

What Do Managers Do?

Suggestive Evidence and Potential Theories About Building Relationships*

R. Gibbons, R. Henderson, N. Repenning, and J. Sterman**

MIT

April 24, 2007

Please note that this draft contains only Section II from the outline below.

Persistent Performance Differentials?

- I. Sketching Performance Leadership
- II. Microeconomic Evidence (written with N. Beaulieu)
- III. Dimensions of Potential Theories (developed with J. Rivkin)

What Do Managers Do (to Create PPDs)?

- IV. Managers in Economic Models?

Building and Changing Relationships

- V. What the Folk Theorem Does Tell Us
- VI. Thick Descriptions and “Grounded” Theories (channeling D. Kreps)
- VII. Elemental Theories (summarizing S. Chassang (and G. Ellison-R. Holden?))
- VIII. Applied-theory Modeling and Testing

Interim Conclusion and Next Steps

- IX. Organizational Capabilities?

* To appear in *Handbook of Organizational Economics* (R. Gibbons and J. Roberts, eds.), Princeton University Press, 2008.

** We are grateful to: Nancy Beaulieu and Jan Rivkin for many contributions; David Kreps, Sendhil Mullainathan, John Roberts, and seminar audiences at Berkeley, Harvard, Michigan, NBER, and Washington University in St. Louis for helpful comments; and MIT’s Program on Innovation in Markets and Organizations for financial support.

Please note that this draft contains only Section II from the outline above.

II. Microeconomic Evidence of Persistent Performance Differences among Seemingly Similar Enterprises

N. Beaulieu, R. Gibbons, and R. Henderson
April 20, 2007

In this section we begin to survey the microeconomic evidence of persistent performance differences (PPDs) among seemingly similar enterprises (SSEs). Attempts to document such PPDs among SSEs face three broad challenges. First, there is the challenge of similarity: it is often difficult to control for heterogeneity in both inputs and outputs. Second, there is the challenge of persistence: it is often difficult to distinguish permanent from transient or spurious measured performance differentials. Third, there is the challenge of generalizability: datasets that begin to offer solutions to the first two challenges are often small and focused.

Because of these three challenges, no one paper (indeed, perhaps no one literature) can fully resolve all the concerns a reader might have. Our approach is therefore to present a wide range of studies, in an attempt to construct a mosaic that supports a central theme. This range of studies includes data drawn from different industries, data analyzed with different methodologies, levels of analysis from a plant to a corporation, and performance differences both between and within firms.

Because this range of studies is quite broad, there are more published papers than we can discuss here. Furthermore, thanks to ongoing improvements in both data and empirical methodologies, new papers that bear (directly or indirectly) on our topic appear regularly. We would welcome hearing about existing and future work that we have omitted in this first attempt at a survey.

To impose some order on this wide range of studies, we have grouped the evidence into four categories: large-sample studies of profitability, large-sample studies of productivity measured in dollars, studies of productivity measured in physical units, and studies relating management practices to productivity. No one category resolves every concern, but each provides a useful piece of our mosaic. The central theme that we see emerging from this mosaic is that persistent performance differences do exist among seemingly similar enterprises, and these performance differences are related in part to differences in the internal structures and processes of organizations.

II.A Large-Sample Profitability Studies

Over the last two decades, large-sample studies of the intertemporal patterns of firm performance have played a central role in the spirited debate between IO economists and

scholars sympathetic to the resource-based view of the firm. These studies have sought to determine which has a larger influence on the performance of an enterprise: the external workings of the market or the internal workings of the firm? Answers to this question have typically been generated by analyses of variance and covariance in firm performance. Multiple studies conducted on different samples provide evidence of significant and stable differences in performance between firms in the same industry.¹ Though the studies differ in the exact proportion of the variance attributable to corporate entities, business segments, and business units, these studies find that firm-level effects dominate industry-level effects for the economy overall and for the manufacturing sector in particular (**Cubbin and Geroski, 1987; Rumelt, 1991; McGahan and Porter, 1997; Claver, Molina, and Tari, 2002; Mauri and Michaels, 1998; Hawanini, Subramanian, and Verdin, 2003**).

The literature on performance differences within and between industries was initiated by **Schmalensee (1985)** with analyses of variance in return on assets using a single year of business-unit performance data produced by the Federal Trade Commission. With just one year of data, it was impossible to identify stable performance differences between business units, though a corporate effect could be estimated from covariation in performance among business units in the same corporation. Based on prevailing IO theory, Schmalensee used industry market share as a proxy for the corporate effect and found that less than 1% of variation in ROA could be explained by variation in corporate market share. Furthermore, this corporate effect was exactly offset by a negative covariance between corporate market share and the industry effect. In the end, Schmalensee's analysis explained just 20% of the variation in ROA, and all of this effect was attributable to industry membership.

Schmalensee's cross-sectional data limited the analyses he could conduct. Using four years of the same FTC data, **Rumelt (1991)** found that industry effects could explain between 8% and 18% of the variance, depending on the sample and the model used for analysis.² Exploiting the panel data, Rumelt estimated that business unit effects accounted for 34% to 47% of the variation in ROA (again, depending on the sample and model). Corporate effects (covariation in performance of business units belonging to the same corporation) ranged from 0% to 18%.

Rumelt's estimates of the proportion of variance in firm performance explained by business-unit effects were replicated in multiple studies using different data sets and different models. Using Compustat data, **McGahan and Porter (1997, 1999)** significantly extended the research of Rumelt by using longer panels of data, including non-manufacturing firms, and explicitly modeling the temporal persistence of shocks through an autoregressive error term.³ While demonstrating that Rumelt's findings were

¹ In these studies, performance is measured in multiple ways including ROA, EVA (adjusting ROA for cost of capital) and total market value.

² In the ANOVA analyses, the results also depended on the order in which variables were entered.

