

Learning from Peers

Tat Y. Chan, Jia Li, and Lamar Pierce

Abstract. We study the impact of peers on coworker productivity growth using four years of individual cosmetic sales data from a Chinese department store. Learning in our retail setting is not trivial, since the sales process cannot be standardized and requires the workers to understand how to execute customized sales techniques that fit with heterogeneous and often unobservable customer needs. We find that the relative productivity of peers, within and across firm boundaries, is the major driving force behind a salesperson's learning curve. In contrast, the impact of the salesperson's own learning-by-doing is negligible and there is no evidence of forgetting. We exploit the existence of firm boundaries and two sales tasks of different difficulty in our data to find evidence consistent with two learning mechanisms: (1) learning by observing the sales techniques of high-skilled peers; and (2) direct teaching by these superior peers. Our paper suggests that both the organizational learning curves and inter-organizational knowledge spillovers observed in past studies have micro-foundations in individual peer-based learning.

I. Introduction

A well-developed literature in economics and management studies organizational learning curves. The primary focus of this research has been learning-by-doing (Arrow 1962), with empirical work identifying cost decreases with the cumulative production of aircraft (Alchian 1963; Benkard 2000), ships (Rapping 1965; Thompson 2001), trucks (Argote and Epple 1990), chemicals (Lieberman 1984), and semiconductors (Hatch and Mowery 1998).¹ This literature has also identified depreciation of knowledge, or organizational forgetting, in periods of low production in multiple empirical settings (Argote et al. 1990; Darr et al. 1995; Benkard 2000, 2004; Thompson 2007). These findings have led to important studies in the industrial organization literature demonstrating the implications of learning-by-doing and forgetting in shaping the dynamic competitive strategies of firms and their consequences for industry structure (e.g. Cabral and Riordan 1994; Besanko et al 2010).

These empirical findings, however, identify learning curves at the organizational or group level. We have little empirical evidence about the underlying individual mechanisms that generate this phenomenon, a shortfall frequently noted by those studying learning (Adler and Clark 1991; Argote 1999; Lapré et al. 2000). While one hypothesized mechanism is individual workers' learning-by-doing through experimentation and observation of own outcomes, economic theory on social learning suggests that peer-based learning mechanisms may play a critical role in worker productivity growth (Ellison and Fudenberg 1993, 1995; Bala and Goyal 2001; Young 2009). This peer-based learning may involve a combination of observing coworkers' practices and the active teaching of coworkers (Arrow 1994), with both mechanisms facilitating knowledge transfer between coworkers in an organization. Despite abundant

¹ See Thompson (2010) for a more extensive discussion of this literature.

empirical evidence of peer-based learning in education (Hoxby 2000; Sacerdote 2001; Carrell et al. 2009; Carrell and West 2010), crime (Bayer et al. 2009), economic development (Bandiera and Rasul 2006; Duflo et al. 2011; Kremer and Miguel 2007; Conley and Udry 2010), and the family (Huang et al. 2009), direct evidence of knowledge transfer among individual workers within organizations remain rare.²

In this paper we dissect the separate roles of peer-based learning and individual learning-by-doing (and forgetting) in worker productivity change. Our primary interest is to understand the importance of peer-based learning for individual workers and their organizations. We extend our investigation to study how peer-based learning may also occur across firm boundaries. This simultaneous examination of peer-based learning both within and across firms is motivated by more macro-level evidence of knowledge transfer across firms (Argote et al. 1990; Gruber 1998), between stores under a common franchisee (Darr et al. 1995), and across products (Benkard 2000; Thornton and Thompson 2001) or shifts (Epple, Argote, and Murphy 1996) within the same firm. Yet like other industrial learning research, these works do not identify individual-level mechanisms. The advancement of this research is critical due to its immediate implications for phenomena such as productivity growth (Ghemawat and Spence 1985; Lucas 1988), open source software (Lerner and Tirole 2002), agglomeration effects (Jaffe et al. 1993; Zucker et al. 1999), and economic development (Mostafa and Klepper 2011).

Our empirical application is cosmetic sales in a Chinese department store, an ideal setting for our research for two primary reasons. First, multiple manufacturers employ salespeople at co-

² Bercovitz and Feldman (2008) find evidence consistent with learning by observing that peers impact academic scientists' adoption of invention disclosure procedures. Studies linking improved team performance with team tenure also suggest that worker knowledge transfer is important for productivity gains (Edmondson, Bohmer, and Pisano 2001; Pisano, Bohmer, and Edmondson 2001; Huckman, Staats, and Upton 2009). In addition, a substantial literature in sociology and management has inferred knowledge transfer through the network structure of workers (Reagans and Zuckerman 2001; Ingram and Roberts 2000; Reagans and McKeivily 2003).

located counters on the same retail floor, allowing us to observe peer-based learning both within and across firms. More importantly, the complicated job task of our salespeople (instead of simpler and more standardized tasks in farms or factories) makes on-job learning more important. We conceptualize the sales task as a stochastic production function with parameters linking the output (sales) with input variables (selling activities) unknown to workers. In order to maximize sales, a new worker has to learn the true value of the parameters. While these true parameter values could in principle be achieved by experimenting with input variables and observing the outcome (Easley and Kiefer 1988), such a learning-by-doing process is likely to be extremely slow, because the sales task involves high-dimensional parameters: selling to each customer may represent a completely different process involving an entirely new set of parameters. A skilled salesperson must know how to identify unique and unobservable customer needs, match these customer needs with the right cosmetic products, and convince the customer to purchase these expensive products, often without direct evidence of product efficacy. Furthermore, the short-term cost of failure (i.e. lost sales commissions) may dissuade the worker from experimenting with input variables and eventually learning the true parameter values.³

We argue that for complicated production processes like ours, a more effective learning strategy is to observe the practice of peers, within and outside firms, and to seek direct instruction from peers on the different optimal techniques to different customers. The value of such peer-based learning is not identical across all coworkers, since observing the successful practice of a more productive worker should convey more information on the true parameter values in the production function than observing failures from lesser peers. The value of direct teaching by peers is similarly asymmetric. The existence of different cosmetic products (as we

³ Easley and Kiefer (1988) also showed that under general conditions the worker's beliefs will converge to a limit distribution that may not be concentrated at the true parameter values.

will explain later) and firm boundaries in our setting allows us to dissect the two mechanisms of peer-based learning. While some knowledge may be learned through directly observing other salespeople, more difficult products to sell may require active teaching from peers (Arrow 1994), a process unlikely across firm boundaries.

We build a reduced-form empirical model allowing an individual's productivity in any period to depend on her knowledge accumulated from past interaction with peers, her own learning-by-doing, and forgetting. Assuming that the worker will allocate the optimal level of input variables to maximize expected sales, better knowledge of the parameter values in the production function will lead to better input decision and consequently higher sales. Peer-based learning therefore can be inferred from the dynamic change in the worker's sales as a function of the variation in the pool of peers. This variation comes from the high turnover rate we observe in the data. We further exploit other variations in data that can reasonably be treated as exogenous shocks to help identify the learning. First, frequent leaves of absence by workers for vacation and personal reasons (e.g. sick or maternity leave) provide additional fluctuations in peers with whom the worker works. Furthermore, cosmetic counters are abruptly relocated on the retail floor in the middle of our sample due to construction surrounding the department store, immediately changing the stock of peers at adjacent counters for each worker. This shock helps us to identify the impact of learning from peers at competing counters.

