Working Paper 23-061

Organizational Responses to Product Cycles

Achyuta Adhvaryu
Vittorio Bassi
Anant Nyshadham
Jorge Tamayo
Nicolas Torres

Organizational Responses to Product Cycles

Achyuta Adhvaryu
University of California, San Diego

Vittorio Bassi
University of Southern California

Anant Nyshadham
University of Michigan

Jorge Tamayo
Harvard Business School

Nicolas Torres
Good Business Lab

Working Paper 23-061

Funding for this research was provided in part by Harvard Business School.

Organizational Responses to Product Cycles*

Achyuta Adhvaryu† Vittorio Bassi‡ Anant Nyshadham§

Jorge Tamayo¶ Nicolas Torres‖

March 21, 2023

Abstract

Product cycles entail the mass production of new – and often increasingly complex – products on a regular basis. How do firms manage these changes? We use granular daily data from a leading automobile manufacturer to study the organizational impacts of introducing new models to the auto assembly line. We show that the numbers of vehicles and parts do not change after a new model is introduced; the main change is a large, discontinuous increase in *new* parts. The product cycle thus necessitates dealing with new complex problems: we accordingly show that defects increase substantially after the production change, then decrease to prior levels over about three weeks. We next ask how the firm's organization facilitates this problem-solving. We find that the firm trains and promotes mid-level employees to manage the production of new parts, but waits to back-fill mid-level positions until after defect rates recover. That is, the firm reduces the distance – in terms of knowledge layers – between front-line workers, who are dealing with these new tasks, and managers further up the hierarchy, who have the necessary knowledge to solve complex problems. This organizational response is unique to the product cycle: when the firm increases the scale of production of existing models, it adds layers to the hierarchy without increasing average skills, as predicted by canonical models of knowledge-based hierarchies. We develop an extension to this theory that reconciles our novel empirical results on product cycle responses.

JEL Codes: D22, M12, M53, O3
Keywords: product cycles, organizational behavior, knowledge hierarchies, training, worker skills, auto manufacturing, Argentina

*We would like to thank Daron Acemoglu, Robert Clark, Larry Katz, Amit Khandelwal, Andrea Prat, Esteban Rossi-Hansberg, Raffaella Sadun, Jesse Shapiro, Eric Verhoogen, Tom Wollmann, as well as seminar participants at UCSC, HBS and the 2022 IPA-CDEP Entrepreneurship and Private Sector Development Working Group Meeting for insightful comments. All errors are our own.

†UC San Diego, NBER, J-PAL, Good Business Lab & BREAD; email: aadhvaryu@ucsd.edu.

‡University of Southern California, J-PAL, BREAD, CEPR & IGC; email: vbassi@usc.edu.

§University of Michigan, NBER, J-PAL, Good Business Lab & BREAD; email: nyshadha@umich.edu.

¶Harvard Business School, Digital Reskilling Lab - The Digital, Data, and Design Institute at Harvard; email: jtamayo@hbs.edu.

‖Good Business Lab; email: nicolas.t@goodbusinesslab.org.

1 Introduction

Product cycles are a fundamental feature of innovation, growth, and survival in many industries (Grossman and Helpman, 1991a,b, 1994; Krugman, 1979; Vernon, 1966). Firms develop new generations of products to remain profitable when the production of current generations becomes imitable. This phenomenon often entails the mass production of new – and often increasingly complex – products on a regular basis. In the automotive industry, for example, the Toyota Corolla, the Honda CR-V, and the Ford F-150 – indeed, most mass-market cars – are completely redesigned every five or six years to keep up with competition. The product cycle for electronics is even shorter, with significant upgrades made to staple products like smartphones and personal computers every 2 to 3 years (Bayus, 1994).

The frequency and intensity of product cycles demand substantial adaptation from upstream suppliers or subsidiaries, who must deal with new and increasingly complex parts and processes. We focus on two main ways in which these firms may respond. First, they may increase the knowledge of their workers to manage these changes through training provision: a long literature has studied the degree to which firms engage in on-the-job training (Acemoglu and Pischke, 1998, 1999; Becker, 1962), with recent empirical studies identifying large potential returns to such training (Adhvaryu et al., 2018; Espinosa and Stanton, 2022; Sandvik et al., 2020). Second, firms may change their internal organizational structure to facilitate problem solving: a parallel literature has modeled and empirically tested predictions on how the distribution of knowledge within the firm might change as the firm re-optimizes under new production goals (Caliendo et al., 2020, 2015; Caliendo and Rossi-Hansberg, 2012; Garicano, 2000; Garicano and Rossi-Hansberg, 2006).[1]

We draw from both these strands to study how firms manage production changes arising from the product cycle. We focus on the production of new models of automobiles. Leveraging

[1] A third possibility is of course that firms do not respond to product cycles in any significant way, keeping organizational structures and the intensity of training investments constant and relying primarily instead on the speed of learning by doing. A large literature has highlighted the important role of learning by doing in organizations (Adhvaryu et al., 2019; Arrow, 1962; Irwin and Klenow, 1994; Levitt et al., 2013).

daily administrative data from a leading global auto manufacturer and using event study and discontinuity-based methods, our core contribution is to trace out the organizational impacts of product cycles, in terms of the resulting impacts on training provision and the internal structure of the organization.

Our data comes from an Argentinian subsidiary plant of the auto manufacturer. We begin by showing that demand (number of vehicles) and number of total parts do not change in the short term after a new model is introduced; the main change is a large, discontinuous increase in new parts: after a model change, about 13-15% of the parts that need to be assembled are new. The production of new models thus necessitates substantial acquisition and distribution of new knowledge. As a result, we show that defects per vehicle spike immediately by up to 2 standard deviations after the model change, and decrease to their prior level over approximately a three-week period. This increase in defects created by new product introduction, though it is temporary, is indeed costly to the firm, as it results in an unrecoverable reduction in cars produced.

We next ask how changes in the firm's organizational decision-making facilitate the problem-solving required to bring defects back down to a minimum in such a short time. Using granular data on the hierarchical structure of working groups, we find that when a new model is introduced on the assembly line, the firm promotes mid-level employees to higher level positions to manage the production of new parts, but waits to back-fill mid-level positions until after defect rates recover to pre-model-change levels. That is, we find a significant reduction in the number of hierarchical layers within working groups, which decrease by about 1 layer (down from an average of about 6) in the first three weeks after the change. In doing so, the firm reduces the distance – in terms of knowledge layers – between front-line workers, who are dealing with these new tasks, and managers farther up the hierarchy, who have the necessary knowledge to help solve these new complex problems. This promotion is achieved through training provided internally by the firm, which improves the ability of higher level managers to solve the new complex problems arising on the production line. As a result, the aggregate level of skill capital within the firm increases. In sum, we document that both changes in organizational structure and training provision are important instruments used by the firm to deal with the increased complexity of production and bring defects down to pre-change levels fairly quickly.

We then contrast these results with the impacts of increases in quantity produced of the same model, for which we show the organizational response is a monotonic, permanent increase in both employment and knowledge layers, consistent with prior evidence from manufacturing in high-income countries (Caliendo et al., 2020, 2015). Importantly, the increase in knowledge layers is not accompanied by an increase in average skill capital through training when the change involves producing more of the same model.

We interpret our empirical results in light of the canonical models of knowledge hierarchies within the firm of Garicano (2000) and Caliendo and Rossi-Hansberg (2012). Exactly as in our results, as well as the empirical results of Caliendo et al. (2015) and Caliendo et al. (2020), the canonical theory predicts that an increase in the volume produced of the same model should result in an expansion of the hierarchy through the addition of more knowledge layers: as production expands, adding layers allows the firm to focus the scarce time of skilled managers on solving only the most complex problems. Importantly, however, this canonical theory *does not* predict the reduction in knowledge layers that we observe in response to the product cycle (i.e., the introduction of new models). In the canonical model, the increase in problem complexity resulting from model changes should yield again an increase in layers, as each worker can solve a smaller fraction of the problems (which are now more complex).

We therefore propose an extension to the canonical theory that is able to reconcile the effects on the organizational hierarchy and knowledge investments through training that we document as a response to model changes. To motivate the extension, we first show in the data that when a new model is introduced, the firm invests in improving its in-house training programs by adding courses and bringing in new trainers. We then modify the canonical theory, in line with this empirical observation, to allow firms to make a costly investment in increasing the efficacy of training, and show that this can lead the firm to optimally reduce the number of knowledge layers as a response to the introduction of new models. Intuitively, as the firm can train its workers more efficiently in problem solving, managers farther up the hierarchy are able to solve more problems, and fewer knowledge layers are thus needed to produce. We show that allowing for endogenous investment in the productivity of training can reconcile our empirical results on the impact of model changes on the internal organization of the firm. Importantly, we show that volume changes still lead to an increase in the number of knowledge layers in this modified version of the theory, which is again in line with our empirical findings.

We aim to contribute to the literature documenting the importance of management and organizational design in determining the productivity and growth of firms (Aghion et al., 2014; Bloom et al., 2013, 2010, 2016; Bloom and Van Reenen, 2007, 2010). In particular, we focus on the importance of knowledge flows and organizational flexibility in dealing with external stimuli. Previous work in this area has focused quite fruitfully on the effects of demand shocks (Caliendo et al., 2020, 2015; Garicano, 2000). We build on the elegant theory developed in this and earlier work (most directly, the foundational work of Garicano (2000) and Caliendo and Rossi-Hansberg (2012)) to demonstrate how product cycles – and the ideas of increasing complexity and learning embodied within them – generate quite clear, and distinct, movements within the knowledge hierarchy. We use granular data (at the shift-by-day as opposed to yearly level as in previous work) to document the surprising flexibility of organizational structure in response to product cycles. We uncover that organizational structure can evolve very rapidly in response to the introduction of new models, and that these substantial changes undertaken by the firm would be overlooked when relying on less granular data. We build an extension to the canonical theory that is critical to reconciling the theory's predictions with the empirically observed patterns related to the impacts of the product cycle on manufacturing firms.

Intuitively, there are dynamic elements and adjustment costs at play in this setting. In the short run, firms are confronted with changes in complexity arising from new models that require new knowledge. However, adjustment costs render quantity fixed in the short run. In line with this, the firm first retrains existing workers and reorganizes teams to deal with the higher complexity at the current quantity produced. After this phase, the firm can choose to expand production in the longer term if it wants. That is, when there is this adjustment cost to changing quantity, knowledge (training) and layers are substitutes in the short run, but complements in the long run. Accordingly, we also show that increases in volume lead to increases in layers, consistent with the original theoretical predictions and previous empirical results from firms in France and Portugal.

In documenting the role of firm-provided training in dealing with product cycles, we also contribute to a large literature on the returns to on-the-job training. This literature has highlighted that training provision to workers can have high returns for the organization (Adhvaryu et al., 2018; Espinosa and Stanton, 2022; Sandvik et al., 2020), as well as that firms may face challenges to allocating training efficiently within the organization (Adhvaryu et al., 2022; Sandvik et al., 2022). Our contribution is to show how training enables firms to deal with increased product complexity, and to highlight its interaction with changes in organizational structure in helping firms adapt to product cycles.[2]

We also add to the understanding of the economic consequences of product cycles. Much of this literature has focused on consumer-facing firms' decisions to innovate and the implications for trade of the product cycles created by continual innovation and imitation (Grossman and Helpman, 1991a,b, 1994). Less attention has been paid to suppliers, who play an increasingly important role in the execution of product cycles. In nearly all industries in which product cycles feature heavily, consumer-facing firms either contract with suppliers or have their own operations in the "global south". This means that the firms that actually deal with the continual increase in complexity generated by innovation are different from the firms actually doing the R&D. Understanding the behavior of these suppliers and their ability to adapt to

[2] A related literature studies the allocation of workers to mangers within organizations (Adhvaryu et al., 2020; Fenizia, 2022; Frederiksen et al., 2020; Limodio, 2021).

product cycles is thus of first-order interest to understand the global patterns of trade that are created by continual innovation.

The rest of the paper is organized as follows: Section 2 describes the setting of our study and the data used for estimation. In Section 3 we describe the empirical strategy, and Section 4 shows the results on the organizational response to discrete changes in both the models and the quantity of cars produced. In Section 5, we introduce a model to reconcile our results with standard models of knowledge hierarchies within the firm. Section 6 concludes. Additional details are in the Online Appendix.

2 Context, Data and Descriptives

In this section, we describe the setting where our study takes place and the data used for estimation.

2.1 Context, Organization of Production, and Product Cycles

We partnered with a leading global auto manufacturer subsidiary plant located in Argentina. The plant began operating in the 1990s and now produces more than 140,000 cars per year, employing over 5,000 workers. Around two-thirds of the production is exported to different countries in Latin America, and the rest is dedicated to the local market.

2.1.1 Production Process

Production takes place in a production line setting. The production line is made up of eight sectors: Press, Welding, Painting, Frame & RX Axle, Engines, Resin, Assembly, and Quality Check. The different parts of a production unit (e.g., a car) are produced in parallel by different sectors of the production line: chassis and car-body components are manufactured and connected by Press and Welding respectively, the rest of the parts are manufactured by Frame & RX Axle, Engines, and Resin. Finally, these different parts are all assembled together in the Assembly sector. Assembly is the most labor-intensive sector, representing 30% of the employees on the production line. It is also one of the most delicate phases of production, with 75% of the defects per vehicle occurring at this stage of the production process.

2.1.2 Organization of Labor, Knowledge Hierarchies and Training

Production line workers are divided into *working groups* within each sector of the line. The plant has two *shifts* per day (morning and afternoon) with each working group operating in one of the two daily shifts.

Employees within each working group are organized in a clear production hierarchy, characterized by different knowledge layers. Workers are assigned to the different layers depending on the tasks they perform and the skills they have accumulated. Employees can accumulate skills by receiving in-house training programs provided by the firm. A diagram of the hierarchical structure of working groups is shown in Table 1.

There are three main employee categories within the hierarchy. The first are Front-line Operators (FL), who engage in production tasks on the production line. Above FL workers are Mid-line Operators (ML), who are tasked with identifying production problems and closely supervising and helping FL workers whenever problems or bottlenecks occur. Finally, above ML workers are Superiors (S), who are in charge of supporting and managing the overall working group performance. If a FL worker faces an issue with the assembly process, one of the ML workers makes a quick diagnosis of the issue and offers a solution to the FL worker. In case the issue exceeds the knowledge of the ML worker, the problem is then directed to the Superior.[3]

Within each main employee category (i.e., within the FL, ML and S categories), employees are then divided into different *layers*, depending on their skills and training received. FL workers are divided into three possible layers of New Front-line Operators (New FL), four

[3]Superiors are supervised by Line Superiors, who are tasked with overseeing multiple working groups.

layers of Regular Front-line Operators (Reg FL), and one layer of Mid Front-line Operators (Mid FL). New FL are the lowest level of the hierarchy, comprising workers who have recently started and have received only basic training. Reg FL workers are more experienced and skilled, having typically spent at least a year on the production line and having completed at least five training programs, and Mid FL are the highest skilled within the FL category, having more than two years of experience and having gone through 13-14 training programs. ML and S workers are subdivided into four layers each, again based on their experience and training. So, in total, there are 16 possible employee layers that a working group can have – eight within FL workers, four within ML workers, and four within S workers. Each of these 16 layers is associated with a different level of training and therefore knowledge acquired by the workers in that layer.

The hierarchical structure is pyramidal. The typical working group has about 20 workers, divided into about 16 FL workers, 3 ML workers and 1 S worker. Each ML worker typically supervises a team of FL workers. This structure allows FL workers to get support quickly from ML workers, who can then get support from superiors if needed, avoiding production disruptions.[4]

The firm provides in-house training to its employees, and through such training employees can be promoted to higher positions in the working group and advance their professional careers. Training is provided in-house through training programs administered by trainers. Training is primarily focused on management, problem solving and alignment of production line activities with company targets rather than on practical skills to operate on the production line. More advanced training programs place a greater emphasis on problem solving and management.

Most new employees start as FL workers with the lowest levels of basic skills (i.e., they start in the first layer as New FL Operator 1). Workers are promoted as they gain experience, complete specific training programs, and develop new job skills required for each layer, as

[4]A given working group does not necessarily have an employee from each of the 16 layers.

Table 1: Hierarchical Structure of Working Groups

	Employee layer	Tasks/Responsibilities	Cumulative training
S	Superior 4 Superior 3 Superior 2 Superior 1	Support and manage overall working group performance	17 to 20 training programs
ML	Mid-line Operator 4 Mid-line Operator 3 Mid-line Operator 2 Mid-line Operator 1	Identify production problems; closely supervise front-line workers to ensure production standards are met	14 to 17 training programs
FL	Mid Front-line Operator	Engage in production tasks; support other front-line operators	13 to 14 training programs
	Reg. Front-line Operator 4 Reg. Front-line Operator 3 Reg. Front-line Operator 2 Reg. Front-line Operator 1	Engage in production tasks	12 to 13 training programs 10 to 12 training programs 8 to 10 training programs 5 to 8 training programs
	New Front-line Operator 3 New Front-line Operator 2 New Front-line Operator 1	Engage in production tasks; learn basic skills	4 to 5 training programs 1 to 4 training programs 0 to 1 training programs

Note: Table 1 shows a diagram of the hierarchical structure of working groups. For each employee layer, the diagram shows (i) typical tasks and responsibilities; and (ii) the cumulative number of training programs typically received by workers in a given layer.

shown in Table 1. For instance, as the New FL gain experience and skills they can get promoted first to the higher layers within the New FL category (e.g., from New FL Operator 1 to New FL Operator 2), and then across categories (e.g., from New FL Operator 3 to Reg. FL Operator 1) all the way up the hierarchy.

2.1.3 Product Cycles and Changes in Production Volume

In our analysis, we exploit two types of discrete changes to the production process. First, we exploit the fact that the product cycle is characterized by the frequent changes in model variants over time. Over our sample period, the company manufactures two auto models, which are continually updated through new model variants, usually every two or three years. Most changes tend to be made to the bumpers, lights, grille, wheels, and color options. As we will document in our data, such model changes result in a significant increase in the number of new parts that have to be produced and assembled, with resulting changes in the skills needed in production and the types of problems that arise on the assembly line. Second, we exploit the fact that the plant expanded its production volume from around 80,000 cars per year to more than 140,000 in seven years, between 2011 and 2018. This expansion occurred progressively through a sequence of discrete "jumps" in the number of cars that the plant was asked to produce by the manufacturer. Such volume changes result in a substantial increase in the number of cars that have to be produced, but not in the complexity of what needs to be produced since the model does not also change at the same time.

