Key Textile Management Practices over Time

2.3.2 Results

The authors estimate intention to treat (ITT) effects using the following model:

$$OUTCOME_{i,t} = aTREAT_{i,t} + bDURING_{i,t} + c_t + d_i + e_{i,t}, \tag{3}$$

where $OUTCOME$ is one of the performance metrics (quality, inventory, output, and total factor productivity).[3] $TREAT_{i,t}$ takes the value of one for the treatment, while $DURING_{i,t}$ takes the value of one for the treatment plants during the six-month window from the start of the diagnostic phase until the end of implementation phase. c_t are a full set of weekly time dummies to control for seasonality, and d_i are a full set of plant dummies that were included to control for differences between plants. The parameter a is the ITT, which is the average impact of the implementation in the treated plants, and b shows the short-term impact during the implementation.

[3]TFP is defined as log(value added) - 0.42*log(capital) - 0.58*log(labor), where the factor weights are the cost shares for cotton weaving in the Indian Annual Survey of Industry (2004-5); capital includes all physical capital (land, buildings, equipment, and inventory); and labor is production hours.

Table 2. Table 2 summarizes the ITT effects on performance in quality, inventory, output and TFP using the fixed effects model. Given the small cross-sectional sample size, the authors implement permutation tests whose properties are independent of sample size. The estimates suggest statistically significantly improvements in quality, reductions in inventory, and increasing TFP. The authors calculate a roughly $325,000 per plant increase in profits per year. Given the cost of consulting is $250,000, the one year rate of return is 130%.

TABLE II
THE IMPACT OF MODERN MANAGEMENT PRACTICES ON PLANT PERFORMANCE

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable	Quality defects	Inventory	Output	TFP	Quality defects	Inventory	Output	TFP
Specification	ITT	ITT	ITT	ITT	Weeks of treatment	Weeks of treatment	Weeks of treatment	Weeks of treatment
Intervention$_{i,t}$	−0.564**	−0.245**	0.090**	0.154*				
	(0.235)	(0.117)	(0.037)	(0.084)				
During implementation$_{i,t}$	−0.293**	−0.070	0.015	0.048				
	(0.137)	(0.093)	(0.031)	(0.056)				
Cumulative treatment$_{i,t}$					−0.032**	−0.015**	0.006***	0.009**
					(0.013)	(0.005)	(0.002)	(0.004)
Small sample robustness								
Ibragimov-Mueller (95% CI)	[−1.65,0.44]	[−0.83,-0.02]	[0.05,0.38]	[−0.014,0.79]				
Permutation test (p-value)	.001	.060	.026	.061				
Time FEs	127	127	127	127	127	127	127	127
Plant FEs	20	18	20	18	20	18	20	18
Observations	1,807	2,052	2,393	1,831	1,807	2,052	2,393	1,831

Figure 6. Figure 6 plots the trend of the Quality Defect Index (QDI) over the experiment. A reduction in QDI means improved quality. The quality improves significantly and rapidly from about week 5 onward, which was the beginning of the implementation phase following the initial one-month diagnostic phase. The control firms also show reductions in defects that are smaller and more delayed. The time trends are similar for inventory and TFP.