³ The FTC line-of-business data differ from the Compustat data on a few dimensions. First, there are only 4 years of FTC data from the 1970s; Compustat has a much longer panel (15+ years) from a more recent time period (1986 onwards). Second, the FTC data contain information at the business-unit level and only for manufacturing firms; the Compustat data are drawn from firms' 10K reports, cover all economic sectors

generalizable beyond the specific model and dataset he used, M/P found that the relative importance of business-unit and industry effects varied tremendously depending on the economic sector in which the firm competed. Analyses of business performance in non-manufacturing sectors (wholesale and retail trade, agriculture and mining, services, transportation, lodging and entertainment) indicated that a firm's industry affiliation could explain a large percentage (29% to 64%) of the variation in performance. Business-segment effects were estimated to be small for non-manufacturing industries (2% to 10%) with the exception of lodging/entertainment and services (19% and 33% of variance, respectively).

There are multiple potential explanations for McGahan and Porter's finding of differences in the relative importance of firm-level effects by sector, and some of these explanations may be substantively important for the study of PPDs. First, this finding could be an artifact of the industry classification system if industries in the manufacturing sector are more homogenous than industries in the service sector. Second, it could be that underlying differences in the product or the production process contribute to the feasibility of establishing persistent performance differences (*e.g.*, TQM might differ for a physical product versus a service). Third, it might be easier to patent a physical product than a service (*e.g.*, an iPod vs. internet airline booking). Fourth, patterns of business conglomeration might differ systematically across sectors (*i.e.*, economies of scale and scope).

While the claim of stable business-segment performance effects appears uncontroversial (at least for some sectors), there is much less agreement in the literature on the importance of corporate effects on performance. When the sample is restricted to include only multi-segment (*i.e.*, diversified) firms, corporate affiliation exceeds industry affiliation in the percentage of business-segment variation explained (**Brush, Bromiley, and Hendrickx, 1999; Roquebert, Phillips, and Westfall, 1996**). These studies find corporate effects explaining approximately 18% of total variance, while the proportion of variance explained by business-segment effects remains large (about 33%). In addition, the Roquebert *et al.* study finds that the magnitude of the estimated corporate effect is decreasing in the number of business segments represented under each corporation; going from 4 business segments to 6 business segments cuts the variance explained by corporate affiliation in half (18% to 9%). It is unclear whether these findings would hold up in analyses of individual economic sectors. In M/P analyses of non-manufacturing sectors, corporate affiliation explained nearly as much of the variation in business-segment performance as industry affiliation; however, the sample for this study includes a large number of single-segment firms (5212 out of 7003) and for these firms, the corporate effect is constrained to be zero. M/P also find a negative covariance between corporate and industry effects.

Differences in estimated PPDs relating to different levels of aggregation within the firm (*i.e.*, business units, business segments, corporate entities) also offer some potential insight into the sources of PPDs. The original studies of Rumelt and McGahan-Porter are

except financial, and contain data reported at the business segment level (a higher level of aggregation than business unit). Most US studies conducted after Rumelt's use Compustat data.

not directly comparable because they use different data sets. Rumelt specifies effects for both the corporate entity and the business unit. In the Compustat data used by M/P, the lowest level of aggregation is the business segment (often encompassing multiple business units). When M/P ran Rumelt's model on manufacturing firms in the Compustat data, they found smaller firm effects. It could be that aggregation to the business segment obscures heterogeneity among the constituent business units, which when combined with variation in the composition of business segments across firms, might lead to attenuation of stable performance differences. This variation in performance among business units within the same business segment (or corporate entity) raises intriguing questions about replication of successful business strategies within firms, not just the imitation of successful strategies between firms.

Hawanini, Subramanian, and Verdin (2003) simultaneously confirmed, challenged, and extended the empirical results in this literature. Using a new data source, the authors confirmed the dominance of stable firm effects (in this case, measured at the corporate level) in explaining the variance in business performance over time. In previous studies, researchers had relied on accounting measures of profitability as the measure of firm performance; Hawanini *et al.* show that these past results are robust to the use of different performance measures that more closely accord with value creation (specifically, economic profit and market capitalization). Finally, the authors test whether previous findings on the relative importance of firm and industry effects are sensitive to the exclusion of extreme performers in each industry. To conduct this test, they dropped the top two and the bottom two performers in each industry from the sample and re-estimated the models. In terms of the percentage of total variance explained, firm effects fell by 35 to 54% in the restricted sample, whereas industry effects increased by 100 to 300%. The ratio of the variance explained by firm relative to industry effects also fell. These findings suggest that persistent performance differences at the firm level, separate from industry affiliation, may be confined to a small number of firms in each industry.

At least two studies have moved beyond variance decomposition to investigate the association between organizational strategy and stable firm effects (a topic we consider in more detail in Section II.D). **Hansen and Wernerfeldt (1989)** used a cross-sectional dataset comprised of 60 Fortune 1000 firms encompassing 300 lines of business to assess the relative importance of industry and firm factors in explaining firm profitability. The authors regressed firm profits on economic measures (industry profitability, market share, firm size) and organizational measures (*e.g.*, emphasis on human resources, emphasis on goal accomplishment). The results indicate that organizational factors explain about twice as much variation in performance (profitability) compared to economic factors and that these two sets of factors appear to have independent effects on firm profitability. In a similar spirit, **Mauri and Michaels (1998)** examine whether, within industry, firms are differentiated in their "core strategies," or alternatively whether variation in core strategy is closely associated with industry affiliation. The authors use intensity of advertising and R&D expenditures, respectively, as proxies for core marketing and technology strategies. Analyses of variance decomposition suggest that industry affiliation explains the larger proportion of variance in core strategy investments while firm effects account for more of the variation in financial performance.

In summary, the studies describing the decomposition of firm-level profitability into industry, corporate, and business-unit effects suggest that persistent performance differences exist between comparable business entities. In multiple studies, using different data sets collected over different time periods, scholars have consistently attributed roughly one third of the total variation in profitability to stable firm or business-unit effects. While inferences from the early studies in this literature were limited by small and narrow samples, the critical findings were replicated in subsequent studies based on longer panels and covering firms from multiple economic sectors. It is also noteworthy that 35% to 55% of the performance variation could not be explained by the models employed. Finally, the finding by Hawawini *et al.* that a small number of extreme performers in each industry account for most of the intra-industry variation in profitability warrants further investigation.