It is important to note that the plant has no discretion as to whether or when to implement model and volume changes: such decisions are made by the manufacturer; the plant is tasked only with executing production. Our interest is precisely in studying the organizational responses put in place by the plant as a response to such product cycles and volume changes.

2.2 Data and Descriptives

We use data for the Assembly sector as that is the most labor-intensive sector and accounts for most of the defects per vehicle in the plant. Our data comes from three main sources. The first is data on productivity and employment at the shift-day level. The second is daily data on the composition of each working group, in terms of the number of employees by layer. The third is a record of the exact dates when the introduction of new model variants and changes in the production volume took place.

2.2.1 Productivity and Employment

We have access to daily productivity data for each shift on the Assembly sector for eight years, from January 2012 to February 2019. The data include the number of cars produced per day for each model and the number of defects per 100 vehicles produced (DPV). The DPV is a key performance indicator that the plant uses to monitor productivity for each shift-day.

Panel A of Table 2 presents summary statistics for our two key productivity variables of interest at the daily level.[5] The plant is large, producing around 410 cars per day, and

[5] Productivity data is originally at the day-shift level. To go to the daily level, we sum the number of cars

there is significant variation across days in the number of cars produced. We standardize the number of defects per vehicle using the mean and standard deviation of the full sample, so that our DPV measure has mean 0 and standard deviation 1.[6] Our data shows that the incidence of DPV is associated with a sizeable decrease in the number of cars produced per day, thus confirming that DPV is an appropriate measure of productivity in this context. More precisely, Appendix Figure A2 displays the coefficients of a Distributed Lag Model where we regress the number of cars produced per day on lags of an indicator variable equal to 1 if the plant experienced a DPV above the 75th percentile on a given day. The figure shows a contemporaneous decrease of 20 cars the same day the plant experiences high DPV. The negative effect fades quickly, but there is no evidence of a significant increase in number of cars produced in subsequent days, thus confirming that high DPV on a particular day is associated with an overall permanent reduction in the total number of cars produced.

We also have data on employment in the Assembly sector over the same time period. The data includes information on the date of hiring (and exit) from the plant, whether the worker is present at the factory on any given day, current position and historical data on previous positions, as well as date and type of training program received by each worker. Panel B of Table 2 presents basic summary statistics on the number of employees in Assembly, where information on employment is averaged at the weekly level to minimize measurement error related to short-term absences from the plant.[7] There are around 2,800 employees each week: 2,300 front-line employees, 400 mid-line employees and 100 superiors, confirming the pyramidal structure described above.

and average the defects per vehicle across the two shifts within the same day.

[6] To protect the confidentiality of the plant, we were not allowed to disclose mean and standard deviation of the DPV measure. Appendix Figure A1 shows the distribution of (standardized) DPV-day observations, confirming that there is substantial variation in defects per vehicle across days.

[7] Note that such measurement error is not an issue with DPV measures as they are key daily KPIs for the plant and so are carefully recorded in the administrative data. Our data on employment instead records the number of employees present at the factory on a given day, and so is subject to potential measurement error driven by short-term absenteeism.

Table 2: Descriptive Statistics on Productivity and Employment

	Mean	SD
Panel A: Productivity, daily level		
Number of Cars	413.32	110.86
DPV	0.00	1.00
Number of observations	1590	
Panel B: Employment, weekly level		
Number of Employees	2804.25	488.73
Front-line Employees (FL)	2286.97	411.94
Mid-line Employees (ML)	408.12	55.36
Superior Employees (S)	106.70	25.52
Number of observations	316	

Note: Panel A shows descriptive statistics on productivity at the daily level, while Panel B on employment averaged at the weekly level, both for the Assembly sector. DPV is standardized using the mean and standard deviation of DPV of the full sample.

2.2.2 Working Groups

We have available unusually rich information on the production hierarchy of the firm. Our data allows us to track the composition of each working group for two years, between January 2017 to February 2019. That is, we know the precise layer each employee belongs to (out of the possible 16 layers described in Table 1) and we know their cumulative training received. This enables us to identify promotions across layers within each employee category (e.g., New FL 1 to New FL 2) and across categories (e.g., Mid FL Operator to ML Operator 1). Our data on the production hierarchy within the firm is thus even more granular than the data used in recent papers studying production hierarchies in French and Portuguese organizations (Caliendo et al., 2020, 2015).[8]

[8] For instance, Caliendo et al. (2015) identify five separate hierarchical layers in their French production data, and Caliendo et al. (2020) are able to identify eight hierarchical layers in Portuguese production data. In contrast, we identify 16 layers. Working with only one firm makes it possible to gather our very granular data, although of course this comes at the expense of not being able to analyze responses across many firms, as these other papers do.

Table 3 shows descriptive statistics on the structure of working groups at the weekly level.[9] The average size of each working group is about 20, which includes around 16 FL workers, 3 ML workers, and 1 S worker.[10] On average there are about eight layers of skills observed across groups: five within FL, two within ML, and one within the S worker category.

Table 3: Descriptive Statistics on Working Groups and Employee Hierarchies

	Mean	SD
Panel A: Composition, weekly level		
Number of Employees	20.25	8.23
Number of FL Employees	16.47	7.56
Number of ML Employees	2.89	1.27
Number of S Employees	0.89	0.31
Number of Layers	7.80	1.99
Number of FL Layers	5.02	1.65
Number of ML Layers	1.89	0.76
Number of S Layers	0.89	0.31
Number of observations	13,280	
Panel B: Change in layers, weekly level		
Change in Number of Layers	0.015	0.349
Change in Number of FL Layers	0.011	0.303
Change in Number of ML Layers	0.003	0.141
Changes in Number of S Layers	0.002	0.035
Number of Observations	13,060	

Note: The information presented in Panel A is averaged at the working group-shift-week level. The changes in layers in Panel B are computed as the average change in the number of layers in the working group at the weekly level.

Panel B reports the mean and standard deviation of the week on week change in the number of layers of each working group over the sample period. Working groups are not static: while the number of layers tends to increase on average over the sample period, there is

[9] In the data, a working group is defined as a set of workers with the same Superior each day. We again average information at the weekly level to smooth out measurement error arising from short-term worker absences.

[10] There are some periods in which Mid-line workers act as Superiors. Such cases are limited however to about 11 percent of working group-shift observations, which do not have a Superior.

substantial variation in the evolution of the hierarchical structure of working groups: working groups both add and remove layers over our sample period. Consequently, it is precisely this organizational flexibility that we seek to analyze in the rest of the paper, exploiting our granular data on the production hierarchy of the firm.

2.2.3 Model and Volume Changes

We observe seven discrete changes in the model variants produced during our sample period, as a result of product cycles. We call these "Model changes". We also observe five discrete changes in production volume, as a result of the overall expansion of production described above. We call these "Volume changes".

On average, with each model change, 13% of the car parts are modified, so that the production line needs to deal with such new parts. However, the number of cars to be produced does not change. In contrast, volume changes represent a discrete increase in the number of cars produced, which jump up by 42% on average, relative to the month before the change. However, volume changes do not bring about new parts, as the model that is produced does not get updated, so that the production line needs to deal with more of the same tasks. Additional summary statistics on the model and volume changes that we exploit are reported in Appendix Table A1.[11]

Model and Volume changes are therefore very different in nature: Model changes require dealing with new parts, but without a change in overall quantity produced; Volume changes require producing a higher quantity of the same product. Our data allows to study organizational responses to both types of events within the same context.

[11] Two of the five model changes and two of the seven volume changes take place in the period between January 2017 to February 2019 when we have data for working groups.

3 Empirical Strategy

Our aim is to study how the internal organization of the firm and training provision change in response to the Model and Volume changes described above. We begin the empirical analysis by showing how Model changes impact productivity (measured by the occurrence of DPV) and how the organization of production changes in response, in terms of the hierarchical structure of the working groups and the training provided to employees. Then, we contrast the impacts of Model changes to those of Volume changes to highlight the different effects of needing to deal with new tasks, as opposed to needing to conduct more of the same tasks. In doing so, we highlight the unique implications of product cycles for the internal organization of the firm.

To answer these questions, we exploit the high-frequency nature of our data to implement event studies and regression discontinuity specifications around the exact time of such changes. The event studies allow us to trace out the dynamics of the impacts. The discontinuity in time specifications allow us to provide a useful summary point estimate of the impacts averaged over the time period into consideration. Next, we describe both sets of specifications.

Our null hypothesis is that the firm does not react to Model or Volume changes in terms of organizational structure or training provision, so that only learning by doing takes place. If this is the case, we should find no discrete change in organizational structure or training provision after Model or Volume changes. Our alternative hypothesis is that the firm responds to such changes, in which case we should expect a sharp effect on organizational structure or training provision after the Model or Volume changes.

3.1 Event Studies

For our key productivity outcome of DPV, which is available at the level of each shift in each day, we estimate the following specification:

$$Y_{st} = \theta_s + \gamma_m + \gamma_y + \sum_{k \geq -b, k \leq b, k \neq -1} \beta_k D_t^k + \delta X_t + \epsilon_{st} \tag{1}$$

where Y_{st} is the outcome measured in shift s and day t. D_t^k is an indicator of the distance with respect to the (Model or Volume) change, where k indicates the distance to the event on day t. In our dataset we know the exact day when the event takes place. b represents the bandwidth (window of time) considered in the event study. Following Calonico et al. (2014) we find that the optimal bandwidth for Model changes when using the DPV measure as outcome is 40 days (8 five-day working weeks) before and after the event. We use this bandwidth throughout for both Model and Volume changes and so set $b = 40$.[12] Our specification controls for shift fixed effects θ_s, year and month fixed effects γ_m and γ_y, to net out any shift-specific effects and to control flexibly for time trends and seasonality. Additionally, we control for (linear) distance to other events X_t (e.g., when looking at a given Model change, we control for distance to all other Model and Volume changes). Finally, standard errors are clustered by week-shift, to allow for correlated shocks over time within the same shift, and across shifts within the same week.

For the outcomes at the level of individual working groups (e.g., composition of the working groups), we estimate a specification similar to equation 1 but with information at the working group-week level:

$$Y_{iw} = \theta_i + \gamma_m + \gamma_y + \sum_{k \geq -b, k \leq b, k \neq -1} \beta_k D_w^k + \delta X_w + \epsilon_{iw} \tag{2}$$

where Y_{iw} is the outcome of group i in week w. D_w^k is an indicator for the distance with respect to the event, where k indicates the distance to the event at week w. We use the same optimal bandwidth as in equation 1, and so in this case $b = 8$ as information is at the weekly level. The specification further controls for working group fixed effects θ_i, month and year fixed effects γ_m and γ_y, and (linear) distance to other Model and Volume changes in weeks

[12] We use the same bandwidth for Model and Volume changes to increase comparability of the results across the two types of events. The optimal bandwidth is similar for Volume changes (25 days).

X_w. Standard errors are clustered at the week-working group level to allow for correlated shocks within working groups over time, and across working groups within the same week.

Our parameters of interest are the β_k. For $k \geq 0$ (i.e., for time periods corresponding to the time of the event or later) we interpret β_k as the causal effect of the Model or Volume change on the different outcomes at time k. All β_k parameters are relative to the outcomes in the time period immediately before the event $k = -1$, which is the excluded category in these regressions.

Identification relies on the assumption that, after controlling for month and year fixed effects, any discrete changes in the outcomes observed just after the event are not due to underlying (residual) trends in the outcome variables. The high-frequency nature of our data and our ability to focus on a narrow time window around the events reassures us regarding the validity of this assumption. The focus on a narrow time window reduces the possible concern that changes in the outcome variables after the Model or Volume changes are due to other events affecting trends in productivity or organizational structure (and we are controlling for distance from other Model and Volume changes). As described in Section 2, the plant has no discretion over whether to implement changes in models or volume; it simply has to execute production. Of course, there is still the possibility that such changes are communicated to the firm in advance, which may lead to anticipation effects. The availability of high-frequency data before the event allows us to test for the presence of such possible anticipation effects. Lack of significance of the estimated β_k in the time periods before the event (i.e., $k < -1$) would provide evidence in support of the limited role of any anticipation effects or other trends in potentially biasing our results.

3.2 Regression Discontinuity in Time

In addition to the event study specifications described above, we also estimate regression discontinuities in time. Doing so allows us to estimate the average effect of the event in our time window, thus providing a useful summary point estimate. This method also has the advantage of pooling coefficients across time periods, thus improving precision.

For the DPV outcome, we estimate the following specification:

$$Y_{stw} = \theta_s + \gamma_m + \gamma_y + \beta_1 I(0 \leq dis_w \leq 3)_w + \beta_2 I(4 \leq dis_w \leq 7)_w + f(dis_t) + \delta X_t + \epsilon_{stw} \quad (3)$$

where Y_{stw} is the outcome measured in shift s, day t and week w. $I(0 \leq dis_w \leq 3)_w$ is an indicator for being four weeks (we define a week as five working days) after the event occurring at time t. $I(4 \leq dis_w \leq 7)_w$ is an indicator for being more than four weeks after the event at day t. Recall that the optimal bandwidth for the event studies was 40 days, which guides our definition of the indicator functions. We control for shift fixed effects θ_s, month and year fixed effects γ_m and γ_y. Additionally, we control for a function of the distance to the event $f(dis_t)$ in days, where we choose the linear function, and for the (linear) distance in days to all other Model and Volume changes X_t. Standard errors are clustered by distance to the event-shift level, again allowing for correlation of the shocks both within shift over time and across shifts in the same time period.

For working-group level outcomes, we estimate a specification similar to equation 3 but where the level of observation is a working group in a week:

$$Y_{iw} = \theta_i + \gamma_m + \gamma_y + \beta_1 I(0 \leq dis_w \leq 3)_w + \beta_2 I(4 \leq dis_w \leq 7)_w + f(dis_w) + \delta X_w + \epsilon_{iw} \quad (4)$$

where Y_{iw} is the outcome for working group i in week w. $I(0 \leq dis_w \leq 3)_w$ is an indicator for being within four weeks after the event at week w. $I(4 \leq dis_w \leq 7)_w$ is an indicator for being more than 4 weeks after the event. We control for working group fixed effects θ_i, month and year fixed effects γ_m and γ_y, and for a linear function of the distance to the event $f(dis_w)$ as well as distance to all other events X_w in terms of weeks. Standard errors are clustered by distance to the event-working group level.

Our parameters of interest are β_1 and β_2, which capture the average effect of each event in the two narrow windows of time. This allows us to speak to the dynamics of the Model and Volume changes, while still providing summary point estimates for the average effect of each event.

4 Results

We present the impacts of Model and Volume changes, in turn. For both, we first validate the nature of the event by studying impacts on number of cars produced and number of new parts. Then, we report impacts on defects per vehicle, to confirm that both types of changes have real implications for productivity. Finally, we present impacts on the structure of working groups and training provision, to study the organizational responses the plant puts in place to bring down defects per vehicle and how they differ across the two types of changes.

4.1 Product Cycles

4.1.1 Validating Model Changes

Table 4 reports the results of regressions like equation 3 but where the outcome variables are total number of cars assembled per day, total number of parts assembled per day, and the share of parts that are new relative to the last model variant produced just before the Model change.[13]

Columns 1 and 2 show that Model changes do not result in a significant change in the number of cars nor in the total number of parts assembled. The coefficients for both outcomes are small in magnitude relative to the mean. Column 3 instead shows that Model changes result in an increase of 13-15% in the share of parts that are new over the two months after the change. The increase in the share of new parts is stable, in line with the factory now

[13] The level of observation in Table 4 is a day and the sample size is 567 because there are 7 Model changes and for each event we choose a bandwidth of 40 days around the event, so 81 days in total for each event.

producing a new model variant. Table 4 confirms that Model changes result in an increase in complexity of production as a substantial amount of new parts need to be assembled.

Table 4: Impact of Model Changes on Production

	(1) Total Cars	(2) Total Parts (Millions)	(3) Share New Parts
0-3 weeks	13.94	-0.210	12.19***
	(29.59)	(0.178)	(2.695)
4-7 weeks	35.94	-0.105	15.75***
	(30.97)	(0.184)	(3.310)
Observations	567	567	567
Obs. Level	Day	Day	Day
Mean	372	2	0

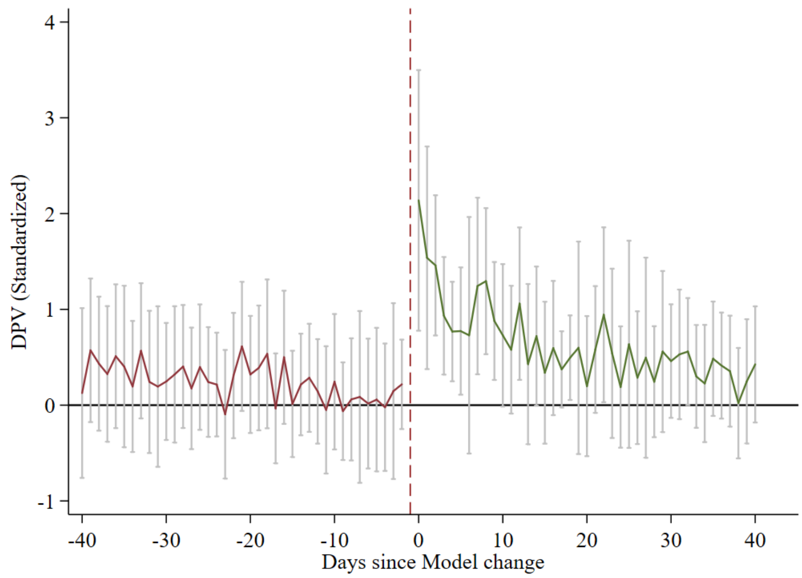
Note: Standard errors clustered by distance to the Model change. Number of observations: 81 days x 7 Model changes. Total parts are expressed in millions. Share of new parts is the percentage of new parts introduced in each model change relative to those used in the previous variant of the model. Car production is reported by the plant at the daily level. We control for month and year fixed effects as well as a linear function of distance to the Model change and distance to every other Model and Volume change in the data. * p<0.1, ** p<0.05, *** p<0.01

4.1.2 Impacts on Productivity

Figure 1 reports the results of an event study specification following equation 1 with DPV (standardized) as outcome. The Figure shows that there is a discrete jump in defects per vehicle right after the Model change. Defects increase by about 2 SD right after the Model change, and then they slowly come down, reverting back to the pre-shock level after about three-four weeks.[14] The Figure also shows that daily DPV does not exhibit any significant pre-trend in the period up until the *exact day* of the Model change. This confirms that any anticipation effects are not first order, thus validating the identification assumptions.