Using four years of individual sales data for 92 salespeople working for 11 co-located counters, we find that the role of learning-by-doing is very small in the productivity improvement of workers. The impact from forgetting is also insignificant. Learning from peers either at the same counter or from adjacent competing counters dominates the learning process. For instance, by working alone for 180 hours in the first employment month, a new salesperson

only improves her sales by 0.5%. In contrast, if she constantly works with a within-counter peer with twice her productivity, her sales will grow by 23%. Such a peer at an adjacent competing counter would generate productivity growth of 5%. With whom a salesperson works fully captures the large variation of productivity growth across new workers in the data. Relative peer productivity also influences knowledge transfer across firm boundaries, suggesting that worker interactions may be the source for the knowledge transfers across firms, locations, and product categories documented in past studies.

We further explore the mechanism of peer-based learning by comparing the magnitude of such effects in two sub-categories with different sales task difficulties – skin care and makeup. Skin care products, as credence goods, are more difficult to sell than makeup products. We find that for skin care products, cross-counter learning is far weaker than within-counter learning, while for makeup products the two effects are similar. We argue that workers can glean knowledge from observation on simpler tasks, thereby making cross-counter peers a viable source for learning. On the other hand, teaching from peers, which is unlikely across counters, is essential for more complicated tasks. We also find that peer-based learning at the firms using individual-based compensation systems is different from this process at those using team-based compensation, suggesting that financial incentives may induce workers to invest different levels of time and effort toward learning from or actively teaching their peers.

Finally, we investigate whether our results could be driven by the endogenous selection of workers into brand-based counters. Through reduced-form regressions that specifically control for firms selecting workers based on learning capability, we find results largely consistent with our main findings. We further examine other alternative explanations for our results, including customer loyalty, reflection problems, and peer effects in work ethic. We present more evidence

that dispels these alternative hypotheses and supports the existence of peer-based learning in the data.

Understanding the underlying mechanisms in organizational learning curves has critical implications for firm policy and market equilibrium. If individual learning-by-doing and forgetting are the main mechanisms, firms can improve their competitiveness over time by gaining market share through aggressive pricing or mergers and acquisitions, shaping industry concentration in the long-run. However, for complicated production tasks it may be more effective for firms to design a work environment that encourages worker interaction that facilitates learning from successful or experienced peers. The competitive advantage of a firm therefore will rely more on organizational decisions such as job assignments, team formation, and turnover of workers, rather than the production scale or market share. Understanding mechanisms of individual worker learning is therefore critical to the literature of industrial organization and personnel economics.

The structure of the paper is as follows. Section 2 discusses the empirical setting of the paper. In Section 3 we develop a model of worker learning. Section 4 presents the results. Section 5 concludes this paper.

II. Empirical Setting

Our empirical setting is cosmetic sales in a department store in a large metropolitan area in Eastern China. This department store is one of the largest in China in both sales and profit and has 15 major brands in its cosmetics department, with each occupying a counter in the same floor area. These brands hire their own workers to promote and sell their products, while paying the

department store a share of their revenues. The cosmetics floor area effectively becomes an open market, with multiple firms competing for customers in a shared space.⁴

We observe cosmetic sales for 11 of the 15 counters over a four-year period (January 1, 2003 – December 31, 2006). Descriptive statistics for these brands are summarized in Table 1. The counters vary both in average price and total revenue. The 11 brands in our data use two different compensation systems: team-based commissions (TC) and individual-based commissions (IC). Four of the brands use team-based commissions and pay each worker a monthly salary of 900-1000 Chinese Renminbi (CNY)⁵ plus 0.5% of the monthly total counter sales. The other seven brands use individual-based commissions. In these counters, workers are given a monthly salary of CNY800-900 plus 2% of personal monthly sales.

Table 1: Descriptive Statistics of Cosmetic Brands

	Compensation System (IC/TC)	Annual Sales Revenue (CNY in Thousand)	Average Price (CNY)			Average Transaction Size (Units)
			All Products	Skin Care Products	Makeup Products	
Brand 1	IC	3742.86	131.28	140.38	106.96	1.83
Brand 2	TC	3602.28	106.64	119.67	93.88	1.54
Brand 3	TC	3145.11	85.28	89.69	79.59	1.57
Brand 4	TC	1196.08	127.76	141.75	91.96	1.71
Brand 5	TC	682.02	47.84	55.23	44.39	1.60
Brand 6	IC	1058.30	130.00	143.68	103.68	1.56
Brand 7	IC	1752.44	150.80	164.19	110.83	1.34
Brand 8	IC	525.51	108.40	115.11	73.33	1.71
Brand 9	IC	1054.94	118.08	133.51	74.20	2.06
Brand 10	IC	763.89	136.24	143.79	100.39	1.66
Brand 11	IC	725.77	109.76	114.40	102.68	1.48

In each counter, salespeople work in one of three overlapping shifts during the seven days per week that the department store is open: first shift from 9am to 3pm, second shift from 12pm to 6pm, and third shift from 3pm to 9pm. Workers typically rotate shifts every day, which

⁴ This setup is similar to inside contracting systems historically prevalent in manufacturing (e.g Buttrick 1952; Williamson 1980; Bucheli et al. 2010).

⁵ One US dollar was about CNY7.6 during our observation period.

ensures that each salesperson will work with a variety of peers at the same counter and across counters.

