LEAN WORK

EMPOWERMENT AND EXPLOITATION IN THE GLOBAL AUTO INDUSTRY

Edited by Steve Babson

Wayne State University Press Detroit

9

The International Assembly Plant Study: Philosophical and Methodological Issues

JOHN PAUL MACDUFFIE and FRITS K. PIL

The International Assembly Plant Study was the central research project in MIT's International Motor Vehicle Program (IMVP), and was featured prominently in the IMVP's summary book, *The Machine that Changed the World*. Its findings of large performance differentials across and within the U.S., Europe, and Japan, and its conclusion that the best-performing plants achieved their competitive advantage by following a lean production (as opposed to a mass production) approach have been highly visible and influential, both in the general business press and in the auto industry. Academic scrutiny has also been abundant and varied, ranging from insightful to uninformed. This essay is intended to provide more information about the goals and scope of the International Assembly Plant Study, and to describe the philosophical and methodological approach to performance measurement we have taken. In doing so, we also respond to critiques of the study raised in a number of settings, including the 1993 Lean Production and Labor conference in Detroit.[1]

The International Assembly Plant Study (IAPS) started on a small scale but grew over time to become an international project of tremendous scope. It began in 1986 when John Krafcik at MIT undertook a careful comparison of productivity differences among four plants—NUMMI (the

GM-Toyota joint venture), GM-Fremont (the closed plant that became the facility for NUMMI), GM-Framingham, and Toyota-Takaoka—using a methodology to correct for differences in plant characteristics. Eighteen months later, Krafcik's master's thesis at MIT reported performance differentials for a sample of thirty-eight plants and investigated some early hypotheses about the determinants of performance. From 1988 to 1990, John Paul MacDuffie joined Krafcik in expanding the sample to seventy plants from twenty-four companies and sixteen countries, and in developing a longer survey to collect more detailed data on such factors as level and type of technology, product mix complexity, manufacturing policies, work organization, and human resource policies. The much-publicized results on assembly plant performance are based on analyses of 1989–1990 data from this large sample.[2]

A second round of data is being gathered in 1993–1994 by MacDuffie, now at Wharton, and Frits Pil. These new data, from a sample that will be larger than the first round, will permit us to capture the dynamic aspects of change over time in both performance and production system characteristics. In this round, we have also had the opportunity to expand greatly the set of topics and issues researched, as well as to gather more in-depth data on issues studied in the last round. Some observers have failed to recognize the evolution of the IAPS over the past eight years, and persist in using the early stages of the research (back to 1986) as their reference point for the entire project.

The IAPS has a narrow but deep focus—it examines only assembly plants, but collects data on all aspects of plant operations, ranging from measures of technology and product complexity to measures of manufacturing policies and human resource practices; in the new round, this list expands to include supplier relations, design factors, and accounting systems. In this paper, we will primarily emphasize how we assess performance. Most coverage of the IAPS focuses on only one of the performance measures we use—labor productivity, as measured by the standardized comparison of labor hours required per car—so the bulk of this essay will address the philosophical and methodological issues associated with this measure.

However, from the start the IAPS has emphasized the importance of using multiple performance indicators as a way of eliminating biases that might result from using just one such measure. For example, we believe it can be very misleading to examine labor hours without also considering the quality of vehicles that are built. Thus, we will explain briefly the other performance measures we are using. Finally, we will touch on our measurement strategy for the independent variables—technology, manufacturing policies and human resource practices, product complexity—that help ex-

plain performance. After describing the IAPS methodology in all of these areas, we will address various critiques of the study.[3]

Performance Measures

All of the IAPS performance measures are plant-level measures. In the first round, we had two key indicators of plant performance: labor productivity and customer-perceived vehicle quality. In the second round, we added two other performance measures: first-time-through capabilities of different departments in the plant, and various indicators of employee well-being and satisfaction. A measure of environmental performance may also be developed.