[14] Appendix Table A2 confirms the results in Figure 1 by reporting that Model changes lead to an increase in DPV of about 0.75 SD in the first four weeks after the shock, with the effect coming down to zero after that.

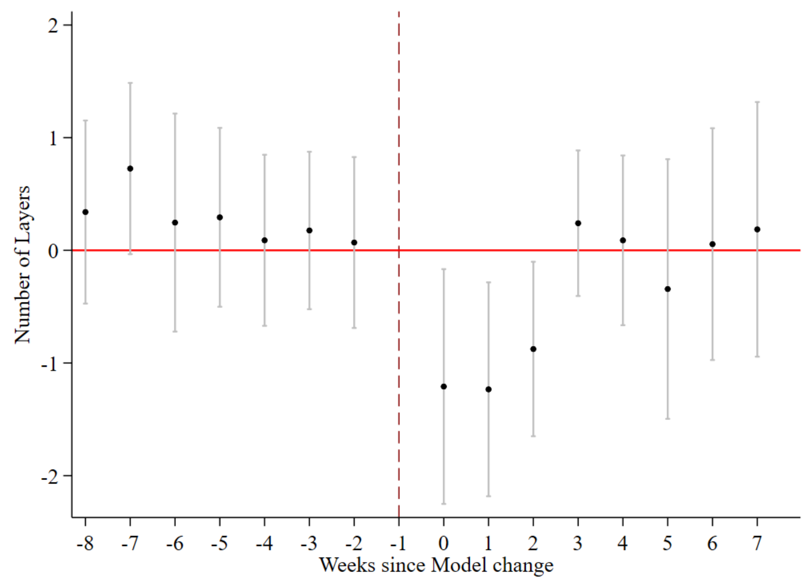
Figure 1: Event Study of Model Changes on Productivity



Note: Figure 1 shows the effect of Model changes on DPV in a time window running from 40 days before the event to 40 days after the event. DPV is computed as number of defects per 100 vehicles, and is standardized using the mean and standard deviation of the full sample. We control for month, year, and shift fixed effects. We also control for a linear function of distance to the Model change and to all other Model and Volume changes. Standard errors are clustered by distance to the event-shift level. 95% confidence intervals are reported. Number of observations: 2 shifts x 81 days x 7 events.

The results in Figure 1 show that there are real negative productivity implications of the introduction of new model variants as a result of product cycles. At the same time, the plant is able to bring these back down fairly quickly. Next, we study which organizational responses to the Model changes the plant puts in place and how these enable the firm to bring back down the level of defects and increase productivity. The null hypothesis is that the firm does not put in place any active response to Model changes, so that the only driver of the reduction in DPV is learning by doing taking place on the Assembly line. The alternative hypothesis is that the firm complements any such learning by doing with changes in organizational structure and training provision.

Figure 2: Event Study of Model Changes on Number of Layers



Note: Figure 2 shows the effect of model changes on the number of layers within working groups in a time window running from 8 weeks before the event to 8 weeks after the event (where the week of the Model change is labelled as week 0 on the x-axis). Number of layers is defined as the number of separate positions present in a working group. We control for month, year, and group fixed effects. We also control for a linear function of distance to the Model change and to all other Model and Volume changes. Standard errors are clustered at week-working group level. 95% confidence intervals are reported. Number of observations: 220 working groups x 16 weeks x 2 events.

4.1.3 Organizational Responses to Product Cycles

Figure 2 reports the estimation results of equation 2 with the number of layers per working group (averaged at the weekly level) as the dependent variable. We see that Model changes result in a sudden *decrease* in the number of layers in the weeks immediately following the event. More precisely, the number of layers decreases on average by about one layer, and this effect is stable and significant in the first three weeks after the shock, after which the number layers reverts back to pre-shock levels. The dynamics of the impacts show a remarkable similarity to Figure 1 in that the effect on the number of layers dissipates after about four weeks since the Model change. These results are validated in the corresponding regression discontinuity in time specifications following equation 4 and reported in Table 5: column 1 shows a negative and significant effect of about 0.7 layers in the first four weeks after the shock, corresponding to a 10% decrease, which then becomes substantially smaller (and less significant) in the second four weeks.

Table 5: Impact of Model Changes on Number of Layers and Average Training

	(1) Num. Layers	(2) Avg. Training
0-3 weeks	-0.666***	0.315***
	(0.124)	(0.073)
3-7 weeks	-0.285	0.095
	(0.190)	(0.113)
Observations	7,040	7,040
Obs. Level	Group-Week	Group-Week
Mean	6.373	12.841

Note: Standard errors clustered by distance to event and working group. Number of observations: 220 working groups x 16 weeks x 2 events. Number of Layers is defined as the number of separate positions present in a working group. Avg. training is the average number of training programs received by the employees in each working group. We use as controls month, year and group fixed effects. We also control for a linear function of distance to the Model change and to all other Model and Volume changes. * p<0.1, ** p<0.05, *** p<0.01

The negative effect on layers means that the plant is reducing the number of separate skill levels within working groups after a Model change. To understand how this change in organizational structure comes about, in Table 6 we study impacts on total number of employees. The Table reports results of regressions like equation 4 at the shift-week level, with various outcome variables related to employment. We find no evidence that Model changes lead to a change in the number of employees in the Assembly sector, nor in the number of hires or separations, in the four weeks after the event: all coefficients are small in magnitude relative to the mean and far from being significant. This indicates that the changes in number of layers must be taking place as a result of the *reallocation* of existing workers across layers, rather than by increasing or downsizing the workforce. This result is sensible in that the total number of cars that need to produced has not changed.[15]

[15] Appendix Table A4 confirms that also the number of working groups and the size of the average group are unaffected by Model changes.

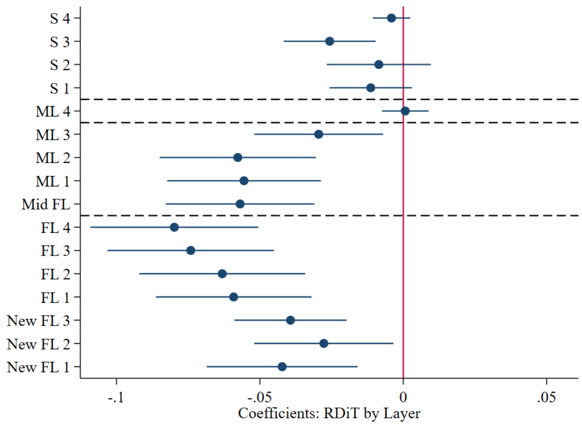
Table 6: Impact of Model Changes on Employment

	(1) Employment	(2) Hires	(3) Separations
0-3 weeks	3.894	-1.172	-0.367
	(21.73)	(4.855)	(0.758)
4-7 weeks	10.85	15.30	0.720
	(40.86)	(12.46)	(1.260)
Observations	64	64	64
Obs. Level	Shift-Week	Shift-Week	Shift-Week
Mean	1175	21	2

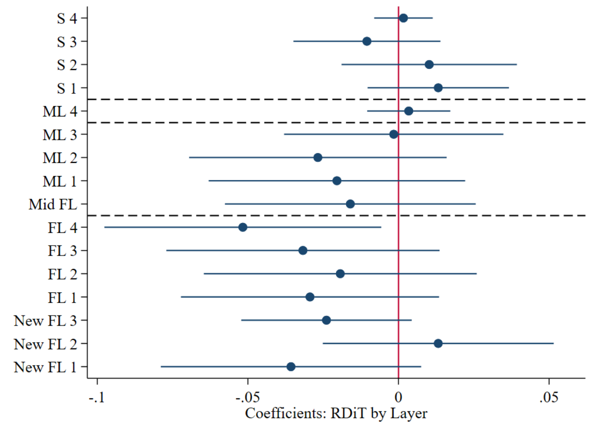
Note: Standard errors clustered by distance to event and shift. Number of observations: 2 shifts x 16 weeks x 2 events. For more details on the definition of the layers see Table 1. We control for month, year, and shift fixed effects. We also control for a linear function of distance to the Model change and to all other Model and Volume changes. * $p<0.1$, ** $p<0.05$, *** $p<0.01$

Finally, in Figure 3 we study where in the hierarchy the effect on layers is concentrated. We do so by running a series of regressions like equation 4 with as dependent variables dummies for whether the working group has at least a worker in a given layer in a given week. Focusing first on panel (a), which looks at the effect in the first four weeks after the event, we notice that the negative effect on the probability of having a given layer is found throughout the hierarchy, but is larger in the *middle* of the hierarchy: positions in the upper Reg. FL are the ones less likely to be present right after a Model change. Panel (c) then reports impacts on the share of workers in the working group by layer. This shows a reallocation of workers away from lower-middle layers and towards higher level positions, so that the share of upper-FL workers decreases, and the share of upper ML and S workers increases. That is, the plant is promoting workers from the lower-middle ranks to the higher ranks, and is waiting to back-fill the lower-middle ranks positions. As promotions in the plant are associated with receiving additional training, the plant is increasing the overall stock of human capital of working groups. Indeed, column 2 of Table 5 confirms that as a result of Model changes the average worker has completed an additional 0.3 training programs in the four weeks after the change.

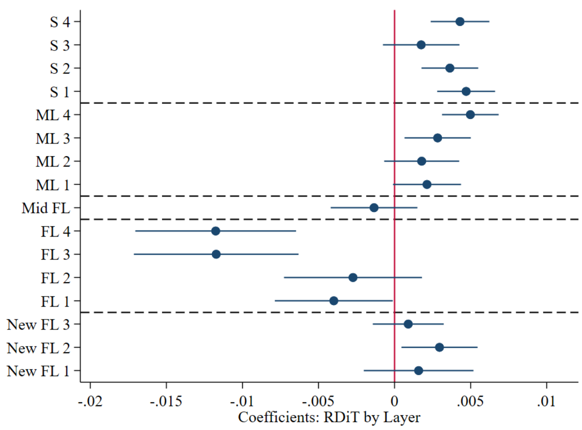
Figure 3: Impact of Model Changes on Working Group Structure and Employee Hierarchies



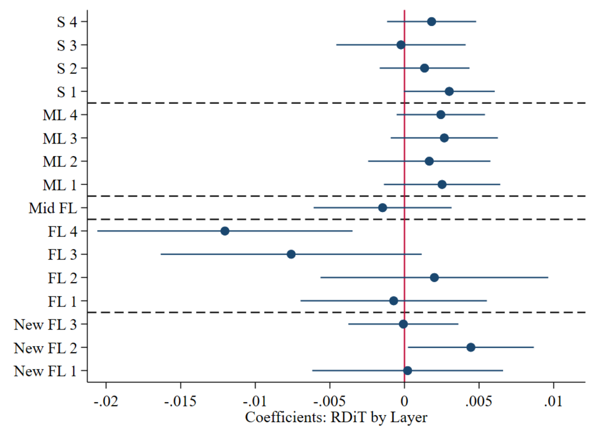
(a) Layer dummies, 0-3 weeks



(b) Layer dummies, 4-7 weeks



(c) Share by layer, 0-3 weeks



(d) Share by layer, 4-7 weeks

Figures 3a and 3b show the effect of Model changes on the probability of having a given layer in the group at 0-3 weeks and 4-7 weeks post-shock, respectively. Figures 3c and 3d show the effect of Model changes on the share of workers in the working group by layer at 0-3 weeks and 4-7 weeks post-shock, respectively. For more details on the definition of the layers see Table 1. Each coefficient is estimated from a separate regression. We control for month, year, and group fixed effects. We also control for a linear function of distance to the Model change and to all other Model and Volume changes. Standard errors are clustered by distance to Model change and working group. 95% confidence intervals are presented in the figure. Number of observations: 220 working groups x 16 weeks x 2 events.

In doing so, the plant is compressing the hierarchy and reducing the distance, in terms of knowledge layers, between workers further up the hierarchy, who have been trained to possess the necessary knowledge to solve complex problems, and front-line workers working on physical production, who are coming against new problems as a result of the introduction of new parts. As described in Section 2, training focuses primarily on management and problem solving, especially for workers farther up the hierarchy. These results then indicate that through such training provision, the firm increases the ability of upper level workers to help front-line workers solve the new complex problems they come up against.

These results are notable in that they show how the firm combines both in-house training and changes in organizational structure to adapt to product cycles. Our granular data allows us to document the surprising flexibility of the organization and to uncover a substantial and immediate response to the introduction of new models, which would largely be missed with less granular data aggregated at the yearly level for example.

These results highlight how an increase in the complexity of the problems that need to be solved can lead to a compression of the knowledge hierarchy through a reduction in organizational layers. This is a novel result, which stands in contrast to the literature studying how the hierarchical structure of firms changes with an expansion of production. As we will discuss in more detail in the next section, this literature tends to find that as firms grow larger and there is a larger number of problems to be solved, the number of knowledge layers *increases*, as this allows to focus the scarce time of skilled managers only on the most complex problems (Caliendo et al., 2020, 2015). Our unique setting allows us to highlight instead organizational responses to an increase in the complexity of what needs to be produced, while keeping total quantity produced constant. We show how the organizational response is very different when the complexity of production increases, but quantity produced does not change.

In the next section, we exploit discrete increases in the volume of production to show that when production expands but the complexity of problems does not change, this leads to an increase in overall employment and in the number of layers. Thus, we can replicate the finding in the literature regarding organizational responses to the expansion of production. In the final section, we then develop a model which extends classic models of hierarchies (Caliendo and Rossi-Hansberg, 2012) to help reconcile why the number of layer decreases with Model changes but increases with Volume changes.

Finally, Figure 3 shows that the effect on layers is short-lived (in line with Figure 2).

Panel (b) shows that the negative effect on the probability of having a given layer becomes smaller and largely insignificant 5-8 weeks after the shock. Panel (d) shows that impacts on the share of workers by layers are also more muted and less precisely estimated after 5-8 weeks. This result is sensible in that it shows that once working groups have learned the new tasks and the negative productivity impacts following the introduction of new model variants have been addressed, working groups again revert to the initial hierarchical structure.[16] In sum, these results document the surprising flexibility of large organizations and how this allows them to respond to product cycles.

4.2 Contrasting Product Cycles with Volume Changes

We now contrast the organizational response to product cycles documented in the previous section with the response of the firm to sudden and sharp increases in the volume of cars that need to be produced. In short, we find that while volume changes lead to a quantitatively similar spike in the incidence of defects per vehicle, the organizational response is very different: the plant hires more entry-level workers and adds more layers to working groups, so that the distance – in terms of knowledge layers – between front-line workers and supervisors farther up the hierarchy increases, and the average skill level in the firm decreases, due to the inflow of entry-level workers.

These results for volume changes are in line with the literature on knowledge hierarchies (Caliendo et al., 2020, 2015), which tends to find that as firms expand production, new layers are added to the hierarchy, because the higher volumes allow the firm to better focus the talent of highly skilled managers on solving only the more complex problems. Contrasting the results of volume changes with those of model changes allows us to highlight how the response

[16] Table 6 provides some evidence on the channels that allow this readjustment: in particular, column 2 shows a positive and large point estimate on hires 5-8 weeks after the shock, although not significant. This result suggests that the plant then hires additional workers to eventually back-fill lower rank positions and get closer again to the initial hierarchical structure and average training level. As there is turnover throughout the hierarchy of working groups because workers leave the plant or are reassigned to other sections of the production line, this can explain why there is no overall impact on the number of workers and size of working groups even though more workers are hired eventually.

to product cycles is instead very different: when workers have to deal with new tasks rather than produce more of the same, the firm trains existing workers and compresses the hierarchy, reducing the distance between more highly skilled workers and front-line employees.

Moving on to the impacts of volume changes, we first of all validate in Table 7 that volume changes result in a sudden and sizeable increase in the number of cars produced and, consequently, in the number of parts that have to be assembled: the total number of cars produced per day jumps up by about 42% and the number of parts increases by 37% after the volume change, and these effects on production are long-lasting. Since by definition the model produced on either side of volume changes does not change, the types of parts that have to be assembled are the same as before: the plant suddenly needs to produce more of the same.