II.B Large-Sample Productivity Studies

It is widely appreciated that profitability and other accounting-based measures of performance are imperfect proxies for productive efficiency. For example, higher profitability may be achieved through lower input prices or higher output prices, rather than through higher volume of output created per unit of standardized inputs. In our investigation of PPDs, we are interested in performance differences related to the internal workings of the firm, not those attributable to firms' market power in input or output markets. In this sub-section, we review empirical studies of productivity that control not only for differences in firms' industry affiliations but also for the inputs used in production.⁴

II.B.1 Firm-Level Heterogeneity in Productivity

In the late 1970s, Griliches and Mairesse set out to study the effects of research and development on productivity at the firm level. These analyses required them to assemble the first large-sample data sets containing detailed plant-level data on inputs and outputs. Though not the primary focus of their research, they discovered a surprisingly large amount of between-firm heterogeneity in the data (*e.g.*, in deflated sales, number of employees, physical plant, and R&D capital stock) even after accounting for the plant's industrial sector and adjusting for labor inputs (**Griliches and Mairesse, 1981, 1982, 1985; Griliches, 1986**). For example, in 12 years of data from a sample of 133 large US firms (mostly in manufacturing), over 70% of the variability in deflated sales per employee occurred between firms compared to within firms over time, and 90% of the variability in a measure of the R&D capital stock per employee was attributable to

⁴ In most studies in this literature, the productivity of a firm or establishment is measured as a residual – what is left over after prices and quantities of inputs have been accounted for. This residual is typically referred to as Total Factor Productivity (TFP) and is a primary object of analysis in economic growth theory. Index procedures are used to adjust TFP for the changing prices of inputs over time. While TFP is conceptually easy to understand, there is no universally agreed upon method for computing it. See Hulten (2000) for a discussion of the origins of TFP, alternative methods for its computation, and implications for empirical research.

differences between firms (**GM 1981, Table A1**). This heterogeneity carried over into their econometric analyses as well: estimated parameters from a simple model of the production function revealed a large amount of between-firm variability in the slope coefficient for R&D capital (**Griliches and Mairesse, 1988**).⁵ In the quest to better understand this apparent heterogeneity, they estimated a separate production function for each firm and aggregated the results to compute the implied distributions of the constant term and the capital slope coefficient. They found that the implied firm-level variability in these parameters was robust to a variety of analytic approaches and exceeded the researchers' expectations by at least a factor of 2.⁶

Griliches and Mairesse based most of their empirical studies on data from France, Japan, and the United States. As more micro-data sets were assembled and analyzed, it became clear that this heterogeneity was characteristic of firms in many developed economies. In a study of scale economies and market power in 14 Norwegian manufacturing industries, **Klette (1999)** found significant and persistent productivity differences between plants operating in the same industry. In 13 out of 14 industries, the estimated variance of the distributions of firm fixed effects was significantly different from zero (Table III). Also using data from Norway (paper and pulp, chemical, and basic metals industries), **Biorn (2002)** estimated a four-factor production function by industry on four years of panel data and allowed for heterogeneity in the intercept term and the scale and input coefficients. Biorn found that 72-84% of the gross disturbance was due to heterogeneity in the random intercepts and that 82-91% of the gross disturbance was accounted for by the combination of the random intercepts and random coefficients (input and scale elasticities). Using census data on the Columbian clothing and apparel manufacturing industries, **Van Beisebroeck (2004)** documented substantial dispersion in firm-level productivity – 50% of firms on average had productivity levels less than 32% of the median or 35% above the median. The author compared the estimated productivity dispersion (measured by interquartile ranges) implied by different econometric models and concluded: “The ranges are very similar which is remarkable because the methods rely on very different calculations and assumptions” (p. 21). In Taiwanese data, **Aw, Chen and Roberts (2001)** found that entering and exiting firms were less productive than surviving firms, though the entrants themselves were quite heterogeneous with respect to productivity: For most industries, in the year that they entered, new firms that ultimately survived were significantly more productive than new firms that ultimately failed (see Table 6). In eight out of nine manufacturing industries, productivity increased

⁵ Griliches and Mairesse estimated a simple model of the Cobb-Douglas production function in which the log of output per labor input was regressed on a constant term and a measure of capital employed per labor input.

⁶ “There are two numbers worth keeping in mind: our prior expectation about a reasonable heterogeneity in the true β 's is a standard deviation of 0.1. The observed standard deviation at the individual level is about 0.5 (0.4 in first differences). A direct subtraction of an estimate of the sampling variance (the Swamy method) results in a “residual” estimate of the true dispersion σ_β between 0.4 in levels and 0.2 in first differences, the latter still being about twice as high as our prior expectations (pp. 21-22).” Griliches and Mairesse, 1988.

over the 10-year study period, however this increase in productivity was not in general accompanied by a narrowing of the productivity distribution (see Table 3 and Figure 1).⁷

Though initial research in this field focused on the productivity of capital (both physical and R&D) and economies of scale, later studies explored variability in labor productivity as a potential explanation for heterogeneous firm performance. In their study of labor productivity in Israeli firms, **Griliches and Regev (1995)** found that most of the increases in labor productivity over time occurred within firms instead of being the result of firm entry and exit. However, the authors noted that: “the bulk of productivity differences across firms is not accounted for by the regression and that such ‘firm’ effects, or unobservable factors, are a relatively permanent aspect of the firm and an important part of the story ... In terms of variance components, the unobserved firm factors account for over 85% of the residual variance” (p. 197). **Haltiwanger, Lane and Spletzer (1999)** used Census data to explore the potential relationship between firm productivity (both in levels and growth rates) and workforce composition. They began by confirming in their data previous findings of heterogeneity in firm-level measures of productivity: “After removing year and two-digit industry means, our measure of labor productivity varies widely across firms: the interquartile range is 55 log points in levels, and 39 log points for the four-year difference [growth]. Productivity levels are significantly positively correlated over time” (p. 96).