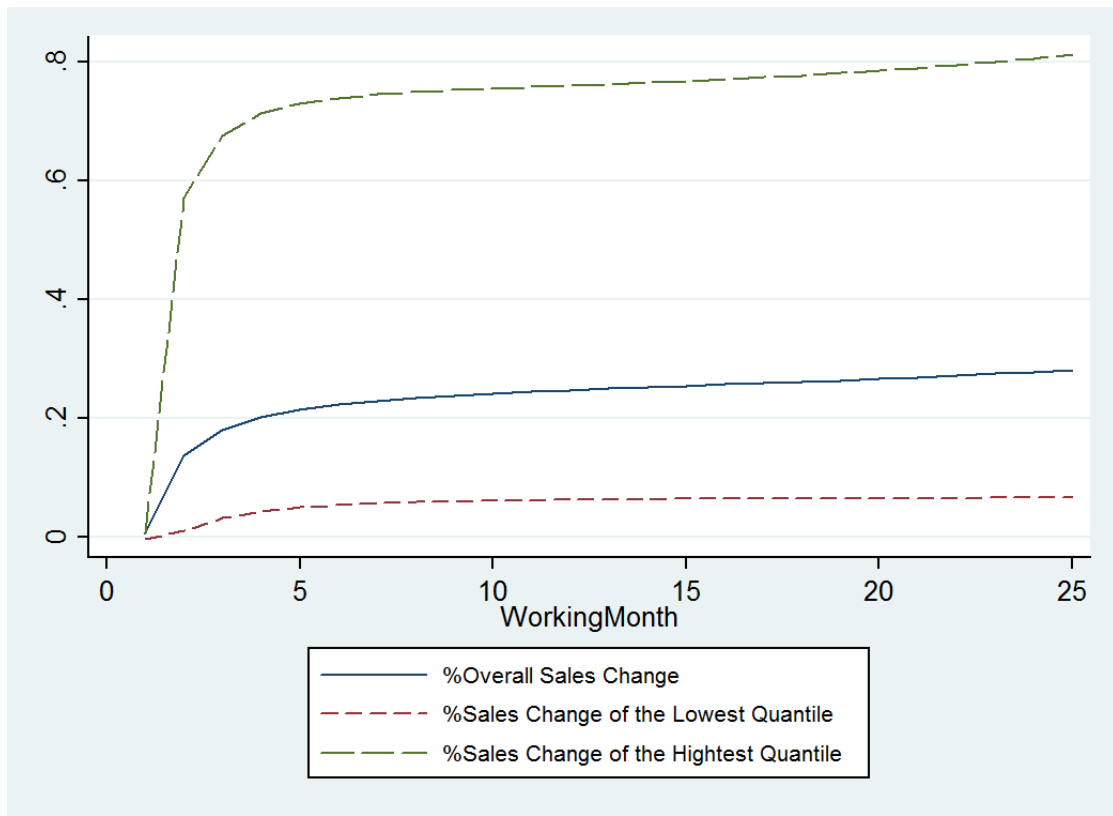
Given the average item price of about CNY114 (see Table 1), cosmetics are luxury products for most Chinese consumers. Selling cosmetics is a far more challenging task than selling other categories such as groceries or clothing. Customers have unique product needs; even for the same product (e.g. lipstick) there are numerous colors and scents that customers evaluate differently. A newly hired salesperson has to learn about various types of products and their attributes as well as how they match with customers with unique needs. She must also be patient but persuasive when talking to customers who may want to know every detail about product attributes. Interviews with the store manager reveal that the selling process can last an entire hour before a customer makes a decision.

There are two major cosmetic categories – skin care and makeup. Skin care products are typically more expensive than makeup products (average prices at CNY124 and CNY89, respectively). While the benefits of makeup in improving appearance can be immediately verified by customers, the value of skin care products can only be observed through long-term usage. Further, for most Asian customers the benefits sought from skin care are the “whitening” and “smoothing” of skin, which may not be obvious even after long usage. In this sense most of the skin care products can be classified as credence goods (Darby and Karni 1973). The information asymmetry of credence goods provides considerable incentives for opportunistic seller behavior (Emons 1997), a problem that makes customers very cautious in the buying process. A good salesperson has to have good selling techniques so that she can convince customers that the product is worthy of its high price. Our interviews with cosmetic salespeople

and managers in both China and the US consistently reveal that selling skin care products is a more challenging task than selling makeup products.

When an individual salesperson with a unique ID completes a sale, the cashier records the identity of that salesperson, the product identity, quantities, prices, and time of the sale. This careful sales tracking provides the store with detailed information about every cosmetic sale for each of its brands at the individual level and allows us to infer the change of sales productivity over time for each worker. Our data exhibit considerable changes in worker productivity over time, especially for newly-hired ones. Figure 1 plots the productivity growth rate of new workers in our data in their first two years of service. The middle curve is the non-parametric smoothed “learning curve” averaged over all workers, showing that the growth occurs mainly in the first 6 months of employment, from which one may infer as learning-by-doing. However, there is a large variation across workers, and for the same worker across months (see the smoothed curves in Figure 1 representing the lower and upper quantiles of productivity change of all workers in each month). This variation may be due to the heterogeneity of learning ability between workers, or due to the difference in the learning process as a result of the variation in the pool of peers in different months.

Figure 1: Productivity Growth of New Workers



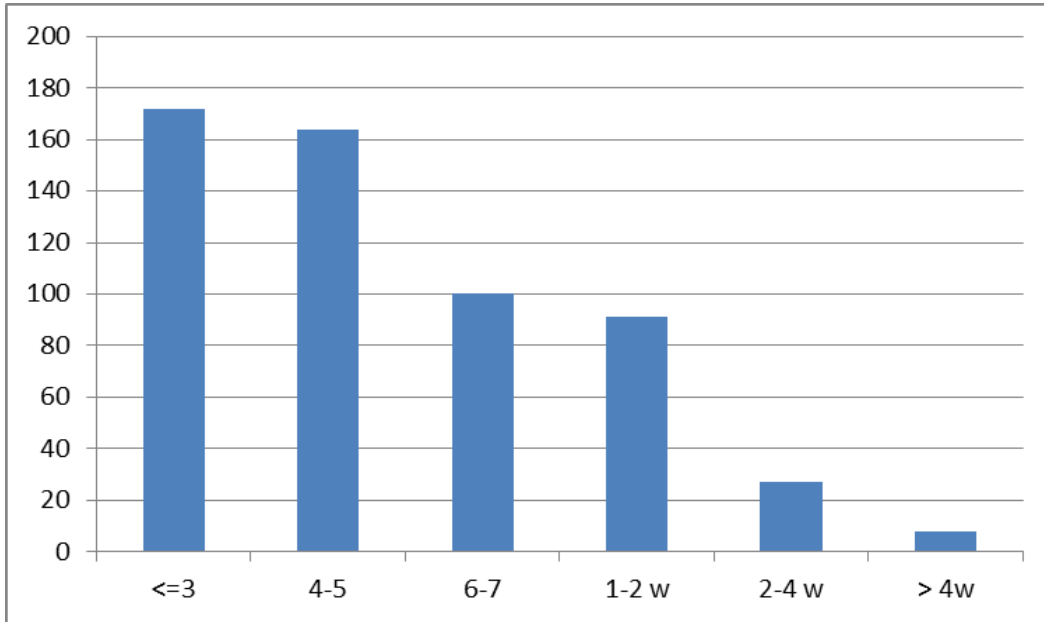
We will discuss some unique features in the data that allow us to identify peer-based learning. The first is the departure of existing workers and the entry of new workers. Ninety-two female salespeople worked for the 11 brands, with 44% turnover among workers during the entire sample period. The store manager offered us various reasons for attrition: some were promoted to management positions, some left for better paying jobs, and others left for different personal or family reasons. Table 2 provides basic information about individual cosmetics sales teams. For some brands, the sum of entering and exiting salespeople is larger than the total number of observed salespeople (e.g. Brand 2) because some salespeople hired during the sample period leave the job before the end of the sample period. The last column reports the average tenure length across counters truncated from both the left and right of the data.