Table 7: Impact of Volume Changes on Production

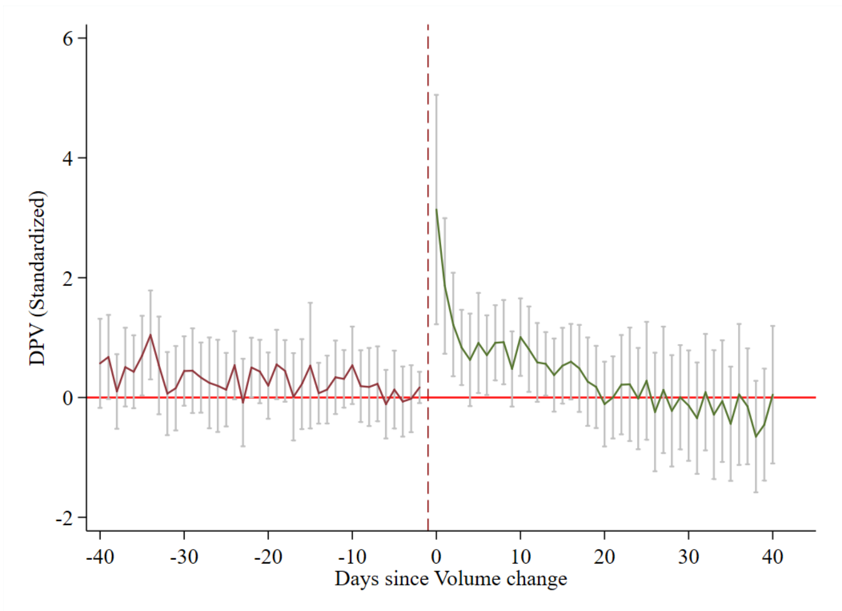
	(1) Total Cars	(2) Total Parts
0-3 weeks	160.1***	0.749***
	(49.60)	(0.222)
4-7 weeks	230.2***	1.086***
	(56.73)	(0.254)
Observations	405	405
Obs. Level	Day	Day
Mean	377	2

Note: Standard errors clustered by distance to Volume change. Number of observations: 81 days x 5 events. We use as controls month and year fixed effects. Total Parts in millions. Cars production is reported by the plant at daily level. We control for a linear function of distance to the Volume change and distances to the other Volume and Model changes. * p<0.1, ** p<0.05, *** p<0.01

The change in volume leads to a sudden and large increase in defects per vehicle, as shown in Figure 4. DPV shoot up by about 2-3 SD in the immediate aftermath of the change, and then come down to pre-shock levels in about three weeks. The similarity with Figure 1 is remarkable: both volume and model changes lead to a similar short-term reduction in productivity.[17]

However, the organizational response of the firm to the increase in the volume of production is very different: Figure 5 shows that volume changes lead to a sudden *increase* in the number of layers within working groups, which go up by just under 1 layer (from a mean of about 5 layers). This increase is persistent for several weeks after the volume change, which is consistent with the persistent increase in the volume that needs to be produced, as shown in Table 7. Column 1 in Table 8 shows the corresponding discontinuity in time specifications, finding results in line with Figure 5, in terms of both magnitude and persistence.

Figure 4: Event Study of Volume Changes on Productivity



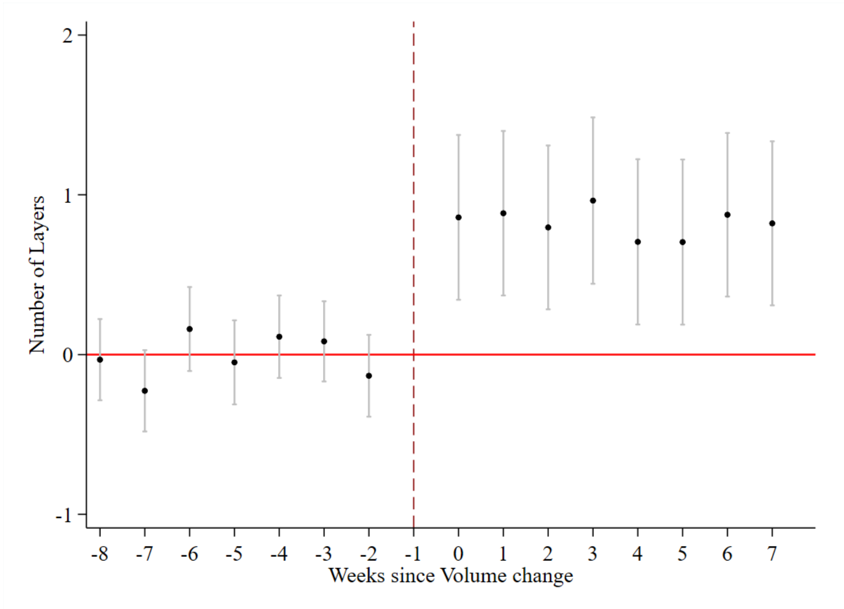
Note: Figure 4 shows the effect of Volume changes on DPV in a time window running from 40 days before the event to 40 days after the event. DPV is computed as number of defects per 100 vehicles, and is standardized using the mean and standard deviation of the full sample. We control for month, year, and shift fixed effects. We also control for a linear function of distance to the Volume change and to all other Volume and Model changes. Standard errors are clustered by distance to the event-shift level. 95% confidence intervals are reported. Number of observations: 2 shifts x 81 days x 5 events.

Next, we ask how the increase in layers is generated. Table 9 shows that as a response to volume changes the firm hires about 24 additional workers per shift-week in the first four weeks after the Volume change (column 2) while separations do not increase significantly (column 3). The result is that overall shit-level employment increases by about 80 workers

[17]Appendix Table A5 shows the corresponding discontinuity in time estimates.

over the two months following the volume change, or a 7% increase over a mean of about 1,100 workers per shift (column 1). As per company policy, the plant hires primarily entry level workers. This is confirmed in column 2 of Table 8, which shows that the average level of training in working groups decreases, precisely due to the inflow of entry-level workers.[18]

Figure 5: Event Study of Volume Changes on Layers



Note: Figure 5 shows the effect of Volume changes on the number of layers within working groups in a time window running from 8 weeks before the event to 8 weeks after the event (where the week of the Volume change is labelled as week 0 on the x-axis). Number of layers is defined as the number of separate positions present in a working group. We control for month, year, and group fixed effects. We also control for a linear function of distance to the Volume change and to all other Volume and Model changes. Standard errors are clustered at week-working group level. 95% confidence intervals are reported. Number of observations: 220 working groups x 16 weeks x 2 events.

Finally, we study where in the hierarchy the increase in layers is coming from: Figure 6 shows that the increase in number of layers is coming throughout the hierarchy but is more pronounced for FL workers (panel (a)). As a result, the organization becomes more heavy in the lower-middle part (panel (c)). This is again consistent with the decrease in the average

[18] Appendix Table A7 shows that the increase in employment is associated with an increase in the number of working groups in the Assembly sector, but not with an increase in the size of pre-existing working groups. In addition, Appendix Table A8 shows that despite the increase in employment, volume changes still lead to an increase in the number of cars and number of parts per employee as well as per working group. This then justifies why we see an increase in the number of layers within working groups: the firm reorganizes production to deal with the increase in the volume of work per employee. In line with this, we still find a positive and persistent effect on the number of layers when restricting the sample to pre-existing working groups only (Appendix Figure A3).

training level of working groups documented in Table 8. In line with the sustained impacts on the number of layers shown in Figure 5, panels (b) and (d) of Figure 6 show that the impact on the organizational structure is sustained.

Table 8: Impact of Volume Changes on Number of Layers and Average Training

	(1) Num. Layers	(2) Avg. Training
0-3 weeks	0.850***	-0.432***
	(0.098)	(0.060)
4-7 weeks	0.706***	-0.374***
	(0.117)	(0.072)
Observations	7,040	7,040
Obs. Level	Group-Week	Group-Week
Mean	5.284	13.459

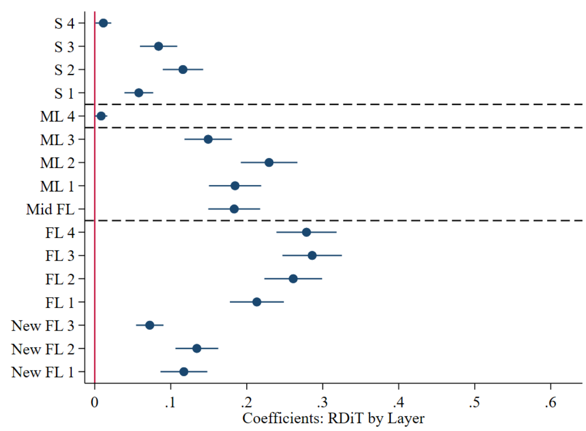
Note: Standard errors clustered by distance to event and working group level. Number of observations: 220 working groups x 16 weeks x 2 events. Layers is defined as the number of separate positions present in a working group. Avg. training is the average number of training programs received by the employees in each working group. We use as controls month, year and group fixed effects. We control for a linear function of distance to Volume changes and distances to the other Volume and Model changes. * p<0.1, ** p<0.05, *** p<0.01

Table 9: Impact of Volume Changes on Employment

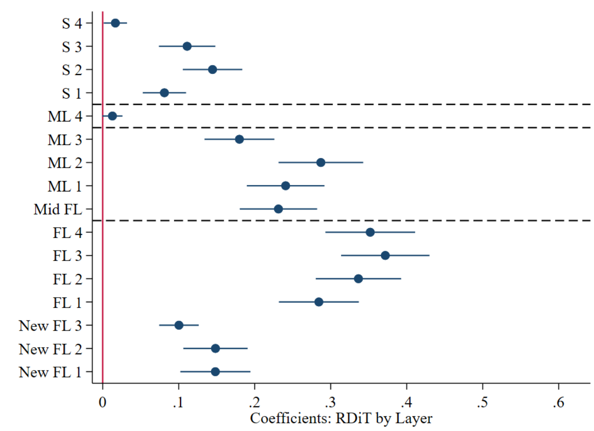
	(1) Employment	(2) Hires	(3) Separations
0-3 weeks	29.46	24.08**	1.67
	(34.32)	(9.63)	(1.50)
4-7 weeks	82.08***	-0.58	0.25
	(28.60)	(2.48)	(0.36)
Observations	64	64	64
Obs. Level	Shift-Week	Shift-Week	Shift-Week
Mean	1094	10	1

Note: Standard errors clustered by distance to Volume change and shift. Number of observations: 2 shifts x 16 weeks x 2 events. We use as controls month, year, and shift fixed effects. We control for a linear function of distance to Volume changes and distances to the other Volume and Model changes. * p<0.1, ** p<0.05, *** p<0.01

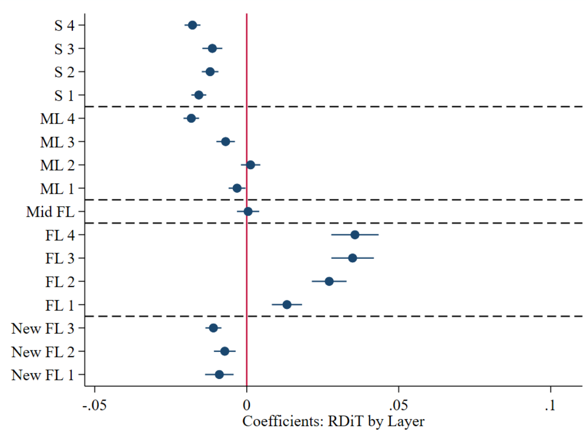
Figure 6: Impact of Volume Changes on Working Group Structure and Employee Hierarchies



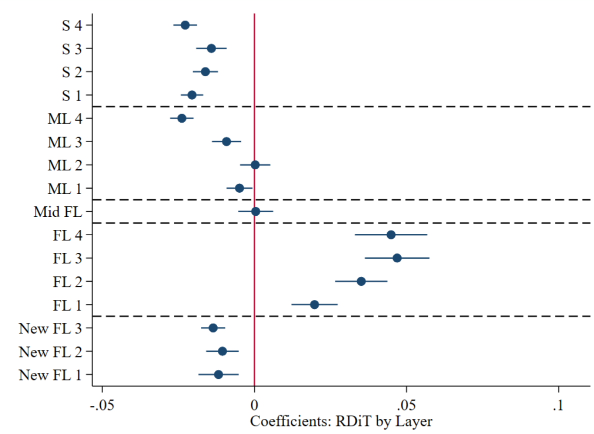
(a) Layer dummies, 0-3 weeks



(b) Layer dummies, 4-7 weeks



(c) Share by layer, 0-3 weeks



(d) Share by layer, 4-7 weeks

Figures 6a and 6b show the effect of volume changes on the probability of having a layers in the group at 0-3 weeks and 4-7 weeks post-shock, respectively. Figures 6c and 6d show the effect of volume changes on share of workers by layers at 0-3 weeks and 4-7 weeks weeks post-shock, respectively. For more details on the definition of the layers see Table 1. Each coefficient is estimated from a separate regression. We control for month, year, and group fixed effects. We also control for a linear function of distance to the Volume change and to all other Volume and Model changes. Standard errors clustered by distance to the Volume change and working group. 95% confidence intervals are presented in the figure. Number of observations: 220 working groups x 16 weeks x 2 events.

Figure 6 shows a remarkable contrast with Figure 3: while model changes lead to workers being trained and promoted away from mid-level positions and towards higher level positions, so that the organization becomes both top and bottom heavy (thus reducing the distance between highly skilled workers and front-line employees), volume changes lead to an organization that is heavier in the middle and lower part. This comparison again highlights how different and unique are organizational responses to product cycles with respect to responses to volume changes, which have been extensively studied in the literature (Caliendo et al., 2020, 2015).

5 Discussion

The findings in Section 4 show that the there is a discrete jump in defects per vehicle right after the Model or Volume change. However, we find that a) when workers have to deal with new tasks, the firm trains existing workers and compresses the hierarchy, and b) Volume changes lead to a different organizational response: the plant hires more entry-level workers and adds more layers to working groups. In this section, we discuss how our results can be reconciled with the literature on hierarchies in organizations (Caliendo et al., 2020, 2015; Caliendo and Rossi-Hansberg, 2012).

5.1 Interpreting Our Results Through Existing Models

We consider as benchmark the model in Caliendo and Rossi-Hansberg (2012). We briefly summarize the key elements of the model, and then discuss whether this model can explain our pattern of empirical results.

In this model, organizations are composed by two types of agents: workers and managers. To generate output, workers in the organization need to solve a problem drawn from a cumulative distribution F, with decreasing density ($f' < 0$). Solving problems requires knowledge. A realization z of a problem implies that in order to solve that problem, the worker needs to have acquired a set of knowledge that includes z as an element. If the worker solves the problem, the production possibility becomes A units of output.

If the worker cannot solve the problem, she asks a manager one layer above for a solution. Then, the manager spends h units of her time listening to the worker's problem (i.e., h is the cost of communication) and solves the problem if her set of knowledge includes z. If the manager cannot solve the problem, she can communicate it to another manager one layer above her. This process continues until the problem is solved or the problem reaches the only agent in the highest layer of the organization, namely the entrepreneur. To achieve a set of knowledge $[0, z]$ for a given agent (which means that the agent can solve any problem within the interval $[0, z]$), the firm must pay wcz, that is, the cost of knowledge is wc per unit of knowledge, where w is the wage and c the training cost.

For simplicity we assume that problems are drawn from the exponential distribution, that is, $F(z) = 1 - e^{-\lambda z}$ for a given $\lambda > 0$. Note that as $f(z) = \lambda e^{-\lambda z}$ is strictly decreasing in z, agents at the bottom of the organization learn the most common problems, while agents in higher layers learn rarer problems. Also note that as λ decreases, the frequency of complex problems increases.

Given a production level q, a wage w and a training cost c, each firm solves an organizational problem that involves choosing the optimal number of layers L, the optimal amount of workers/managers at each layer, and the amount of knowledge these workers and managers acquire. The firm does so by solving two minimization problems, given q, w and c:

1. (Layers problem) It chooses the optimal number of layers L in order to minimize the cost of producing q units of output while paying a wage w to each worker and manager;

2. (Workers and Knowledge problem) Given the number of layers L, it chooses the optimal amount of workers/managers at each layer and the optimal amount of knowledge they acquire to produce q units of output, given a training cost, c.

Using this general framework, we study how the organization of the firm changes endogenously in response to Model and Volume changes.

We begin from considering Volume changes, as this maps more closely to related studies that consider how the organization of the firm changes in response to an increase in the quantity produced. The model predicts that an increase in the volume produced should result in an increase in the number of layers and in the number of employees per layer, which is exactly what we see in the data. The intuition is that when production increases, the number of problems increases too, and so each worker solves a smaller fraction of the problems. As a response, the firm can either: (i) hire new workers and increase the number of layers, so that new managers deal with rarer problems, or (ii) increase the number of employees and knowledge in all layers without an increase in the number of layers. Note that since the firm needs to produce more, the model predicts that new workers will be hired regardless of the impact on the hierarchical structure. Our empirical evidence is consistent with (i) in that we document a sharp and sustained increase in the number of layers and employees as a response to volume changes. Therefore, the model in Caliendo et al. (2015) can reconcile the impact of Volume changes that we document empirically.

Turning to Model changes, we note that this is not something that the literature on hierarchies in organizations has explored before. Nevertheless, we start by discussing whether the model in Caliendo et al. (2015) can reconcile the results on Model changes that we document.

As shown before, every time a new model is introduced, the complexity of the problems increases as the share of new parts in the car increases. We model this as a reduction in λ. A lower λ decreases the production level q_L at which the firm would find it optimal to move from L to $L+1$ layers, thus *increasing* the number of layers and workers per layer for a fixed level of production. Intuitively, if problems become more complex, given a stock of knowledge, each worker solves a smaller fraction of the problems (see Proposition 1 in the Appendix). Thus, to keep production constant, there are two options: (i) increase the number of layers and workers to deal with the more complex problems, or (ii) increase the number of employees and knowledge in all layers, without an increase in the number of layers. That is, the model predicts that Model changes should result in an increase in the number of workers and either no impact or a *positive* impact on the number of layers (depending on the size of the shock). We instead find no impact on the number of workers and a *negative* impact on the number of layers. The model of Caliendo et al. (2015) therefore cannot reconcile the impact of Model changes that we document.[19]

Next, we discuss possible model extensions that could help reconcile our results on Model changes, and present supporting empirical evidence.

5.2 Extending Existing Models: The Role of Training

Training plays a key role in determining the equilibrium levels of knowledge at different levels of the hierarchy in our setting and in many similar manufacturing settings. We know from anecdotal evidence that every time the complexity of the routine tasks increases due to a Model change, the partner firm *improves* the training programs (in terms of both content and trainers) to train workers more efficiently in problem solving and managerial skills. We interpret this as an *increase in the productivity of training*, which in the model would correspond to a reduction in the training cost c. One advance we make to the current state of the theory that is relevant for our setting is to incorporate the role of training and endogenizing the equilibrium level of training the firm offers at different levels of the hierarchy.