II.B.2 Persistent Differences in Productivity

While the research discussed in the previous sub-section provides evidence of substantial productivity differences between firms at a point in time, most of these studies did not establish persistence of the productivity differentials over time. To test for stability (persistence) in their estimates, **Griliches and Mairesse (1988)** split their panel data into two equal time periods (6 years) and estimated production-function parameters for each firm in the two different samples. The authors found a low correlation between the (capital) slope coefficient estimates from the two samples, suggesting that these estimates were not stable over time; they concluded that the sparseness of their model and limitations of their data made it unlikely that they could detect stable differences in the slope coefficient between firms over time. However, for the constant term, the authors found strong evidence of stable heterogeneity over time, providing support for the claim of persistent productivity differences among firms.⁸

The increasing availability of large samples of firm-level micro data enabled IO economists to investigate industry dynamics and the sources of aggregate productivity growth. A major theme in this literature has been to quantify the extent to which industry productivity growth can be ascribed to external restructuring (entry and exit) compared to internal restructuring (productivity growth by incumbents). For PPDs, the relevant issue is whether cross-sectional heterogeneity comes about as a byproduct of industrial

⁷ The cross-sectional dispersion in productivity is found to be even greater in Korea (Aw, Chung, and Roberts, 2002).

⁸ The stability and heterogeneity results for the constant term were provided in footnote 10. The authors provided neither measures of the heterogeneity in the constant term nor measures of its stability over time.

evolution or is persistent over time. **Bailey, Hulten, and Campbell (1992)** documented productivity differences among U.S. manufacturing plants using the Longitudinal Research Database: “As we have examined this data, we have been impressed by the diversity among plants and among industries ... Both in the level of and rate of change in productivity, plants manifest significant differences” (p. 188). Moving beyond measures of cross-sectional dispersion, these authors were among the first to compute transitional matrices for a large representative sample to describe the movement of plants in the productivity distribution over time. To populate these matrices, the authors computed plant-level total factor productivity (TFP) for plants in 23 industries at four points in time (1972, 1977, 1982, and 1987). Decomposing TFP growth, they find that plants entering and exiting made negligible contributions to aggregate productivity growth. In periods of economic growth (1972-1977 and 1982-1987), intra-firm growth in productivity contributed more to aggregate industry growth than market share reallocations among ongoing concerns; the reverse was true during the period of overall productivity decline (1977-1982).

To examine persistence of productivity differences over time, Bailey *et al.* assigned each plant to a quintile of the productivity distribution in each of the survey years. The transition measures provide substantial evidence of persistence at the top of the productivity distribution: 75% of plants in the top quintile in 1972 remain in the top two quintiles in 1977.⁹ Persistence is less strong, but still evident, in the bottom of the productivity distribution: 48% of plants in the bottom quintile in 1972 appeared in the bottom two quintiles in 1977. Extending the time period to a 10-year interval resulted in a modest decrease in observed persistence: fifty-eight percent of plants in the top quintile in 1972 remain in the top two quintiles in 1982; 54% of plants in the bottom quintile in 1972 remain in the bottom two quintiles in 1982.

In a similar vein, **Disney, Haskel, and Heden (2003)** studied manufacturing establishments in the United Kingdom and found that establishments in the top decile of the productivity distribution were 156% more productive than establishments in the bottom decile.¹⁰ To test whether the observed cross-sectional heterogeneity was attributable to market selection of high-productivity firms and new entry of firms with a mix of productivity levels, the authors computed the productivity dispersion for a fixed cohort of firms over time. In 1982, the 90-10 spread for the 1982 cohort ranged from 1.48 (ln TFP) to 1.62 (ln labor productivity). Five years later, in 1987, these spreads for the firms remaining from the 1982 cohort had fallen to 0.81 and 1.43 (ln TFP and ln LP, respectively), but were not close to zero. Even four years further on, in 1991, the heterogeneity for the firms remaining from the 1982 cohort had fallen only a bit further, with 90-10 spreads now 0.66 and 1.28, respectively. Thus, after almost a decade, substantial heterogeneity persisted among those firms that had not exited from a fixed initial cohort of firms.

⁹ In the transition matrix, plant productivities are weighted by their employment). With unweighted data, the percent of 1972 top-quintile plants remaining in the top two quintiles in 1977 is 47%.

¹⁰ Productivity was measured according to real gross output per hour of manual labor (LP). This 90-10 productivity advantage is reduced to 91% when the log of total factor productivity (TFP) was used as the productivity measure.

Foster, Haltiwanger, and Krizan (2002) found similar results for firms in the retail trade industry. The authors employed quintennial census data (1987-1992-1997) to compute establishment-level productivity adjusted for 4-digit industry effects. They found that more than 26% of firms observed in the top quintile in 1987 remained in the top quintile a decade later in 1997. This top quintile survival percentage is even more remarkable in light of the fact that nearly 40% of the firms in the top quintile in 1987 had gone out of business by 1997.

II.B.3 Measurement and Estimation Issues

As evidence of heterogeneity began to mount, economists voiced concern that the observed heterogeneity might be simply an artifact of how inputs and outputs were measured. By necessity, nearly all large-sample studies at the plant- or firm-level use productivity measures that are computed from gross revenues instead of physical output measures. It was therefore plausible that the estimated dispersion in productivity might be due in large part to inter-firm variability in prices. To investigate this concern, **Mairesse and Jaumandreu (2005)** added firm-level output price data and capacity utilization data to their productivity regressions for French and Spanish manufacturing firms. Independent of the estimation technique, the authors found that using real (individual-firm output-price deflated) or nominal output measures mattered very little to their estimates of scale and input elasticities. In fact, using a different data set, **Foster, Haltiwanger, and Syverson (2005)** found that dispersion in the physical output-based measure of productivity was *greater than* the dispersion computed from revenue-based measures of productivity. While not the primary focus of this paper, Foster *et al.* also document persistence of both measures of productivity: annual autocorrelation coefficients from firm-level regressions ranged from 0.75 to 0.8.

In their study of the telecommunications equipment industry, **Olley and Pakes (1996)** brought to light the problems of selection and sample biases that typically plague longitudinal studies of productivity. These biases result from interfirm productivity differentials that are unobserved by the econometrician but condition the firms' investment and exit decisions. A primary contribution of the paper is the illustration of a new technique to adjust for selection and endogeneity (of the firm's choice of inputs) using the firm's capital investment as a proxy for unobserved firm-specific productivity. Their analyses of industry dynamics during the time of deregulation suggest that aggregate productivity growth came about as a result of the reallocation of capital to high-productivity firms: "the productivity growth that followed regulatory change seemed to result from the downsizing (frequently the shutdown) of (often older) unproductive plants, and the disproportionate growth of productive establishments (often new entrants)" (p. 1266). These findings are consistent with a theory developed by **Jovanovic (1982)** in which firms learn over time about their draws from the productivity distribution.

