

NBER WORKING PAPER SERIES

COMPLEMENTARITIES IN AUTOMOBILE PRODUCTION

Johannes Van Biesebroeck

Working Paper 12131

<http://www.nber.org/papers/w12131>

NATIONAL BUREAU OF ECONOMIC RESEARCH

1050 Massachusetts Avenue

Cambridge, MA 02138

March 2006

I would like to thank two anonymous referees for extremely useful suggestions and the participants at the EC2 conference in Marseille, the 19th UBC Summer Conference on IO, the 2004 IIOC and CEA meetings for comments, in particular Brian Viard and Frank Wolak. Financial support by SSHRC under the New Economy program and by a CFI/OIT New Opportunities grant is gratefully acknowledged. Contact: University of Toronto, IPA Suite 707, 140 St. George Street, Toronto, ON M5S 3G6, Canada. Tel.: 416-946-5795; Fax: 416-978-5519; E-mail: [johannes.vanbiesebroek@utoronto.ca](mailto:johannes.vanbiesebroek@utoronto.ca). The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

©2006 by Johannes Van Biesebroeck. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Complementarities in Automobile Production

Johannes Van Biesebroeck

NBER Working Paper No. 12131

March 2006

JEL No. L23, L11, L62, M11, C33

**ABSTRACT**

The number of different car and light truck models produced in North America has increased enormously over the last decades. The data suggests that producing this increased variety of vehicles is associated with a productivity penalty. We show that manufacturers can adopt complementary activities to reduce this penalty. Flexible technology, defined as the ability to assemble models derived from different “platforms” on the same assembly line, and bringing previously outsourced activities in-house are two such activities that we identify. Both are costly themselves, in terms of lower productivity, but they reduce the cost of producing greater variety. The results are robust to controlling for the endogeneity of the adoption decisions using activity-specific instruments, as proposed by Athey and Stern (2003).

Johannes Van Biesebroeck

Department of Economics

University of Toronto

Institute for Policy Analysis, Suite 707

140 St. George Street

Toronto, ON M5S 3G6

CANADA

and NBER

[johannes.vanbiesebroeck@utoronto.ca](mailto:johannes.vanbiesebroeck@utoronto.ca)

# 1 Introduction

The organizational design of firms is likely to affect performance in crucial ways. In the automotive industry there is a lot of scope for such organizational decisions to matter. By the very nature of the product, coordination problems are crucial. Each vehicle is made up of a myriad of components which cannot all be produced by the same firm. The end product is differentiated in many dimensions. The minimum efficient scale of production is large, possibly changing over time, and differs for many activities that make up the assembly process. Moreover, the industry is highly capital intensive and unionized which is likely to make adjusting to a new environment more difficult.

A growing literature has focused on the presence of complementarities, defined as a positive dependency between pairs of activities. One of the most influential studies of complementarities, Milgrom and Roberts (1990), referred extensively to the automotive industry.<sup>1</sup> Data limitations and the importance of unobservables make it a nontrivial task to distinguish between interdependent returns and the impact of an omitted variable on several adoption decisions, see Athey and Stern (2003). The distinction matters crucially to isolate the impact of an activity holding all else constant. In particular, by studying the adoption decisions of other firms, we would like to learn what return a firm can expect from adopting a new activity itself.

We specifically look at the impact of an expanded model lineup on plant-level productivity. The variety of models for sale (and produced) in North America has increased tremendously in the last decades. While the increase could to some extent be the response to innovations in production technology, it is generally viewed as a response to changing consumer demand and the cause of severe production difficulties. While some firms may have a comparative advantage in producing a greater variety of vehicles at minimal extra cost, other (less capable) firms follow their lead to preserve market share. In an industry with high fixed costs this is not necessarily irrational, but these firms will potentially incur much greater cost increases or productivity declines. The following quote illustrates the industry's sentiment:

“All that model proliferation, of course, increased costs. [...] Today, through

---

<sup>1</sup>Macduffie (1995) and Helper (1997) are empirical studies investigating complementarities between human resource activities in the automotive industry. Novak and Stern (2003) study complementarities between the outsourcing decisions of major components. Jovanovic and Stolyarov (1997) illustrate a theory of asynchronous use of complementarity technologies under learning by doing, drawing on evidence from the automotive industry.

bitter lessons drenched in red ink, automakers are learning to consolidate product development and even production, while producing vehicles that are different enough in style to suit finicky consumers. [...] Today we have insufficient volume on products, so we're not taking advantage of the economies of scale that we could." *Automotive News* (1996)

Increasing the number of models produced in a single plant is the first 'activity' we consider. Flexible technology and insourcing (the reverse of outsourcing) are two 'activities' that have the potential to reduce the cost of assembling the increased variety. Both are widely debated trends in the automotive and many other manufacturing industries.<sup>2</sup> Note that the potential complementarities are of an unusual form. Rather than joint adoption boosting the positive return of adoption in isolation, in our application each activity would lower productivity on its own, but adopting them jointly potentially reduces the productivity penalty. We do not model the implicit benefits of producing greater variety, but refer to Petrin (2002) for an illustration of the (demand-side) benefits of the introduction of the minivan.

We will investigate the potential complementarities between these three activities on the universe of all car and light truck assembly plants in North America over the 1994–2004 period. The empirical literature has used two approaches to study the potential interdependence of activities. Some studies look directly at the clustering of adoption decisions, while others look at the joint impact of activities on a performance variable, like costs or productivity.<sup>3</sup> As pointed out by Athey and Stern (2003), many empirical studies finding evidence for complementarities do not allow for unobservables that might be correlated with the adoption decision of each activity.

In some cases, one can convincingly claim that observable characteristics absorb the majority of firm-specific effects that determine adoption. This is arguably the case in the setting of Ichniowski et al. (1997). The performance measure is narrow (output of a single production line); production technology is similar across observations (homogenous steel finishing); fixed-effects are included; and differences in adoption are likely to be driven by variations in the fixed costs of adoption (for example knowledge about best practice or trust between labor and management varies by firm). While automobile assembly shares some similarities with the previous situation, the production process is a lot more

---

<sup>2</sup>For flexibility see Gal-or (2002) or Norman (2002) and for outsourcing see Girma and Gorg (2004).

<sup>3</sup>Examples of the first approach are Arora and Gambardella (1990), Novak and Stern (2003), Miravete and Pernías (2005), and Åstebro, Colombo, and Seri (2005). Ichniowski, Shaw, and Prenzushi (1997), Macduffie (1995), and Helper (1997) are examples of the second approach.

complicated. Even though the general technology is similar across observations, firms (plants) are likely to be heterogeneous in terms of their ability to implement and adjust to innovative activities. Given that firms vary widely in their observed productivity levels, we cannot rule out that unobserved differences in capabilities are important.<sup>4</sup>

An alternative approach is to draw inferences from the restrictions that complementarities place on the distribution of the error terms. Miravete and Pernías (2005) develop a maximum likelihood estimator that explicitly incorporates correlation in the unobservables that enter the adoption decision of each activity. Functional form assumptions on the return function and the distribution of unobservables allow them to derive different areas of support consistent with joint adoption under the existence or absence of complementarities.<sup>5</sup>

Finally, Athey and Stern (2003) demonstrate that one can construct a true test for complementarities using *activity-specific* instruments, “i.e. observable factors impacting the adoption of individual practices but independent of other practices *and* measured performance. (p. 3)” Novak and Stern (2003) follow this approach and derive instruments for the outsourcing decision of major components in a sample of luxury cars. They argue that information on pre-existing in-house sunk investments or a limited in-house capacity for system  $A$  are valid instruments for the outsourcing decision on system  $A$  in an equation that explains outsourcing of some other system  $B$ .

We follow this last approach to deal with potentially correlated unobservables determining the different adoption decisions. For model variety, we adopt the logic used in differentiated product demand estimation and use the number of varieties produced by competing firms as *activity-specific* instrument. Once we control for the market segment of the plant’s output, this variable has no role anymore explaining productivity. The distance from the geographical center of the industry is a negative predictor for the degree of outsourcing and it should not belong in the productivity regression after controlling for country-fixed effects. For flexible technology, we use the method of shift relief as instrument as it reflects labor-management relations that influence the expected success of flexible production methods, but is otherwise not found to explain productivity once

---

<sup>4</sup>The more general nature of the activities in Arora and Gambardella (1990)—R&D agreements and stock acquisitions—similarly warrants more caution in drawing conclusions from the documented tendency for joint adoption. The activities could be complementary in use, i.e. their returns are interdependent, but other explanations cannot be ruled out.

<sup>5</sup>Åstebro et al. (2005) use a similar approach to distinguish between complementarities and what they call strong one-step-ahead non-causality or strong simultaneous independence.

nationality of the owner is controlled for.

We are not aware of any previous study that uses activity-specific instruments to estimate the return function directly. Existing studies that control for simultaneity have studied the adoption decisions directly. In contrast, we infer complementarities from the effect of interactions of activities on productivity. In our application, this approach has empirical and theoretical advantages. Most importantly, firms can acquire the capability for an activity, e.g. flexible technology, but only put it in practice when necessary. Adoption is then distinct from the actual use of the activity. While this is problematic for an equation explaining (the timing of) adoption, the return function is unaffected by such distinction. The benefit of flexible technology or outsourcing will be more relevant when a large variety is produced, which varies year by year. The approach extends straightforwardly to continuous activity choices. For example, outsourcing decisions are observed for almost 30 tasks. A continuous outsourcing index is more intuitive than an all or nothing choice.

Moreover, recent theoretical models have shown that complementary inputs might be upgraded at different dates if upgrades involve learning (Jovanovic and Stolyarov 1997) or fixed costs (Jovanovic and Stolyarov 2000). The latter paper concludes more generally that “complementarity does not necessarily imply comovement, not even for a single decision maker (p. 1)”. While adoption decisions of complementary activities might not be clustered under these circumstances, the joint impact on productivity will still exist.

The remainder of the paper is organized as follows. After introducing the data set in Section 2, we discuss the three activities—model proliferation, flexible technology, and in/outsourcing—and the instruments in detail in Section 3. The empirical model underlying the investigation of complementarities is introduced in Section 4. Results of the analysis with robustness checks are presented in Section 5. Section 6 concludes.

## 2 Data

The plant-level information comes from *The Harbour Report North America*, published in 1980, 1981, 1989, and annually from 1994 onwards. Statistics are calculated using a uniform methodology from information supplied by the firms, supplemented with plant visits by representatives of Harbour Consulting. Firms voluntarily agreed to provide information to benefit from the productivity benchmarking exercise Harbour performs. The principle output of the benchmarking, the calculation of the number of hours it

takes each plant to assemble a vehicle, will be the dependent variable in our analysis. While quoted regularly in the trade press and daily newspapers, to our knowledge this information has not previously been used in the economics literature.

To guarantee coverage of the universe of plants in North America and a uniform definition of variables over time, the sample period is limited to 1994–2004.<sup>6</sup> All car and light truck assembly plants in the U.S., Canada, and Mexico are sampled, but some observations are dropped because of missing values. We estimate that the final sample covers 95% of the plants in the industry, accounting for an even higher share of output. Coverage is less complete for engine, transmission and stamping plants, especially in earlier years, and these plants are not included in the analysis.

A disadvantage relative to using data collected by the U.S. Bureau of the Census, as in Van Biesebroeck (2003), is the absence of capital stock information. The main advantage is the ability to include Canadian and Mexican plants and rich information on the types of products assembled in each plant. Relative to the data set constructed from plant surveys by the International Motor Vehicle Program, see Macduffie (1995), the benefit of The Harbour Report information is the complete coverage of the North American industry and the time dimension in the panel.

The performance measure used as dependent variable is the standard measure of (the inverse of) labor productivity in the industry: hours-per-vehicle (*h<sub>pv</sub>*).<sup>7</sup> For this to be an informative productivity measure, one is implicitly assuming that other inputs are fixed over time and across plants or vary proportionally to output. While obviously a strong assumption, it is not entirely implausible for the industry. Similar to the situation in Ichniowski et al. (1997), firms share the same technology (a moving assembly line), the production process follows the same steps (welding together stamped panels, painting the body, and assembling all components to the vehicle) and final products are made up of the same set of components. The scope for substituting between different inputs is clearly limited. It is unlikely to be controversial to assume that materials and intermediate inputs vary proportionally to output. Van Biesebroeck (2003) finds for U.S. assembly plants that

---

<sup>6</sup>Foreign producers were gradually added to the sample as they agreed to participate and the range of information collected has also expanded over time.

<sup>7</sup>Prior to 1998 a slightly different productivity measure was calculated: workers-per-vehicle (*w<sub>pv</sub>*). This measure ignored daily fluctuations in production and converted employment to full-time equivalent workers. Because both measures are available in 1998 and 1999, all *w<sub>pv</sub>* statistics can be converted to *h<sub>pv</sub>* using a conversion factor that varies by owner (firm) and country of location (of the plant). In all regressions we include a pre-1998 dummy. In the robustness checks we perform the analysis excluding all pre-1998 observations and find virtually identical results.

the sales-material ratio varies very little over time or between plants.