Labor productivity. We measure labor hours per vehicle, standardizing for vertical integration, product size, and option content, and some design factors, and adjusting for actual work time and absenteeism. The productivity methodology is described in greater detail below.

Quality. The importance of quality as a factor affecting the vehicle purchases of U.S. consumers became evident during the competitive struggles between U.S. and Japanese producers in the 1980s. Vehicle quality is the outcome of many factors, including the quality of components received from suppliers and the design of the product. However, the assembly plant has a tremendous impact on a customer's perception of quality and overall satisfaction, particularly in "fit and finish" (the alignment of body parts, the appearance of the paint job, the care taken with assembly tasks, and the effective performance of all components), areas in which Japanese competitors appear to have a distinct advantage over U.S. companies.

To assess the impact of the assembly plant on quality, we needed a measure that could differentiate those problems the plant could control from those that originated outside of the assembly plant's domain, such as certain design or supplier-related problems. With the generous support of J. D. Power and Associates, we were able to develop such a measure. Every fall, J. D. Power surveys new car owners based on a random sampling methodology of new car registrations in the United States. All customers are surveyed at the same time, and at the survey date they have all owned their cars for three months. They are asked very detailed questions about their early experience with their new cars, with over one hundred different problem categories from which to choose.[4]

We start with these data, aggregated across all vehicles built at a plant, and construct a "defects per one hundred vehicles" measure that includes

only those problems that are under the direct control of the assembly plant—for example, problems related to body finish, paint finish, squeaks and rattles, water leaks, and some electrical problems. Other problems related to the transmission, engine, steering and handling, and some electrical components (e.g., difficulty tuning the radio) that are attributable to vehicle design or component suppliers are excluded from our measure.

There is a subset of plants that do not sell vehicles in the United States. For these, we use correlation mapping with internal corporate data and non-U.S. market data to derive a measure of quality that is analogous to our J. D. Power-based measure.

First-time-through capabilities. In addition to the customer-based measure of quality, we want to know about a plant's ability to build high-quality products without extensive rework before a vehicle is shipped to the final customer. For that purpose, we now collect information on first-time-through capabilities for the body, paint, and assembly areas of a plant. First-time-through capability refers to the percentage of cars that are built right the first time around—cars that require no rework once the car leaves the main line.

Employee well-being and satisfaction. Here we emphasize plant-level measures that provide some indication about the attitudes and experiences of production employees. In the first round of IAPS, we measured labor turnover and absenteeism. In the second round, we added measures of injury rates, including the number of incidents and the severity in terms of days of work lost.

Environmental performance. A separate project at MIT seeks to evaluate the environmental performance of assembly plants in terms of energy usage, emissions from the paint plant, and the generation and processing of toxic wastes. Data from this project will be linked to our database when the environmental impact assessment is complete.

The Methodological Strategy for Measuring Productivity

The original elements of the productivity measure were developed by John Krafcik. To ensure comparability over time, we have retained Krafcik's methodology, with some fine-tuning. Krafcik's aim was to make an "apples to apples" comparison among assembly plants by adjusting for, or measuring the influence of, factors affecting productivity that vary across plants. The measurement strategy associated with this goal had several key components: (1) develop independent measures rather than asking

for data from company records; (2) focus on labor productivity; (3) emphasize physical rather than financial measures of productivity; and (4) develop a productivity measure that adjusts for factors independent of firm choices about the plant's production system, for example, degree of vertical integration, product size, and complexity.[5]

Independent Measures of Productivity. While the use of independent measures of assembly plant productivity requires both the development of a methodology and the labor-intensive task of data collection, it is clearly the only way to achieve the desired goal of an "apples-to-apples" comparison. Companies develop their own extensive measures of plant productivity, but these are idiosyncratic. We have looked at data like these from many companies and found that there is no way to integrate them to permit systematic and accurate comparisons across companies. Even financial-based measures of performance, which would appear to be most directly comparable, vary greatly due to differences in internal accounting systems. This is the case even for companies located within the same country. Comparisons of financial data across companies are further complicated by the problems of interest-rate differentials, depreciation practices, and exchange-rate choices. Finally, companies are generally unwilling to release detailed plant-level financial data.