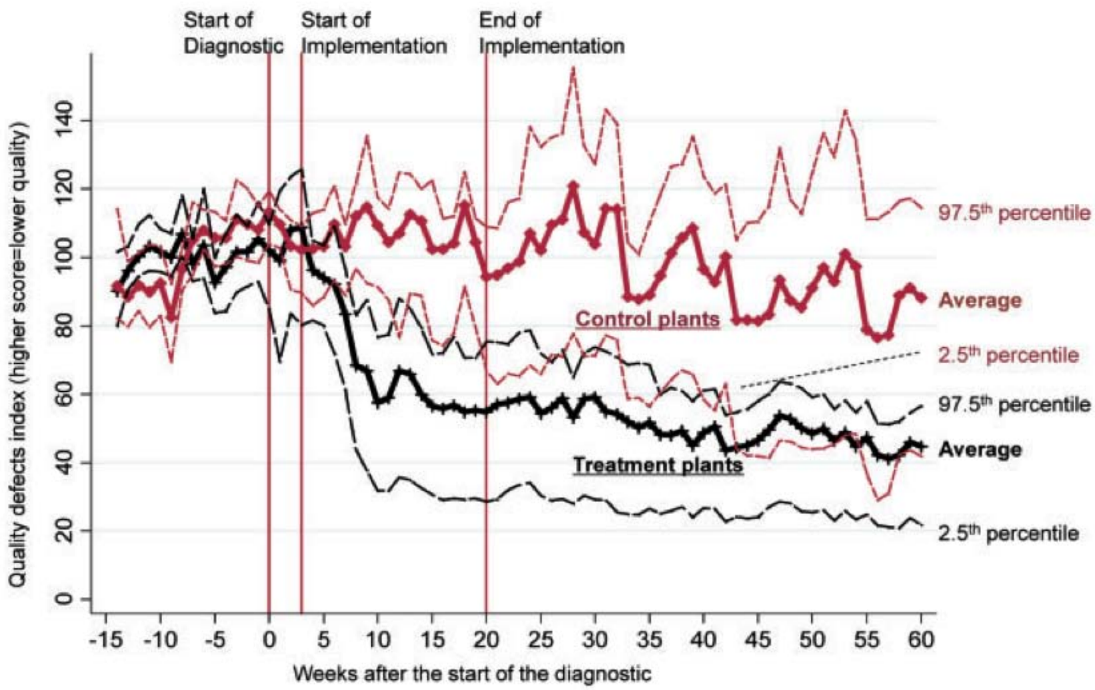

FIGURE VI

Quality Defects Index for the Treatment and Control Plants

Table 3. Table 3 presents estimates of the long-run effect of management practices. The authors focus on the number of plants as a long-run performance indicator because it is easy to measure. The treatment appears to have led to plant expansions. The treatment firms also spread these management improvements from their treatment plants to other plants they owned, providing some revealed preference evidence on their beneficial impact.

TABLE III

LONG-RUN IMPACT OF THE EXPERIMENT ON FIRM SIZE AND DECENTRALIZATION

	Firm size			Delegation to plant management		
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	No. of plants	No. of plants	No. of plants	*z*-score	*z*-score	*z*-score
Sample	Industry	Experiment	Industry	Industry	Experiment	Industry
Time period	2011	2008–2011	2008–2011	2011	2008–2011	2008–2011
Management$_{i,t}$	1.040*			0.597†		
	(0.563)			(0.370)		
Male family members$_{i,t}$	0.210***			0.010		
	(0.065)			(0.042)		
Posttreatment$_{i,t}$		0.217*	0.259**		0.103**	0.171***
		(0.122)	(0.110)		(0.049)	(0.035)
Plant manager related$_i$				0.423***		
				(0.150)		
Plant manager tenure$_i$				0.014**		
				(0.007)		
Small sample robustness						
Permutation tests (*p*-value)	n/a	0.21	0.02	n/a	0.12	0.001
Time FEs	n/a	3	3	n/a	3	3
Plant/Firm FEs	n/a	17	121	n/a	28	128
Observations	107	68	468	120	108	499

2.3.3 Speculating on the reasons for non-adoption

To understand the root causes for the non-adoption of each practice after the diagnosis phase, the consultants conducted a series of analyses to assess whether the cause is an external constraint such as labor regulations, or "incorrect information." The authors also separate the practices into common practices (those that 50% or more of the plants were using before the experiment, like quality and inventory recording) and uncommon practices (those that less than 5% of the plants were using in advance, like quality and inventory review meetings, which are derived from the Japanese-inspired lean manufacturing revolution and are now standard across North America, Japan, and northern Europe but not in developing countries.)

Table 4. Table 4 explores how the reasons for non-adoption evolve before and after the treatment. For all practices, external constraints never appear to be the reason for non-adoption. For common practices, incorrect information is usually the constraint. For uncommon practices, lack of information is the major initial barrier, but can be rapidly overcome. The authors argue that another possible constraint is that owners are time constrained or procrastinate to adopt a practice. Finally, plant managers (versus director-owners) may have limited control over factory management and limited incentives to improve performance because promotion is not possible.