We want to argue that investments in fixed capital are also relatively similar across plants. Infrastructure investments are usually provided by local or state jurisdictions in order to ‘level the playing field’ in the bidding war to attract FDI, see Appel Molot (2005). Using plant-level Census data (which contains capital stock information), Van Biesebroeck (2003) estimates different capital coefficients for ‘lean’ or ‘mass’ technology plants. The estimated capital elasticities are very similar, 0.136 and 0.106, and not significantly different even at the 10% level, even though differences in the operation of the plants or productivity growth were large. In each regression, we include country dummies. Only if the capital-labor substitution varies by plant within the same country will the labor productivity measure be misleading. Wages are especially low in Mexico and substitution of labor for capital is most likely to occur there. In a robustness check we omit the Mexican observations and find very similar coefficient estimates of all variables of interest. Another robustness check includes plant-fixed effects to absorb capital stock differences between plants and results are again extremely similar. Finally, to capture technological innovations embodied in the capital stock we also experimented with the inclusion of information on the year the assembly plant was last retooled, but that variable was always insignificant.<sup>8</sup>

For comparability reasons, the *Harbour Report* presents the *h<sub>pv</sub>* comparisons by segment. We will include segment dummies in each regression to account for the complexity of the vehicle produced, e.g. a plant assembling luxury cars has on average a higher *h<sub>pv</sub>* than a plant assembling compact cars.

While the physical productivity comparison embodied in the *h<sub>pv</sub>* statistics is of obvious interest to identify complementarities in production, it would be useful to have an idea of the wider importance of this performance measure. Unfortunately, no other plant-level information is publicly available. One possibility would be to adjust *h<sub>pv</sub>* for the value of the vehicle produced. Unfortunately, a crucial aspect of the analysis is that models are sold in different configurations or styles, discussed below. We do not observe the breakdown of production by variety, while prices vary a lot (up to 100%). At the firm-level, the *Harbour Report* calculates each year a comparison of North American pre-tax profit per vehicle. In most years the ranking of firms is identical to the *h<sub>pv</sub>* ranking, as proof of its importance.<sup>9</sup> At the same time, the relationship between *h<sub>pv</sub>* and financial success

---

<sup>8</sup>Retooling is defined as a capacity change of more than 10% or the introduction of a vehicle from a different segment. The short duration of the panel makes this variable censored for almost 50% of plants.

<sup>9</sup>For example, in 2003 the only discrepancy is that Toyota is estimated to require 2% less *h<sub>pv</sub>* than



makes it also more likely that unobservables influence both productivity and successful adjustments to new activities.

The most crucial explanatory variables in the analysis are the number of varieties produced, flexibility in production, and the extent of outsourcing. Our preferred measure of model proliferation is the sum of the number of body styles and chassis configurations produced in the plant; for flexibility it is the number of platforms produced per production line;<sup>10</sup> and for outsourcing it is the average of all task-specific outsourcing dummies.

An instrument for the number of varieties produced in each plant is the average number of varieties produced in plants that produce vehicles for the same market segment and are owned by competing firms. Instruments for the adoption of flexible technology are the size of the plant in square footage (area) and whether shift relief is ‘mass’ or ‘tag’.<sup>11</sup> Finally, the extent of outsourcing is instrumented by the distance from each plant to the midpoint of the automotive industry in North America and to the industry’s midpoint within the plant’s country. Distances are calculated from the plants’ longitudes and latitudes and change slightly over time as the centers shift.<sup>12</sup> In the next Section, we discuss each of the three activities in detail, illustrate the changes over time, and provide details on the variable construction and the intuition for the instruments.

Control variables included in all regressions include location dummies (U.S., Canada, or Mexico); ownership dummies (U.S. or foreign-owned)<sup>13</sup>; a year trend and a dummy for the pre-1998 period to control for the change in measurement of the dependent variable; segment dummies ((sub-)compact cars, mid/full-size cars, sport/speciality cars, luxury cars, SUVs, pickup trucks, minivans, and full-size vans)<sup>14</sup>; and scale, measured by the logarithm of production capacity, calculated using a constant line rate and the regular shift pattern followed during the year. Summary statistics for all variables are in Table 1.

⇒ [Table 1 approximately here] ⇐

---

Honda, on average, while it makes lower profits per vehicle (compare tables on pages 30 and 150).

<sup>10</sup>A production line is defined as the average number of body and assembly lines per plant.

<sup>11</sup>Under mass relief, all workers in the plant change shifts at the same time, while the assembly line is stopped for 10–15 minutes. Under tag relief, individual workers relieve the worker on the previous shift they are replacing.

<sup>12</sup>For U.S. plants the exact location is taken from the U.S. Environment Protection Agency’s web site <http://www.epa.gov/air/opar/auto/>. For Canadian and Mexican plants, we use the center of the town where plants are located.

<sup>13</sup>Joint venture plants are treated as foreign-owned.

<sup>14</sup>In the rare cases where vehicles from more than one segment are produced, the dummies are weighted by production shares.

## 3 Three Trends in Automobile Production

### 3.1 Model Proliferation

The total number of car and light truck models for sale in North America increased from 185 in 1974 to 320 in 2004 (statistics in Table 2). In the U.S. the increase was even more pronounced, from 133 to 282 models, a more than doubling of the number of products. The median consumer, who replaces his vehicle every five years, had 25 models added to his choice set with every purchase. Such an evolution is typical for a mature industry; to retain market power firms differentiate their products.

⇒ [Table 2 approximately here] ⇐

Three more facts stand out from Table 2. First, the number of car models leveled off around 1994, while light truck models are being added at an accelerated pace. This is consistent with the relative maturity of the two segments; the use of light trucks as passenger vehicles is relatively recent. The latest wave of ‘crossover’ vehicles, bridging the gap between the two segments, keeps adding to the proliferation of models. Second, the number of models is not the entire story. The average number of variations per model—combinations of different engines, transmissions, drivetrains, and body styles—also increased over time, especially for light trucks. It adds to the manufacturing complexity. In the analysis we will use chassis configurations and body styles as measures of variety. Simply counting ‘names’ (models) risks double-counting vehicles that are merely rebadged clones and/or omitting different vehicles that are marketed under the same model-name.

Third, the number of models produced in North America has kept pace with the number of models offered for sale, while the number of assembly plants did not increase. As a result, the average number of models produced in each plant necessarily increased.<sup>15</sup> Total sales increased over time, but not to the same extent as the increase in models. The average production run for models produced in North America decreased noticeably, illustrated by the black line in Figure 1. This trend is most pronounced for cars (grey solid line), for which the average production run almost halved from an average of 138,000 units in 1974–76 to 75,000 by the end of the period (2002–04). The dashed line illustrates the declining sales per model (for the U.S.), which includes imported vehicles.

---

<sup>15</sup>The fraction of models produced domestically even increased slightly from 49% to 52%. Note that the definition of an assembly plant differs somewhat in the aggregate statistics, from *Ward’s Automotive Yearbook*. In the plant-level sample used in the analysis, from *The Harbour Report*, some plants report their car and light truck operations separately, even though they share the same production site.

⇒ [Figure 1 approximately here] ⇐

Furthermore, looking at averages might be deceiving. Even though 174 cars and 90 light trucks were sold in the United States in 2002, the combined market shares of the ten most popular vehicles were 32% and 42%, respectively. Total sales of 8.1 million cars over 174 different models averages 46,600 units per model, but the ten largest sellers average 259,000 units, while the remaining 164 models sell on average only 33,600 apiece. Manufacturers need to produce several of these models side-by-side in the same plant to reach efficient scale, which is traditionally put at approximately 200,000 vehicles per year.

The extent of model proliferation observed in the sample roughly matches the aggregate evolution. Simply summing the models produced in all plants in the sample, the first two columns of Table 3, shows only a modest increase from 172 models in 1994 to 175 in 2004. However, a large number of models (43) were initially produced in more than one plant, a practice that declined over time. Omitting double-counted models, last two columns of Table 3, the sample statistics approach the aggregates in Table 2. The remaining differences reflect that a few plants are missing from our sample.

Three other measures of product variety are also listed in Table 3: platforms, body styles, and chassis configurations. While firms define a platform in slightly different ways, it can be understood as “a set of common product and manufacturing standards related to a vehicle type. *Ward’s AutoWorld* (2001)”<sup>16</sup> The definition of body styles should be self-explanatory and chassis configurations indicate variations in drivetrain, i.e. engine, transmission, and which wheels are driven. These measures are arguably more reliable indicators of variety than models as they correspond directly to physical similarities or differences between vehicles.

⇒ [Table 3 approximately here] ⇐

The number of platforms that vehicles are based on increased much less than any of the other measures of variety. This reflects a deliberate strategy of manufacturers to base several models on a single platform. Platform sharing makes it easier to produce different models on the same assembly line, discussed in detail in the next Section. In 1994, the simple sum of body styles equaled the number of models. On average, each

---

<sup>16</sup>Firms often use different terminology. After its failed platform strategy in the 1980s General Motors now prefers the term architecture, while Volkswagen stresses the sharing of modules and systems, as opposed to common dimensions of the chassis or the frame.

model was produced in a single body style in each assembly plant; omitting duplication, each model was available in 1.10 styles. Variety in body styles was achieved by producing models in more than one plant, the case for 31% of models. By 2004, the ratio of body styles to models had increased to 1.29 (by plant) and 1.27 overall. The pattern for chassis configurations is similar, but less pronounced; total configurations increased by 20% from 1994 to 2004 versus a 44% growth rate for body styles.

Finally, Figure 2 illustrates the evolution of the average number of body styles and chassis configurations produced in each plant, normalized by the 1994 value. We do not discard duplicates because we want to capture variety at the plant-level (as opposed to the industry-level). Both averages clearly increased, again more for styles (38%) than for configurations (22%). The increase is broad-based, not driven by outliers. Only 10% (24%) of plants produce less body styles (chassis configurations) in 2004 than in 1994. This understates the increase as surviving plants already produced more variety in 1994. The black line plots the sum of configurations and styles, which is our preferred measure of variety. The average for the sample increased by 30% over the sample period; the average increase for survivors (plants observed all eleven years) is 35%.

⇒ [Figure 2 approximately here] ⇐

A complication for our analysis is that firms could decide to increase variety primarily in their best-performing plants, where they might expect a lower productivity penalty. As new models are introduced to achieve product differentiation, decisions will depend on choices of competing firms. We construct an instrument for plant-level variety exploiting these competitive interactions and the fact that different models produced in the same plant tend to be very similar.<sup>17</sup> If product introductions are endogenous, the average variety of competing firms in each market segment will enter a firm's decision of how much variety to produce itself. At the same time, other firms' variety will be independent of the unobservable we are concerned about, a plant-specific productivity term. For each plant-year we construct the average number of varieties produced by plants of other firms, limited to models that compete in the same segment, as instrument for its own variety.<sup>18</sup> The instrumenting strategy is similar to the use of characteristics of competitors' vehicles as an instrument for price in differentiated goods demand estimation, see for example Berry et al. (1995).

---

<sup>17</sup>We discuss the practice of basing different models on the same platform in the next Section.

<sup>18</sup>For plants that produce models in different segments, we average the competitors' variety over the segments.

## 3.2 Flexible Technology

The model proliferation discussed so far has lead all firms to produce several models in the same assembly plant. The oldest trick in the book, introduced by Alfred P. Sloan at General Motors in the 1920s, is to sell mechanically similar cars under different name-plates. Models differed mainly in appearance, standard features, and trim level and it was straightforward to develop and assemble them together. All firms have mastered such a ‘platform stretching’ strategy, even though they do not all use it to the same extent. It allowed firms, as recent as the 1970s, to combine production of high volume vehicles that share a platform in branch assembly plants across the country, see Rubenstein (1992) for a history. These trends are consistent with the patterns in Table 3: the number of platforms increased less than the number of models.

The more recent evolution of this practice is depicted in Figure 3A. We plot the cubic time trend on the average number of models per platform across all North American assembly plants for each the Big Three U.S. firms (including DaimlerChrysler), the combined Japanese firms (Toyota, Nissan, and Honda), and the group of joint venture plants.<sup>19</sup> Only the Japanese firms have noticeably stretched their platforms further over the sample period. To some extent it is merely driven by the addition of models that always shared the same global platform to their North American plants as they expanded capacity. The joint venture plants, on the other hand, have sharply reduced platform stretching as experience taught them that rebadging Japanese vehicles under a U.S. brand was not very profitable.<sup>20</sup> For the U.S. firms there is no real trend although the variation across firms and plants is large, which we will be able to exploit in the analysis.

The pressure to differentiate vehicles further, reflected by the increasing numbers of body styles and chassis configurations in Table 3, reduced the similarity between models derived from the same platform and increased manufacturing complexity. General Motors in North America and Volkswagen in Europe also ran into financial problems as consumers switched towards their cheaper brands (Chevrolet or Seat), given that those vehicles were often mechanically similar to models offered under premium brands (Buick or Audi). In order to successfully market vehicles derived from the same platform, the models had to

---

<sup>19</sup>NUMMI is a joint venture between GM and Toyota; CAMI between GM and Suzuki; AutoAlliance between Ford and Mazda; Diamond Star between Chrysler and Mitsubishi (sold to Mitsubishi when Chrysler faced financial difficulties); and Subaru-Isuzu is an all-Japanese joint venture.

<sup>20</sup>The Geo Metro, Geo Prizm, Ford Probe, nor Eagle Talon could come close to the sales success of their Japanese siblings, the Suzuki Swift, Toyota Corolla, Mazda MX6, and Mitsubishi Eclipse, to name only a few examples.

be sufficiently differentiated. Moreover, the emergence of new segments and the growing importance of light trucks reduced the possibility to base models on the same platform.