Publicly available information has different problems. For example, Harbour and Associates publishes a yearly report on assembly plant productivity that uses public information on the number of vehicles produced and the number of employees at a plant to develop a simple "hours per vehicle" measure. These data are available for all plants, can be obtained without company permission, and can be tracked over time. However, they in no way allow for true comparability across plants, since not even the most rudimentary adjustments for product differences or level of vertical integration are made.

Labor Productivity Focus. One common observation about many advanced manufacturing industries is that labor costs, as a percentage of total costs, have dropped dramatically over time, primarily due to automation. While this is generally true of auto assembly, particularly in the highly automated welding and painting departments, it is also true that final assembly of automobiles remains one of the most labor-intensive tasks of any advanced manufacturing setting. Furthermore, while direct labor costs do shrink as automation increases, it is less clear whether indirect (or overhead) labor costs decrease; some argue that they increase. Thus labor productivity, particularly if broadened as in the IAPS methodology to include direct, indirect, and salaried employees, is clearly a critical measure of assembly plant performance.[6]

While it would be advantageous to use a broader measure of productiv-

ity—for example, Total Factor Productivity, encompassing the full range of inputs to the production process (capital, materials, and energy as well as labor)—it is hard to get these data and impossible to ensure their comparability across countries. Furthermore, recent studies of company-level productivity differences in the automotive industry, comparing U.S. and Japanese companies, have found tremendous variation in labor productivity but nearly equivalent levels of capital productivity. Explaining the labor productivity variation becomes the intriguing research question. Finally, labor productivity is by far the most relevant productivity measure for a study that examines how human and technical capabilities are organized.[7]

Physical measures of labor productivity. As noted above, labor productivity is defined in this study as the hours of actual working effort required to build a vehicle at a given assembly plant. Thus it focuses on the physical conversion of labor inputs into outputs, and does not tell us directly about differences in the cost structure of plants. However, by focusing on effort rather than cost, the productivity measure is not affected by the problems with financial data noted above, nor by wage differentials or differences in national employment policies, which might influence a labor cost comparison.

For similar reasons, the "hours per vehicle" calculation is adjusted for absenteeism in order to exclude those people added to the payroll to cover for absent employees. This adjustment is warranted because absenteeism may be influenced by the type and availability of social benefits covering various absences as well as norms and customs specific to a particular region. Thus, only the number of working hours of people who actually build cars on a given day are included for the productivity calculation.[8]

Adjustments to Productivity. In order to develop a productivity measure that is consistent across plants and that captures the true differences in capabilities of the plants (rather than idiosyncratic differences resulting from factors outside of the plant's control), we make several adjustments. As described above, we extract absenteeism from our labor hours per vehicle figures. The other primary adjustments to the productivity calculation are for vertical integration and product differences. Both of these factors are largely independent of choices made about the plant's production system.

To adjust for differing levels of vertical integration at plants, the productivity methodology considers only a set of "standard activities" that are common to virtually every plant in the world. Some plants make their own body stampings, while many more receive stampings from a supplier plant. Therefore, stamping is not included as a standard activity. Many subassemblies, typically shipped to the assembly plant in completed form (e.g., seats, wire harnesses, fuel pumps) are also excluded from the stan-

dard activities. We further exclude activities related to the production of knock-down kits and other components intended for use at other plants.

Adjustments for product differences include size, option content, and product manufacturability. Since a large vehicle requires significantly more effort to assemble than a small vehicle, adjustments are made to a standard vehicle size. Likewise, installing options requires time. Some options are extremely time intensive (e.g., air bags, sunroofs, etc.), and plants that produce vehicles with high labor-intensive option content suffer a productivity handicap. Therefore, we adjust to a standard option content.