TABLE IV

REASONS FOR THE NONADOPTION OF THE 38 MANAGEMENT PRACTICES (% OF ALL PRACTICES), BEFORE AND AFTER TREATMENT

Nonadoption reason	Group	Management practice type	Timing relative to treatment					
			1 month before	1 month after	3 months after	5 months after	7 months after	9 months after
Lack of information (plants never heard of the practice before)	Treatment	Common	3.3	3.2	0.5	0	0	0
	Treatment	Uncommon	64.0	19.1	2.9	1.5	0	0
	Control	Common	1.9	0	0	0	0	0
	Control	Uncommon	67.8	23.7	22.0	22.0	22.0	22.0
Incorrect information (heard of the practice before but think it is not worth doing)	Treatment	Common	30	22.4	15.4	15.2	14.4	14.4
	Treatment	Uncommon	30.9	50.7	50.7	49.3	49.3	47.1
	Control	Common	18.5	18.5	18.5	18.5	18.5	18.5
	Control	Uncommon	27.1	52.5	50.9	50.9	49.2	49.2
Owner time, ability, or procrastination (the owner is the reason for nonadoption)	Treatment	Common	1.1	0.8	0.5	0.8	1.6	0.8
	Treatment	Uncommon	3.7	13.2	13.2	13.2	13.2	14.0
	Control	Common	3.7	3.7	3.7	3.7	3.7	3.7
	Control	Uncommon	3.4	20.3	18.6	18.6	18.6	18.6
Other (variety of other reasons)	Treatment	Common	0	0	0	0	0	0
	Treatment	Uncommon	2.1	1.5	1.5	2.2	2.2	2.2
	Control	Common	0	0	0	0	0	0
	Control	Uncommon	0	0	0	0	0	0
Total nonadoption	Treatment	Common	34.6	26.4	16.3	16.0	16.0	15.2
	Treatment	Uncommon	98.5	84.6	78.2	66.2	65.1	63.2
	Control	Common	25.1	22.2	22.2	22.2	22.2	22.2
	Control	Uncommon	98.3	96.6	91.5	91.5	89.8	89.8

Later in the paper, the authors speculate on why better-managed firms have not taken over the market. The authors argue that an important reason may be that owners are unwilling to decentralize management because of low trust of managers and poor law enforcement for cases like manager stealing. For firms that cannot delegate management control, they are able to expand beyond the size that could be managed by a single director only if other family members are available to share executive responsibilities (captured by the statistically significant coefficient on number of male family members in Table 3.)

3 Take-aways

My goal in covering this topic is the following: while there have been excellent recent advances on these topics in fields including macro (*e.g.* Hsieh and Klenow (2009)) and industrial organization (*e.g.* Syverson (2004a)), there are important, interesting, and open questions that would benefit from rigorous applied microeconomics research of the type that the methods and techniques we've covered in 14.662 leave you well-suited to pursue. As I highlighted up front, although most of the research in this area has been applied to "firms" and "managers" in a traditional sense, thinking about ways of porting over these ideas and methods to other areas - schools, hospitals, the organization of scientific labor, etc. - seems like a very fruitful area. One recent paper that is a great example is the recent paper on hospital productivity by Amitabh Chandra, Amy Finkelstein, Adam Sacarny, and Chad Syverson: `http://economics.mit.edu/files/8500`.

References

Bartelsman, Eric and Mark Doms, "Understanding productivity: Lessons from longitudinal microdata," *Journal of Economic Literature*, 2000, *38* (3), 569–594.

Bertrand, Marianne and Antoinette Schoar, "Managing with style: The effect of managers on firm policies," *Quarterly Journal of Economics*, 2003, *118* (4), 1169–1208.

Bloom, Nicholas and John Van Reenen, "Measuring and explaining management practices across firms and countries," *Quarterly Journal of Economics*, 2007, *122* (4), 1351–1408.

_ and _ , "Human resource management and productivity," in Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics Volume 4B*, 2011, pp. 1697–1767.

_ , Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts, "Does Management Matter? Evidence from India," *The Quarterly Journal of Economics*, 2013, *128* (1), 1–51.

Foster, Lucia, John Haltiwanger, and Chad Syverson, "Reallocation, firm turnover, and efficiency: Selection on productivity or profitability?," *American Economic Review*, 2008, *98* (1), 394–425.

Griliches, Zvi, "The discovery of the residual: An historical note," *Journal of Economic Literature*, 1996, *34* (4), 1324–1330.

Hsieh, Chang-Tai and Peter Klenow, "Misallocation and manufacturing TFP in China and India," *Quarterly Journal of Economics*, 2009, *124* (4), 1403–1448.

Syverson, Chad, "Market structure and productivity: A concrete example," *Journal of Political Economy*, 2004, *112* (6), 1181–1222.

_ , "Product substitutability and product dispersion," *Review of Economics and Statistics*, 2004, *86* (2), 534–550.

_ , "What determines productivity?," *Journal of Economic Literature*, 2011, *49* (2), 326–365.

Walker, Francis, "The source of business profits," *Quarterly Journal of Economics*, 1887, *1* (3), 265–288.