⇒ [Figure 3 approximately here] ⇐

When a plant is forced to produce radically different models, it can do so by adding assembly lines to a plant. Approximately one out of six observations in the sample has more than one assembly line, more than one in five in 1997 and 1998. Alternatively, manufacturers started to assemble models derived from different platforms on the same production line. This is how we will define flexible technology.<sup>21</sup> Figure 3B plots the trend in the number of platforms per line for the same five (groups of) firms. In three cases there is a strong upward trend in flexibility. GM and especially the Japanese firms and joint ventures produced more platforms per line in 2004 than in 1994. In sharp contrast with the tendency for platform stretching in Figure 3A, line sharing shows divergent trends for different firms. Both Ford and DaimlerChrysler had hardly adopted the practice in 2004. The fraction of plants producing more than one platform per line ranges from a low of 4% in 1996 to a high of 24% in 2004.

It should be stressed that the end goal is to produce a greater variety of vehicles per assembly line in an efficient way and flexible technology is merely a means to that end. Being able to produce models from different platform on the same line is certainly an advantage, but the ability to stretch the platform, e.g. deriving radically different vehicles from it, would preempt the need for flexibility, defined this way. The opposite evolution of the joint ventures in Figures 3A and 3B is consistent with this. The two measures of flexibility tend to be negatively correlated for all plants. Only the Japanese firms are exploiting both dimensions of flexibility (in different plants). In the analysis we use platforms per line as measure of flexibility, but we control for models per platform.

Flexibility also changes over time. For example, Honda assembled the Civic (compact car) and Acura CL (luxury car) on the same line in its East Liberty, OH plant (1996–1999), but only until a new line was installed at its Alliston, ON plant. Currently, each of its four North American plants is able to produce nearly every model in its lineup with only minor tooling changes, even though the company rarely uses this flexibility, *Automotive News* (2002). This underscores a crucial aspect of our application. Even if a plant has the capability for flexible production, it will not necessarily exploit it every

---

<sup>21</sup>Widely advertised examples of models derived from different platforms that are assembled on the same line are the Honda Civic and Accord in East Liberty, OH, the Chrysler Caravan minivan and Pacifica SUV in Windsor, Canada and the BMW Z4 roadster and X5 SUV in Spartanburg, SC.

year and it might not show up in our data set. If complementarities are inferred from the clustering of adoption decisions, one has to somehow distinguish between *acquiring* and *exploiting* the capability. If we identify complementarities directly from the joint effects on productivity, this distinction is not a problem.

Instruments for flexibility are the inside area of the plant and the way shift relief is organized. Both are assumed to be pre-determined. Most plants are several decades old and their size was determined long before the capital equipment was installed or work practices determined. Size is important as only larger plants have the option to duplicate assembly lines, an alternative approach to producing more variety. The method of shift relief, *tag* or *mass*, is determined at the plant-level through negotiations with the labor union. Controlling for ownership, shift relief is unlikely to be correlated with unobserved plant-level productivity, but as it reflects how labor and management have chosen to interact they are likely to influence the adoption of flexibility. The production complexity under flexible technology requires more autonomy and problem solving skills of employees. The presence of *tag* relief, whereby individual workers relieve each other, signals the level of trust between workers and supervisors.

### 3.3 In/Outsourcing

Another aspect of firm organization that has received a lot of attention recently is the make-or-buy decision. Outsourcing, especially to low-cost overseas destinations, is the most recent buzz word in the business press. A major difficulty to study this phenomenon is a lack of data, as outsourcing decisions are considered highly confidential. One exception is Novak and Stern (2003) who use detailed information on nineteen automobile model-years, but cannot reveal the identities of the firms involved.

Official statistics, aggregated up from plant-level information, can be used to track one measure of outsourcing: the share of materials (the sum of raw materials, intermediate inputs, and purchased business services) in final sales. The most detailed industry classification with comparable information for the three North American countries is “Motor Vehicles” (SIC 371). The ratio trended up until the late 1980s and declined subsequently, see Figure 4. At the very end of the period the materials-sales ratio increased again, but in 2001 the ratio in Canada and the U.S. was the same as in 1972. For the sub-sector “Motor Vehicles and Passenger Car Bodies” (SIC 3711), data from the U.S. Bureau of the Census reveals the same pattern of increase and decline, but changes are magnified (dotted black line).

⇒ [Figure 4 approximately here] ⇐

The absence of a clear time trend is not entirely surprising. The range of materials and components each assembly plant receives from other plants is relatively similar. Much of the changes in sourcing are likely to happen at the firm instead of the plant-level. The measure in Figure 4 does not distinguish between outsourcing to other plants owned by the same firm or to independent suppliers. Aggregate industry statistics for the U.S. do indicate a growing importance of the “Motor Vehicle Parts & Accessories” (SIC 3714) sub-sector, with sales growth of 9% per year over the last 15 years, compared to only 3% growth for SIC 3711. Unfortunately, no information is available at the firm-level.

The plant-level data from *Harbour* contains information on another aspect of the outsourcing decision: plants report which of 29 tasks they perform on-site. While distinct from the more commonly studied make-or-buy decision of components, this provides information on a dimension where assembly plants are likely to vary more.<sup>22</sup> At one extreme, the Oshawa GM plant performed none of the tasks on the list in-house in 1994, while the Georgetown Toyota plant performed all but two tasks on-site prior to 1996. The average fraction of tasks outsourced is 64% and the standard deviation is 17%. The index in Table 1 refers to insourcing, defined as one minus the fraction of outsourced activities, to facilitate the interpretation of coefficient estimates later on.

Over time, outsourcing is relatively constant, see the first column in Table 4. In 1994, 36% of all activities were performed in-house and in 2004 the fraction was 35.6%, after a temporary increase to a maximum of 38.5% in 1998. The pattern is virtually identical for the entire sample or the balanced sample of plants that survived from 1994 to 2004, which excludes several foreign transplants that enter after 1994. Compositional effects—transplants tend to perform more activities in-house initially before they establish relationships with local suppliers—seem to be minor. Similarly as for flexibility, the average hides important cross-sectional variation. For example, Van Biesebroeck (2006) shows that there are large differences by ownership.

The set of tasks can be split into a first group of administrative and manufacturing tasks which are usually outsourced and a group of assembly tasks that are more likely to be performed in-house. *Harbour* adjusts the *h<sub>pv</sub>* statistics for tasks in the first (but not the second) group that are insourced. Statistics in the second and third column indicate that even though the average level differs markedly for the two groups, there is no real

---

<sup>22</sup>As with the aggregate statistics, we cannot distinguish between outsourcing to independent suppliers or to other divisions of the same firm.



time trend for either.

Using the detailed responses for individual tasks, we can look at changes in outsourcing more directly. In columns 5 to 7 we indicate how many plants brought one or more tasks in-house, newly outsourced a task, or did not change any of their outsourcing choices. While no change is the dominant choice in all years, in seven of the ten years plants that increased insourcing outnumber plants that increased outsourcing.<sup>23</sup> In columns 8 to 10 we sum all the changes, in or out, over plants. For example, in 1995 sixteen plants newly insourced a total of twenty three tasks, while thirteen plants newly outsourced a total of fifteen tasks. The net effect of tasks performed in-house was an increase by eight. In most years the net change is positive and the cumulative change over the entire period and all plants is for an average of 3.5 tasks per year to be moved in-house. Outsourcing decisions were changed more often in the earlier period. In 1997, eighty changes are made and insourcing is twice as common as outsourcing. In 1998, one hundred changes are made and insourcing is 50% more likely.

⇒ [Table 4 approximately here] ⇐

Of the three activities considered, simultaneity between unobserved productivity and outsourcing is most likely. Firms are expected to perform less activities in-house if they know a plant is particularly unproductive. Instruments for outsourcing are the distances from each plant to the center of the industry, calculated by year both for the entire North American industry and by country.<sup>24</sup> Plants located farther from the center will find it harder or simply more costly to find suppliers. After controlling for the country (which will also soak up fixed border effects) and ownership, there is no reason to believe distance is correlated with unobserved productivity. The location of older plants in the sample is predetermined relative to the current outsourcing decision—several sites predate World War II—while the location of newer plants is generally considered to be mostly determined by shipping costs for the final vehicle and labor market conditions.<sup>25</sup> While both distances are negatively related to outsourcing, the distance to the country-specific center proved to be the strongest predictor.

---

<sup>23</sup>The first two categories are not exclusive.

<sup>24</sup>The center for North America is (on average) 62 miles south-west of the Louisville, KY Ford plant. The U.S. center is approximately 43 miles south-east of the Subaru-Isuzu plant in Lafayette, IN. From 1994 to 2000, it moved almost 30 miles south. The Canadian center is closest to the Cambridge Toyota plant, but distances never go beyond 340 miles; the Mexican center is closest to Silao, Guanajuato. An attempt was made to control for travel-time, but given that most distances are large and plants are always adjacent to highways, differences were minor and coefficient estimates were identical.

<sup>25</sup>Shipping finished vehicles is considerably more expensive than components, see Rubenstein (1992).

## 4 Model

### Complementarities

Complementarities are defined as the joint impact of activities on some return function, productivity in our case. To fix ideas consider the following simple model. Total labor hours required to assemble a vehicle can be divided into the direct labor input on each vehicle and externalities on other types of vehicles produced in the same plant, which can be positive or negative:

$$\text{Hours} = \sum_{i=1}^N \left[ a Q_i + \sum_{j \neq i}^N b Q_j \right], \quad (1)$$

$i$  and  $j$  index the  $N$  varieties assembled in the same plant. The marginal labor requirement for each extra vehicle is  $a$  hours and additionally it increases (or decreases) the labor requirements on all vehicles of a different type by  $b$ .

Both the direct effect and the spillovers depend on other characteristics of the plant:

$$a = a(\text{Flexibility, Insourcing, Scale, Other controls})$$

$$b = b(\text{Flexibility, Insourcing, Other?}).$$

Assuming that each of these variables is measured continuously,<sup>26</sup> we expect:

$$\frac{\partial a}{\partial \text{Flexibility}}, \quad \frac{\partial a}{\partial \text{Insourcing}} \geq 0, \quad \frac{\partial a}{\partial \text{Scale}} \leq 0, \quad \text{and} \quad (2)$$

$$\frac{\partial b}{\partial \text{Flexibility}}, \quad \frac{\partial b}{\partial \text{Insourcing}} \leq 0. \quad (3)$$

Flexibility makes the entire production process more complicated and will, *ceteris paribus*, lead to higher direct labor requirements. The benefit will be to lower labor requirements of other varieties produced in the same plant—captured by the negative derivative of the  $b$  function. Doing more tasks in-house will straightforwardly increase labor requirements, but it can generate useful expertise which reduces the spillovers on other types of vehicles, possibly even reducing their labor requirements.<sup>27</sup> Through the  $a$  function, we

---

<sup>26</sup>It would be straightforward to extend the model to allow for discrete adoption of activities. This would require the necessary assumptions for the productivity function to be supermodular in the different activities, as in Athey and Stern (2003) and Miravete and Pernías (2005). Given that the model only serves an illustrative purpose, it is limited to the continuous case and we can define complementarity simply in terms of cross-partial derivatives, as in Arora and Gambardella (1990).

<sup>27</sup>One mechanism would be fixed costs to outsourcing relationships that are task and vehicle-type specific (e.g. managing deliveries), while the fixed costs to perform an in-house task could be not or to a

control explicitly for scale economies (assumed to be positive) and other effects. Other factors, maybe scale, could affect spillovers, but we do not have a strong prior on the direction of those effects. Finally, producing different varieties in the same plant generates on average positive spillovers, i.e. it increases labor requirements for other vehicles:  $b(\overline{\text{Flexibility}}, \overline{\text{Insourcing}}) \geq 0$ .<sup>28</sup>

Rewriting equation (1) as

$$\frac{\text{Hours}}{Q} = (a - b) + b N,$$

using ‘Variety’ for  $N$ , and substituting linear specifications for the  $a$  and  $b$  functions,<sup>29</sup> we obtain

$$\begin{aligned} hpv = & \alpha_V \text{Variety} + \alpha_F \text{Flexibility} + \alpha_I \text{Insourcing} & (4) \\ & + \left[ \beta_{VF} \text{Flexibility} + \beta_{VI} \text{Insourcing} \right] \times \text{Variety} \\ & + \alpha_{FI} \text{Flexibility} \times \text{Insourcing} + \alpha_S \text{Scale} + \text{Other controls.} \end{aligned}$$

The dependent variable, hours per vehicle, is the inverse of labor productivity. The  $\alpha$  coefficients capture direct effects. Flexible production, insourcing of tasks, and producing a large variety are all expected to be associated with productivity penalties, i.e. have a positive impact on  $hpv$  ( $\alpha_V, \alpha_F, \alpha_I \geq 0$ ). Scale economies in automobile production are expected to be positive ( $\alpha_S \leq 0$ ). The  $\beta$  coefficients capture the reduction in productivity penalty associated with variety if flexibility or insourcing is increased. The main prediction we want to test is that both activities generate negative effects ( $\beta_{VF}, \beta_{VI} \leq 0$ ). Finally, both activities could be complementary themselves, if the cross-partial derivative is negative ( $\alpha_{FI} \leq 0$ ).

### Estimation

We now describe how complementarities can be identified in the presence of unobservables. Consider the following generalization of the return function in equation (4)<sup>30</sup>

$$\frac{1}{\text{Productivity}} = \sum_j \left[ \alpha_j + \sum_{k \neq j} \frac{1}{2} \beta_{jk} Y_k + \xi_j \right] \times Y_j + \text{Controls} + \epsilon. \quad (5)$$

---

lesser extent type specific (e.g. training workers).

<sup>28</sup>If adoption of activities is discrete, we assume that  $b(0, 0) \geq 0$ .

<sup>29</sup>We assume that the direct effects dominate spillovers, such that  $\text{sign}(\partial a / \partial X) = \text{sign}(\partial(a - b) / \partial X)$ .