The design of a product can certainly affect how efficiently a car can be built. We adjust for the manufacturability of the design in the welding and paint departments. We do not make adjustments to productivity based on the impact of design in the assembly area (other than product size and option content corrections), because we have yet to find a good measure for this. The measure reported in *The Machine that Changed the World* was based on corporate reputations for manufacturability, as perceived by a small sample of respondents—hardly valid as data about a particular product at a particular plant.[9]

Ideally, a product manufacturability measure would be based on careful "teardown" data using a consistent methodology. We are exploring the possibilities of gaining access to such data in the second round. An alternative we are also pursuing is to measure aspects of the interaction between the assembly plant and product engineers during the product development process. The assumption is that manufacturability will be better the earlier the assembly plant sees the blueprints for a new model, the earlier plant personnel (from engineers to production workers) are involved in design decisions about prototypes, the more suggestions about manufacturability are communicated from the plant to designers, and the more the plant is involved in the building of pilot vehicles.

In lieu of the ideal measure of manufacturability, we study the effect of design on productivity by using product design age as a proxy variable. The strengths and weaknesses of this measure are described below.

Measuring Factors that Affect Performance

In addition to developing performance measures and collecting the necessary data, we also investigate the influence on performance of factors affected by firm choices about the plant production system, such as the

level and type of technology, product mix, manufacturing policies, work organization, and human resource policies.

These are treated as independent variables, rather than being used to adjust the productivity measure directly. We saw an advantage in minimizing the number of direct adjustments made to the productivity calculation, since once such an adjustment is made it cannot be investigated through multivariate analyses of the determinants of performance. For example, we collect detailed information on the automation in a plant. With respect to flexible automation, we look at the number of robots by controller type, axes of motion, location in production process, and primary function within that process. We could use this measure to adjust the productivity measure, so that "hours per vehicle" would reflect the same level of robotics. Instead, we find it far more valuable to examine the impact on productivity of the use of robotics at different plants, given the potential variation in the technology strategy underlying capital investment in robotics and the effectiveness with which new technologies are implemented.

The "design age" variable is another good example. Product design age is defined as the weighted average number of years since a major model introduction for each of the products built at a given plant. We do not use the variable to adjust the productivity measure for two reasons. First, there are competing hypotheses about the impact design age should have on productivity. One hypothesis depends on a "learning curve" argument—that the longer you build a particular design, the more you learn and the better your productivity. The alternative hypothesis argues that newer product designs are more likely to have been conceived with ease of assembly in mind than older products, and thus could be produced more efficiently. We wanted to discover which of these effects on productivity was dominant.

Second, we recognized that design age, by itself, would tell us little about manufacturability across companies. There is no reason to expect that two products from different companies, each with three-year-old designs, would be similar in manufacturability. However, looking across the sample, those companies that have a rapid product development cycle (of four years, as in Japan, versus six to eight years, as in Europe) will have, on average, younger designs. If one assumes that a rapid product development cycle requires a great deal of concurrent engineering, during which the design and manufacturing functions interact very intensively, then a younger design age may well be associated with more manufacturable designs.[10]

In addition to focusing on factors affected by firm choices about the plants' production system, the development of survey questions about independent variables (predictors of performance) followed two additional guidelines: (1) Only measure policies and practices that can potentially be

implemented in any plant in the international sample; that is, avoid measuring things exclusively associated with one company or country. (2) Draw upon fieldwork observations to choose which policies and practices to measure, emphasizing those that help differentiate best among different models of production organization. Let us look at each of these guidelines in more detail.