We first explore empirically how training investments by the firm differ for Volume and Model changes, to substantiate the claim that the firm increases the productivity of training in response to Model changes. Table 10 shows that the firm increases substantially the number of *different* courses provided *in house* and the number of trainers per employee after a new model is introduced (Panel A). These results are consistent with the anecdotal evidence provided by the partner company that the firm makes investments to improve the training programs. It also again highlights training as an important margin that enables the firm to adapt to Model changes and resolve defects quickly, consistent with the empirical evidence documented in Section 4. Reassuringly, we do not see any impact when the firm changes the

[19] For simplicity, in this section, we consider a static model. However, a dynamic version of the model of Caliendo and Rossi-Hansberg (2012) where workers learn by working on new tasks would still not reconcile the empirical results presented in Section 4: a Model change would still lead to a (temporary) increase in either the number of workers or the number of layers. Then, as workers learn the new tasks, the stock of knowledge would increase in each layer, leading to a subsequent reduction in the number of layers or the number of workers. Instead, we document an immediate *decrease* in the number of layers after the Model change, with no impact on the number of workers.

scale of production (Panel B). This is expected as the complexity of the problems does not change with Volume changes.

Table 10: Impact of Model and Volume Changes on Training Investments

	(1) Courses	(2) Trainers
	Panel A: Model Changes	
0-3 weeks	2.743***	3.173***
	(0.443)	(0.668)
4-7 weeks	4.311***	5.345***
	(1.123)	(1.469)
Observations	112	112
Mean	6.255	8.055
	Panel B: Volume Changes	
0-3 weeks	-1.588	-2.213
	(0.943)	(1.668)
4-7 weeks	-0.466	-0.607
	(1.535)	(2.169)
Observations	80	80
Mean	6.100	8.175

Note: Standard errors clustered by distance to event. Courses is the number of courses taught each week and Trainers is the number of trainers teaching each week. We use as controls month and year fixed effects. We also control for a linear function of distance to the event (i.e., Model or Volume change) and to all other events in the data. Panel A shows the effect for Model changes and Panel B shows the effect for Volume changes. Number of observations Panel A: 7 events x 16 weeks. Number of observations Panel B: 5 events x 16 weeks. * p<0.1, ** p<0.05, *** p<0.01

To reconcile the above results with the previous theoretical literature we start by asking: *What are the admissible ratios $\Delta c : \Delta\lambda$ that imply that the firm does not optimally increase the number of layers when the complexity of tasks increases?* In the Appendix, we show that the partial derivatives $(\partial q_L/\partial\lambda)|_{\vec{p}} > 0$ and $(\partial q_L/\partial c)|_{\vec{p}} < 0$ quantify the opposing effects on the intersecting point q_L (i.e., the quantity q at which the firm finds it optimal to switch from L to $L+1$ layers) of reducing λ or c. The former constitutes the negative impact of the firm having to face more complex problems – which pushes towards an increase in number of layers for a given quantity of production – while the latter is the favorable scenario where the firm can train its workers more efficiently – which pushes towards a *reduction* in the number of layers. In that sense, Proposition 2 in the Appendix provides the minimum investment in reducing training costs c required to balance the effect of problem complexity.

Motivated by the observations discussed above, but being mindful that firms usually can not freely change the cost of training c, in the Appendix we develop a variation of the model in Caliendo and Rossi-Hansberg (2012) where the firm can endogenously invest in reducing the training cost c but this comes at a fee, or penalty $\mathcal{P}$.[20] Intuitively, this extension captures the idea that when the complexity of production increases, the firm may find it optimal to invest in its training programs to improve the productivity of training, consistent with the evidence in Table 10.

In this version of the model with endogenous training cost, for each c the firm has to establish the optimal distribution of knowledge $\{z_L^l\}_{l=0}^L$, and then it selects the optimal training cost c, given the fee $\mathcal{P}$. Proposition 3 in the Appendix shows that in this version of the model there is a neighborhood of λ where an increase in complexity (i.e., a marginal reduction in λ) compresses knowledge layers in the organization.

In Appendix B3, we provide an example to illustrate the results presented in Proposition 3. Our simulations show that there is a set of production levels where the firm reduces the amount of layers every time the complexity of the tasks increases (e.g., a Model change), increasing the amount of knowledge within each layer.[21] Note that this is not the case for Volume changes as the number of layers generally increases for large volume increases, even when we allow the option to invest in training by reducing the training cost c (see Appendix Figure B2).

Intuitively, there are dynamic elements and adjustment costs at play here. In the short run, firms are confronted with these changes in complexity arising from Model changes that require new knowledge. However, it is inefficient for the firm to hire new workers who would

[20] We assume that $\mathcal{P}$ is decreasing with respect to c and independent of the production levels q.

[21] In other cases, the firm operates with the same amount of layers when there is a Model change. Note that this result depends on the magnitude of the increase in complexity.

require even more training to deal with this rise in complexity, especially given that quantity is fixed in the short run. Instead, the firm can first, quickly retrain existing workers and reorganize teams to deal with the higher complexity at the current quantity produced.

To make this intuition clear, rather than propose an explicitly dynamic framework, we differentiate fixed factors in the short run and long run adjustment. With some abuse of notation (since L is not continuous), we see:

$$\frac{dq}{d\lambda}\Big|_{q=\bar{q}} = 0 = \frac{\partial q}{\partial z}\frac{\partial z}{\partial \lambda} + \frac{\partial q}{\partial L}\frac{\partial L}{\partial \lambda} \tag{5}$$

and hence

$$\frac{\partial q}{\partial z}\frac{\partial z}{\partial \lambda} = -\frac{\partial q}{\partial L}\frac{\partial L}{\partial \lambda}. \tag{6}$$

We know q increases in z and L; so equation 6 shows that an increase in training to deal with the new problem distribution will go together with a reduction in the number of layers. After this phase, the firm can choose to expand production in the longer term if it wants. That is, when there is this adjustment cost to changing quantity, knowledge (training) and layers are substitutes in short run, but complements in the long run.

This discussion shows that taking into account the "technology of training" and how firms might endogenously decide to invest in improving it can be important for understanding organizational responses to the introduction of more complex problems such as those resulting from product cycles. In considering the role of the training technology we advance the recent theoretical literature on knowledge hierarchies, which has mostly considered organizational responses to changes in quantity produced rather than in the complexity of what needs to be produced (Caliendo et al., 2020, 2015).

6 Conclusion

Focusing on the automotive sector as a prototypical example of an industry with product cycles, we study how the organization of the firm changes in response to these cycles. To do so, we combine granular administrative data on production, employee hierarchies and training provision from an Argentinian subsidiary plant of a leading global auto manufacturer with event study and discontinuity-based methods.

We find that demand (number of vehicles) and number of total parts do not change in the short term after a new model is introduced. The main change is a large, discontinuous increase in new parts. Accordingly, the production of new models necessitates dealing with new complex problems. Indeed, we show that defects per vehicle increase substantially after the production change, and decrease to their prior level over a period of about 3 weeks. We then show that dealing with product cycles involves a compression of the knowledge hierarchy: the firm trains and promotes mid-level employees to top level positions to manage the production of new parts but waits to back fill mid-level positions until after defect rates recover to pre-model-change levels. In doing so, the firm reduces the distance – in terms of knowledge layers – between front-line workers, who are dealing with these new tasks, and managers further up the hierarchy, who have the necessary knowledge to solve the new complex problems that arise as a result.

These results for model changes contrast interestingly with the impacts of an increase in quantity produced (i.e., volume changes). In this case we show the organizational response is a monotonic, permanent increase in both employment and knowledge layers, consistent with prior evidence from manufacturing in high-income countries (Caliendo et al., 2020, 2015).

We provide what is to our knowledge the first study of the way in which organizational structure responds to model changes like those that occur regularly in product cycles. Our findings show how large suppliers in the "global south" are highly flexible in their internal organization of labor, and how this allows them to adapt and respond to the ever increasing complexity of production necessary to remain competitive in global product markets.

References

Acemoglu, D. and Pischke, J.-S. (1998). Why do firms train? theory and evidence. *The Quarterly journal of economics*, 113(1):79–119.

Acemoglu, D. and Pischke, J.-S. (1999). Beyond becker: Training in imperfect labour markets. *The economic journal*, 109(453):112–142.

Adhvaryu, A., Bassi, V., Nyshadham, A., and Tamayo, J. A. (2020). No line left behind: Assortative matching inside the firm. Technical report, National Bureau of Economic Research.

Adhvaryu, A., Kala, N., and Nyshadham, A. (2018). The skills to pay the bills: Returns to on-the-job soft skills training. Technical report, National Bureau of Economic Research.

Adhvaryu, A., Murathanoglu, E., and Nyshadham, A. (2022). On the allocation and impacts of managerial training.

Adhvaryu, A., Nyshadham, A., and Tamayo, J. A. (2019). Managerial quality and productivity dynamics. Technical report, National Bureau of Economic Research.

Aghion, P., Bloom, N., and Van Reenen, J. (2014). Incomplete contracts and the internal organization of firms. *The Journal of Law, Economics, & Organization*, 30(suppl_1):i37–i63.

Arrow, K. J. (1962). The economic implications of learning by doing. *The Review of Economic Studies*, 29(3):155–173.

Bayus, B. L. (1994). Are product life cycles really getting shorter? *Journal of product innovation management*, 11(4):300–308.

Becker, G. S. (1962). Investment in human capital: A theoretical analysis. *Journal of political economy*, 70(5, Part 2):9–49.

Bloom, N., Eifert, B., Mahajan, A., McKenzie, D., and Roberts, J. (2013). Does management matter? evidence from india. *The Quarterly Journal of Economics*, 1(51):51.

Bloom, N., Sadun, R., and Van Reenen, J. (2010). Does product market competition lead firms to decentralize? *American Economic Review*, 100(2):434–38.

Bloom, N., Sadun, R., and Van Reenen, J. (2016). Management as a technology? Technical report, National Bureau of Economic Research.

Bloom, N. and Van Reenen, J. (2007). Measuring and explaining management practices across firms and countries. *The Quarterly Journal of Economics*, pages 1351–1408.

Bloom, N. and Van Reenen, J. (2010). Why do management practices differ across firms and countries? *The Journal of Economic Perspectives*, pages 203–224.

Caliendo, L., Mion, G., Opromolla, L. D., and Rossi-Hansberg, E. (2020). Productivity and organization in portuguese firms. *Journal of Political Economy*, 128(11):4211–4257.

Caliendo, L., Monte, F., and Rossi-Hansberg, E. (2015). The anatomy of french production hierarchies. *Journal of Political Economy*, 123(4):809–852.

Caliendo, L. and Rossi-Hansberg, E. (2012). The impact of trade on organization and productivity. *The quarterly journal of economics*, 127(3):1393–1467.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82(6):2295–2326.

Espinosa, M. and Stanton, C. T. (2022). Training, communications patterns, and spillovers inside organizations. Technical report, National Bureau of Economic Research.

Fenizia, A. (2022). Managers and productivity in the public sector. *Econometrica*, 90(3):1063–1084.

Frederiksen, A., Kahn, L. B., and Lange, F. (2020). Supervisors and performance management systems. *Journal of Political Economy*, 128(6):2123–2187.

Garicano, L. (2000). Hierarchies and the organization of knowledge in production. *Journal of political economy*, 108(5):874–904.

Garicano, L. and Rossi-Hansberg, E. (2006). Organization and inequality in a knowledge economy. *The Quarterly journal of economics*, 121(4):1383–1435.

Grossman, G. M. and Helpman, E. (1991a). Endogenous product cycles. *The Economic Journal*, 101(408):1214–1229.

Grossman, G. M. and Helpman, E. (1991b). Quality ladders and product cycles. *The Quarterly Journal of Economics*, 106(2):557–586.

Grossman, G. M. and Helpman, E. (1994). Endogenous innovation in the theory of growth. *Journal of Economic Perspectives*, 8(1):23–44.

Irwin, D. A. and Klenow, P. J. (1994). Learning-by-doing spillovers in the semiconductor industry. *Journal of political Economy*, 102(6):1200–1227.

Krugman, P. (1979). A model of innovation, technology transfer, and the world distribution of income. *Journal of political economy*, 87(2):253–266.

Levitt, S. D., List, J. A., and Syverson, C. (2013). Toward an understanding of learning by doing: Evidence from an automobile assembly plant. *Journal of political Economy*, 121(4):643–681.

Limodio, N. (2021). Bureaucrat allocation in the public sector: Evidence from the world bank. *The Economic Journal*, 131(639):3012–3040.

Sandvik, J., Saouma, R., Seegert, N., and Stanton, C. (2022). Should workplace programs be voluntary or mandatory? evidence from a field experiment on mentorship.

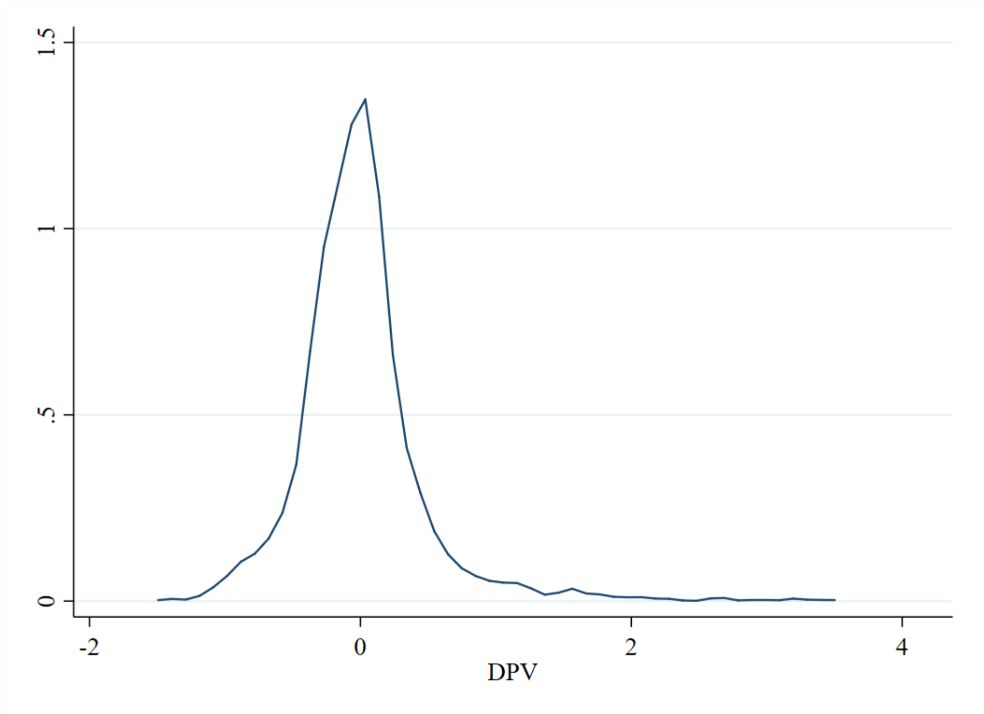
Sandvik, J. J., Saouma, R. E., Seegert, N. T., and Stanton, C. T. (2020). Workplace knowledge flows. *The Quarterly Journal of Economics*, 135(3):1635–1680.

Vernon, R. (1966). International investment and international trade in the product cycle. *The Quarterly Journal of Economics*, 80(2):190–207.

Online Appendix

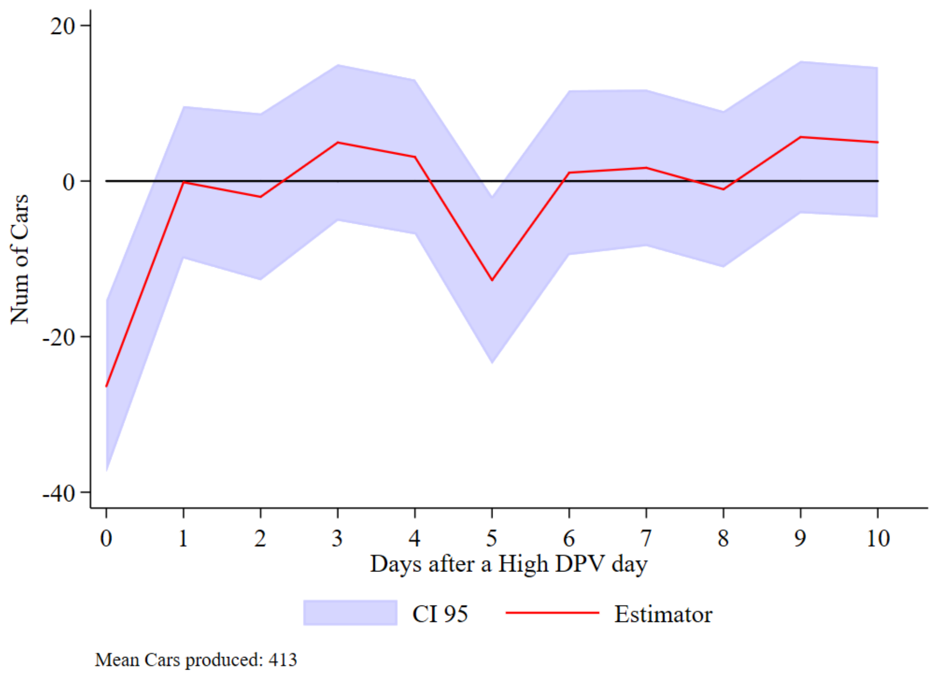
A Tables and Figures

Figure A1: Dispersion in Daily DPV



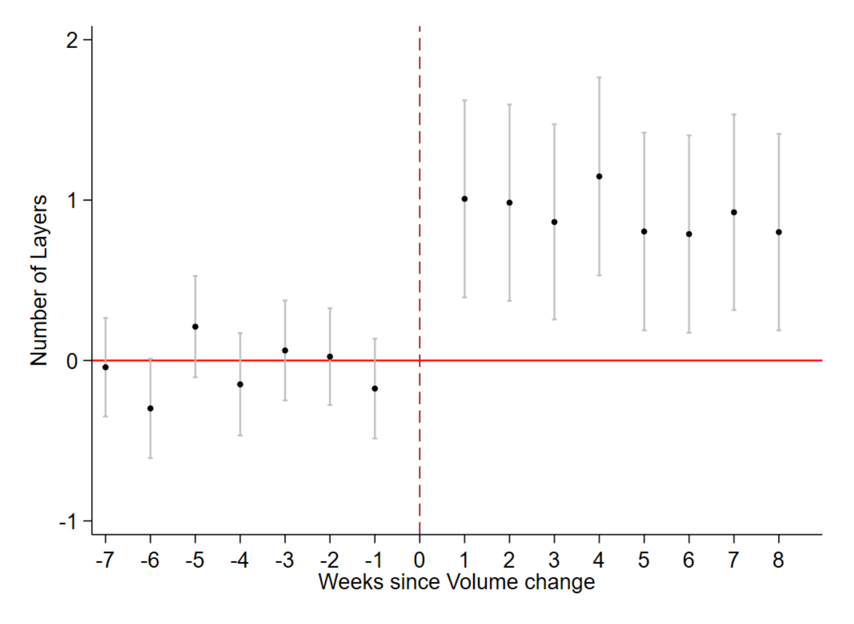
Note: Figure A1 plots the distribution of DPV-day observations, pooling across all days in the data. DPV is a standardized variable with mean 0 and standard deviation 1.