<sup>30</sup>Eq. (5) is similar to the firm’s payoff function on p. 368 in Arora and Gambardella (1990); the objective function on p. 16 in Novak and Stern (2003); or the profit function (7) in Miravete and Pernías (2005).

Each activity  $j \in \{V, F, I\}$  has a direct effect on productivity ( $\alpha_j$ ), but the total effect will differ if other activities  $k$  are adopted jointly ( $\beta_{jk}$ ). For example, producing a larger variety is expected to reduce productivity ( $\alpha_V > 0$ ), but this reduction might be mitigated by adopting complementary activities like flexible technology ( $\beta_{VF} < 0$ ) or insourcing of tasks ( $\beta_{VI} < 0$ ). Part of the return to each activity is assumed to be unobservable ( $\xi_j$ ) and these returns could be correlated across activities. A set of exogenous control variables and an i.i.d. error term complete the estimation equation.

We are primarily interested in identifying the complementary effect of different activities, the  $\beta$  coefficients. It is by now widely understood that estimating equation (5) by least squares will produce biased estimates if firms make endogenous adoption decisions, see Arora (1996) or Athey and Stern (2003). If the unobservable return to each activity is correlated, for example because they all depend on an omitted variable like managerial ability, the activities will look like complements even if they do not influence each other's return. This will happen when one studies the adoption decisions directly or looks at the joint impact on productivity.

Assuming all activities are continuous, a firm will choose the level of  $Y_j$  by equating the marginal cost associated with  $Y_j$ —part of which is the impact on productivity (multiplied by  $\lambda$  to make units comparable)—to the marginal benefit:

$$\lambda \frac{\partial \text{Productivity}^{-1}}{\partial Y_j} + MC(Y_j) = MB(Y_j).$$

For simplicity, we assume marginal costs and benefits are linear in  $Y_j$  and  $X_j$ , an exogenous shifter to the cost or benefit of activity  $j$ . This generates the following first order condition for activity  $j$

$$\alpha_j + \sum_{k \neq j} \beta_{jk} Y_k + \xi_j = \gamma_j^0 + \gamma_j^y Y_j^* + \gamma_j^x X_j + \eta_j. \quad (6)$$

In matrix form, the system of first order conditions for all activities is

$$B Y^* = \Gamma^0 + \Gamma X + \xi + \eta,$$

where  $Y = [Y_V, Y_F, Y_I]'$ , and similarly for all other vectors. The  $\Gamma^0$  vector groups the constant terms and  $\Gamma$  is a diagonal matrix. Solving for the endogenous activities gives

$$Y^* = \Gamma^{0'} + B^{-1} \Gamma X + B^{-1} \xi + \eta'. \quad (7)$$

The B matrix has  $-\gamma_j^y$  on its diagonals and  $\beta_{ij}$  ( $i \neq j$ ) off-diagonal elements. Unless complementarities are entirely absent, the optimal level of each of the activities will depend on each of the three unobservable returns in the  $\xi$  vector.

Two straightforward estimation procedures are possible. We can estimate equation (6) for activity  $j$  using the excluded benefit or cost shifters for the other activities  $k$  as instruments for  $Y_k$ . Novak and Stern (2003) follow this approach to estimate complementary effects of outsourcing decisions on different components of a car. Note that it is crucial to use activity-specific instruments, as all exogenous variables that influence the choice of activity  $j$  are included in the equation.

Alternatively, we can estimate the return function (5) directly, instrumenting the endogenous  $Y$  variables with the exogenous cost and benefit shifters. Given that the coefficients enter linearly, a two-step approach works: first estimate the system of reduced form equations (7) and replace the endogenous variables in (5) with their predicted values ( $\hat{Y}$ ). Alternatively, we can estimate the productivity equation using instrumental variables, using interactions of the exogenous variables as additional instruments. The advantage of this latter approach is that we can use a GMM estimator to gain efficiency and test for overidentifying restrictions.

The empirical strategy of estimating the return function instead of the first order conditions has a number of advantages. First, it is robust to spurious correlation in adoption costs. Second and particularly important for our application, the level of some activities is harder to change than for other. For example, tasks brought in-house when variety was high, can take some time to outsource when variety falls. While the first order conditions for some activities might sometimes be slack, this should not change the productivity effects. Third, the only way to identify the direct effects, the  $\alpha$  coefficients, is from the return function. While these are often of lesser interest, in our application complementarities are expected to be of an unusual form: direct effects of activities on productivity are negative, but interaction effects positive. To find evidence for this, we have to estimate the productivity equation directly.

Athey and Stern (2003) investigate explicitly under what assumptions activity-specific instruments will be able to identify complementarities. For example, one has to rule out unobservables that affect the joint return of adoption of different activities—this would be  $\beta_{jk} = \beta'_{jk} + \xi_{jk}$  in equation (5). To some extent this is merely definitional, all joint effects of activities could be called complementarities. The rigorous framework in Athey and Stern (2003) distinguishes between joint returns that result from unobservables and true complementarities that accrue even in the case of an exogenous shift in adoption.

## 5 Results

### *Productivity regressions*

The dependent variable in all regressions is hours per vehicle (*h<sub>pv</sub>*), the inverse of labor productivity. The crucial stylized fact—a productivity penalty for producing variety—is documented in Table 5 using a number of different measures for variety: platforms, models, body styles, chassis configurations and the sum and product of styles and configurations. All coefficient estimates are positive and significantly different from zero, indicating lower productivity if higher variety is produced. As expected, producing more platforms is especially costly. The relative magnitudes seem reasonable and are similar to the extent to which varieties share the same production line. The preferred measure, ‘styles + configurations’, leads to only a moderate productivity penalty: 21 minutes per vehicle for each extra unit of variety.

In each of the subsequent regressions we use the same controls as in Table 5. Productivity growth is positive. The negative sign on the pre-1998 dummy indicates that productivity growth is slowing down. The dummy is included primarily to capture the change in the dependent variable, but it is rarely significant. The negative coefficient on capacity indicates increasing returns to scale. Productivity is slightly higher in Canada than in the U.S. (the excluded category), requiring on average 1.5 fewer hours per vehicle, while the lower productivity in Mexico is more pronounced. Foreign-owned plants, which includes joint ventures, are more productive. Once we instrument for activities, see below, the point estimate on the foreign dummy is estimated very imprecisely.<sup>31</sup> Models in different market segment differ substantially in their labor input requirement, with luxury cars and (mini)vans being most labor intensive (the excluded category is compact cars).

⇒ [Table 5 approximately here] ⇐

The negative coefficient estimates on the time trend indicates substantial productivity growth, the time to assemble a vehicle falls by 15 to 45 minutes per year. In addition to the linear effects of variety and time, each regression also contains the interaction of the two variables. The productivity penalty associated with variety is falling for each of the measures. Possibly this trend is explained by the adoption of activities that make it more economical to produce variety, such as flexible technology and insourcing of tasks so they can be shared between models. This is the focus in the remainder of the analysis.

---

<sup>31</sup>Once Mexican plants are dropped or activities are measured discretely, the coefficient becomes significant again and is indistinguishable from the OLS estimate.

The first column in Table 6 contains the least squares estimate of equation (4). For this and all following regressions, the coefficient estimates on the control variables are reported in Table A.1 in the Appendix. All coefficients have the expected sign. The linear coefficients on variety, flexibility, and insourcing are estimated positively and coefficients on all interactions are negative. Increasing the use of flexible technology reduces productivity directly, the linear “Flexibility” term is positive, but makes it less costly to add variety.<sup>32</sup> Similarly, bringing activities in-house requires more hours to produce a vehicle, *ceteris paribus*, but the increase in hours for added variety is reduced (although the point estimate is not significant). The two activities reinforce each other. Increasing flexibility lowers the productivity cost of insourcing, and vice versa. The increased complexity when different platforms are assembled on the same line seems to reduce the number of hours that can be saved by outsourcing.

As discussed before, one cannot take these results as evidence for complementarities in production as they can merely reflect correlated unobservables in the adoption decisions. In the other columns of Table 6 we explore how sensitive the results are to alternative exogeneity assumptions. The implicit assumption in column (1) is that all activities are exogenous. In columns (2) to (4), we instrument for flexibility, for outsourcing, and for both. In column (5) all activities, including variety, are treated as endogenous. Instruments were discussed in detail in the respective Sections (3.1–3.3): for flexibility we use plant size and type of shift relief, for insourcing distance to the industry center in North America and within the country, and for variety we use variety in competing plants.

Tests for validity of the instruments are reported below the coefficients. The marginal explanatory power of the instruments in the different first stage regressions varies somewhat across variables and specifications, but the p-value is never above 5%.<sup>33</sup> For each endogenous variable (or product of endogenous variables) we can reject that the instruments have no explanatory power. Overidentification restrictions can never be rejected either, which to some extent merely reflects the imprecision of some coefficient estimates.

When instrumenting, the absolute values of the point estimates and the standard errors tend to increase, but all signs are unchanged. The interactions make it impossible to predict the bias one would expect using least squares if adoption is endogenous. We take heart that in every single equation the signs of all coefficients correspond to our

---

<sup>32</sup>A dummy to control for the practice of basing several models (more than two) on a platform is included (reported in Table A.1.), as platform stretching is an alternative to flexibility as we defined it.

<sup>33</sup>To improve precision the coefficients are estimated by GMM and not by the two-step approach that underlies the test.

expectations and that the interactions between variety and both activities are significantly negative in columns (4) and (5) where we instrument for both flexibility and insourcing.

At the bottom of Table 6 we report the results of Hausman specification tests for each model versus the models in columns (4) and (5). The test statistic in column (4) suggests that one cannot reject that variety is exogenous, as the p-value is greater than 0.99. Comparing the two columns in Table 6 with those in Table A.1 suggests that changes tend to be much smaller for the control variables. Limiting the test to the six variables of interest, the three activities and their interactions, the test statistic is only slightly reduced and the p-value is still 0.92. Hence, we will treat variety as exogenous in the subsequent regressions, consistent with the general view within the industry, as discussed in the introduction. Similar specification tests, against columns (4) or (5), for endogeneity of flexibility or insourcing or both, find stronger but no overwhelming evidence that these activities are endogenous. Erring on the conservative side, we will continue to instrument for these activities.

Comparing the preferred results in column (4) with the OLS estimates, we find that the direct negative effect of both activities on productivity has increased. To the extent that activities are chosen endogenously, firms are successfully implementing them in plants that do not suffer as large a (direct) productivity penalty from the activities. The same mechanism is consistent with a reduction in the absolute value of the coefficient on the variety\*flexibility interaction when we instrument for insourcing (or additionally for variety in column (5)). Only exploiting the exogenous variation in the insourcing decision, as captured by distance, lowers the estimated complementarity between insourcing and variety.

The interactions with flexibility, on the other hand, become larger and highly significant when we instrument. Results now suggest that as flexibility is introduced because of space constraints and not because the plant is expected to master it well, the direct productivity penalty is greater (as expected), but the complementary effect to reduce the productivity penalty for variety is enhanced (contrary to expectation). If variety is exogenous but positively correlated with flexibility, the underestimate of the direct effect of flexibility in (1) could lead to an upward bias for the interaction between flexibility and variety. If variety is chosen endogenously, the apparent complementarity between variety and flexibility in (4) merely picks up the decision to add variety to a plant that is expected to handle it well. The reduction in the interaction coefficient from (4) to (5) would be consistent with this.



⇒ [Table 6 approximately here] ⇐

So far we have treated all observations as independent, without exploiting the panel structure of the data set. Given that unobserved heterogeneity is a concern when trying to identify complementarities, we would ideally estimate the model in first differences. The problem is that the instruments hardly vary over time. Taking first differences leads to extremely imprecise coefficient estimates for most variables. Moreover, the theoretical model in Jovanovic and Stolyarov (2000) illustrates that joint adoption is not a requirement for complementarities. Estimating the model in first differences would identify complementarities solely from the joint effects at the time of adoption. If one wanted to adopt such an identification strategy, it would be more straightforward to estimate the first order conditions for optimal activity choice directly, see below.

As an alternative, we include plant-fixed effects in the model and estimate as before. This way we remove the time-invariant component of productivity from the analysis. We also estimate the model with firm-fixed effects, as capabilities that enter the adoption decisions are likely to be shared between plants owned by the same firm. While there are 92 plants in the sample, they are owned by only 11 firms, conserving degrees of freedom. Finally, we also estimate the model using random effects, allowing the error terms to be correlated across plants.

Surprisingly, not a single coefficient estimate changes sign. While estimation precision is reduced, the variety\*insourcing interaction remains significant at the 5% or 10% level in each specification and the variety\*flexibility interaction is only insignificantly different from zero when random effects are included. The direct coefficient estimates are also similar, which is especially surprisingly in column (8), where they are entirely identified from the variation over time. The estimated productivity penalty for variety is 5.5 with plant-fixed effects, only slightly above the penalty of 4.9 estimated earlier. This modest increase is consistent with the assumption that plant-level variety is mostly exogenous. Otherwise, the change in the coefficient estimate would likely be larger using cross-time instead of cross-plant variation. In contrast, identifying insourcing mostly from variation over time increases the point estimate considerably.

We include the tests for validity of the instruments at the bottom. While we had expected the instruments not to have much explanatory power in the first stage, especially when plant-fixed effects are included, this does not turn out to be the case. Every single endogenous variable (or interaction of endogenous variables) is well predicted by the instruments.