In a similar vein, **Eslava, Haltiwanger, Kugler and Kugler (2004)** show that persistent productivity differences remain even after correcting measurement problems (by using

data on quantities instead of revenues) and adjusting for sample and selection biases (by using an instrumental-variables technique similar to that proposed by Olley and Pakes). Using a rich panel data set on Columbian manufacturing plants, the authors investigate the effects of market reforms on productivity dispersion and the efficiency of market reallocation. They find a large amount of dispersion in plant-level productivity both before market reforms (std. dev. / mean = 55) and after market reforms (std. dev. / mean = 20). The authors measure persistence as the coefficient on a one-year lag variable; for total factor productivity, the coefficient on the lag variable equals 0.90 prior to market reforms and 0.94 after market reforms. The authors note that the standard deviation of productivity increases considerably following market reforms, even though many low-productivity firms exit.

This review of micro-econometric studies from the productivity literature reaffirms the finding from profitability studies of the existence of PPDs among similar business entities. Productivity analyses have yielded persistent firm or establishment effects that could not be eliminated by adjustment for input mix, the application of different estimation methods, functional-form assumptions, or more detailed data. Markov-style analyses of productivity distributions over time indicate that membership in the top decile or quintile of the distribution has been remarkably stable over five- and even ten-year periods. Though these productivity studies come much closer than profitability studies to capturing the theoretical conception of productive efficiency, they nonetheless face measurement and estimation issues. However, empirical studies designed to address these issues do not provide sufficient evidence to reject PPDs; in some cases, refinements of data and methodology have led to increases, rather than decreases, in estimates of intra-industry productivity dispersion.

II.C Productivity Studies with Physical Output Measures

In large samples of diverse firms, it is difficult to rebut the concern that the estimated productivity dispersion is an artifact of the way productivity is measured (*e.g.*, as a residual of the total factor productivity equation – see footnote 4) or of unmeasured differences in the products the firms produce or the markets they serve (including differences in prices). To address these concerns, in this sub-section we discuss studies in which productivity is measured in units of physical output (or rates of physical output produced per unit of time). In some of these studies, the sample is comprised of multiple plants or establishments belonging to the same firm or firms. In these intra-firm studies, we have increased confidence that observed productivity differences are not derived from unaccounted differences in products, physical plant, equipment, or production technology (as traditionally conceived).

II.C.1 Raw or Adjusted Output Measures

In a series of detailed field studies (**Macher and Mowry, 2003; Hatch and Mowry, 2002; Appleyard, Hatch, and Mowrey, 2000**) Mowrey and colleagues investigated the effects of various organizational practices (*e.g.*, personnel, information technology,

knowledge management, and research and development) on productivity in semiconductor manufacturing. In these studies, performance was measured by defect rates, yield, and cycle time, and by improvements in these measures (as a proxies for organizational learning). In two separate studies, the authors found that, following the introduction of new technology, initial defect rates varied by a factor of five. Furthermore, these initial differences did not disappear over time: “The average within-firm coefficient of variation for yield improvement was 1.44, while the between-firm coefficient of variance (sic) was 4.42. Similarly, the average within-firm coefficient of variation for cycle time was 0.66, while the between-firm coefficient of variation was 2.14. Our performance metrics thus appear to capture firm-specific differences that are both substantial and enduring, consistent with the arguments of other researchers who examine inter-organizational differences in rates of learning” (Macher and Mowry, 2001: 19).¹¹

Two studies in the healthcare industry document performance differences in heart-surgery outcomes among hospitals. **McClelland and Staiger (1999)** found substantial and persistent inter-hospital variation in mortality and readmission rates using a panel dataset spanning 11 years and containing approximately 580,000 patient observations in 3954 hospitals. For example, the average hospital 30-day mortality rate for heart attack patients in 1984 was 19% and the standard deviation exceeded 9% after controlling for differences in patient characteristics and observable hospital characteristics. The authors also found evidence of persistence in these performance measures: the estimated one-year lag coefficient on mortality was 0.887. **Huckman and Pisano (2005)** examine in-hospital mortality following coronary artery bypass graft (CABG) surgery performed in Pennsylvania hospitals during 1994 and 1995. Their patient-level data cover more than 38,000 procedures performed by 203 surgeons operating at 43 hospitals. They compute the mean risk-adjusted hospital mortality rate as 3.1% and the standard deviation as 1.9%, yielding a 67% coefficient of variation.

To compare the efficiency of team production to more traditional manufacturing processes, **Dunlop and Weil (1996)** collected data from 42 business units in the United States producing a narrow range of apparel products. The raw performance data on lead-time (the total time from ordering inputs to finished products) and operating profits (as a percentage of shipments) exhibit substantial inter-unit variability (coefficients of variation of .80 for lead time and .69 for operating profits). In a similar spirit, **Arthur (1994)** reduces product diversity even further in a study of 29 steel mills, but still finds substantial coefficients of variation in labor efficiency (.43) and scrap rates (1.0).¹²

¹¹ The coefficient-of-variation statistics were listed in a pre-publication (2001) version of Macher and Mowery, 2003. In the published version, the statistics were replaced with the following descriptive text: “Significant variance among the firms in the sample is displayed in Figure 1, with initial defect densities of the worst performing fabs roughly five times that of the best performing fabs. Figure 1 also displays the rate of improvement in defect density for each fab displayed. Although all semiconductor manufacturers improve in this manufacturing performance measure, the slopes of these production ‘learning curves’ vary greatly among the individual production facilities” (Macher and Mowery, 2003: 400). A similar description of initial cycle times and rates of cycle time improvement also appears on page 400.

¹² The author measures labor efficiency as the number labor hours used to produce one ton of steel.