Figure A2: DPV and Plant-level Productivity Losses



Note: We plot the coefficients of a Distributed Lag Model of order 10. We define a High DPV day as a day when DPV goes over the 75th percentile in our sample. We control for a quadratic trend, year and month FE, and first lag of number of cars produced. The cumulative effect of the High DPV occurrence over the 10 days is -20.85 cars (with robust standard error 6.31). Confidence Intervals are computed using robust standard errors.

Figure A3: Event Study of Volume Changes on Layers for Pre-Existing Groups



This Figure shows the effect of volume changes from 8 weeks before the event to 8 weeks after the event on the number of layers, limiting the sample to pre-existing working groups only. Number of layers is defined as the number of positions in the working group. Standard errors clustered at weekly-shift level. 95% confidence intervals are presented in the figure. We control for month, year, and group fixed effects. We also control for a linear function of distance to the Volume change and to all other Volume and Model changes. Number of observations: 167 working groups x 16 weeks x 2 events.

Table A1: Descriptive Statistics on Model and Volume Changes

	Model Changes				Volume Changes			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Num of Cars	366.55	112.33	170.71	522.40	388.75	92.70	312.70	537.78
Num of Parts	1,421,760.00	834,783.50	1,838.32	2,416,469.00	1,790,505.00	505,012.30	1,425,545.00	2,577,422.00
Num of New Parts	189,047.50	174,787.90	69.96	533,349.20	0	0	0	0
Share of New Parts	0.13	0.11	0.04	0.37	0	0	0	0
Num of Models	1.44	0.74	1.00	3.10	1.26	0.52	1.00	2.19
Num of New Models	1.12	0.10	1.00	1.25	0	0	0	0
Number of Events				7				5

Note: The information presented comes from shift-day level information of production from 2012 to 2019. The Model and Volume changes information is the average of daily information in the month after the change happens.

Table A2: Impact of Model Changes on Productivity

	(1) DPV
0-3 weeks	0.745***
	(0.134)
4-7 weeks	0.198
	(0.186)
Observations	1,134
Obs. Level	Shift-Day
Mean	0.000

Note: Standard errors clustered by distance to event and shift. Number of observations: 2 shifts x 81 days x 7 events. Productivity measures are reported by the plant at the shift-day level. DPV is the number of defects per 100 vehicles, and is standardized. We control for month, year, and shift fixed effects. We also control for a linear function of distance to the Model change and to all other Model and Volume changes. * p<0.1, ** p<0.05, *** p<0.01

Table A3: Impact of Model Changes on Group Composition

	(1) Layers	(2) Layers New FL	(3) Layers FL	(4) Layers Mid FL	(5) Layers ML	(6) All FL/ML	(7) ML/S
0-3 weeks	-0.666***	-0.111***	-0.282***	-0.061***	-0.455***	-0.525***	-0.155***
	(0.124)	(0.026)	(0.053)	(0.013)	(0.084)	(0.120)	(0.044)
4-7 weeks	-0.285	-0.044	-0.146*	-0.027	-0.217*	-0.244	-0.004
	(0.190)	(0.041)	(0.083)	(0.021)	(0.129)	(0.190)	(0.069)
Observations	7,040	7,040	7,040	7,040	7,040	7,040	7,040
Obs. Level	Group-Week	Group-Week	Group-Week	Group-Week	Group-Week	Group-Week	Group-Week
Mean	6.373	0.871	2.797	0.495	4.163	5.026	1.920

Note: Standard errors clustered by distance to event and working group. Number of observations: 220 working groups x 16 weeks x 2 events. Layers is defined as the number of positions in a working group. FL: Front-line worker. New FL: FL who recently started and has received only basic training. Mid FL: FL in line for promotion to ML. ML: Mid-line worker. S: superior. We use as controls month, year and group fixed effects. We control for a linear function of distance to Model change and distances to all other Model and Volume changes. * p<0.1, ** p<0.05, *** p<0.01

Table A4: Impact of Model Changes on Number of Groups and Groups size

	(1) Num of Groups	(2) Group Size
0-3 weeks	-0.786	0.039
	(1.703)	(0.240)
4-7 weeks	-0.789	0.105
	(2.304)	(0.349)
Observations	64	64
Obs. Level	Shift-Week	Shift-Week
Mean	58	20

Note: Standard errors clustered by distance to event and shift. Number of observations: 2 shifts x 16 weeks x 2 events. We use as controls month, year, and shift fixed effects. We control for a linear function of distance to Model change and distances to the other Model and Volume changes. Number of groups is the number of working groups in each shift. Group size is the average number of employees in each working group. * p<0.1, ** p<0.05, *** p<0.01

Table A5: Impact of Volume Changes on Productivity

	(1) DPV
0-3 weeks	0.490***
	(0.085)
4-7 weeks	0.326***
	(0.083)
Observations	810
Obs. Level	Shift-Day
Mean	0.000

Note: Standard errors clustered by distance to event and shift. Number of observations: 2 shifts x 81 days x 5 events. We use as controls month, year, and shift fixed effects. We control for a linear function of distance to the Volume change and distances to the other Volume and Model changes. * p<0.1, ** p<0.05, *** p<0.01

Table A6: Impact of Volume Changes on Group Composition

	(1) Layers	(2) Layers New FL	(3) Layers FL	(4) Layers Mid FL	(5) Layers ML	(6) All FL/ML	(7) ML/S
0-3 weeks	0.850***	0.128***	0.358***	0.0733***	0.559***	0.610***	0.240***
	(0.098)	(0.018)	(0.043)	(0.010)	(0.065)	(0.094)	(0.033)
4-7 weeks	0.706***	0.115***	0.288***	0.074***	0.474***	0.493***	0.190***
	(0.117)	(0.022)	(0.052)	(0.013)	(0.078)	(0.115)	(0.042)
Observations	7,040	7,040	7,040	7,040	7,040	7,040	7,040
Obs. Level	Group-Week	Group-Week	Group-Week	Group-Week	Group-Week	Group-Week	Group-Week
Mean	5.284	0.686	2.314	0.415	3.414	4.158	1.694

Note: Standard errors clustered by distance to Volume change and working group level. Number of observations: 220 working groups x 16 weeks x 2 events. Layers is defined as the number of positions in a working group. FL: Front-line worker. New FL: FL who recently started and has received only basic training. Mid FL: FL in line for promotion to ML. ML: Mid-line worker. S: superior. We use as controls month, year and group fixed effects. We control for a linear function of distance to the Volume change and distances to the other Volume and Model changes. * p<0.1, ** p<0.05, *** p<0.01

Table A7: Impact of Volume Changes on Number of Groups and Groups size

	(1) Num of Groups	(2) Groups size	(3) New Group size	(4) Old Groups size
0-3 weeks	3.765**	-0.298	21.074***	1.724
	(1.737)	(0.381)	(0.669)	(2.805)
4-7 weeks	3.624*	-0.372	21.104***	1.253
	(1.949)	(0.615)	(0.714)	(4.124)
Observations	64	64	64	64
Obs. Level	Shift-Week	Shift-Week	Shift-Week	Shift-Week
Mean	54	20	0	19

Note: Standard errors clustered by distance to Volume change and shift. Number of observations: 2 shifts x 16 weeks x 2 events. We use as controls month, year, and shift fixed effects. We control for a linear function of distance to the Volume change and distances to the other Volume and Model changes. Number of groups is the number of working groups in each shift. Group size is the average number of employees in each working group. * p<0.1, ** p<0.05, *** p<0.01

Table A8: Impact of Volume Changes on Cars per Employee/Group and Parts per Employee/Group

	(1) Cars per Emp	(2) Cars per Group	(3) Parts per Emp	(4) Parts per Group
0-3 weeks	0.0371***	0.890***	194.4***	4,628***
	(0.00747)	(0.133)	(37.18)	(660.0)
4-7 weeks	0.0471***	1.058***	249.4***	5,583***
	(0.00942)	(0.180)	(46.37)	(881.3)
Observations	162	162	162	162
Obs. Level	Day	Day	Day	Day
Mean	0.256	5.025	1211.198	23750.390

Note: Standard errors clustered by distance to Volume change. Number of observations: 81 days x 2 events. Cars per Emp: number of cars produced over number of workers by day. Cars per group: number of cars produced over number of working groups per day. Parts per Emp: number of parts used in produced cars over the number of employees by day. Parts per Group: number of parts used in produced cars over the number of working groups by day. We use as controls month and year fixed effects. We control for a linear function of distance to the Volume change and distances to the other Volume and Model changes.
* p<0.1, ** p<0.05, *** p<0.01

B Model

We use the model of c to understand how product cycles affect the organization of the firm (i.e., optimal production structure of the firm like the number of layers, number of production workers and the knowledge they acquire at each layer). As we mentioned before, every time a new model is introduced, the share of new parts in the car increases, increasing the complexity of the problems solved by the workers. Anecdotal evidence shared by the partner firm and the empirical evidence presented in Section 4, suggest an increase in the stock of knowledge and a reduction in the number of layers every time the company faced a "model change." In this section, we explore under what conditions this anecdotal evidence can be rationalized by the model and is optimal for the firm. We contrast these results with the impacts of a positive volume change that increases quantity produced, for which we show the organizational response is a monotonic, permanent increase in both employment and management layers, consistent with prior evidence from manufacturing in high-income countries (Caliendo et al., 2020, 2015).

Layers Problem: Suppose that a firm pays a wage w to each of its workers and wishes to produce q units of output. The firm chooses the optimal number of layers L in order to minimize the cost of producing q units of output while paying a wage w to each worker and manager. The firm solves

$$C\left(q;w\right) \equiv \min_{L\geq 0}\left\{C_L\left(q;w\right)\right\}, \tag{7}$$

where $C(q;w)$ denotes the minimum variable cost of producing q units of output and $C_L(q;w)$, the minimum cost of producing q units of output with an organization with $L+1$ layers, it is defined by (10) below.

Workers and Knowledge Problem: Suppose that a firm has chosen an organization with $L+1$ layers. The amount of workers the firm hires at the lowest layer ($l=0$) is denoted by

n^0_L, and the knowledge they acquire is denoted by z^0_L. At an intermediate layer l $(0 < l < L)$, it hires n^l_L managers, each one with knowledge z^l_L. Since there is only one entrepreneur in the firm, then $n^L_L = 1$. z^L_L denotes the entrepreneur's knowledge.

If the firm hires n^0_L workers at the lowest layer, each of which possess knowledge z^0_L, then each of these workers is capable to solve a fraction $F(z^0_L)$ of the problems that the firm faces. The fraction of unsolved problems $1 - F(z^0_L)$ is left for the next layer, $l = 1$. Note that managers at layer $l = 1$ spend a fraction h of their unit of time listening to the workers' problems, which implies that each manager can deal with at most $\frac{1}{h}$ problems. It follows that n^1_L must be proportional to the amount of unsolved problems they can deal with, i.e.,

$$n^1_L = h n^0_L \left(1 - F\left(z^0_L\right)\right). \tag{8}$$

Note that as the cost of communication h increases, n^1_L increases. Similarly, the amount of managers at layer l $(l > 1)$ must be proportional to the amount of unsolved problems at that point,

$$n^{l+1}_L = n^l_L \left(1 - F\left(z^l_L\right)\right) \text{ for all } 0 < l < L. \tag{9}$$

Given a sequence of knowledge $\{z^l_L\}^L_{l=0}$, equation (9) gives us an *evolution law* for the population within the firm. Note that, since $n^L_L = 1$, given a sequence of knowledge $\{z^l_L\}^L_{l=0}$, the evolution law completely determines the values of n^l_L for $0 \leq l < L$. It follows that the firm only has to find the optimal knowledge sequence $\{z^l_L\}^L_{l=0}$ that allows it to produce q units of output. That is, the firm solves

$$C_L\left(q;w\right) \equiv \min_{\left\{n_L^l, z_L^l\right\}_{l=0}^{L} \geq 0} \sum_{l=0}^{L} \overbrace{n_L^l}^{\text{labor}} \underbrace{w\left(cz_L^l+1\right)}_{\text{wages}}, \tag{10}$$

$$\text{s.t. } A \cdot F\left(Z_L^L\right) n_L^0 \geq q,$$

$$n_L^l = n_L^0 h e^{-\lambda Z_L^{l-1}},$$

where $F\left(z\right) = 1 - e^{-\lambda z}$, for $0 < l < L$, and $n_L^L = 1$. Here, $Z_L^L \equiv \sum_{l=0}^{L} z_L^l$ represents the cumulative knowledge of the firm. Note that in (10), the firm is minimizing the cost of the labor plus the cost of educating the workers.

B1 Stock of Knowledge

Proposition 1. *If a firm wants to increase its cumulative workforce knowledge by adding a new layer, such that* $Z_{L+1}^L - Z_L^{L-1} = \varepsilon$, *for some* $\varepsilon > 0$, *then* $z_L^L > z_{L+1}^{L+1}$ *and* $z_L^l > z_{L+1}^l$ *for* $0 \leq l < L$.

System (13) provides explicit formulas to determine the knowledge at every layer, and Proposition 1 uses this information to depict how a firm redistributes its total knowledge when it changes layers. More specifically, when $\varepsilon \to 0$ this transference of knowledge results into a more efficient organizational structure, since the firm would not have to invest in increasing its cumulative knowledge directly. Instead, by disclaiming less information per layer it will be able to afford additional layers, and even reduce its average costs for levels of production q large enough. Furthermore, the proof seen in the Appendix B4 shows how the knowledge of the workers at higher layers is the most affected by moving from L to $L+1$ layers, while the entrepreneur is the one that gives up the least amount of knowledge with the transition.

B2 Model Changes

Using this general framework, we study how the organization of the firm changes endogenously in response to product cycles. As we mentioned before, every time a new model is introduced, the complexity of the problems increases as the share of new parts in the car increases, which we model as a reduction in λ. A lower λ, decreases the production level q_L at where the firm should move from L to $L+1$ layers (Caliendo and Rossi-Hansberg, 2012). However, if the firm is uninterested in changing its number of layers, it must invest in modifying other parameters in order to balance λ's impact. From anecdotal evidence, investing in reducing the training cost c seems like a plausible option, so we asked: *What are the admissible ratios* $\Delta c : \Delta\lambda$ *that prevent the firm from increasing layers?*[22]

To do so, define the production level where the average cost of the firm working with L or $L+1$ layers intersect as a function of λ and c ($i.e.$ $q_L := q_L(\lambda, c)$). Therefore, the unitary vector $\vec{v}_- := -\alpha\hat{\lambda} - \sqrt{1-\alpha^2}\,\hat{c}$ (or $\vec{v}_+ := -\alpha\hat{\lambda} + \sqrt{1-\alpha^2}\,\hat{c}$) for $\alpha \in [-1,1]$ encodes the directional derivative of q_L at $\vec{p} := (\lambda_0, c_0)$ as

$$D_{\vec{v_-}} q_L(\vec{p}) = -\alpha \left(\frac{\partial q_L}{\partial \lambda}\right)\bigg|_{\vec{p}} - \sqrt{1-\alpha^2}\left(\frac{\partial q_L}{\partial c}\right)\bigg|_{\vec{p}}.$$

We support the usage of $\vec{v}_-$ over $\vec{v}_+$ because empirical tests suggest that $\partial q_L/\partial\lambda > 0$ and $\partial q_L/\partial c < 0$. Since model changes imply a drop in λ, the firm should restrict itself to $\alpha \geq 0$. For this case,

$$D_{\vec{v_-}} q_L(\vec{p}) \geq 0 \quad \text{if} \quad 0 \leq \alpha^2 \leq \left(\frac{\partial q_L}{\partial c}\right)^2 \left[\left(\frac{\partial q_L}{\partial \lambda}\right)^2 + \left(\frac{\partial q_L}{\partial c}\right)^2\right]^{-1} =: \kappa_\alpha, \quad \kappa_\alpha \leq 1.$$

Hence, the admissible directions on the third quadrant of the λc-plane in which q_L increases belong to the interval $\mathcal{D} := [0, \sqrt{\kappa_\alpha}]$. As a consequence, the firm can choose any $\alpha \in \mathcal{D}$ and establish a $\Delta c : \Delta\lambda$ ratio of $\sqrt{1-\alpha^2} : \alpha$ aiming to maintain or even reduce its optimal number of layers.

[22] We consider Δ as the absolute change of a variable, that is $\Delta\lambda := |\lambda_1 - \lambda_0|$ and $\Delta c := |c_1 - c_0|$.

The appropriate sign of $\partial q_L/\partial\lambda$ and $\partial q_L/\partial c$ might vary depending on the operating point $\vec{p}$. To verify which one is the case, we also provide explicit formulas for these partial derivatives in the proof of Lemma 1, in Appendix B4.

Lemma 1. *For any $L > 1$, if there exists a unique z_L^L that satisfies system (13), then $\partial q_L/\partial\lambda$ and $\partial q_L/\partial c$ can be explicitly and uniquely determined.*

In particular, if q_L presents the usual behavior at $\vec{p}$,[23] then $D_{\vec{v}_-} q_L(\vec{p})$ increases for $\alpha \to 0$, while $D_{\vec{v}_-} q_L(\vec{p}) \to 0$ for $\alpha \to \kappa_\alpha$, and thus, the firm might be tempted to select very small values of α. However, the new operating point $\vec{p_1} = (\lambda - \Delta\lambda, c_0 - \Delta c)$ must have positive coordinates and since Δc is inversely correlated to α, the firm has to be aware of not choosing an α small enough for $\Delta c > c_0$. We formalize this analysis in the following proposition.