The results in Table 7 suggest that it is highly unlikely that the complementarities identified earlier are just picking up unobserved heterogeneity. At the same time, the relative size of the effects change. In column (4), the primary effect of insourcing is to lower the direct cost of flexibility. In the other columns, the effect of insourcing on the productivity penalty for variety is enhanced, at the expense of the variety\*flexibility term, and most of the insourcing\*flexibility interaction effect disappears. While the results still indicate that both activities lower the productivity penalty for variety, potential complementarities between the two activities makes it hard to determine their relative importance.

Measuring the activities continuously facilitates the theory, but makes it harder to interpret the coefficient estimates. For example, the productivity impact of the flexible technology depends on the number of varieties produced and on the extent of outsourcing. To illustrate the impact of flexibility and insourcing on the cost of producing variety, the predicted  $h_{pv}$  as a function of variety is plotted in Figures A.1a and A.1b in the Appendix using coefficient estimates from columns (1) and (4). While the coefficient estimates are too imprecise to use them to make predictions, we include these figures to illustrate the economic significance of the estimates. In addition, the figures also indicate that large changes in point estimates do not necessarily translate in large  $h_{pv}$  differences as opposing changes on direct and interaction effects will partially offset one another.

Evaluating all variables at their sample mean (solid black line),  $h_{pv}$  is predicted to rise more quickly with variety using the GMM estimates (bottom panel) than using the OLS estimates (top panel). If a plant produces with average flexibility but insources 67% of activities instead of the average of 36%, adding variety tends to lower the predicted hours (grey short-dashed line). This level of insourcing corresponds to the 90th percentile. Evaluating  $h_{pv}$  at the 90th percentile of flexibility (1.5 platforms per line) and the mean level of outsourcing (grey long-dashed lines) has similar effects. The higher direct effects for the GMM estimates shift the curves upward farther, while the more negative variety\*flexibility interaction is also noticeable. Finally, the joint effect of adopting both activities eliminates most of the productivity penalty for variety, the white lines become almost flat, but especially under the GMM results this would be profitable only for very high levels of variety (the 90th percentile is 10 styles + configurations).

### Robustness checks

In Table 8, we present a number of robustness checks using the same estimation method as in column (4) of Table 6, which is repeated for convenience. Given that the dependent

variable is always positive, we can estimate the productivity regression in logarithms.<sup>34</sup> Results in column (9) indicate that direct effects are still positive and interaction effects negative, with little change in the precision of the estimates. Evaluated at the sample mean, adding one extra variety to the plant increases labor requirements by 5%. Increasing both flexibility and insourcing by one standard deviation, would eliminate the productivity penalty almost entirely. In this case, the coefficient on scale, 1.269, has a direct interpretation. This level of scale economies is consistent with Friedlaender et al. (1983) who find 1.23 for the most recent year in their sample (1979) using multifactor productivity. Hence, the assumption that capital stock differences between plants are relatively unimportant seems not unreasonable.

As mentioned earlier, the calculation of the dependent variable differed pre-1998 and we used two overlapping years to construct a conversion factor. A dummy variable was included to soak up remaining differences, but it is rarely significant. To control for the change in a more flexible way, we re-estimate the model dropping all observations prior to 1998, results are in column (10). The only noticeable change is that the direct effect of outsourcing becomes much smaller and insignificant. The coefficients of interest, the interaction coefficients of variety with insourcing and flexibility, are virtually unchanged and still significantly negative.

In spite of our motivation for the use of the *h<sub>pv</sub>* variable, one might be concerned that it incorporates differences between plants in capital intensity or differences in the degree of component outsourcing (as opposed to ‘tasks’, for which we control explicitly). Differences are likely to be largest between Mexican and other plants because of wage differences. Omitting the Mexican plants from the sample, results in column (11), changes several point estimates, but all the qualitative findings go through. The largest change is for the three coefficients related to insourcing. Dropping the Mexican plants removes much of the time variation from the instruments for variety (distances to the industry center) and makes both instruments much more alike. The estimates related to flexible technology, which could be a more capital intensive way of operating, change much less.

Instead of measuring flexibility and insourcing as continuous variables, one can also measure them discretely: adopt or don’t adopt. For flexibility, we define an adoption dummy that switches from zero to one if more than one platform is produced per line. For insourcing, we define a dummy indicating whether a plant performs more activities in-house than the median plant in that year. Variety is still measured continuously as

---

<sup>34</sup>Eq.(1) could be specified with multiplicative spillover effects on the labor requirements of other models.

the sum of body styles and chassis configurations. While the coefficient estimates in columns (12) are not comparable to other columns, the absolute magnitudes indicate very large reductions in the productivity penalty when the activities are adopted. Under the restrictive functional form we use,  $h_{pv}$  even goes down with increased variety, for flexible plants or those that insource most. Still, the very high direct effects of both activities makes adopting only profitable for plants that produce most varieties.<sup>35</sup>

Finally, results in column (13) add scale interacted with variety to the model. In each regression we included a direct effect of scale (estimated negatively and highly significant), we now allow scale to influence spillovers as well. I.e. scale could be one of the ‘other’ factors in the  $b$  function in equation (1). We find that indirect effects of scale hardly matter at all. The other coefficient estimates are virtually unchanged, standard errors are very similar, and the coefficient on the variety\* $\text{scale}$  interaction term is small and insignificant.

⇒ [Table 8 approximately here] ⇐

### Optimal activity choices

Finally, we estimate the first order conditions for the three activities, equation (6), directly. We follow the approach in Novak and Stern (2003) using the cost or benefit shifters for activities  $k$  as instruments for  $Y_k$  in the equation for  $Y_j^*$ . The main disadvantage of this identification approach, relative to estimating the return function directly, is that all first order conditions have to hold with equality all the time for the relationship to be stable. If one activity is more easily changed than another, e.g. because of differences in sunk or fixed costs, the equation will be misspecified. As mentioned earlier, we also cannot identify the direct effects this way, but we can test for complementarities.

Results in Table 9 show that the main prediction of the analysis is confirmed. Variety enters with a positive sign in the equations describing the flexible technology and insourcing choice. Variety at competing plants is used as instrument for own variety in both equations. Results in column (1), show that variety at competing plants is a very strong predictor of own variety. The effects of flexibility and insourcing on each other and for insourcing on variety choice are insignificant. In contrast with most of our earlier findings, they enter each other’s equation with a negative coefficient.

---

<sup>35</sup>With discrete activities, the model is now in the same form as that discussed in Athey and Stern (2003). In our estimating equation, their test statistic for complementarities between insourcing and flexibility would be measured directly by the  $\beta_{FI}$  coefficient. The large and significantly negative estimate suggests the presence of complementarities.

⇒ [Table 9 approximately here] ⇐

## 6 Concluding remarks

On November 21, 2005 GM announced that it would eliminate a total of 30,000 jobs across North America by 2008 and the next February it announced a 2005 loss of \$8.6b, the second-highest in its history. In between, GM introduced a slate of new vehicles at the Detroit Autoshow, raising the number of models it will sell in 2006 in the U.S. alone to 72. Ford's restructuring plan called for 25,000 layoffs, but stressed the large investments in flexibility that are underway. Its Oakville assembly plant will receive a \$1b investment to make it fully flexible, following similar investments in its Kansas City plant and a new truck plant in its home town, Dearborn. Meanwhile, Toyota continues to prosper in North America. On average, it adds one model per year to its North American production lineup, it performs more tasks within its assembly plants than any other firm, and it has mastered flexibility, which shows up most clearly in its capacity utilization, averaging 99% over the last five years (the industry average was 85%).

While these are merely some examples of the interplay between model proliferation, flexible technology and outsourcing, the analysis has documented more systematically that the three activities are complementary in the production of automobiles. In particular, flexibility and insourcing can lower the productivity penalty associated with increased variety. Plants that are able to assemble vehicles derived from different platforms on the same production line face a lower productivity penalty when they increase variety. Similarly, plants that outsource less activities face a diminished productivity cost to increasing variety. Some results suggest that flexible technology and insourcing are themselves complements and that their cost of adoption, in terms of lost productivity, is reduced when they are adopted together. The crucial interaction coefficients are not always estimated very precisely and the point estimates cover a wide range, but the effects are extremely consistent across different specifications.

The nature of these complementarities differs from previous results. These activities do not increase productivity in their own right with an extra boost for joint adoption. Rather, increasing one activity tends to *decrease* productivity, but this reduction is diminished if other activities are increased as well. In particular, while flexibility and performing more tasks in-house increases direct labor requirements, they reduce the marginal labor requirement associated with increases in variety. Even absent heterogeneity in unobserv-

able returns or costs, implementing these activities will be worthwhile for some plants, but not for all.

On the methodological side, we are the first study to find complementarities in a framework that studies the joint impact on a performance variable using the activity-specific instrumenting approach advocated in Athey and Stern (2003). The main advantages over studies that look at clustering in adoption decisions is that we also recover the direct effects on productivity and that the model is robust to learning and sunk or fixed costs of adoption. A major disadvantage is that instrumenting for interactions of endogenous variables results in less stable point estimates.

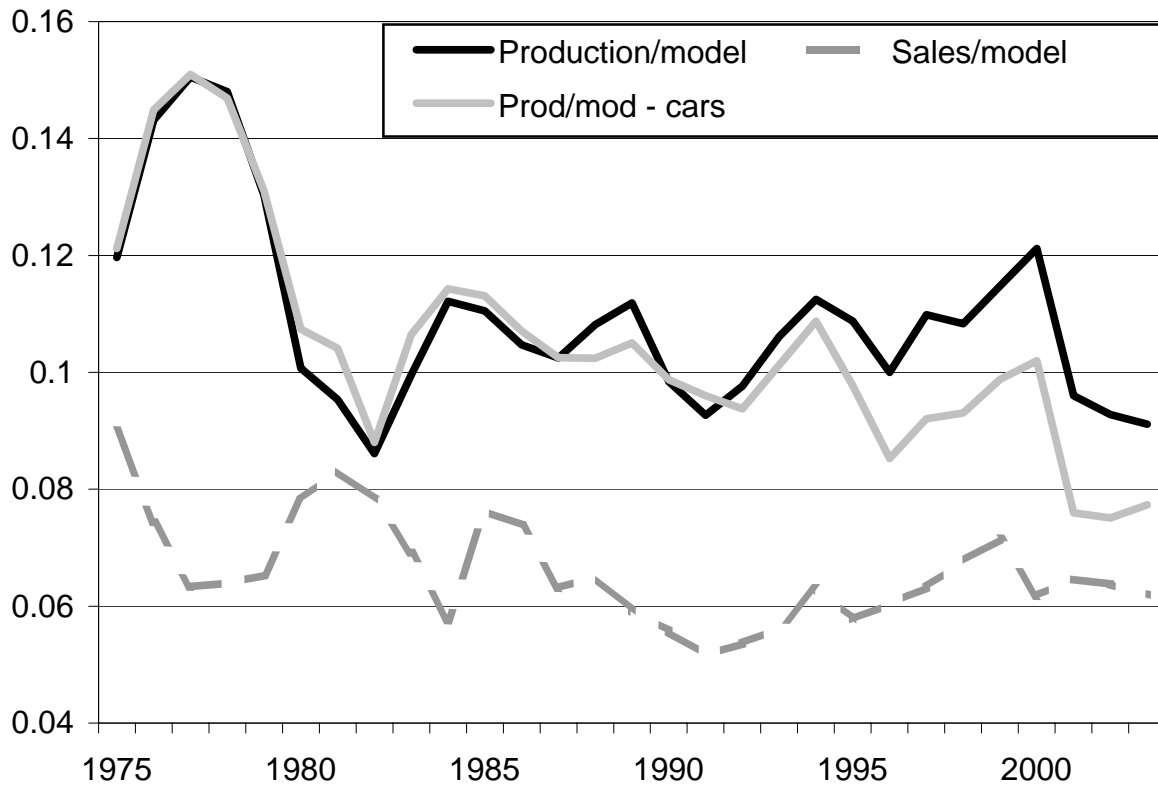
## References

- Appel Molot, M. (2005). Location Incentives and Inter-state Competition for FDI: Bidding Wars in the Automotive Industry. In L. Eden and W. Dobson (Eds.), *Governance, Multinationals and Growth*. Edward Elgar.
- Arora, A. (1996). Testing for Complementarities in Reduced Form Regressions: A Note. *Economic Letters* 50(1), 51–55.
- Arora, A. and A. Gambardella (1990, June). Complementarity and External Linkages: The Strategies of the Large Firms in Biotechnology. *Journal of Industrial Economics* 37(4), 361–79.
- Åstebro, T., M. G. Colombo, and R. Seri (2005, March). The Diffusion of Complementarity Technologies: An Empirical Test. Working Paper.
- Athey, S. and S. Stern (2003, June). An Empirical Framework for Testing Theories about Complementarity in Organizational Design. Working Paper.
- Automotive News* (1996, June 26). Once, a Ford Was a Ford; Today, It Has Many Shapes, Sizes and Names, (by Matt Roush). *Crain Communications Inc.*
- Automotive News* (2002, October 14). Honda Leads the Pack in Plant Flexibility (by Lindsay Chappell). *Crain Communications Inc.*
- Berry, S., J. Levinsohn, and A. Pakes (1995). Automobile Prices in Market Equilibrium. *Econometrica* 63(4), 841–90.
- Friedlaender, A. F., C. Winston, and K. Wang (1983). Costs, Technology, and Productivity in the U.S. Automobile Industry. *Bell Journal of Economics* 14(1), 1–20.
- Gal-or, E. (2002, October). Flexible Manufacturing Systems and the Internal Structure of the Firm. *International Journal of Industrial Organization* 20(8), 1061–96.
- Girma, S. and H. Gorg (2004, November). Outsourcing, Foreign Ownership, and Productivity: Evidence from UK Establishment-Level. *Review of International Economics* 12(5), 817–32.
- Harbour Consulting (2003). *The Harbour Report: Competitive Assessment of the North American Automotive Industry*. Rochester, MI: Harbour Consulting.
- Helper, S. (1997, May). Complementarity and Cost Reduction: Evidence from the Auto Supply Industry. *NBER Working paper*, No. 6033.
- Ichniowski, C., K. Shaw, and G. Prennushi (1997, June). The Effects of Human Resource Management Practices on Productivity: A Study of Steel Finishing Lines. *American Economic Review* 87(3), 291–313.