Measure practices that are potentially universal. Our thinking with respect to the measurement of company policies and practices was twofold. First, we customized our questions as much as possible to the assembly plant context. This reduced the possibility of confusion, particularly with an international set of respondents, and increased the reliability of responses. Second, we avoided measuring policies and practices that were unique or idiosyncratic to a particular nation or company, and would therefore function as a dummy variable for that nation or country. For example, with respect to Japanese employment practices, we did measure the use of teams, quality circles, and job rotation, since these are commonly found in other countries as well as Japan. We did not ask questions about such practices as the *nenko* wage system, the *satei* personnel evaluation system, enterprise unions, or lifetime employment—features that are more specific to the Japanese context.

Differentiating among models of production organization. From our fieldwork, we were aware of important differences in the model of production organization used in different plants. We wanted to measure those policies and practices that captured this variation in the sample. For example, the level of inventory—both of parts and of work in process—was an important indicator of different philosophies regarding the role of buffers in the production system, so we included several questions related to inventory. The use of "on-line" work teams and "off-line" problem-solving groups was another important differentiating factor, revealing different views about the role of worker involvement in the production process. Conversely, we did not include questions where there would be little or no variation within the sample. For example, all of the plants in our sample had a moving assembly line, so it made no sense to investigate variation for this variable.

Critiques of the Study

Because of the high visibility of the IMVP project, and particularly of the findings about assembly plant performance differentials, the Interna-

tional Assembly Plant Study has garnered a lot of attention in recent years. Several critiques have emerged, each of which falls into one of the following categories: (1) misunderstandings about what was measured and what adjustments were made; (2) disagreements about whether the measurements and adjustments were adequate to achieve true comparability across plants; and (3) arguments that any effort to measure plant performance (or a plant's production system) are doomed to fail because such measurements can never be accurate. Earlier sections of this essay will serve, hopefully, to correct simple misunderstandings. Here, we first address disagreements about what is measured and the adjustments that are made.

Critique: By focusing on "hours per vehicle," the study essentially blames direct labor workers for poor plant performance, neglecting the role of management decisions at the corporate level that affect capital investment, what products are built in a plant, and what organizational policies are followed.

Answer: Our measure of "hours per vehicle" includes not only direct labor hours but also indirect and salaried labor hours, so it reflects the inefficiency (or efficiency) of all employees in a plant. The adjustments made in the productivity methodology provide for an "apples to apples" comparison for many factors determined at the corporate level, such as the level of vertical integration, product size, and option content. The independent variables we measure almost entirely reflect management decisions, from the level of capital investment to what manufacturing or human resource policies to follow. How direct labor employees are managed in the context of the overall production system is the important factor for plant performance, not necessarily the personal characteristics of those employees.

Critique: The study doesn't adjust adequately for capacity utilization. A plant that is operating under capacity will obviously perform more poorly than one at full capacity, while plants doing overtime to run over capacity will appear to have superior performance.

Answer: We adjust for capacity utilization in two ways. First, for plants in an over-capacity situation, we exclude all overtime hours from the calculation of hours worked. We also exclude any temporary workers that may be brought on to reach higher production levels. Thus, over-capacity production has no effect on the productivity calculation. Second, for under-capacity situations, we look to see whether the downturn in production is short-term or long-term. If it is short-term, we ask the plant to give us data from their last period of regular (i.e., steady state) production. If the plant has operated well under its rated capacity for a long time, we take its actual production level as its de facto capacity, under the assumption that the plant has had the opportunity to adjust labor inputs downward to match the

wer production level. This assumption is plausible with respect to labor oductivity, while it would not apply if we were studying capital produc-ity.

Critique: The study doesn't adjust adequately for the difference be-een mass market and luxury vehicles. Luxury vehicles are so complex, ey must take much longer to assemble.

Answer: Luxury vehicles benefit from two adjustments to the produc-ity calculation. First, luxury vehicles are generally bigger, so the adjust-ent of the Product Size Factor reduces their hours relative to plants with aller cars. Second, luxury vehicles have many more options than mass arket vehicles. The Option Content adjustment we make focuses on elve different options, many of which appear only in more expensive rs: for example, power windows, doors, seats, sun roofs, and anti-lock akes. The combination of these two adjustments can reduce hours per car a luxury plant by over 20 percent.