Table 2: Descriptive Statistics of Cosmetics Sales Teams

	Total # of Salespeople during the Sample Period	# of Entering Salespeople	# of Exiting Salespeople	Average Working Days
Brand 1	13	8	5	749.22
Brand 2	14	9	7	373.31
Brand 3	7	3	3	1084.89
Brand 4	6	3	2	1169.31
Brand 5	9	6	4	697.78
Brand 6	5	2	2	802.59
Brand 7	9	5	4	602.19
Brand 8	5	2	2	561.52
Brand 9	11	8	6	361.19
Brand 10	5	2	2	1087.49
Brand 11	8	5	4	541.17
Mean	8.36	4.82	3.73	730.06

The law requires that workers on average have two days off per week. The department store is open seven days per week, and weekends and public holidays such as New Year and Chinese New Year are usually its busiest time. It cannot let workers off during these days, nor do salespeople want to given their commission-based pay. The workers in our data typically continuously work for long periods without days off, redeeming their accumulated vacation for longer breaks. A salesperson on average works 181 hours per month with a standard deviation of 61 hours. Across the entire sample period, we observe 562 times a worker is absent for more than 2 days. The average absence time is 4.5 days with a standard deviation of 5.6, and the maximum is 63 days. Figure 2 provides the empirical distribution of workers' days of leave-of-absence. Based on our understanding the longer leaves are mainly used for vacation trips or maternity leave. We also find from data that these leaves are distributed over various months in a year. This data feature will provide a nice variation in the pool of peer workers within and cross counters that helps to identify peer-based learning.

Figure 2: Empirical Distribution of Worker Leave-of-Absence



Finally, from the beginning of our sample (January 1, 2003) until the end of October 2004, the cosmetics department occupied the east gate area of the store, which used to be the main customer entry point. The floor plan and location of the counters in this period is presented in Figure 3(a), with TC counters in blue and IC counters in red. On November 1, 2004, all cosmetics counters were relocated to the west gate area. This was because the construction surrounding the store caused the west gate to replace the east gate as the store's main entry point. Figure 3(b) presents the floor plan and location of the cosmetics counters after the relocation. Such a relocation offers another exogenous variation in the pool of cross-counter competing peers – each salesperson after relocation faces a new set of peers from competing counters.

Figure 3(a): Cosmetics Floor Layout Before Relocation

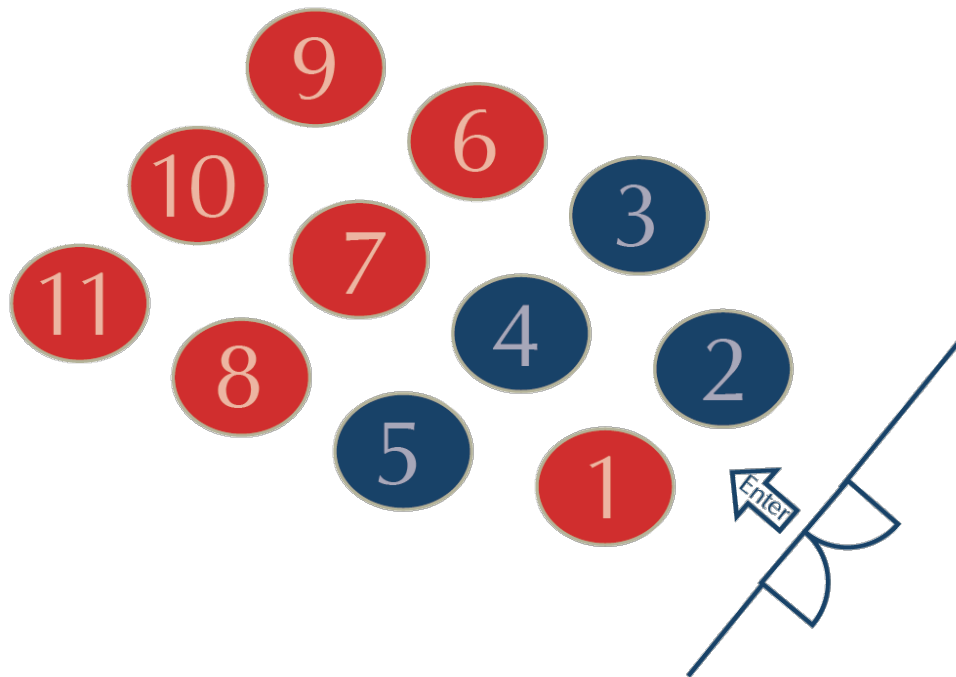
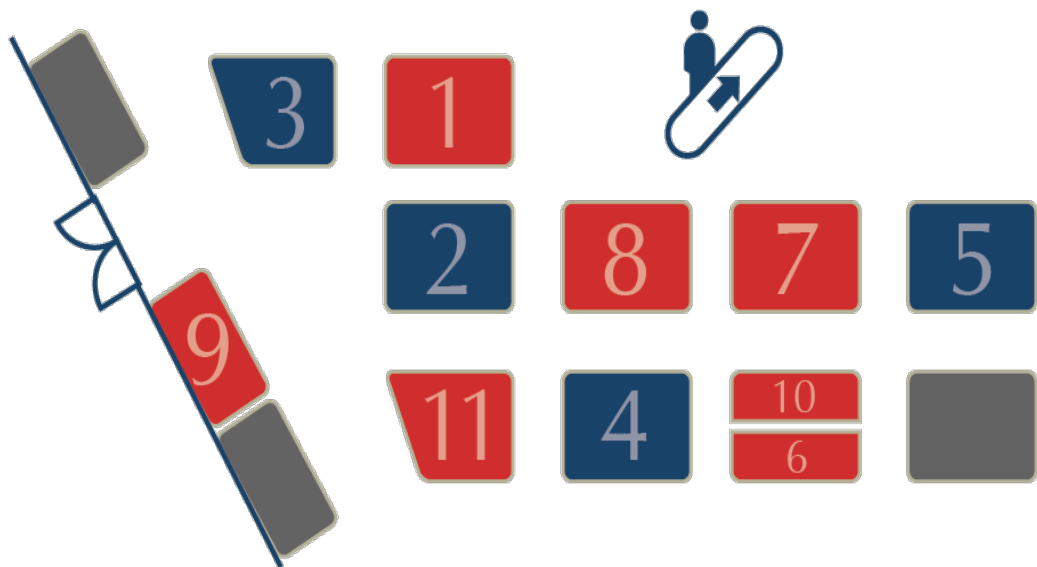


Figure 3(b): Cosmetics Floor Layout After Relocation



Red Counters: IC ; Blue Counters: TC ; Grey Counters: not in dataset

III. An Empirical Model of Worker Learning

To examine the role of peer-based learning among workers, we build a model that identifies how co-located salespeople in our data influence the productivity of one another through repeated interactions, accounting for their own learning-by-doing and forgetting. Suppose that there are I salespeople working for J counters in the store in month t . We assume that for each salesperson i her average hourly sales revenues, Y_{it} , is a function of her productivity in the month, \hat{y}_{it} , and a (row) vector of covariates that may affect sales, \mathbf{Z}_{jt} , including year (Year 2 through Year 4), month (February – December), day of week (Monday through Saturday), and brand indicators as follows:

$$Y_{it} = \hat{Y}_{it} \cdot e^{\mathbf{Z}_{jt}\beta} \cdot \eta_{it} \quad (1)$$

where η_{it} is a (positive) error term representing demand shocks, such as store promotion of cosmetics or other product categories, that may affect the store traffic. Let $y_{it} = \ln(Y_{it})$, $\hat{y}_{it} = \ln(\hat{Y}_{it})$ and $\eta_{it} = \ln(\eta_{it})$. The above equation can be re-written as

$$y_{it} = \hat{y}_{it} + \mathbf{Z}_{jt}\beta + \eta_{it} \quad (2)$$

The focus of our model is the evolution of the salesperson's productivity \hat{y}_{it} .