- Jovanovic, B. and D. Stolyarov (1997, January). Learning, Complementarities and Asynchronous Use of Technology. *NBER Working paper*, No. 5870.
- Jovanovic, B. and D. Stolyarov (2000, March). Optimal Adoption of Complementary Technologies. *American Economic Review* 90(1), 15–29.
- Macduffie, J. P. (1995, January). Human Resource Bundles and Manufacturing Performance: Organizational Logic and Flexible Production Systems in the World Auto Industry. *Industrial and Labor Relations Review* 48(2), 197–221.
- Milgrom, P. and J. Roberts (1990, June). The Economics of Modern Manufacturing: Technology, Strategy, and Organization. *American Economic Review* 80(3), 511–528.
- Miravete, E. J. and J. C. Pernías (2005). Innovation Complementarity and Scale of Production. *Journal of Industrial Economics* forthcoming.
- Norman, G. (2002, July). The Relative Advantages of Flexible versus Designated Manufacturing Technologies. *Regional Science and Urban Economics* 32(4), 419–45.
- Novak, S. and S. Stern (2003, November). Complementarity Among Vertical Integration Decisions: Evidence from Automobile Product Development. Working Paper.
- Petrin, A. (2002). Quantifying the Benefits of New Products: The case of Minivan. *Journal of Political Economy* 110(4), 705–29.
- Rubenstein, J. (1992). *The Changing U.S. Auto Industry: A Geographical Analysis*. New York: Routledge.
- Van Biesebroeck, J. (2003, January). Productivity Dynamics with Technology Choice: An Application to Automobile Assembly. *Review of Economic Studies* 70(1), 167–198.
- Van Biesebroeck, J. (2006). Trends and complementarities in the Canadian Automobile Industry. In Z. Chen and M. Duhamel (Eds.), *Industrial Economics and Performance in Canada*, pp. forthcoming. Ottawa: Industry Canada.
- Ward's AutoWorld* (2001, March 1). Rethinking Platform Engineering. *Primedia*.

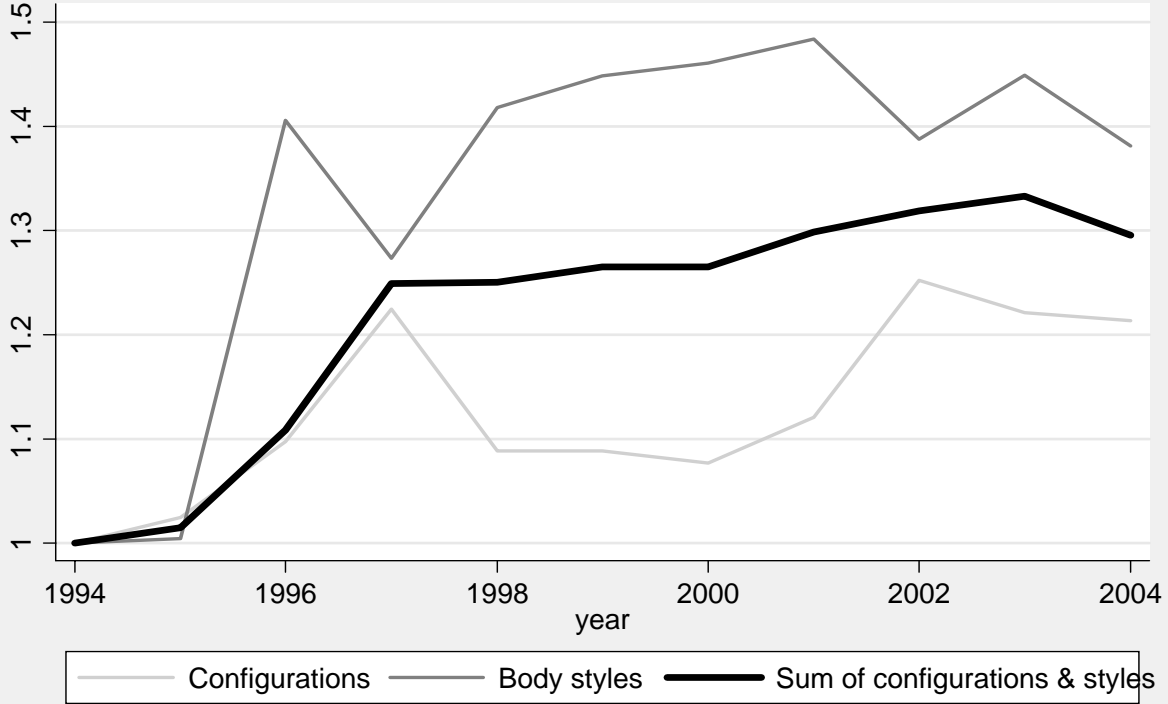


**Figure 1: Decreasing sales and production per model**



Source: Own calculations based on statistics from *Ward's Automotive Yearbook*

Figure 2: Model variety produced per plant has increased over time (1994=1)



Note: Average variety across all plants, where variety is defined as body styles, chassis configurations, or the sum of both

Source: Own calculations based on plant-level data from *The Harbour Report* (various years)

Figure 3A - Flexibility 1 : Models per platform (firm average)

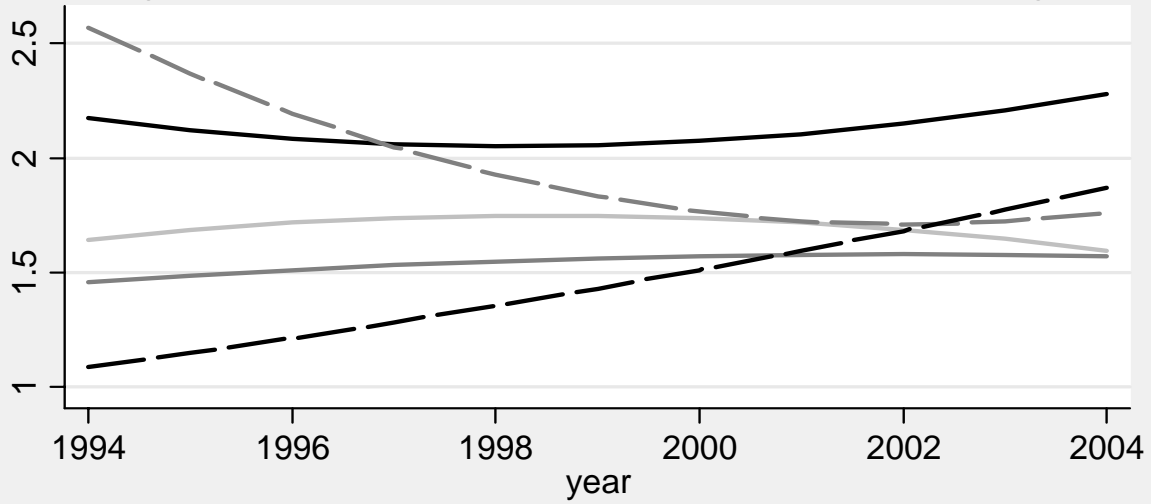
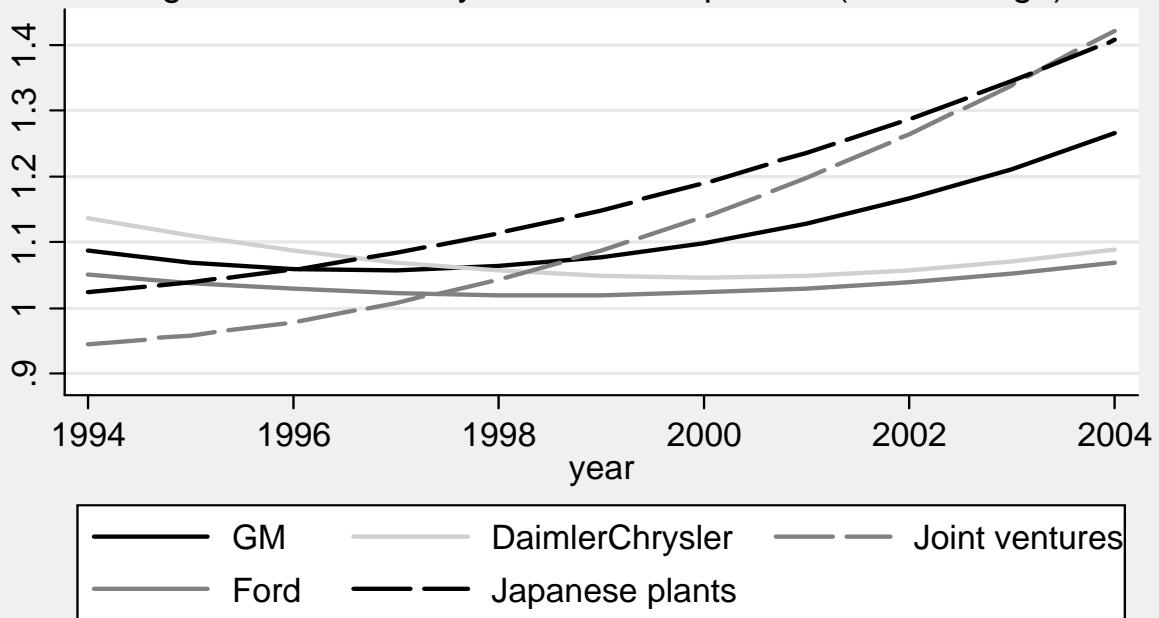
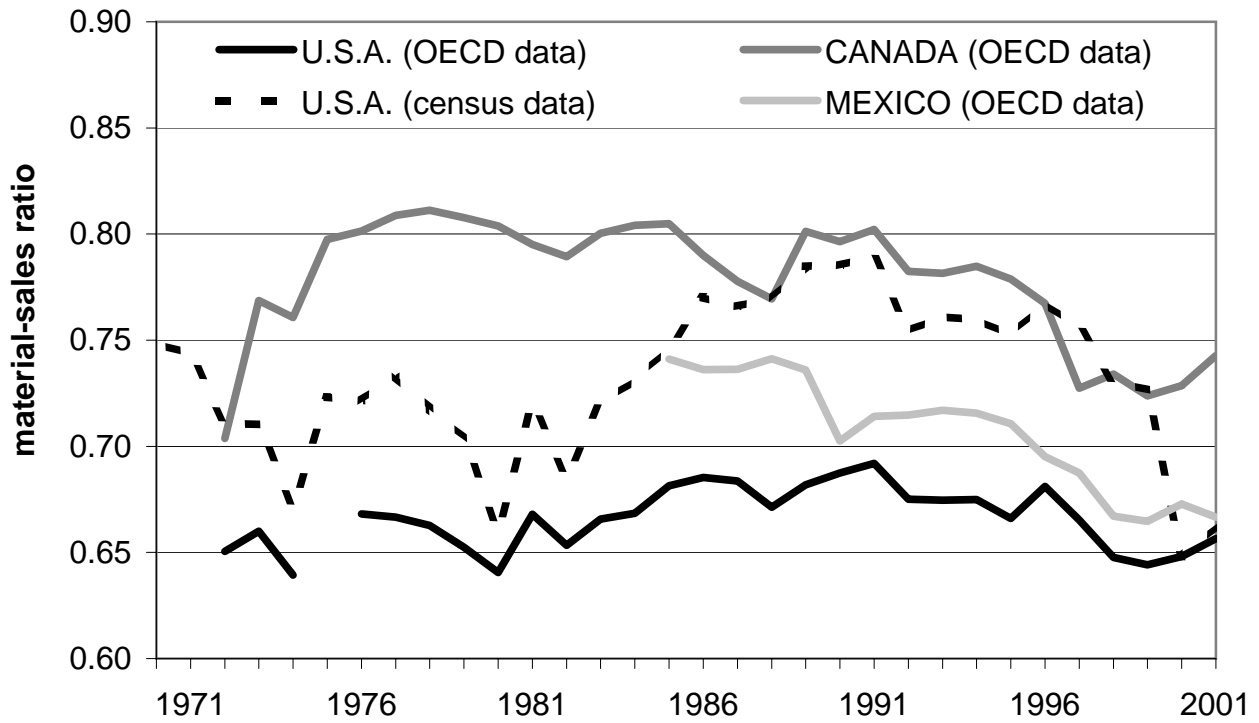


Figure 3B - Flexibility 2 : Platforms per line (firm average)