Critique: The study doesn't adjust adequately for design factors. Plants th well-designed products that are easy to assemble will have a signifi-nt performance advantage over those with poorly-designed products.

Answer: As noted above, we make adjustments related to design fac-s in the weld and paint areas as part of the productivity methodology, d also look at product design age, which is likely to be moderately corre-ed with manufacturability. Ideally, one would like to have detailed tear-wn data on all products in the world, evaluated for "ease/difficulty of sembly" with a consistent methodology. We do not have access to such ta, and therefore cannot adjust adequately for the influence of design ferences on productivity in final assembly. However, we believe that the propriate response to this measurement problem is to look for indicators interaction between manufacturing and product designers during the oduct development process. How many months before product launch are gineers, managers, and workers from an assembly plant given a chance to ok at product blueprints? How many months before launch are they con-ted on the design of process equipment? Are prototype and/or pilot vehi-s built in the assembly plant which will actually produce them? The erative hypothesis here is that earlier and more intense involvement of rsonnel from the assembly plant in the design process will be associated th more manufacturable designs. We are gathering data of this kind in second round of the assembly plant study.

Critique: Plants that have multiple body shops or assembly lines are e to handle a complex product mix more easily than plants that must x platforms and models on a single line. This would make them appear re productive.

Answer: The measure of Model Mix Complexity we use adjusts for

the number of body shops and assembly lines. For example, a plant making two different platforms on two assembly lines would have the same Model Mix Complexity as a plant making one platform on one assembly line (or three platforms on three assembly lines.) With this measure, we can examine the impact of true model mix complexity (i.e., multiple platforms and models on the same line) on productivity and quality.

Critique: The organizational practices measured by the study are simply proxy variables for particular countries. For example, plants that use teams are obviously Japanese plants, and so the "teams" variable functions as a dummy variable for Japan. What this variable picks up, therefore, has nothing to do with teams, and everything to do with unique aspects of Japan (e.g., low cost of capital, work ethic of Japanese workers, the multi-tier supplier system).

Answer: As noted above, we took great care to avoid measures that would simply vary at the national level. Work teams may be commonly used in Japan, but they have been found in U.S. and European assembly plants for more than a decade. The same holds true for every manufacturing and human resource practice we measure. We apply this policy as much to Europe and the U.S. as to Japan. For example, we do not measure the use of works councils (applicable only in Europe) or the number of three-step grievances (applicable only in the U.S.).

Critique: Plants will provide erroneous information in the survey to make their performance look good.

Answer: This risk exists for any study. However, we have several reasons for being confident that there is minimum risk plants would do this deliberately, and that we catch most or all problems resulting from misunderstandings, poor translations, and so forth.

First, we keep all plant names completely confidential, and do not write reports or prepare graphs that allow any company or plant to be identified. This is essential for us to receive data from the auto companies—data that is often viewed as highly sensitive or even proprietary. However, it also means that a plant will not gain any public relations benefit from misrepresenting their performance to us. More importantly, plants provide us with data because, in return, we visit each participating plant, no matter where it is located, and provide in-depth feedback. This includes meetings with managers and union representatives to provide a detailed analysis of how the plant compares to others in its region, and other regions of the world on a series of practices and philosophies. Gathering the data we require is extremely time-intensive, and the plants view it as an opportunity for learning. Providing phony data would weaken the usefulness of the feedback we provide them—both because their performance would be

mischaracterized, and also because the regional averages that we give them for comparison purposes would be biased.

We do extensive follow-up research with plants to insure that the data are accurate, including three or four pages of questions after receipt of the survey, and extensive phone and fax exchanges. During our plant visits, we also do on-site data verification of a set of key responses. Lastly, we have a systematic method for analyzing responses for internal consistency within the plant, among plants of the same company, and for the same plant over time. This allows us to pinpoint questionable responses and ensure they are corrected.