II.C.2 Intrafirm Studies

To better understand the processes underlying knowledge creation and transfer, **Chew, Bresnahan, and Clark (1990)** undertook a study of 40 operating units of a single commercial food division of a large multi-divisional corporation. Despite the fact that these 40 operating units were very similar along multiple dimensions (*e.g.*, all were located in the U.S., employed low-skill labor, used the same technology, and produced similar products for similar customers), the researchers found a three-fold difference in productivity across units (in a standardized measure of meals produced divided by labor and capital inputs). Even after using regression analysis to adjust for local labor-market characteristics, size of the local product market, unionization, age of equipment, product quality, and local monopoly, the top-ranked unit was twice as productive as the bottom-ranked unit.

Organizational researchers frequently employ firm fixed effects in their regression analyses either to control for time-invariant firm-specific elements of performance or to delve more deeply into the internal structures and processes that could plausibly generate performance variation. An example of the former is the study by **Argote et al. (1990)** in which the authors examine the volume of Liberty ships produced in 16 separate shipyards over a five-year time period during World War II. Due to special circumstances, these data present an unusual opportunity to study performance differences and organizational learning: “A large number [of ships] were produced – 2708. A standard design was adopted and produced with minor variation in all of the yards. Parts were standardized, procured by a central authority, and distributed to the yards. All of the yards producing the Liberty Ship began production during 1941 or 1942. The Liberty Ship was the first ship to be produced in any of the yards, and it was the only ship produced by the yards during a significant part of the war. The vast majority of the workers employed in the Emergency Shipyards had no prior experience in shipbuilding” (p. 142). Using tonnage of vessels produced per month as a dependent variable, the authors estimate a production function that controls for labor and capital inputs as well as the cumulative experience and accumulated knowledge in the yard. The authors include a yard-specific dummy variable for each of the 16 shipyards. They do not report the coefficients for these dummy variables but remark in a footnote: “A joint test of the null hypothesis that there are no yard-specific effects is rejected at a very high significance level ($p < 0.001$), so important yard-specific effects appear to be present” (footnote 5, p. 144).

In a longitudinal study of research and development in the pharmaceutical industry, **Henderson and Cockburn (1996)** measure productivity in terms of the number of important patents secured by the firm. The unbalanced panel is comprised of data on R&D productivity at the research-program level for 10 firms over 20 years on average. With up to 39 research programs per firm, the authors are able to examine both intra- and inter-firm productivity differences. In their firm-level Poisson regressions (where the dependent variable is the number of patents held by the firm), the authors show that the addition of firm fixed effects results in an increase in the log likelihood function that is equivalent to increasing the R-squared statistic from 0.29 to 0.79; thus, a substantial amount of the total variation in research productivity occurs between firms and is

persistent over time. When the regressions are estimated at the research-program level and fixed effects are included for research classes, firm fixed effects remain significant, which indicates that differences across firms in the composition of research programs cannot explain all of the persistent differences in research productivity among the firms.

In the section we have examined studies that measured productivity in physical units. The primary contribution of such studies to assessing the potential existence of PPDs is the elimination of unmeasured variation in prices (and, in some studies, product quality) as an alternative explanation for observed performance differences. Collecting data for studies of this sort often requires the researcher to spend time immersed in the production process. The knowledge gained by this field experience and embodied in the written research product gives us confidence that the performance differences documented in these studies are real. Also, even though each study is necessarily within a narrow industry, PPDs in the form of physical productivity differences have been documented in a wide range of industries, including semiconductors, apparel manufacturing, hospitals, steel-mini-mills, ship building, pharmaceutical research and high-precision machining. Finally, such PPDs have been documented within firms as well as between firms. This intra-firm variation in productivity prompts questions concerning the obstacles to replication of successful practices within organizations, including whether and how these internal obstacles are different from the obstacles to imitation of successful practices between organizations.

II.D Productivity Differences and Management Practices

The studies summarized thus far, when taken together, suggest the existence of persistent performance differences that cannot be explained by differences in industry affiliation, output prices or qualities, or input prices or qualities (where inputs are conceived as the hard physical quantities that typically appear as arguments in a production function, such as capital and labor). The argument for such persistent performance differences could be made more compelling by documenting that the internal workings of firms differ and that these internal differences are systematically related to performance differences. Scores of scholars from a variety of fields (at least sociology, psychology, organizational behavior, and operations management, as well as economics) have contributed to a large literature on the effects of management practices on firm productivity and profitability. In this subsection we review a few of these studies that illuminate the extent to which productivity differences between and within firms can be explained by differences in observable management practices.

In order to establish a systematic link between management practices and firm performance, it is first necessary to demonstrate variation in the use of such management practices. **Bloom and van Reenan (2006)** do just that in a recent paper reporting the results of an establishment-level survey of management practices in 732 manufacturing firms located in four countries. The authors present data on 18 management practices grouped into four categories (operations, monitoring, targets, and incentives) and each firm's implementation of each of these management practices is scored on a scale of 1 to

5. There is substantial within-country variation in management practices: “About 2% of the overall variation in firms’ average management scores is across countries, 42% is across countries by three-digit industry, and the remaining 56% is within country and industry” (p. 36). Furthermore, a firm’s management scores are highly correlated with its total factor productivity, profitability, Tobin’s Q, sales growth and survival rate.

As a complement to cross-sectional data covering multiple management practices in multiple industries in multiple countries, **Ichniowski, Shaw and Prennushi (1997)** collected panel data on both physical output and a detailed set of management practices, which of course necessitated collecting a much smaller sample within a much narrower industry context. In this “insider econometric” study, the authors collected monthly data on eight human-resource practices (such as incentives, training, teamwork, screening, and communication) from 36 finishing lines in 17 steel minimills. For such finishing lines, the performance measure is simple: is the line running or not? In the raw data, ISP find substantial differences in productivity (measured as the percentage of the month that the lines was running) between lines with different bundles of human resource (HR) practices. Ordinary least squares regression confirmed these significant differences among lines with different HR bundles; however, these marginal effects were reduced somewhat by the addition of technology controls, suggesting complementarity between production technology and HR bundles. A fixed-effects model indicated significant improvements in productivity among those establishments implementing enhanced HR bundles during the period of study. It is worth noting however, that the R-squared statistics in both the OLS regressions and the fixed-effects regressions were relatively low, indicating that a substantial amount of inter-line variation in productivity, and changes in productivity among switchers (in the fixed-effects model), could not be explained by this detailed econometric model of a homogenous production process.