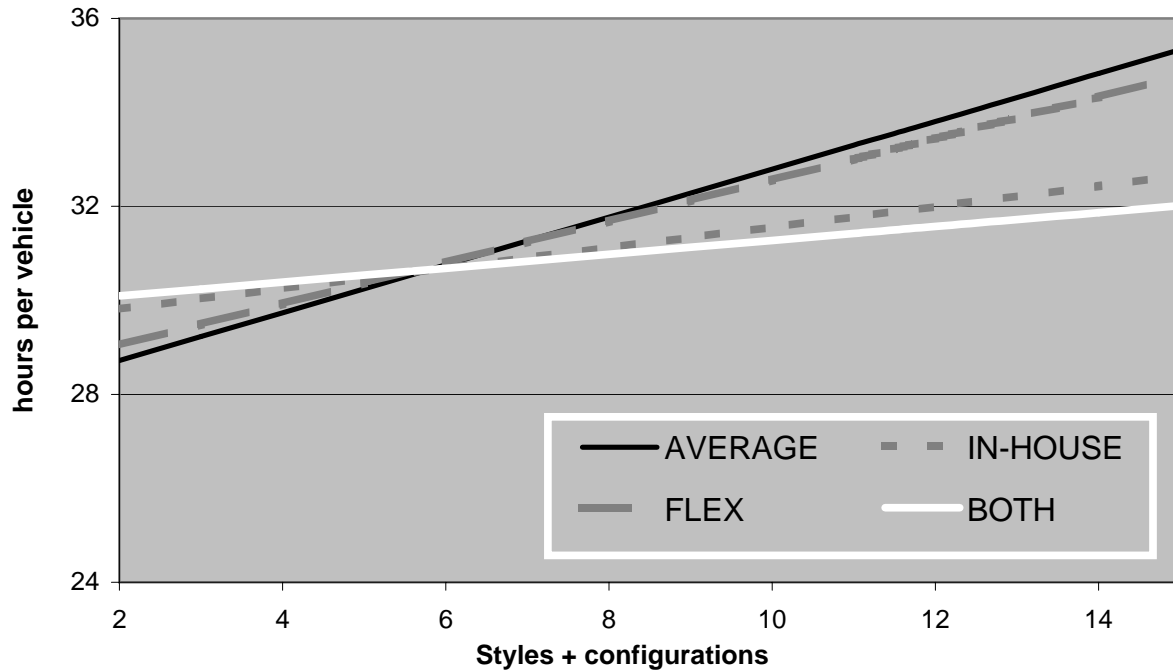
Note: Cubic time trend on the models per platform (top panel) or platforms per production line (bottom panel), estimated separately for all North American plants of the respective firms or groups of firms. Source: Own calculations based on plant-level data from *The Harbour Report* (various years)

**Figure 4: Outsourcing trends in North America**  
**(average material-sales ratio at the plant-level)**

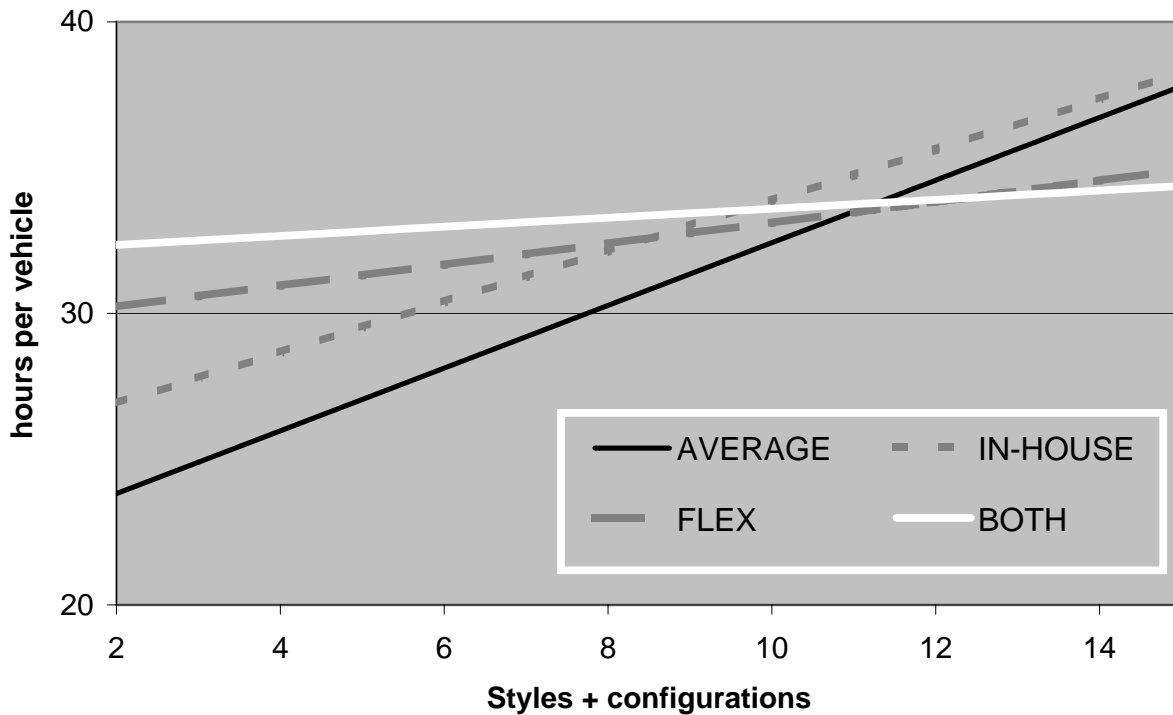


Note: Materials-sales ratio for the "Motor vehicles" industry (SIC 371) based on statistics from the OECD STAN database (solid lines) and for the subsector "Motor vehicles and passenger car bodies" (SIC 3711) using data from the U.S. Bureau of the Census (dashed line).

**Figure A.1a: Predicted hours per vehicle using OLS estimates**



**Figure A.1b: Predicted hours per vehicle using GMM estimates**



Note: Own calculations based on coefficient estimates from column (1) in Table 6 (top panel) and column (4) in Table 6 (bottom panel).

**Table 1: Summary statistics (1994-2004)**

	Mean	Std. Dev.	Min	Max
<i>Dependent variable:</i>				
Hours per vehicle ( <i>h<sub>pv</sub></i> )	30.13	11.41	15.69	108.51
<i>Activities:</i>				
Number of chassis configurations	2.67	2.46	0.75	23.10
Number of body styles	2.84	2.21	1.00	16.00
Number of (configurations + styles)	5.51	4.08	1.75	39.10
Number of models	2.11	1.04	1.00	6.00
Number of platforms	1.23	0.54	1.00	5.00
Flexibility index	1.09	0.34	0.50	3.33
In-sourcing index (inverse of outsourcing)	0.36	0.17	0.00	0.92
<i>Instruments:</i>				
Confs + Styles in competing plants	5.41	2.62	2.18	13.73
Distance (from NA center, miles)	517.57	406.20	61.99	1915.07
Distance (from country center, miles)	309.75	314.71	17.45	1918.01
Area	2449690	974434	216345	5501700
Mass relief	0.41	0.49	0.00	1.00
<i>Controls:</i>				
Log capacity	12.15	0.44	9.30	13.07
Pre-1998 dummy	0.28	0.45	0.00	1.00
Canadian plant	0.20	0.39	0.00	1.00
Mexican plant	0.09	0.29	0.00	1.00
Foreign-owned plant	0.17	0.37	0.00	1.00
Number of observations	860			
Number of unique plants	92			

Source: *The Harbour Reports* (various years)

**Table 2: Number of models sold and/or produced in North America has increased substantially**

	1974	1984	sample period	
			1994	2004
Models for sale in North America	185	228	273	320
Models for sale in United States	133	195	238	282
Car models	96	140	164	167
Car variations			468 *	503
Light truck models	37	55	74	115
Light truck variations			558 *	1805
Models produced in North America	90	125	139	165
Assembly plants in North America	68	76	68	64

\* Note: 1996

Source: *Ward's Automotive Yearbook* (various years) and *Ward's Infobank* (2004)

**Table 3: Production variety in the sample**

	simple sum by plant		take out duplicates	
	1994	2004	1994	2004
<b>Platforms</b>	101	98	70	75
<b>Models</b>	172	175	129	162
<b>Body styles</b>	175	235	143	206
<b>Chassis configurations</b>	182	206	147	176

Source: Own calculations based on statistics from *The Harbour Reports* (various years)

**Table 4: Outsourcing indicators in the (plant level) sample**

year	Fraction of tasks performed in-house			Number of plants that			Sum of changes across all tasks/plants			
	all activities (1)	usually outsourced (2)	often in-house (3)	all plants (4)	added tasks (5)	deleted tasked (6)	did not change (7)	net changes (8)	all additions (9)	all deletions (10)
1994	36.0%	10.1%	56.2%	82						
1995	36.7%	9.8%	58.2%	84	16	13	61	8	23	-15
1996	34.7%	10.3%	53.3%	79	15	6	59	9	15	-6
1997	36.5%	11.3%	55.7%	78	29	16	37	27	53	-26
1998	38.5%	12.5%	58.9%	77	29	25	34	19	60	-41
1999	37.1%	12.3%	56.0%	77	20	19	42	-9	23	-32
2000	36.3%	12.0%	54.7%	77	11	19	53	-16	15	-31
2001	36.3%	11.7%	54.9%	78	19	16	46	3	24	-21
2002	35.7%	11.0%	54.6%	77	12	14	53	2	19	-17
2003	35.5%	11.0%	54.1%	76	6	9	64	-5	14	-19
2004	35.6%	11.8%	53.3%	75	14	12	54	-3	23	-26

Source: Own calculations based on statistics from *The Harbour Reports* (various years)



**Table 5: Productivity penalty associated with production of greater variety**

	Dependent variable: hours per vehicle					
	(1)	(2)	(3)	(4)	(5)	(6)
<u>Measures of Variety:</u>						
Platforms	2.522*** (0.564)					
Models		1.612*** (0.299)				
Chassis Configurations			0.564*** (0.135)			
Body Styles				0.437*** (0.158)		
Configurations + Styles					0.353*** (0.082)	
Configurations * Styles						0.038** (0.015)
Year * Variety	-0.239 (0.147)	-0.334*** (0.077)	-0.067 (0.043)	-0.091* (0.048)	-0.055** (0.024)	-0.003 (0.005)
<u>Controls</u>						
Year	-0.667*** (0.234)	-0.241 (0.217)	-0.754*** (0.177)	-0.707*** (0.192)	-0.649*** (0.191)	-0.749*** (0.145)
Pre-1998 dummy	-0.393 (0.948)	-0.461 (0.935)	-0.268 (0.945)	-0.053 (0.959)	-0.076 (0.948)	-0.318 (0.954)
Scale (log Capacity)	-12.192*** (0.731)	-12.077*** (0.723)	11.853*** (0.728)	-12.028*** (0.741)	-11.998*** (0.733)	-11.915*** (0.736)
Canada	-1.343* (0.749)	-1.681** (0.729)	-1.555** (0.741)	-1.738** (0.747)	-1.564** (0.742)	-1.724** (0.747)
Mexico	9.260*** (1.113)	8.802*** (1.103)	9.654*** (1.095)	10.184*** (1.095)	9.683*** (1.095)	10.164*** (1.099)
Foreign-owned	-6.357*** (0.778)	-5.899*** (0.759)	-5.821*** (0.771)	-5.941*** (0.778)	-5.831*** (0.771)	-5.948*** (0.778)
mid-size cars	1.586 (1.025)	0.921 (1.012)	1.385 (1.022)	1.702 (1.034)	1.575 (1.023)	1.511 (1.031)
sport cars	4.319*** (1.510)	4.446*** (1.482)	4.441*** (1.504)	4.518*** (1.521)	4.419*** (1.507)	4.513*** (1.519)
luxury cars	7.361*** (1.628)	7.041*** (1.577)	9.277*** (1.520)	9.833*** (1.530)	9.453*** (1.519)	9.751*** (1.530)
pickup trucks	-0.874 (1.000)	-0.033 (0.994)	-2.429** (1.073)	-1.768 (1.088)	-2.625** (1.104)	-1.600 (1.062)
SUVs	2.833*** (1.074)	2.029* (1.076)	1.849* (1.121)	2.992*** (1.082)	2.273** (1.086)	2.814*** (1.085)
minivans	5.861*** (1.397)	4.955*** (1.391)	5.703*** (1.393)	5.688*** (1.408)	5.604*** (1.396)	5.833*** (1.407)
vans	5.552*** (1.388)	5.517*** (1.366)	4.464*** (1.431)	4.792*** (1.466)	4.063*** (1.460)	5.086*** (1.436)
Constant	1508.3*** 467.5	653.6 433.8	1679.1*** 353.9	1587.4*** 384.1	1471.0*** 383.3	2016.7*** 289.6
Observations	860	860	860	860	860	860
R <sup>2</sup>	0.539	0.552	0.541	0.532	0.540	0.532

Note: Estimation by least squares on the entire sample; standard errors in parenthesis; \* significant at the 10% level, \*\* 5%, \*\*\* 1%