3.1 Modeling Peer-based Learning

We model how the productivity of individual salespeople changes over time through repeated interactions with their peers. In our analysis, a salesperson's peers are defined as all other cosmetics salespeople who either work at the same counter (within-counter peers or inside peers) or work in any adjacent counter (cross-counter peers or outside peers).⁶

⁶ Salespeople at distant counters cannot be directly observed or interacted during work time therefore their techniques are difficult to be learnt. Chan, Li, and Pierce (2010) showed from the same data that cross-counter peer effects rapidly diminish with distance between counters.

Let N_{jh} and $N_{j'h}$ denote the total number of salespeople working in the worker's own counter j and in any adjacent counters j' in any hour h that the salesperson works, respectively. We assume that productivity \hat{y}_{it} in month t is affected by the peer interaction in the previous month $t-1$. Since salespeople could work with different peers in different working hours throughout the month, we begin by specifying the peer interaction at the *hour* level. We assume that \hat{y}_{it} is affected not by the absolute productivity of peers but by the difference between her productivity and peers', because what a worker can learn from peers should also depend on how productive she is. A star salesperson is less likely to learn than a newly hired salesperson, given identical productivity levels of their peers. We further assume that \hat{y}_{it} is affected by the average of those productivity differences in any hour h the salesperson works, represented by

$\frac{\sum_{k \in N_{jh, k \neq j}} (\hat{y}_{k,t-1} - \hat{y}_{i,t-1})}{N_{jh} - 1}$ and $\frac{\sum_{k' \in N_{j'h}} (\hat{y}_{k',t-1} - \hat{y}_{i,t-1})}{N_{j'h}}$, respectively. We then aggregate these hourly

interactions to the *month* level and normalized by the total work hours $H_{i,t-1}$. We start with a baseline model that includes only this peer-based learning (Model 1):

$$\hat{y}_{it} = \hat{y}_{i,t-1} + \theta_1 \sum_h \left\{ \left[\frac{\sum_{k \in N_{jh, k \neq j}} (\hat{y}_{k,t-1} - \hat{y}_{i,t-1})}{N_{jh} - 1} \right] / H_{i,t-1} \right\} + \theta_2 \sum_h \left\{ \left[\frac{\sum_{k' \in N_{j'h}} (\hat{y}_{k',t-1} - \hat{y}_{i,t-1})}{N_{j'h}} \right] / H_{i,t-1} \right\} + \tau_{it} \quad (3)$$

In the specification, \hat{y}_{it} is first assumed to be dependent on the worker's ability in the previous month, $\hat{y}_{i,t-1}$. Parameters θ_1 and θ_2 represent the within-counter and cross-counter peer-based learning effects, respectively, to be estimated from the data. Finally, τ_{it} is an error term representing "learning shocks" that are not captured in our model (e.g. the salesperson may enroll in a training course). The major difference between the learning shocks and demand

shocks (η_t) on sales y_{it} is that the former will be carried over to future periods (as capital of knowledge) and the latter will not.

Given that there are I salespeople, we have I equations (2). Let $\hat{\mathbf{y}}_t = (\hat{y}_{1t}, \hat{y}_{2t}, \dots, \hat{y}_{It})'$ and $\hat{\mathbf{y}}_{t-1} = (\hat{y}_{1,t-1}, \hat{y}_{2,t-1}, \dots, \hat{y}_{I,t-1})'$ be the vectors of their ability in months t and $t-1$, respectively. The equation system can be transformed into a matrix format as follows:

$$\hat{\mathbf{y}}_t = \Theta_{t-1} \hat{\mathbf{y}}_{t-1} + \boldsymbol{\tau}_t \quad (4)$$

where $\boldsymbol{\tau}_t = (\tau_{1t}, \tau_{2t}, \dots, \tau_{It})'$, and Θ_{t-1} is an I by I square matrix with the $[i, s]$ element (assume worker i works in counter j) as

$$\Theta_{t-1}[i, s] = \begin{cases} 1 - (\theta_1 + \theta_2), & \text{if } s = i \\ \theta_1 \sum_h \left[\frac{1\{s \in N_{jh}\}}{N_{jh} - 1} \right] / H_{i,t-1}, & \text{if } s \neq i \text{ and both work at the same counter} \\ \theta_2 \sum_h \left[\frac{1\{s \in N_{j'h}\}}{N_{j'h} - 1} \right] / H_{i,t-1}, & \text{if } s \neq i \text{ and both work at different counters} \end{cases} \quad (5)$$

where $1\{s \in N_{jh}\}$ and $1\{s \in N_{j'h}\}$ are indicators that peer worker s worked during hour h in month $t-1$ (for counter j or any of j 's adjacent counters, respectively).

Let $\mathbf{y}_t = (y_{1t}, y_{2t}, \dots, y_{It})'$ be the vector of the (log) average hourly sales revenues of the I salespeople in month t , and similarly let \mathbf{Z}_t be a matrix with the row vectors \mathbf{Z}_{jt} for all salespeople combined. From equations (1) and (4) we establish a relationship between workers' sales in month t and own and peers' productivity in month $t-1$ as:

$$\mathbf{y}_t = \Theta_{t-1} \hat{\mathbf{y}}_{t-1} + \mathbf{Z}_t \boldsymbol{\beta} + \boldsymbol{\tau}_t + \boldsymbol{\eta}_t \quad (6)$$

We can similarly rewrite $\hat{y}_{t-1} = \Theta_{t-2}\hat{y}_{t-2} + \tau_{t-1}$ and plug into equation (6). Assume that we observe the sales for T months. For any given month $t=2, 3, \dots, T$, we can repeat the iteration $t-1$ times to obtain

$$y_t = \left(\prod_{g=1}^{t-1} \Theta_{t-g} \right) \hat{y}_1 + Z_t \beta + \varepsilon_t \quad (7)$$

where \hat{y}_1 represents the *initial* productivity of all salespeople, which is a vector of parameters we will estimate. For new workers who were hired in the middle of our sample period, this is their ability in the first month in the store. For the workers who had worked in the store before our data started, this is their productivity in the first month of our sample. The error term ε_t is a combination of demand shocks and past learning shocks as follows:

$$\varepsilon_t = \tau_t + \sum_{g=1}^{t-1} \left[\left(\prod_{h=1}^g \Theta_{t-h} \right) \cdot \tau_{t-g} \right] + \eta_t$$

Under such specification ε_t for every salesperson will be serially correlated. This implies that we have to allow for a general structure of heteroskedasticity for ε_t in our estimation.