In a similar vein, but turning from the workforce to informational technology (IT), **Brynjolfsson and Hitt (1995)** estimated production functions for a large sample of Fortune 500 firms. In their econometric specification, they separated IT capital from non-IT capital. The authors then estimated parameters for Cobb-Douglas and translog production functions, both with and without firm fixed effects. For both types of production functions, they found that the addition of firm fixed effects reduced the marginal productivity of IT capital by roughly 50%. They conclude that firms’ IT investments have significant productivity effects and that the remaining variation in productivity (after accounting for IT investment) is largely explained by unobserved but stable firm effects.

In a second study of R&D productivity in the pharmaceutical industry, **Henderson and Cockburn (1994)** used qualitative research methods to develop measures of the firms’ research strategies and organization. This research was prompted by the finding in their companion 1996 paper that idiosyncratic firm effects could account for a large amount of inter-firm variation in research productivity.¹³ Based on structured interview data and internal company documents, the authors coded the following organizational variables for

¹³ The vagaries of the review and publication process are the reason that the time sequence of the research does not match the time sequence of publication.

each research program in each firm: 1) promotion incentives for scientists to publish research, 2) facilitation of intellectual exchange across firm boundaries, 3) concentration of decision-making over resource allocation, and 4) integration of research conducted in different parts of the world. Adding firm dummy variables to the baseline regression (without the organizational variables) increases the R-squared statistic by 76% (from 0.49 to 0.86).¹⁴ In comparison, adding the four organizational variables to the regression (without firm fixed effects) increases the R-squared statistic by 34% (from 0.49 to 0.66). Adding the organizational variables to the regression that includes the firm dummy variables results in practically no change to the R-squared statistic; the firm dummy variables and two of the four organizational variables remain significant. One interpretation of these results is that the organizational variables employed in these analyses can partly (but not fully) explain idiosyncratic differences in R&D productivity among firms in this sample.

In this section, we have described three empirical studies that document systematic relationships between management practices and firm performance. We, and the authors of these studies, recognize that adoption of these organizational practices is not random. To the contrary, the endogeneity of (and documented complementarity among) these organizational practices naturally limits any claims of causality. Nevertheless, we take three important findings from these carefully conducted studies: 1) existence of substantial variation in internal practices, 2) significant correlation between internal practices and performance, and 3) residual performance differences among firms that cannot be explained by variation in internal practices.

¹⁴ The baseline model includes an intercept, a year dummy for 1978, measures of total firm research (size), measures of scope of firm research, time, and an interaction between time and the 1978 year dummy.

References

- Argote, L., S. Beckman, D. Epple, 1990. "The Persistence and Transfer of Learning in Industrial Settings," *Management Science*, vol. 36, No. 2: 140-154 (February).
- Arthur, J., 1994. "Effects of Human Resource Systems on Manufacturing Performance and Turnover," *The Academy of Management Journal*, Vol. 37, No. 3, (June) pp. 670-687.
- Aw, B. Y., X. Chen, and M. J. Roberts (2001). "Firm-level Evidence on Productivity Differentials and Turnover in Taiwanese Manufacturing," *Journal of Development Economics* 66, 51-86.
- Aw, B. Y., S. Chung, and M. J. Roberts, 2002. Productivity, Output, and Failure: A Comparison of Taiwanese and Korean Manufacturers, National Bureau of Economic Research, working paper 8766.
- Bailey, M., C. Hulten and D. Campbell. "Productivity Dynamics in Manufacturing Plants," *Brookings Papers on Economic Activity, Microeconomics (1992): 187-267*.
- Bernard, A., J. Eaton, J. Jensen, S. Kortum, 2003. "Plants and Productivity in International Trade," *American Economic Review*, Vol. 93, No. 4 (September), pp. 1268-1290.
- Biorn, E., K-G. Lindquist, and T. Skjerpen, 2002, "Heterogeneity in the Returns to Scale," *Journal of Productivity Analysis*, Vol. 18 : 39-57.
- Bloom, N., J. Van Reenen, 2006. "Measuring and Explaining Management Practices Across Firms and Countries," National Bureau of Economic Research Working Paper 12216 (May).
- Brush, T., P. Bromiley, and M. Hendrickx, 1999. "The Relative Influence of Industry and Corporation on Business Segment Performance: An Alternative Estimate," *Strategic Management Journal*, Vol. 20, No. 6 : 519-547 (June).
- Brynjolfsson, E., L. Hitt, 1995. "Information Technology as a Factor of Production: The Role of Differences Among Firms," *Economics of Innovation and New Technology*, Vol. 3, No. 4 (Special Issue on Information Technology and Productivity Paradox), pp. 183-200.
- Chew, W., T. Bresnahan, K. Clark, 1990. "Measurement, Coordination, and Learning in a Multiplant Network," in *Measures for Manufacturing Excellence*, edited by R. Kaplan, Boston: Harvard Business School Press.

Claver, E., J. Molina, and J. Tari, 2002. "Firm and Industry Effects on Firm Profitability: A Spanish Empirical Analysis," *European Management Journal*, Vol. 20, No. 3 : 321-328 (June)

Cubbin, J. and P. Geroski, 1987. "The Empirical Renaissance in Industrial Economics," *The Journal of Industrial Economics*, Vol. 35, No. 4 : 427-442 (June)

Davis, S., J. Haltiwanger, L. Katz, R. Topel, 1991. "Wage dispersion between and within U.S. Manufacturing Plants, 1963-1986," *Brookings Papers on Economic Activity, Microeconomics* vol. 1991 (1991), pp. 115-200.

Disney, R., J. Haskel and Y. Heden, 2003. Restructuring and productivity growth in UK Manufacturing," *The Economic Journal*, Vol. 113 (July), 666-694.

Dunlop, A., D. Weil, 1996. "Diffusion and Performance of Modular Production in the U.S. Apparel Industry," *Industrial Relations*, Vol. 35, No. 3 (July), pp. 334-355.