**Table 6: Productivity regressions under different exogeneity assumptions**

Estimation method: (instrumented)	Dependent variable: hours per vehicle				
	OLS (1)	GMM (Flex) (2)	GMM (Insource) (3)	GMM (Ins, Flex) (4)	GMM (Ins, Flex, Var) (5)
Variety (Configurations + Styles)	1.618*** (0.436)	2.729** (1.345)	2.424** (0.959)	4.915*** (1.625)	2.261 (3.056)
Insourcing (1-Outsourcing)	9.455*** (2.860)	7.453* (4.467)	11.584 (8.345)	13.203*** (3.955)	13.072*** (3.498)
Flexibility (Platforms per line)	2.600 (2.012)	15.233 (10.557)	14.529** (5.633)	22.413*** (6.755)	14.627** (7.389)
Variety * Insourcing	-1.622*** (0.488)	-1.272 (0.866)	-1.540 (1.580)	-0.638** (0.280)	-0.828** (0.344)
Variety * Flexibility	-0.292 (0.273)	-1.654 (1.143)	-1.639** (0.749)	-3.845*** (1.544)	-1.358 (1.079)
Flexibility * Insourcing	-1.139*** (0.334)	-1.243*** (0.348)	-24.424** (11.710)	-21.321*** (7.329)	-17.518*** (5.887)
R <sup>2</sup>	0.554	0.531	0.453	0.141	0.100
F-test for instrument validity: F(k,860-k) test statistic (p-value in brackets)					
Variety					13.74 (0.00)
Insourcing			3.68 (0.00)	2.70 (0.00)	4.44 (0.00)
Flexibility		1.91 (0.05)		3.19 (0.00)	2.10 (0.02)
Variety * Insourcing			2.07 (0.03)	3.00 (0.00)	1.83 (0.04)
Variety * Flexibility		2.42 (0.01)		5.00 (0.00)	5.95 (0.00)
Flexibility * Insourcing		16.21 (0.00)	2.31 (0.03)	2.03 (0.02)	3.31 (0.00)
Hansen J statistic for overidentification test of all instruments: $\chi^2(k)$ test statistic (p-value in brackets) (k differs by column)					
		5.56 (0.14)	4.10 (0.25)	9.01 (0.25)	8.09 (0.32)
Hausman specification tests: $\chi^2(20)$ test statistic (p-value in brackets)					
against (5)	41.87 (0.03)	21.98 (0.34)	11.95 (0.92)	3.31 (>0.999)	
against (4)	15.62 (0.74)	12.84 (0.88)	6.74 (0.99)		
Omitting control variables from the Hausman specification test - $\chi^2(6)$ test statistic					
against (5)	22.50 (0.00)	12.58 (0.05)	7.34 (0.29)	1.98 (0.92)	
against (4)	9.54 (0.14)	8.20 (0.22)	5.99 (0.42)		

Note: Variety, flexibility and insourcing are measured as continuous variables. Instruments are the variety of competing firms (for variety), distance to industry center for North America and within the plant's country (for insourcing), and plant area and a dummy for *tag* shift relief (for flexibility). Interactions of instruments are included as well. The same set of controls as in Table 5 are included (reported in Table A.1 in the Appendix). \* significant at the 10%, \*\* 5%, \*\*\* 1%

**Table 7: Productivity regressions controlling for unobserved heterogeneity**

	Dependent variable: hours per vehicle			
	i.i.d. errors	IV (Insourcing & Flexibility)		
		Firm FE	Plant RE	Plant FE
	(4)	(6)	(7)	(8)
Variety	4.915*** (1.625)	6.081** (2.433)	4.268** (1.835)	7.476*** (1.742)
Insourcing	13.203*** (3.955)	28.505*** (8.830)	23.195** (9.709)	20.477** (8.757)
Flexibility	22.413*** (6.755)	10.307** (5.622)	8.535** (4.353)	10.711** (4.714)
Variety x Insourcing	-0.638** (0.280)	-4.401* (2.545)	-4.794* (2.717)	-7.044** (2.523)
Variety x Flexibility	-3.845*** (1.544)	-3.321* (1.888)	-1.393 (1.386)	-2.984*** (1.111)
Flexibility x Insourcing	-21.321*** (7.329)	-14.772* (8.474)	-12.124 (11.254)	-9.300 (8.272)
R <sup>2</sup>	0.141	0.279	0.312	0.619
F-test for instrument validity: F(k,860-k) test statistic (p-value in brackets)				
Insourcing	2.70 (0.00)	2.23 (0.01)	2.37 (0.00)	2.57 (0.00)
Flexibility	3.19 (0.00)	3.11 (0.00)	2.28 (0.01)	2.40 (0.00)
Variety * Insourcing	3.00 (0.00)	4.41 (0.00)	3.33 (0.00)	3.88 (0.00)
Variety * Flexibility	5.00 (0.00)	3.81 (0.00)	3.40 (0.00)	2.50 (0.00)
Flexibility * Insourcing	2.03 (0.02)	3.27 (0.00)	2.88 (0.00)	2.46 (0.00)

Note: Estimation is with GMM, using instruments for insourcing and flexibility. The same set of controls as in Table 5 are included (reported in Table A.1). Results in column (4) treat all errors as i.i.d., in column (6) firm-fixed effects are included, in column (7) errors are allowed to be correlated across plants (random effects estimator), and in column (8) plant-fixed effects are included. \* significant at 10%, \*\* 5%, \*\*\* 1%

**Table 8: Robustness check on complementarities**

	Dependent variable: hours per vehicle					
	full sample	log <i>h<sub>pv</sub></i>	no pre-1998	no Mexico	discrete activities	add scale control
	(4)	(9)	(10)	(11)	(12)	(13)
Variety	4.915*** (1.625)	0.199*** (0.050)	5.333*** (1.715)	11.198*** (3.418)	3.838*** (1.201)	5.382 (3.989)
Insourcing	13.203*** (3.955)	0.991*** (0.347)	10.417 (11.643)	29.887** (12.837)	8.996** (4.353)	13.237*** (5.181)
Flexibility	22.413*** (6.755)	0.563*** (0.157)	19.523*** (6.009)	13.018* (7.391)	32.347** (15.575)	17.026*** (5.891)
Variety x Insourcing	-0.638** (0.280)	-0.025*** (0.008)	-0.426* (0.196)	-1.132*** (0.321)	-4.215*** (1.463)	-0.736** (0.310)
Variety x Flexibility	-3.845*** (1.544)	-0.136*** (0.051)	-5.060*** (1.747)	-3.782* (1.947)	-6.220** (2.837)	-4.102** (1.760)
Flexibility x Insourcing	-21.321*** (7.329)	-0.513** (0.207)	-13.446* (7.887)	-10.008 (14.928)	-6.989** (3.469)	-22.806*** (8.149)
Variety x Scale						-0.064 (0.253)
Number of observation:	860	860	615	780	860	860
R <sup>2</sup>	0.141	0.072	0.104	0.091	0.127	0.141

Note: Estimation is with GMM, using instruments for insourcing and flexibility. The same controls as in Table 5 are included (reported in Table A.1). \* significant at 10%, \*\* 5%, \*\*\* 1%

**Table 9: Direct estimation of the first order conditions for the three activities**

Dependent variable:	Model variety	Flexible technology	Insourcing
	(i)	(ii)	(iii)
<i><u>Endogenous variables</u></i>			
Variety (in own plant)		0.019*** (0.006)	0.021*** (0.007)
Insourcing	2.003 (8.208)	-0.424 (0.407)	
Flexible technology	13.512*** (5.022)		-0.187 (0.154)
<i><u>Instruments</u></i>			
Variety (of competitors)	0.769*** (0.162)		
Area		-0.014 (0.022)	
Shift relief		0.090* (0.045)	
Distance to NA center			-0.035 (0.063)
Distance to country-center			0.117** (0.063)
<i><u>Controls</u></i>			
Scale	0.884*	-0.020	-0.007
Year	-0.197*	0.016*	-0.002
Pre-1998 dummy	-0.804	0.045	-0.007
Canada	-0.163	-0.032	0.030
Mexico	1.029	0.052	0.001
Foreign-owned	-1.969	0.155*	0.159***
Stretched platform	3.559***	-0.228***	-0.057
mid-size cars	-0.571	0.036	0.052**
sport cars	0.662	-0.032	-0.005
luxury cars	-4.500***	0.283***	-0.070
minivans	-1.663	0.111	0.091**
SUVs	-0.574	0.006	-0.038
pickup trucks	1.982**	-0.110**	-0.010
vans	1.638	-0.047	0.019
Constant (U.S. location and ownership)	368.778	-31.110*	5.102
R <sup>2</sup>	0.086	0.105	0.111

Note: Estimation is with GMM using the excluded instruments in each equation as instruments for the included endogenous variables: variety, insourcing and/or flexibility. \* significant at 10%, \*\* 5%, \*\*\* 1%

**Table A.1: Coefficient estimates on the control variables for all regressions in Table 6 to Table 8**

Dependent variable: hours per vehicle													
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Year	-0.922***	-0.957***	-1.528***	-1.159***	-1.436***	-0.625***	-0.820***	-0.898***	-0.038***	-1.232***	-0.889***	-0.587*	-1.155***
	(0.134)	(0.165)	(0.546)	(0.223)	(0.275)	(0.232)	(0.176)	(0.179)	(0.006)	(0.218)	(0.205)	(0.322)	(0.238)
Pre-1998 dummy	-0.024	-0.143	-5.568	-3.366**	-4.137**	-0.273	-0.739	-0.280	-0.112**		-3.020**	-1.317	-3.518**
	(0.945)	(1.065)	(3.669)	(1.625)	(1.615)	(1.462)	(1.096)	(1.032)	(0.044)		(1.302)	(1.522)	(1.780)
Scale	-11.880***	-11.230***	-10.731***	-10.488***	-11.047***	-12.376***	-4.414**	-0.748	-0.269***	-9.741***	-9.452***	-9.580***	-10.395***
	(0.723)	(1.078)	(1.395)	(1.326)	(1.354)	(1.278)	(1.806)	(2.251)	(0.034)	(1.327)	(1.369)	(1.627)	(1.916)
Canada	-1.400*	-1.341**	-3.692***	-4.028***	-2.337**	-3.603***	-1.323		-0.095***	-3.359***	-0.679	-0.179	-4.347***
	(0.736)	(0.669)	(1.292)	(1.104)	(1.186)	(1.297)	(2.470)		(0.030)	(1.176)	(0.709)	(0.995)	(1.201)
Mexico	9.464***	8.722***	11.963***	11.606***	8.780***	11.838***	19.721***		0.297***	5.247***		10.150***	11.980***
	(1.087)	(1.673)	(2.591)	(2.265)	(2.276)	(2.368)	(4.200)		(0.053)	(1.775)		(2.302)	(2.362)
Foreign-owned	-6.721***	-6.948***	15.937	10.672	6.763	0.000	-1.368		0.178	6.391	-8.231***	-8.413*	-12.615
	(0.812)	(0.932)	(10.331)	(7.352)	(5.138)	(0.000)	(10.522)		(0.192)	(7.422)	(2.259)	(4.308)	(8.348)
Stretched platform	2.039***	2.971**	2.208**	2.518*	4.148**	-1.805	0.913	2.294	0.112**	3.350**	3.320***	1.680	1.782
	(0.750)	(1.495)	(1.025)	(1.520)	(1.635)	(2.133)	(2.328)	(2.467)	(0.045)	(1.696)	(0.866)	(1.572)	(1.559)
mid-size cars	1.663	1.370	2.659	1.587	1.462	2.361	7.229**	8.862	-0.033	-1.452	0.560	4.178**	1.908
	(1.014)	(1.061)	(1.843)	(1.640)	(1.570)	(1.581)	(3.623)	(7.192)	(0.044)	(1.721)	(1.325)	(1.724)	(1.905)
sport cars	4.504***	4.982***	4.680**	5.143**	4.617**	4.770**	4.450	-2.644	0.103*	3.050	5.032**	8.740***	5.259**
	(1.488)	(1.743)	(2.318)	(2.266)	(2.208)	(2.318)	(4.120)	(7.585)	(0.058)	(2.556)	(1.972)	(2.769)	(2.383)
luxury cars	9.138***	7.479***	-2.433	-0.957	-2.286	12.718***	10.224*	7.400	-0.047	-4.851	6.059*	3.516	-0.396
	(1.575)	(2.850)	(5.369)	(4.097)	(3.964)	(4.605)	(6.208)	(9.198)	(0.114)	(5.020)	(3.180)	(4.858)	(4.285)
pickup trucks	-2.404**	-1.733	-3.468*	-4.434**	-6.479***	-5.052**	1.859	4.985	-0.134***	-0.463	-3.290**	-5.284***	-4.544**
	(1.100)	(1.373)	(1.986)	(1.783)	(2.008)	(2.193)	(3.042)	(8.776)	(0.051)	(2.096)	(1.456)	(1.576)	(1.980)
SUVs	2.012*	2.161**	-0.742	-0.107	-1.552	4.200**	9.942***	13.693*	-0.014	-0.281	2.486*	4.449**	-0.060
	(1.095)	(0.992)	(1.866)	(1.597)	(1.799)	(1.799)	(3.320)	(8.053)	(0.045)	(1.778)	(1.447)	(1.881)	(1.852)
minivans	4.691***	4.519***	6.179***	4.786**	3.138*	5.857***	10.291*	14.459	0.112**	4.002*	2.897	3.165*	5.501**
	(1.424)	(1.264)	(2.358)	(2.023)	(1.799)	(2.182)	(5.470)	(11.789)	(0.053)	(2.234)	(1.991)	(1.838)	(2.416)
vans	4.442***	5.423***	4.611*	3.429*	-0.212	0.701	12.623***	13.648	0.107*	6.434***	5.434***	3.955**	3.212
	(1.450)	(1.577)	(2.770)	(1.967)	(2.586)	(2.499)	(4.431)	(12.589)	(0.055)	(2.410)	(1.893)	(1.917)	(2.159)
Constant	2005.8***	2057.5***	3175.0***	2415.0***	2984.6***	1366.5***	1663.9***	1766.2***	80.4***	2558.7***	1845.5***	1300.6***	2410.4***
	(269.2)	(327.8)	(1081.7)	(443.1)	(557.4)	(465.5)	(362.3)	(375.0)	(12.7)	(430.0)	(409.2)	(655.9)	(476.5)

Note: Column numbers refer to Tables 6 to 8; sample and estimation methods are indicated in the respective tables; \* significant at 10%, \*\* 5%, \*\*\* 1%