3.2 Modeling Learning-by-Doing and Forgetting

To model learning-by-doing, standard practice uses past production or sales as a proxy for past work experience. However, if we were to only use past sales, a good salesperson with higher productivity will always have higher work experience than her peers, even though they have worked the same number of hours. This will create bias in estimating the effect of learning-by-doing. In a selling environment like ours, a salesperson may improve her skills through repeated interaction with customers, even though this does not always lead to purchases. Therefore, unlike previous studies, we choose work hours as the measurement for past experience.

We assume that a worker's productivity in month t is affected by (the log of) the total number of hours she worked in month $t-1$, $h_{i,t-1}$. We model forgetting as the fraction of productivity lost from last month if the worker did not work. Specifically, equation (3) is redefined to include peer-based learning, learning-by-doing and forgetting (Model 2) as:

$$\hat{y}_i = \gamma \hat{y}_{i,t-1} + \theta_1 \sum_{k \in N_{jh, k \neq j}} \left\{ \frac{(\hat{y}_{k,t-1} - \hat{y}_{i,t-1})}{H_{i,t-1}} \right\} + \theta_2 \sum_{k' \in N_{j'h}} \left\{ \frac{(\hat{y}_{k',t-1} - \hat{y}_{i,t-1})}{H_{i,t-1}} \right\} + \lambda h_{i,t-1} + \tau_i \quad (8)$$

where the parameter γ captures the carry-over of knowledge from last month to the current month. An estimate being significantly smaller than one implies depreciation in the salesperson's productivity, or forgetting. If a worker were on leave for the whole month, her productivity for the following month would be reduced by the proportion of $1-\gamma$. The parameter λ captures the effect of learning-by-doing from the work experience in the previous month.

With the new productivity specification we can also redefine Θ_{t-1} in equation (5): its off-diagonal elements are the same but diagonal element $\Theta_{t-1}[i, i]$ now is $\gamma - (\theta_1 + \theta_2)$. We can rewrite equation (6) as

$$y_t = \Theta_{t-1} \hat{y}_{t-1} + \lambda h_{t-1} + Z_t \beta + \tau_t + \eta_t \quad (9)$$

where $h_{t-1} = (h_{1,t-1}, h_{2,t-1}, \dots, h_{I,t-1})'$. Again by iteration, as in equation (7), we can obtain

$$y_t = \left(\prod_{g=1}^{t-1} \Theta_{t-g} \right) \hat{y}_1 + \lambda \sum_{g=1}^{t-1} \left[\left(\prod_{h=1}^{g-1} \Theta_{t-h} \right) H_{t-g} \right] + Z_t \beta + \varepsilon_t \quad (10)$$

where all other variables and parameters are the same as in equation (7).

3.3 Model Identification

We will discuss in this section those features in the data that allow us to separately identify learning-by-doing, forgetting, and learning from peers in the model. The identification of learning-by-doing and forgetting in each month comes from the accumulation of a salesperson's work experience over time. If learning-by-doing is a key force in changes in

worker productivity, a salesperson's average hourly sales will increase with the accumulated working hours, holding other factors constant. Since learning-by-doing will never reduce work ability, if we observe in our data that after a salesperson worked fewer hours in a month (e.g. for long vacation) her average hourly sales are reduced in the next month, again holding other factors constant, this will be inferred as forgetting in our model.

The identification of peer-based learning from within-counter and cross-counter comes from variation in the pool of peers in the previous month (see equation (8) and (9)). Due to the shift rotation policy of the department store, there is little variation across workers at the same counter in terms of the number of hours working with other peer workers. However, we consistently observe exit and entry from workers for all 11 counters throughout the sample period (see Table 2). This creates the variation in the pool of peers in different months. Suppose, after a star salesperson quits her job, the productivity growth of other salespeople in the same counter in future months decreases. This implies a positive within-counter peer-based learning effect. If the productivity growth of salespeople in other adjacent counters also slows down, we can infer that a positive cross-counter peer-based learning effect also exists. The concern of using the entry and exit of salespeople is that these are endogenous decision variables; however, we have controlled for this potential issue in the model by including brand, year and month fixed effects (in Z_t), and individual peer worker productivities in the previous month y_{t-1} (see equation (8)). Our model identification therefore only requires that entry and exit of peers are exogenous to the residuals ε_t in the estimation model (10) unexplained by these factors.

Our study further exploits two other sources of data variation that can be reasonably assumed to be exogenous shocks to the process of learning. First, the counter relocation in the middle of our sample period brings a change in the pool of outside peers the worker faces and, as

a result, provides her the opportunity to learn by observing how a new group of competitors serve their customers. This new opportunity could be critically important if she had long worked in the store and had exhausted the potential knowledge spillovers from neighboring counters. Second, as discussed before, many salespeople choose to work six or seven days a week in order to save for a longer vacation. These absences also serve as an instrument for identifying the worker learning. For example, a new salesperson who must work alone for more hours because her more experienced peers are on vacation may have a completely different learning process compared with others who can spend their first months observing and learning from peers. We will further investigate the issue of endogenous selection of workers in later section.

The peer-based learning in this study is different from the peer effects identified in the previous literature (e.g. Bandiera et al 2005; Mas and Moretti 2009; Chan et al 2011). The former increases the stock of knowledge of peer workers, providing a long-term impact on productivity, while the latter provides a temporary shock only in that time period. If the sales of other workers in future months systematically decreased after a star salesperson quit her job, this could not be explained by the previously identified temporal peer effects in the literature.

3.4 Model Estimation

In equations (7) and (10), the Θ 's of different months are functions of unknown parameters γ , θ_1 and θ_2 . They multiply themselves over months and interact with the vector of salespeople's initial productivity \hat{y}_1 that are also parameters to be estimated. Thus, both models 1 and 2 are non-linear. If we simultaneously estimate equation (7) or (10) for all parameters via a non-linear least-square approach, the dimensionality problem of the parameter space (the number of parameters is 126 in the simplest model) is very severe. Any numerical algorithm of searching for the optimal parameters will take extremely long to converge and, even when it converges, the

estimates are likely to be local optimum due to the non-linearity nature of the equation system. We adopt the nested optimization procedure proposed by Chan, Li, and Pierce (2011) to solve this problem. This procedure recognizes that the dimensionality problem mainly comes from the 92 parameters in $\hat{\mathbf{y}}_1$, each representing the initial productivity of a salesperson. However, conditional on the three parameters γ , θ_1 and θ_2 , Θ 's can be treated as covariates and hence equations (7) and (10) become linear in \hat{y}_1 . The nested procedure therefore starts by choosing some initial values for γ , θ_1 and θ_2 , computing Θ 's from equation (5), and then estimating \hat{y}_1 (and other parameters including λ and β) via standard linear least-square methods (inner procedure).