Proposition 2. *Suppose $(\partial q_L/\partial\lambda)|_{\vec{p}} > 0$ and $(\partial q_L/\partial c)|_{\vec{p}} < 0$, for $\vec{p} = (\lambda_0, c_0)$, and $\lambda_1 < \lambda_0$. If the firm can invest in decreasing the training cost freely and wants to maintain a production level $q \in [q_{L-1}, q_L]$, then it has to reduce c_0 by at least*

$$\Delta c = \frac{(\lambda_0 - \lambda_1)\sqrt{1-\alpha^2}}{\alpha}, \quad \text{where } \alpha^2 = \left(\frac{\partial q_L}{\partial c}\right)^2 \left[\left(\frac{\partial q_L}{\partial \lambda}\right)^2 + \left(\frac{\partial q_L}{\partial c}\right)^2\right]^{-1}\Bigg|_{\vec{p}} \text{ and } \alpha \geq 0,$$

to avoid increasing its number of layers. Moreover, for any $c' < c_0 - \Delta c$, there exists levels of production q at which the firm opts for an organization with fewer layers.

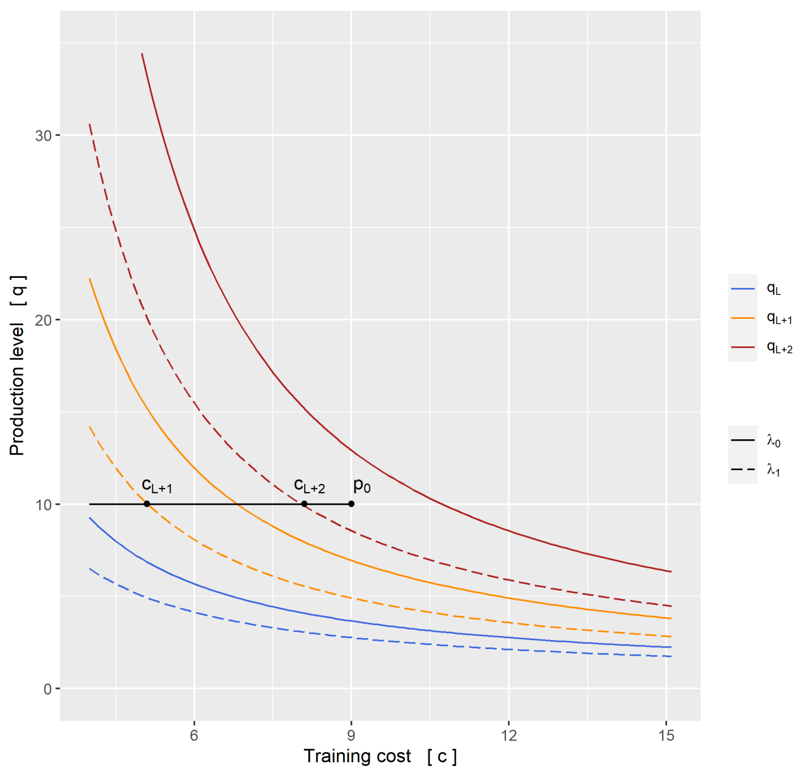
Essentially, the partial derivatives $(\partial q_L/\partial\lambda)|_{\vec{p}} > 0$ and $(\partial q_L/\partial c)|_{\vec{p}} < 0$ quantify the opposing effects on the intersecting point q_L of reducing λ or c. The former constitutes the negative impact of the firm having to face more complex problems, while the latter is the favorable scenario where it can train its workers more efficiently. In that sense, Proposition 2 provides the minimum investment in training costs required to balance the effect of problem complexity.

[23] $\partial q_L/\partial\lambda > 0$ and $\partial q_L/\partial c < 0$.

In addition, there is a simple graphical method, to determine the training costs c that would result in a reduction of layers despite a model change. It consists on the firm calculating the intersecting levels of production q_L for some training costs $c > 0$. Then, for a fixed production q, it identifies the active number of layers L at the original training cost c_0, which will indicate how much cost reduction is needed to reach the q_L or q_{L-1} zone.

For example, in Figure B1, we suppose that the firm maintains a production level $q = 10$, and that it has an original operating point $p_0 = (2, 9)$, with 5 active layers. Then, the model changes from $\lambda_0 = 2$ to $\lambda_1 = 1.5$, and so, if the firm preserves the original training cost $c_0 = 9$ then it has to level up to 6 layers. Conversely, if the firm wants to keep operating with 5 layers, it must reduce its training cost to some point in the interval $(c_4, c_5] := (5.1, 8]$.

Figure B1: Level Curves and Model Changes



Note: The parameters used in the simulation are $L = 3$, $\lambda_0 = 2$, $\lambda_1 = 1.5$, $h = 0.9$, $w = 1$, and $A = 1$.

B3 Endogenous Training cost

Here, we developed a variation for model (7) based on introducing a penalty $\mathcal{P}$ (or a fee) for investing in reducing the training cost, which results in model (11):

$$C\left(q;w\right) = \min_{\{L,\, c\} \geq 0} C_L\left(q;w\right) + \mathcal{P}(c,\lambda,L), \qquad \text{where } C_L\left(q;w\right) \text{ solves (10).} \tag{11}$$

In (11), for each c the firm has to establish the optimal distribution of knowledge $\{z_L^l\}_{l=0}^{L}$, and then it selects the optimal training cost c. In addition, the fee $\mathcal{P}$ is designed to reinforce the effect of training over the one of problem complexity, and reflect that reducing training costs is inexpensive for high values of c but gradually becomes costly. Moreover, since the fee is introduced to counter model changes, we consider penalties $\mathcal{P}$ that are constant for any level of production q. That is, we introduce the following assumption.

Assumption 1.

1. *$\mathcal{P}(c,\lambda,L)$ is decreasing with respect to c.*

2. *$\mathcal{P}(c,\lambda,L)$ is independent of the production levels q.*

Intuitively, choosing the optimal training cost for model (11) is a trade-off between lower costs $C_L(q;w)$ and the price to pay to achieve them represented by $\mathcal{P}$. Ideally, the penalty should be calibrated to be sensitive to problem complexity changes so it can induce sufficient drops in the training costs. Nonetheless, the firm must be wary of extremely costly or volatile penalties, as their effects can overshadow those of the total cost function $C_L(q;w)$, which is the main object of study.

In this paper, we propose the penalty $\mathcal{P}(c,\lambda,L) = (c - (\vartheta\lambda + \vartheta^k L))^{-1}$, where $\vartheta \in \mathbb{R}^+$, $\vartheta > 1$ and $k \in (0,1)$. Using this penalty, we define the auxiliary function

$$\Psi(q,\lambda) = \min_{c\geq 0}\left(C_L\left(q;w\right) + \mathcal{P}(c,\lambda,L)\right) - \min_{c\geq 0}\left(C_{L+1}\left(q;w\right) + \mathcal{P}(c,\lambda,L+1)\right). \tag{12}$$

Suppose that $q_L(\boldsymbol{\lambda})$ is the value of q for which the firm should move from L to $L+1$ layers, for $\boldsymbol{\lambda}$ fixed. Proposition 3 proposes sufficient conditions to guarantee that $q_L(\lambda) < q_L(\lambda_1)$ for $\lambda - \lambda_1 > 0$ sufficiently small.

Proposition 3. *If $\partial\Psi(q_L(\boldsymbol{\lambda}), \boldsymbol{\lambda})/\partial\lambda > 0$, then there is a neighborhood of $\boldsymbol{\lambda}$ where we can parameterize $q_L \equiv q_L(\lambda)$, and this parameterization satisfies $\partial q_L/\partial\lambda < 0$.*

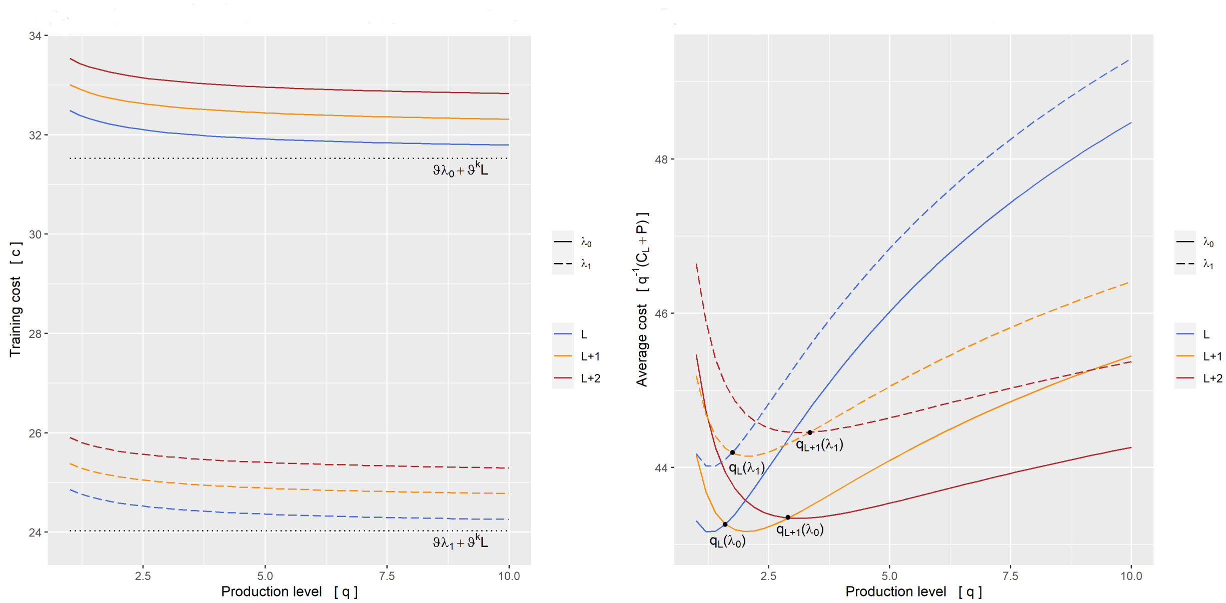
Figures B2a and B2b show the effect of this penalty on the optimal training cost c, and the resulting average cost modelled according to (11). More specifically, in Figure B2a, we see that for every level of production q the optimal training cost drops for a model change, which is a consequence of the $\vartheta\lambda$ component. Also, the training curves tend asymptotically to $\vartheta\lambda + \vartheta^k L$. This separates the optimal training costs for firms with different number of layers, primarily due to the $\vartheta^k L$ component. Notice that k acts as a weight between the problem complexity and layers components of the penalty. That said, the fact that $k \in (0,1)$ is intended to imply that the penalty is more sensitive to λ than to L. Graphically, this is seen in Figure B2a with the cost curves between a firm with different layers being closer that those between a firm facing two different problem complexities.

Besides, with this penalty we guarantee that $c_{L+1} > c_L$ to break $\partial C_L/\partial q$ and $\partial C_{L+1}/\partial q$ apart.[24] These changes seem to lead us to the desired endogenous response as seen in Figure B2b where, despite the model change, the firm in general operates with the same amount of layers, and the particular production levels where it does not is because it has the option to reduce them. More specifically, the production levels where the firm considered can reduce its layers are $q \in [q_L(\lambda_0), q_L(\lambda_1)] \approx [1.6, 1.7]$ or $q \in [q_{L+1}(\lambda_0), q_{L+1}(\lambda_1)] \approx [2.9, 3.3]$.[25]

[24] A graphical intuition about the changes in q_L is that it will move to the right (*left*) if and only if $\partial(C_L + \mathcal{P})/\partial q < \partial(C_{L+1} + \mathcal{P})/\partial q$ ($>$). Furthermore, from the envelope theorem together with the first order conditions for (7) we know that $\partial C_L/\partial q = whc_L e^{\lambda z_L^L}/\lambda A$ for any $L > 1$, and since $\mathcal{P}$ does not depend on q, then $\partial(C_L + \mathcal{P})/\partial q = \partial C_L/\partial q$.

[25] The firm modelled has the parameters: $h = 0.9$, $w = 1$, and $A = 0.5$.

Figure B2: Endogeneous Training Cost



(a) Training cost curves.

(b) Endogenous average cost.

Note: The parameters used in the simulation are: $L = 3$, $\lambda_0 = 2$, $\lambda_1 = 1.5$, $h = 0.9$, $w = 1$, and $A = 1$.

B4 Proofs

Proof of Proposition 1. From Caliendo and Rossi-Hansberg (2012) it follows that (10) is equivalent to solving the system

$$\begin{cases} z_L^L = \frac{1}{\lambda}\ln\left(\frac{A}{Ae^{\lambda Z_L^{L-1}}-hq}\right), \\ z_L^0 = \frac{h}{\lambda}e^{\lambda z_L^L} - \frac{1}{\lambda} - \frac{1}{c}, \\ z_L^1 = \frac{1}{h\lambda}e^{\lambda z_L^0} - \frac{1}{\lambda} - \frac{1}{c}, \\ z_L^l = \frac{1}{\lambda}e^{\lambda z_L^{l-1}} - \frac{1}{\lambda} - \frac{1}{c} \text{ for } 1 < l < L. \end{cases} \tag{13}$$

Also, it follows that

$$z_L^L = \frac{1}{\lambda}\ln\left(\frac{A}{Ae^{\lambda Z_L^{L-1}} - hq}\right) \quad \text{and} \quad z_{L+1}^{L+1} = \frac{1}{\lambda}\ln\left(\frac{A}{Ae^{\lambda Z_{L+1}^{L}} - hq}\right).$$

Therefore, if we define $\varepsilon_L := z_L^L - z_{L+1}^{L+1}$ we have that

$$\varepsilon_L = z_L^L - z_{L+1}^{L+1} = \frac{1}{\lambda}\ln\left(\frac{Ae^{\lambda Z_{L+1}^{L}} - hq}{Ae^{\lambda Z_L^{L-1}} - hq}\right) > 0,$$

since $Z_{L+1}^L - Z_L^{L-1} = \varepsilon > 0$. Now, if $\varepsilon_l := z_L^l - z_{L+1}^l$ for $0 \le l < L$, we obtain that

$$\begin{aligned} \varepsilon_0 &= \frac{h}{\lambda}\left(e^{\lambda z_L^L} - e^{\lambda z_{L+1}^{L+1}}\right) = \frac{he^{\lambda z_L^L}}{\lambda}\left(1 - e^{-\lambda \varepsilon_L}\right) > 0, \\ \varepsilon_1 &= \frac{1}{h\lambda}\left(e^{\lambda z_L^0} - e^{\lambda z_{L+1}^0}\right) = \frac{e^{\lambda z_L^0}}{h\lambda}\left(1 - e^{-\lambda\varepsilon_0}\right) > 0 \quad \text{and} \\ \varepsilon_l &= \frac{1}{\lambda}\left(e^{\lambda z_L^{l-1}} - e^{\lambda z_{L+1}^{l-1}}\right) = \frac{e^{\lambda z_L^{l-1}}}{\lambda}\left(1 - e^{-\lambda\varepsilon_{l-1}}\right) > 0 \quad \text{for } 1 < l < L, \end{aligned}$$

which concludes the proof. □

Proof of Lemma 1. From Caliendo and Rossi-Hansberg (2012) we deduced that the functional form of model (7) is given by:

$$C_L(q,w) = \begin{cases} \frac{wc}{\lambda}\left(\frac{hq}{A}e^{\lambda z_L^L} + (1-e^{\lambda z_L^{L-1}}) + \lambda z_L^L + \frac{\lambda}{c}\right) & \text{for } L>1, \\ \frac{wc}{\lambda}\left(\frac{hq}{A}e^{\lambda z_1^1} + \left(1-\frac{e^{\lambda z_1^0}}{h}\right) + \lambda z_1^1 + \frac{\lambda}{c}\right) & \text{for } L=1, \\ w\left(\frac{c}{\lambda}\ln\left(\frac{A}{A-q}\right)+1\right) & \text{for } L=0. \end{cases} \tag{14}$$

Therefore if q_L is the production level at which the average cost of a firm operating with L and $L+1$ layers intersect for any $L>1$, then q_L satisfies the equation:

$$q_L = \frac{A}{h}\left(\frac{e^{\lambda z_L^{L-1}} - e^{\lambda z_{L+1}^L} + \lambda\left(z_{L+1}^{L+1} - z_L^L\right)}{e^{\lambda z_L^L} - e^{\lambda z_{L+1}^{L+1}}}\right), \tag{15}$$

where z_L^{L-1}, z_L^L, z_{L+1}^L, and z_{L+1}^{L+1}, satisfy the system (13). Notice that the optimal knowledge per layer depends both on λ and c. In particular, the entrepreneur's knowledge also depends on the distribution of knowledge of all the subordinate layers. Hence, the strategy to establish the formula for $\partial q_L/\partial\lambda$ (and *mutatis mutandis* for $\partial q_L/\partial c$) is to first, write $\lambda(\partial z_L^l/\partial\lambda)$ as $K_1^\lambda(z_L^l) + K_2^\lambda(z_L^l)(\partial z_L^L/\partial\lambda)$ for every $l<L$, and subsequently, $\partial z_L^L/\partial\lambda$ as $K_1^\lambda(z_L^L) + K_2^\lambda(z_L^L)(\partial q/\partial\lambda)$. Notice that this is valid for both L and $L+1$. Then, deriving both sides of (15) by λ and replacing $\partial z_L^L/\partial\lambda$ in $\lambda\partial z_L^{L-1}/\partial\lambda$ and $\partial z_{L+1}^{L+1}/\partial\lambda$ in $\lambda\partial z_{L+1}^L/\partial\lambda$ leads to an equation where we can solve $\partial q_L/\partial\lambda$.