Eslava, M., J. Haltiwanger, A. Kugler, and M. Kugler, 2004. "The Effects of Structural Reforms on Productivity and Profitability Enhancing Reallocation: evidence from Columbia," *Journal of Development Economics*, Vol. 75 : 333-371.

Foster, L., J. Haltiwanger, and C.J. Kirzen, 2002. "The Link between Aggregate and Micro Productivity Growth: Evidence From Retail Trade," National Bureau of Economic Research, working paper 9120.

Foster, L., J. Haltiwanger, and C. Syverson, 2005. "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability," Working Paper

Griliches, Z. 1986. Productivity, R&D, and Basic Research at the firm level in the 1970s, *American Economic Review*, Vol. 76, No. 1 :141-154.

Griliches, Z. and J. Mairesse, 1981. "Productivity and R&D at the Firm Level," National Bureau of Economic Research, working paper 826.

Griliches, Z. and J. Mairesse, 1982, Comparing Productivity Growth: An Exploration of French and US industrial and firm data, National Bureau of Economic Research, working paper 961.

Griliches, Z. and J. Mairesse, 1985. "R&D and Productivity Growth: Comparing Japanese and U.S. Manufacturing Firms," National Bureau of Economic Research, working paper 1778.

Griliches, Z. and J. Mairesse, 1988. "Heterogeneity in Panel Data: Are There Stable Production Functions?" National Bureau of Economic Research, working paper 2619.

- Griliches, Z. and J. Mairesse (1998), "Production functions: The search for identification", in S. Ström (ed.), *Econometrics and Economic Theory in the 20th Century: The Ragnar Frisch Centennial Symposium*, Cambridge University Press, 169-203.
- Griliches, Z. and H. Regev, 1995. Firm Productivity in Israeli Industry: 1979-1988. *Journal of Econometrics*, Vol 65:175-203.
- Haltiwanger, J., J. Lane, and J. Spletzer, 1999. "Productivity Differences Across Employers: The Roles of Employer Size, Age, and Human Capital," *American Economic Review*, Vol. 89, No. 2 : 94-98.
- Hansen, G. and B. Wernerfeldt, 1989. "Determinants of Firm Performance: The Relative Importance of Economic and Organizational Factors," *Strategic Management Journal*, Vol. 10, No. 5 : 399-411 (Sept-Oct).
- Hawanini, G., V. Subramanian, and P. Verdin, 2003. "Is Performance Driven by Industry or Firm-Specific Factors? A New Look at the Evidence," *Strategic Management Journal*, Vol. 24 : 1-16.
- Henderson, R., I. Cockburn, 1994. "Measuring Competence? Exploring Firm Effects in Pharmaceutical Research," *Strategic Management Journal*, Vol. 15, Special Issue: Competitive Organizational Behavior (winter), pp. 63-84.
- Henderson, R., I. Cockburn, 1996. "Scale, Scope, and Spillovers: The Determinants of Research Productivity in Drug Discovery," *Rand Journal of Economics*, Vol. 27, No. 1 (Spring), pp. 32-59.
- Huckman, R., G. Pisano, 2004. "The Firm Specificity of Individual Performance: Evidence from Cardiac Surgery," *Management Science*, Vol. 52, No. 4 (April), pp. 473-488.
- Hulten, Charles R., 2000. "Total Factor Productivity: A Short Biography," National Bureau of Economic Research Working Paper 7471 (January).
- Ichniowski, C., K. Shaw, G. Prennushi, 1997. "The Effects of Human Resource Practices on Productivity: A Study of Steel Finishing Lines," *American Economic Review*, Vol. 87, No. 3 (June), pp. 29-313.
- Jovanovic, Boyan. "Selection and the Evolution of Industry." *Econometrica*, 50(3), 1982, pp. 649-70.
- Katayama, H., S. Lu, J. Tybout, 2003. "Why Plant-Level Productivity Studies are Often Misleading, and an Alternative Approach to Inference," National Bureau of Economic Research Working Paper 9617 (April).

Klette, T., 1999. "Market Power, Scale Economies and Productivity: Estimates from a Panel of Establishment Data." *Journal of Industrial Economics*, Vol. 47, No. 4 : pp.451-476 (December).

Macher, J., D. Mowrey, 2003. "Managing' Learning by Doing: An Empirical Study in Semiconductor Manufacturing," *Journal of Product Innovation Management*, Vol. 20, pp. 391-410.

Mairesse, J. and Z. Griliches (1990). Heterogeneity in Panel Data: Are There Stable Production Functions. In P. C. et al. (Ed.), *Essays in Honor of Edmond Malinvaud*, Volume 3, pp. 125–147. Cambridge: MIT Press.

Mairesse, J. and J. Jaumandreu, 2005. "Panel data estimates of the production function and the revenue function: What difference does it make?" working paper.

Mauri, A. and M. Michaels, 1998. "Firm and Industry Effects within Strategic Management: An Empirical Examination," *Strategic Management Journal*, Vol. 19, No. 3: 211-219 (March).

McClellan, M., D. Staiger, 1999. "The Quality of Health Care Providers," National Bureau of Research Working Paper 7327 (August).

McGahan, A. and M. Porter, 1997. "How Much does Industry Matter, Really?" *Strategic Management Journal*, Vol. 18: 15-30, (summer special issue: organizational and competitive interactions).

McGahan, A. and M. Porter, 1999. "The Persistence of Shocks to Profitability," *The Review of Economics and Statistics*, Vol. 81, No. 1 : 143-153, (February).

Olley, G. S, and A. Pakes, 1996. "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, November 1996, Vol. 64, No. 6 : 1263-97.

Roquebert, J., R. Phillips, and P. Westfall, 1996. "Markets vs. Management: What Drives Profitability?" *Strategic Management Journal*, Vol. 17, No. 8 : 653-664 (October).

Rumelt, R.P., 1991: 'How Much Does Industry Matter?' *Strategic Management Journal*, Vol. 12 : 167-185.

Schmalensee, R. 1985: 'Do Markets Differ Much?', *American Economic Review*, 75, 1985, 341-351.

Van Biesebroeck, J., 2004. "Robustness of Productivity Estimates," National Bureau of Economic Research working paper 10303.