The outer procedure is to search for the optimal γ , θ_1 and θ_2 , using standard numerical minimization routines that minimize the sum of squared errors as the criterion function value. In our implementation, we use the Nelder-Mead (1965) simplex method to search for the optimal γ , θ_1 and θ_2 . Since given Θ 's \hat{y}_1 can be computed analytically using the linear method, numerical search is only used in the outer procedure, which is much faster than searching for all parameters in the model simultaneously. Furthermore, given Θ 's the estimate $\hat{\mathbf{y}}_1$ is a unique optimum in minimizing the criterion function value. We find that practically local optima are not an issue in our model estimation – no matter what initial values for γ , θ_1 and θ_2 we use, the procedure always converges to the same optimum.

Since this procedure is only different from the non-linear simultaneous estimation procedure in the numerical implementation, but their criterion functions are the same, estimates obtained from both procedures are equivalent. We can compute the standard errors for all estimates assuming that they are obtained using the simultaneous non-linear least-square

approach. Furthermore, given that ε_t in equations (7) and (10) are possibly serially correlated, we compute the robust standard errors clustered at the individual worker level for the non-linear least-square estimators.

3.5 Asymmetric Learning Models

Our models thus far have several limitations. First, they are “symmetric” in the sense that the θ 's in equation (8) do not differentiate working with superior peers from working with inferior peers. This implies that the positive effect of learning from a peer who on average sells CNY100 more per hour is cancelled by the negative effect of learning from another peer who sells CNY100 less. Second, our baseline models are also silent on the possible different learning processes among new vs. incumbent workers. A salesperson who is newly hired has a lower initial productivity and hence may benefit more from learning. Given higher returns from learning, she will also have a stronger incentive to invest time and effort toward learning from more experienced peers. In this section we extend our baseline models to address these issues.

First, we construct an asymmetric model of peer-based learning allowing the magnitudes of effects from superior peers within and across counters to be different from those from inferior peers (Model 3). The only difference between this asymmetric model and the symmetric model in equation (7) is that for the peer-based learning effects θ_1 and θ_2 , we now estimate two separate effects, θ_g^a and θ_g^b , $g=1, 2$. The former (latter) represents the within- or cross-counter peer-based learning effects from coworkers with higher (lower) productivity. That is, for a focal worker i and her peer worker k , we estimate θ_g^a if $\{\hat{y}_{it} \leq \hat{y}_{kt}\}$ and θ_g^b otherwise. Model 3 consequently has four θ 's instead of the two in Models 1 and 2.

Though the extension is straightforward, the nested non-linear estimation algorithm we adopted to estimate our symmetry models cannot be directly applied to this model. The key to

using the algorithm is that all ability parameters \hat{y}_{it} 's are linear in the y_{it} 's conditional on \mathcal{V} and θ 's. With asymmetric effects, however, \hat{y}_{it} 's now interact with indicator functions $\{\hat{y}_{it} \leq \hat{y}_{kt}\}$ or $\{\hat{y}_{it} > \hat{y}_{kt}\}$. We employ a trick in order to avoid estimating them non-linearly. The key observation is that the indicators $\{\hat{y}_{it} \leq \hat{y}_{kt}\}$ or $\{\hat{y}_{it} > \hat{y}_{kt}\}$ only depend on the productivity ranking for workers j and k in month t . In the estimation, we use the ranking of workers' average sales in the month observed from data to proxy the ranking of ability. Specifically, we use the average hourly sales for all salespeople in month t , \bar{y}_{it} , to construct proxies $\{\bar{y}_{it} \leq \bar{y}_{kt}\}$ or $\{\bar{y}_{it} > \bar{y}_{kt}\}$ for indicators $\{\hat{y}_{it} \leq \hat{y}_{kt}\}$ or $\{\hat{y}_{it} > \hat{y}_{kt}\}$. These two rankings should be consistent with each other if the learning effect does not dominate the main ability difference. Conditional on the ranking, we repeat the nested non-linear algorithm as discussed before to estimate the four θ 's and other parameters. After that, we re-estimate all model parameters simultaneously via the non-linear numerical algorithm, using the estimates we already obtained as the starting values. We find that for all model specifications the algorithm of this second step always immediately converges to the same starting values, indicating that the productivity ranking based on the actual sales revenues is reliable. We also compare the actual sales revenue ranking $\{\bar{y}_{it} \leq \bar{y}_{kt}\}$ and $\{\bar{y}_{it} > \bar{y}_{kt}\}$ with $\{\hat{y}_{it} \leq \hat{y}_{kt}\}$ and $\{\hat{y}_{it} > \hat{y}_{kt}\}$ based on the estimates, and find them consistent for all i and t .

We further extend our asymmetric Model 3 to allow the magnitude of both learning-by-doing and peer-based learning for new workers to be different from those for incumbent workers. To do this we first classify a worker as new for the first three months she worked in the store.⁷

⁷ This classification is based on conversations with the store manager, who suggested that a new worker would take about three months to obtain a full knowledge of products and selling techniques. We also test other specifications (e.g. the first month) and find that results are qualitatively the same.

We also received from the department store the tenure of employment for every worker as of the first month of our data. A worker is classified as an incumbent only if she had worked for longer than three months when the sample period started, otherwise she will be classified as a new worker. We then allow the learning-by-doing parameter λ and the peer-based learning parameters θ_1^a , θ_1^b , θ_2^a and θ_2^b to be different between new and incumbent workers in our model. This becomes our final model – Model 4.

IV. Results

The results from our models are presented in Table 3. Column 1 presents our symmetric model with peer-based learning only. The productivity of peers both within and across counties clearly impacts a salesperson's monthly sales growth. Workers learn at a faster rate when working with high-ability peers both within their own county and at adjacent counties, with the effect from within-county peers being 3-4 times as large. This is likely because salespeople can learn more from within-county peers through closer observation and through soliciting active teaching and advice. Peer teaching is unlikely to occur across counties since they are competing against one another and are compensated based on their sales. Still, the significant cross-county learning effect suggests that by closely locating to a competing county with star salespeople, workers can improve their productivity through the observation of the selling practice of competitors. While sales may be hurt by stronger competition in the short run, the long-term benefit for the county is substantial.

Column 2 presents the results of Model 2 that separates learning-by-doing and forgetting from peer-based learning. Working longer hours in the prior month appears to increase productivity, but the magnitude is very small compared with the peer-based learning coefficient.

The productivity carry-over parameter γ is less than but not statistically different from one, suggesting no forgetting or knowledge depreciation in our empirical context.⁸ This, together with the significant peer-based learning effect, suggests that knowledge transferred from peers will have a long-lasting benefit.