The broadest critique of the assembly plant study is that it is simply impossible to ever achieve "apples to apples" comparisons across plants, so there is no point in trying. Our immediate reaction is that such a view is self-defeating. Naturally, nothing can be learned when one starts with an a priori conclusion that nothing can be learned.[11]

Furthermore, those who take such a stand do not oppose all measurement. Indeed, they offer other data, generally at a company or industry level, in order to refute claims made by the assembly plant study. They assume, for example, that performance at the plant level can be inferred with greater accuracy from publicly-available data at an industry level than from a careful plant-level comparison using independent measures. They argue that "vehicle units produced" is an appropriate measure of output, regardless of whether those units represent trucks, minivans, or cars. They assume that company or even national aggregated information is a more valid way to assess the effects of organizational philosophies, technological policies, and human resource practices on productivity than using precise plant-level data.

These assertions and assumptions are unjustifiable, we believe. There is tremendous variation in the manufacturing requirements for different cars, and even greater differences between cars and trucks. (Trucks are actually much simpler to build, controlling for size.) Using national-level data to draw conclusions about plant-level performance assumes homogeneity across plants in the same country. But we have found huge variation in both performance and in production system practices across companies in the same country or region, as well as across plants in the same company. These variations are among the most interesting phenomena in the entire study because they allow one to assess the impact of different plant-level practices and philosophies holding constant either national or company-level characteristics.

A second response is that the performance differentials identified across and within regions during the assembly plant study are very large. Recall these regional averages in the 1989—1990 data for labor hours per

vehicle: Japanese plants in Japan, 16.8 hours; U.S. plants in North America, 24.9 hours; European-owned plants in Europe, 35.3 hours. Suppose that the assembly plant productivity calculations are off by 10 to 20 percent (we should note that we don't believe they are). Even if this hypothetical measurement error were to favor plants from some regions consistently and hurt plants from other regions, the performance gap across regions would still be large. The same point applies to performance differences within regions. The best Japanese plant in Japan, at 13.2 hours in 1989, was almost twice as productive as the worst Japanese plant, at 25.9 hours. These differentials would not disappear due to measurement error.

We are convinced that the issues addressed by the assembly plant study are too important to be dismissed as "impossible to study." Many valuable insights emerged from the first-round data and these have had a profound impact on the thinking of many managers, union officials, workers, and customers. These insights would not have had much impact, and would not be accorded much legitimacy, if they were contradicted by other similar (i.e., plant-level) data or, more significantly, by the personal experience of people actually working in the auto industry. Many of the plants identified in the study as inefficient by world standards have been able to identify a considerable waste of time, people, equipment, and materials in their operations. While there may be legitimate disagreements about the best ways to improve performance, we believe there has been convergence on the idea that choices about how to organize the production system are critical—more critical than heavy investments in high-tech equipment or better product designs or a better-educated workforce.

We hope that we will be able to continue the International Assembly Plant Study beyond the current round so that it can provide increasingly greater insights into the dynamics of changing production systems and performance improvement over time. We will continue to work to improve the methodologies we use while maintaining comparability with previous data. We look forward to continued valuable feedback and suggestions from all sources, but particularly from those willing to familiarize themselves with the philosophy, research strategy, and methodological choices that have characterized the International Assembly Plant Study throughout its eight-year history.

Notes

1. James Womack, Daniel Jones, and Daniel Roos, *The Machine that Changed the World* (New York: Rawson Associates, 1990).

2. John Krafcik, "Learning from NUMMI," working paper (IMVP, MIT, 1986); John Krafcik, "Comparative Analysis of Performance Indicators at World Auto Assembly Plants" (master's thesis, Sloan School of Management, MIT, 1988); John Krafcik and John Paul MacDuffie, "Explaining High Performance Manufacturing: The International Assembly Plant Study," working paper (IMVP, MIT, 1989); John Paul MacDuffie and John Krafcik, "Integrating Technology and Human Resources for High Performance Manufacturing: Evidence from the World Automobile Industry," in *Transforming Organizations*, ed. Thomas Kochan and Michael Useem (New York: Oxford University Press, 1992).