(i) To determine $\partial q_L/\partial\lambda$:

<u>*Step 1:*</u> From system (13) we establish that

$$\lambda\frac{\partial z_L^l}{\partial\lambda} = \begin{cases} \frac{1}{\lambda}\left(1-he^{\lambda z_L^L}\right) + he^{\lambda z_L^L}z_L^L + h\lambda e^{\lambda z_L^L}\frac{\partial z_L^L}{\partial\lambda}, & \text{for } l=0, \\ \frac{1}{\lambda} + \frac{e^{\lambda z_L^0}z_L^0}{h} + \frac{e^{\lambda(z_L^0+z_L^L)}}{\lambda}\left(\lambda z_L^L - 1\right) + \lambda e^{\lambda(z_L^0+z_L^L)}\frac{\partial z_L^L}{\partial\lambda} & \text{for } l=1, \\ \frac{1}{\lambda} + \sum_{k=1}^{l-1} e^{\lambda Z_L^{[k,l-1]}}z_L^k + \frac{e^{\lambda Z_L^{l-1}}z_L^0}{h} + \frac{e^{\lambda(Z_L^{l-1}+z_L^L)}}{\lambda}\left(\lambda z_L^L - 1\right) + \lambda e^{\lambda(Z_L^{l-1}+z_L^L)}\frac{\partial z_L^L}{\partial\lambda} & \text{for } l<L, \end{cases}$$

from which we obtain the expressions for $K_1^\lambda(z_L^l)$ and $K_2^\lambda(z_L^l)$ for any $l<L$.

Step 2: Since $\lambda z_L^L = \ln(A) - \ln(Ae^{\lambda Z_L^{L-1}} - hq)$ it follows that[26]

$$z_L^L + \lambda \frac{\partial z_L^L}{\partial \lambda} = -e^{\lambda z_L^L}\left[e^{\lambda Z_L^{L-1}}\left(Z_L^{L-1} + \sum_{l=0}^{L-1} K_1^\lambda(z_L^l) + K_2^\lambda(z_L^l)\frac{\partial z_L^L}{\partial \lambda}\right) - \frac{h}{A}\frac{\partial q}{\partial \lambda}\right]. \tag{16}$$

Hence, solving for $\partial z_L^L/\partial\lambda$ leads to

$$K_1^\lambda(z_L^L) = \frac{-1}{K_3^\lambda(z_L^L)}\left[z_L^L + e^{\lambda Z_L^L}\left(Z_L^{L-1} + \sum_{l=0}^{L-1} K_1^\lambda(z_L^l)\right)\right] \quad \text{and} \quad K_2^\lambda(z_L^L) = \frac{he^{\lambda z_L^L}}{AK_3^\lambda(z_L^L)}, \tag{17}$$

where $K_3^\lambda(z_L^L) = \left(\lambda + e^{\lambda Z_L^L}\sum_{l=0}^{L-1} K_2^\lambda(z_L^l)\right)$.

Step 3: Once we derive both sides of (15) by λ we obtain that

$$\begin{aligned}\frac{\partial q_L}{\partial\lambda}K_3(q_L) = K_1(q_L)&\left(e^{\lambda z_L^{L-1}}\frac{\partial(\lambda z_L^{L-1})}{\partial\lambda} - e^{\lambda z_{L+1}^L}\frac{\partial(\lambda z_{L+1}^L)}{\partial\lambda} + z_{L+1}^{L+1} - z_L^L + \lambda\left(\frac{\partial z_{L+1}^{L+1}}{\partial\lambda} - \frac{\partial z_L^L}{\partial\lambda}\right)\right)\\ &+ K_2(q_L)\left(e^{\lambda z_L^L}\frac{\partial(\lambda z_L^L)}{\partial\lambda} - e^{\lambda z_{L+1}^{L+1}}\frac{\partial(\lambda z_{L+1}^{L+1})}{\partial\lambda}\right),\end{aligned}$$

where $K_1(q_L) = e^{\lambda z_L^L} - e^{\lambda z_{L+1}^{L+1}}$, $K_2(q_L) = e^{\lambda z_{L+1}^L} - e^{\lambda z_L^{L-1}} + \lambda(z_L^L - z_{L+1}^{L+1})$, and $K_3(q_L) = A^{-1}h(K_1(q_L))^2$. Substituting $\partial(\lambda z_L^{L-1})/\partial\lambda$ and $\partial(\lambda z_{L+1}^L)/\partial\lambda$ according to *Step 1* and $\partial z_L^L/\partial\lambda$

[26] As an abuse of notation we omit the λ super index in $K_1^\lambda(z_L^l)$ and $K_2^\lambda(z_L^l)$, $0 \le l < L$.

and $\partial z_{L+1}^{L+1}/\partial\lambda$ according to *Step 2*, yields that $\partial q_L/\partial\lambda$ solves the equation:[27]

$$\begin{aligned}
\frac{\partial q_L}{\partial\lambda}K_3(q_L) = {} & K_1(q_L)\, e^{\lambda z_L^{L-1}}\left(z_L^{L-1} + K_1(z_L^{L-1}) + K_2(z_L^{L-1})K_1(z_L^L)\right) \\
& - K_1(q_L)\, e^{\lambda z_{L+1}^L}\left(z_{L+1}^L + K_1(z_{L+1}^L) + K_2(z_{L+1}^L)K_1(z_{L+1}^{L+1})\right) \\
& + K_1(q_L)\left(z_{L+1}^{L+1} - z_L^L + \lambda\left(K_1(z_{L+1}^{L+1}) - K_1(z_L^L)\right)\right) \\
& + K_2(q_L)\left(e^{\lambda z_L^L}\left(z_L^L + \lambda K_1(z_L^L)\right) - e^{\lambda z_{L+1}^{L+1}}\left(z_{L+1}^{L+1} + \lambda K_1(z_{L+1}^{L+1})\right)\right) \\
& + \frac{\partial q_L}{\partial\lambda}K_1(q_L)\left(e^{\lambda z_L^{L-1}}K_2(z_L^{L-1})K_2(z_L^L) - e^{\lambda z_{L+1}^L}K_2(z_{L+1}^L)K_2(z_{L+1}^{L+1})\right) \\
& + \frac{\partial q_L}{\partial\lambda}\lambda K_1(q_L)\left(K_2(z_{L+1}^{L+1}) - K_2(z_L^L)\right) \\
& + \frac{\partial q_L}{\partial\lambda}\lambda K_2(q_L)\left(e^{\lambda z_L^L}K_2(z_L^L) - e^{\lambda z_{L+1}^{L+1}}K_2(z_{L+1}^{L+1})\right).
\end{aligned}$$

(ii) To determine $\partial q_L/\partial c$:

Step 1: From system (13) we establish that

$$\lambda\frac{\partial z_L^l}{\partial c} = \begin{cases} \frac{1}{c^2} + he^{\lambda z_L^L}\frac{\partial z_L^L}{\partial c}, & \text{for } l = 0, \\ \frac{1}{c^2}\left(1 + \frac{e^{\lambda Z_L^{l-1}}}{h} + \sum_{k=1}^{l-1} e^{\lambda Z_L^{[k,l-1]}}\right) + e^{\lambda(Z_L^{l-1}+z_L^L)}\frac{\partial z_L^L}{\partial c} & \text{for } 0 < l < L, \end{cases}$$

from which we obtain the expressions for $K_1^c(z_L^l)$ and $K_2^c(z_L^l)$ for any $l < L$.

Step 2: Since $\lambda z_L^L = \ln(A) - \ln(Ae^{\lambda Z_L^{L-1}} - hq)$ it follows that

$$\lambda\frac{\partial z_L^L}{\partial c} = -e^{\lambda z_L^L}\left[\lambda e^{\lambda Z_L^{L-1}}\left(\sum_{l=0}^{L-1} K_1^c(z_L^l) + K_2^c(z_L^l)\frac{\partial z_L^L}{\partial c}\right) - \frac{h}{A}\frac{\partial q_L}{\partial c}\right].$$

Hence, solving for $\partial z_L^L/\partial c$ leads to

$$K_1^c(z_L^L) = -\frac{e^{\lambda Z_L^L}}{K_3^c(z_L^L)}\sum_{l=0}^{L-1} K_1^c(z_L^l) \quad \text{and} \quad K_2^c(z_L^L) = \frac{he^{\lambda z_L^L}}{\lambda A K_3^c(z_L^L)},$$

[27] As an abuse of notation we omit the λ super index in $K_1^\lambda(z_i^j)$ and $K_2^\lambda(z_i^j)$, for $i \in \{L, L+1\}$ and $j \in \{L-1, L, L+1\}$.

where $K_3^c(z_L^L) = \left(1 + e^{\lambda Z_L^L} \sum_{l=0}^{L-1} K_2^c(z_L^l)\right)$.

<u>*Step 3:*</u> Once we derive both sides of (15) by c we obtain that

$$\begin{aligned} \frac{\partial q_L}{\partial c} \frac{K_3(q_L)}{\lambda} = K_1(q_L) & \left(e^{\lambda z_L^{L-1}} \frac{\partial z_L^{L-1}}{\partial c} - e^{\lambda z_{L+1}^L} \frac{\partial z_{L+1}^L}{\partial c} + \frac{\partial z_{L+1}^{L+1}}{\partial c} - \frac{\partial z_L^L}{\partial c} \right) \\ & + K_2(q_L) \left(e^{\lambda z_L^L} \frac{\partial z_L^L}{\partial c} - e^{\lambda z_{L+1}^{L+1}} \frac{\partial z_{L+1}^{L+1}}{\partial c} \right), \end{aligned} \tag{18}$$

where $K_1(q_L) = e^{\lambda z_L^L} - e^{\lambda z_{L+1}^{L+1}}$, $K_2(q_L) = e^{\lambda z_{L+1}^L} - e^{\lambda z_L^{L-1}} + \lambda(z_L^L - z_{L+1}^{L+1})$, and $K_3(q_L) = A^{-1}h(K_1(q_L))^2$. Substituting $\partial z_L^{L-1})/\partial c$ and $\partial z_{L+1}^L/\partial c$ according to *Step 1* and $\partial z_L^L)/\partial c$ and $\partial z_{L+1}^{L+1}/\partial c$ according to *Step 2*, yields that $\partial q_L/\partial c$ solves the equation:[28]

$$\begin{aligned} \frac{\partial q_L}{\partial c} K_3(q_L) = K_1(q_L) & \left(e^{\lambda z_L^{L-1}} \left(K_1(z_L^{L-1}) + K_2(z_L^{L-1}) K_1(z_L^L) \right) \right) \\ & - K_1(q_L) e^{\lambda z_{L+1}^L} \left(K_1(z_{L+1}^L) + K_2(z_{L+1}^L) K_1(z_{L+1}^{L+1}) \right) \\ & + K_1(q_L) \left(K_1(z_{L+1}^{L+1}) - K_1(z_L^L) \right) + K_2(q_L) \left(e^{\lambda z_L^L} K_1(z_L^L) - e^{\lambda z_{L+1}^{L+1}} K_1(z_{L+1}^{L+1}) \right) \\ & + \frac{\partial q_L}{\partial c} K_1(q_L) \left(K_2(z_L^L) \left(e^{\lambda z_L^{L-1}} K_2(z_L^{L-1}) - 1 \right) + K_2(z_{L+1}^{L+1}) \left(1 - e^{\lambda z_{L+1}^L} K_2(z_{L+1}^L) \right) \right) \\ & + \frac{\partial q_L}{\partial c} K_2(q_L) \left(e^{\lambda z_L^L} K_2(z_L^L) - e^{\lambda z_{L+1}^{L+1}} K_2(z_{L+1}^{L+1}) \right). \end{aligned}$$

□

Proof of Proposition 2*.* Let q be a fixed level of production for a firm operating with L layers (*i.e.* $q \in [q_{L-1}(\lambda_0, c_0), q_L(\lambda_0, c_0)]$). Without loss of generality, consider $q_L(\lambda, c)$ and its directional derivative[29]

$$D_{\vec{v}} q_L(\vec{p}) = -\alpha \left(\frac{\partial q_L}{\partial \lambda} \right) \Bigg|_{\vec{p}} - \sqrt{1-\alpha^2} \frac{\partial q_L}{\partial c}, \quad \text{for} \quad \vec{v} = -\alpha\hat{\lambda} - \sqrt{1-\alpha^2}\hat{c}, \ \alpha \in [0,1].$$

[28] As an abuse of notation we omit the c super index in $K_1^c(z_i^j)$ and $K_2^c(z_i^j)$, for $i \in \{L, L+1\}$ and $j \in \{L-1, L, L+1\}$.

[29] $\hat{\lambda}$ and $\hat{c}$ denote unitary vectors.

Moreover, since $(\partial q_L/\partial\lambda)|_{\vec{p}} > 0$ and $(\partial q_L/\partial c)|_{\vec{p}} < 0$, for

$$\alpha^2 = \left(\frac{\partial q_L}{\partial c}\right)^2 \left[\left(\frac{\partial q_L}{\partial \lambda}\right)^2 + \left(\frac{\partial q_L}{\partial c}\right)^2\right]^{-1}\Bigg|_{\vec{p}}, \tag{19}$$

we have that $D_{\vec{v}} q_L(\vec{p}) = 0$. That is, moving in the direction $\vec{v} = -\alpha\hat{\lambda} - \sqrt{1-\alpha^2}\hat{c}$ does not change the production level at which the firm transitions from L to $L+1$ layers, as long as α satisfies (19). In particular, the vector

$$\vec{w} = \Delta\lambda\ \hat{\lambda} + \frac{\Delta\lambda\sqrt{1-\alpha^2}}{\alpha}\hat{c} = \Delta\lambda\hat{\lambda} - \Delta c\hat{c}, \quad \text{where} \quad \Delta\lambda := \lambda_1 - \lambda_0,$$

has the same direction as $\vec{v}$. Consequently $D_{\vec{w}} q_L(\vec{p}) = D_{\vec{v}} q_L(\vec{p}) = 0$ for α, and if $q < q_L(\lambda_0, c_0)$ then $q < q_L(\lambda_1, c_0 - \Delta c)$, which means that the firm remains with at most L layers. Concluding that $q > q_{L-1}(\lambda_1, c_0 - \Delta c)$ is analogous.

Now, if $c_0 - c' > \Delta c$, then the unitary vector $\vec{v}_{c'} = -\alpha'\ \hat{\lambda} - \sqrt{1-(\alpha')^2}\ \hat{c}$ associated is such that $\alpha' < \alpha$ and thus, $D_{\vec{v}_{c'}} q_L(\vec{p}) > D_{\vec{v}} q_L(\vec{p}) = 0$, since $(\partial q_L/\partial\lambda)|_{\vec{p}} > 0$ and $(\partial q_L/\partial c)|_{\vec{p}} < 0$. This means that by moving in the direction $\vec{v}_{c'}$, the intersection between L and $L+1$ layers occurs at a higher production level. As a consequence, for every $q \in [q_L(\lambda_0, c_0), q_L(\lambda_1, c')]$, the firm would have initially operated with $L+1$ layers and then drop to L layers. □

Proof of Proposition 3. Model 11 is minimizable with respect to c because

$$\lim_{c\to\infty} C_L(q; w) = \lim_{c\to(\vartheta\lambda+\vartheta^k L)^+} \mathcal{P}(c, \lambda, L) = \infty.$$

Therefore, for any pair (q, λ), there exists $c_L(q, \lambda)$ that minimizes Model 11. Figure B2a suggests that this minimum is unique. Nevertheless, if it is not unique, it suffices to take the connected component containing one of those minimums to fully parameterize $c_L \equiv c_L(q, \lambda)$.[30]

[30] This connected component always exists because of the Implicit Function Theorem, and to the fact that $\frac{\partial^2 C_L}{\partial c^2} + \frac{\partial^2 \mathcal{P}}{\partial c^2} > 0$ when evaluated at a minimum.

With this, we can consider the function

$$\Phi_L(q, \lambda) = C_L(q, \lambda, c_L(q, \lambda)) + \mathcal{P}(c_L(\lambda, q), \lambda, L) \tag{20}$$

to be the minimum cost for a firm with L layers, production level q and complexity level $\boldsymbol{\lambda}$. $q_L(\boldsymbol{\lambda})$ then satisfies the equation $\Psi(q_L(\boldsymbol{\lambda}), \boldsymbol{\lambda}) = \Phi_L(q_L(\boldsymbol{\lambda}), \boldsymbol{\lambda}) - \Phi_{L+1}(q_L(\boldsymbol{\lambda}), \boldsymbol{\lambda}) = 0$. Moreover, as $q_L(\boldsymbol{\lambda})$ is the value for which the firm moves from L to $L+1$ layers, then $\frac{\partial \Psi}{\partial q}(q_L(\boldsymbol{\lambda}), \boldsymbol{\lambda}) \geq 0$.

First, we assume that $\frac{\partial \Psi}{\partial q}(q_L(\boldsymbol{\lambda}), \boldsymbol{\lambda}) > 0$. By continuity, there exists a neighborhood $\mathcal{D}_1 \subset \mathbb{R}^2$ of $(q_L(\boldsymbol{\lambda}), \boldsymbol{\lambda})$ for which $\frac{\partial \Psi}{\partial q}(\vec{p}) > 0$ if $\vec{p} \in \mathcal{D}_1$. Therefore, by the Implicit Curve Theorem, we can parameterize q_L as a function of λ for $\lambda \in \mathrm{proj}_2(\mathcal{D}_2)$, with $\mathcal{D}_2 \subset \mathcal{D}_1$. Moreover, the slope of this parameterization is given by

$$\frac{\partial q_L}{\partial \lambda} = -\frac{\partial \Psi}{\partial \lambda} \Big/ \frac{\partial \Psi}{\partial q}.$$

From $\frac{\partial \Psi}{\partial \lambda}(q_L(\boldsymbol{\lambda}), \boldsymbol{\lambda}) > 0$, we can conclude that there is a neighborhood $\mathcal{D}_3 \subset \mathbb{R}^2$ such that $\frac{\partial \Psi}{\partial \lambda}(\vec{p}) > 0$ if $\vec{p} \in \mathcal{D}_3$. This implies that for $\lambda \in \mathrm{proj}_2(\mathcal{D})$, $D := \mathcal{D}_2 \cap \mathcal{D}_3$, the slope of the parameterization $q_L \equiv q_L(\lambda)$ is negative.

On the other hand, if $\frac{\partial \Psi}{\partial q}(q_L(\boldsymbol{\lambda}), \boldsymbol{\lambda}) = 0$, the order of the first derivative that does not vanish must be odd, which implies we repeat the previous argument to the left and right of $\boldsymbol{\lambda}$ to parameterize $q_L \equiv q_L(\lambda)$ and show that the slope of this parameterization is again negative. $\square$