3. More extensive discussion of these variables can be found in other papers: for measures of technology, see John Krafcik, "A Comparative Analysis of Assembly Plant Automation," working paper (IMVP, MIT, 1989), and John Paul MacDuffie, "Beyond Mass Production: Flexible Production Systems and Manufacturing Performance in the World Auto Industry" (Ph.D. diss., Sloan School of Management, MIT, 1991). For measures of manufacturing policies and human resource practices, see John Paul MacDuffie, "Human Resource Bundles and Manufacturing Performance: Flexible Production Systems in the World Auto Industry," unpublished, 1994; for measures of product complexity, see John Paul MacDuffie, Kannan Sethuraman, and Marshall L. Fisher, "Product Variety and Manufacturing Performance: Evidence from the International Automotive Assembly Plant Study," unpublished, 1994.

4. The J. D. Power and Associates Initial Quality Survey sample usually includes between one hundred and four hundred cars produced in each plant that sells products in the United States. Results are strictly confidential.

5. See Krafcik, "Comparative Analysis of Performance Indicators," and John Krafcik, "A Methodology for Assembly Plant Performance Determination," working paper (IMVP, MIT, 1988).

6. Chris Ittner and John Paul MacDuffie, "Exploring the Sources of International Differences in Manufacturing Overhead," manuscript (Wharton School, University of Pennsylvania, 1994); Jeffrey Miller and Thomas Vollmann, "The Hidden Factory," *Harvard Business Review* 63 (Sept.-Oct. 1985): 142–150. A related point is that the assembly plant accounts for only a portion of the total value added in the production of an automobile, given the high cost of both design and components. This critique is not relevant to the assembly plant study per se, which has never claimed to measure performance beyond the plant level. However, we acknowledge that excellent productivity at the assembly plant level won't translate into company-level profitability unless other key pieces in the value chain are equally efficient, pricing is appropriate in relation to the company's cost structure, and customers want to buy the company's products.

7. On Total Factor Productivity, see Robert Hayes and Kim Clark, "Exploring the Sources of Productivity Differences at the Factory Level," in *The Uneasy Alliance: Managing the Productivity-Technology Dilemma*, ed. Kim Clark, Robert Hayes, and Christopher Lorenz (Boston: Harvard Business School Press, 1985). On capital productivity and other factors affecting plant performance, see Michael Cusumano, *The Japanese Automobile Industry: Technology and Management at Nissan and Toyota* (Cambridge: Harvard University Press, 1985), and Marvin Lieberman, Lawrence Lau, and Mark Williams, "Firm-Level Productivity and Management Influence: A Comparison of U.S. and Japanese Automobile Producers," *Management Science* 36 (Oct. 1990): 1193–1215.

8. Since daily working hours will vary based on union contract and national norms, the methodology also adjusts total labor input for a standardized work day.

9. Specifically, to compensate for manufacturability of the design in welding and paint departments, adjustments are made for the number of welds in the design, which affects direct and indirect headcount in the welding department (depending on how automated the weld

process is) and for the amount of joint sealer that must be applied in the paint department, which affects direct labor headcount for this still labor-intensive task. In other words, plants with more welds (spot and arc) or joint sealer than the sample average receive a credit, in the form of a reduction in their hours per vehicle, while plants with fewer welds or joint sealer have their hours per vehicle adjusted up.

10. Kim Clark and Takahiro Fujimoto, *Product Development Performance: Strategy, Organization and Management in the World Auto Industry* (Boston: Harvard Business School Press, 1991).

11. Karel Williams, Colin Haslam, John Williams, Tony Cutler, Andy Adcroft, and Sukhdev Johal, "Against Lean Production," *Economy and Society* (Aug. 1992): 321–354.