Column 3 presents estimation results from our first asymmetric model (Model 3). While the effects from learning-by-doing and forgetting remain unchanged, we observe substantial asymmetry in the effects from superior vs. inferior peers. Superior workers have substantial positive impact on the learning of their peers, with this impact again much stronger within counter than from adjacent counters. In contrast, inferior peers appear to have a very small negative impact on the learning of salespeople.⁹ This may result from the focal salesperson learning bad sales techniques or practices, or from adopting a poor work ethic that has long-term effects. These results also have implications for workers' team formation. Compared to teams consisting of homogeneous salespeople with average productivity, a team with a mix of star salespeople and inexperienced rookies can have a much faster growth in total sales. This finding is consistent the empirical results in Hamilton et al (2003), Mas and Moretti (2009) and Chan et al (2011), but it suggests that the benefit from team heterogeneity can be even stronger in the long run.

Column 4 presents estimation results from Model 4, which separately estimates learning-by-doing and peer-based learning for new and incumbent workers. The results from this model are consistent with Model 3, with the positive effects of peer-based learning from superior

⁸ We note that this result does not imply that workers would never suffer forgetting if they were to leave their jobs for very long time (e.g. years). We are not able to observe such lengthy absence in our data, and thus must remain agnostic on this issue.

⁹ Given that the difference in productivity between an inferior peer worker and the focal worker is negative (see equation 3), a positive coefficient implies the lower the productivity of her peer the lower the productivity of the worker.

within- or cross-counter peers dominating the negative effects from inferior peers. However, the results suggest that the strong learning effects identified in Models 1 - 3 probably come from new workers. Specifically, the learning effect from within-counter superior peers on new workers is about 10 times as large as effect on incumbents.

Our results provide substantial information on the relative importance of each source of learning. Our estimates imply that a new salesperson, after 180 hours of work in the first month, will experience a productivity growth from learning-by-doing of about 0.5%. In contrast, if she also worked at the same counter with a peer of twice her productivity all the time, her productivity would grow by 23%. If this peer were at a competing counter, her productivity would grow by 5%. Given that the worker exit rate is 44% and an average salesperson only works for two years in the data, learning-by-doing alone will not be sufficient to compensate for the high employee turnover. Our findings illustrate that the major contributing factor for the observed worker learning curves is peer-based learning. For a new salesperson, working together with an experienced peer of superior ability greatly enhances her future sales productivity.

The reason for the negligible impact of learning-by-doing is probably due to the nature of cosmetic sales. To become a good salesperson requires not only knowledge of products but, more importantly, how to identify the needs of customers, how to match the right products with their needs, and how to persuade and convince these customers to buy the products. Unique selling processes may be involved in each customer interaction, requiring new sales approaches, techniques, and solutions. These are difficult and also costly to learn from repeated experimentation with new practices; instead, it may be more effective to learn from the experiences of peer workers, either through observation or through their active teaching. We caution that the relative importance and roles of peer-based learning and learning-by-doing may

be different for other job tasks, such as factory and farm production, where fewer parameters are involved in the production process. We also note that most salespeople in our data were required to finish a pre-job training program before working at the counter. The small learning-by-doing effect in our model may indicate that the majority of such learning has been achieved during the training.

4.1 Further Model Extensions and Results on Peer-Based Learning

We further explore how peer-based learning may vary depending on the financial incentives for salespeople. As discussed earlier, inside the department store there are four brands using team-based commissions (TC) and seven using individual-based commissions (IC). We separately estimate the long-term peer effects for TC and IC counters. Specifically, we extend equation (1) to the following:¹⁰

$$\begin{aligned} \hat{y}_{it} = & \gamma \hat{y}_{i,t-1} + (\theta_1^{IC} \cdot 1\{j \in IC\} + \theta_1^{TC} \cdot 1\{j \in TC\}) \sum_h \left\{ \left[\frac{\sum_{k \in N_{jh}; k \neq j} (\hat{y}_{k,t-1} - \hat{y}_{i,t-1})}{N_{jh} - 1} \right] / H_{i,t-1} \right\} \\ & + (\theta_2^{IC} \cdot 1\{j \in IC\} + \theta_2^{TC} \cdot 1\{j \in TC\}) \sum_h \left\{ \left[\frac{\sum_{k \in N_{jh}} (\hat{y}_{k',t-1} - \hat{y}_{i,t-1})}{N_{j'h}} \right] / H_{i,t-1} \right\} + \tau_{it} \end{aligned}$$

in which the variables $1\{j \in IC\}$ and $1\{j \in TC\}$ are indicators that brand j is an IC counter or a TC counter. Parameters θ_1^{IC} and θ_1^{TC} represent the within-counter peer-based learning effects for IC and TC counters, respectively. Parameters θ_2^{IC} and θ_2^{TC} measure the peer-based learning effects from workers at competing counters on salespeople at IC and TC counters, respectively. Columns 1 and 2 of Table 4 report the estimated learning effects. Though the positive effects from within-counter superiors are similar for the two compensations systems, they may reflect two countervailing financial incentives. While team-based compensation provides a stronger incentive for superior workers to actively teach their peers, peers at IC counters have weaker

¹⁰ To focus on peer-based learning effects, here we only discuss the extension of our Model 1. The extension of our other model specifications is similar.

incentives to devote effort to learn. In fact, we observe that learning-by-doing for new workers is higher in IC counters than in TC counters, implying that the stronger financial incentives of IC may increase the worker's effort toward learning. We also observe much stronger cross-counter peer-based learning effects for IC counters, also consistent with higher-powered individual incentives, since cross-counter peers are unlikely to actively teach new workers at competing counters.

Table 3: Estimation Results of Worker Learning Models

	(1)	(2)	(3)	(4)
	Model 1	Model 2	Model 3	Model 4
Carry-Over	1	0.9891 (0.1764)	0.9982 (0.1713)	0.9913 (0.1683)
Overall Within-Counter Peer-based learning	0.0861*** (0.0247)	0.0664*** (0.0179)	--	--
Within-Counter Peer-based learning from Superiors	--	--	0.2511*** (0.0598)	--
Within-Counter Peer-based learning from Inferiors	--	--	0.017*** (0.0071)	--
New Worker Within-Counter Peer-based learning from Superiors	--	--	--	0.3018*** (0.0625)
New Worker Within-Counter Peer-based learning from Inferiors	--	--	--	0.0261*** (0.0089)
Existing Worker Within-Counter Peer-based learning from Superiors	--	--	--	0.0343*** (0.0163)
Existing Worker Within-Counter Peer-based learning from Inferiors	--	--	--	0.0097 (0.0081)
Overall Cross-Counter Peer-based learning	0.0244*** (0.0087)	0.0251*** (0.0089)	--	--
Cross-Counter Peer-based learning from Superior	--	--	0.0536*** (0.0185)	--
Cross-Counter Peer-based learning from Inferiors	--	--	0.0142*** (0.0061)	--
New Worker Cross-Counter Peer-based learning from Superiors	--	--	--	0.0727*** (0.0216)
New Worker Cross-Counter Peer-based learning from Inferiors	--	--	--	0.0129*** (0.0051)
Existing Worker Cross-Counter Peer-based learning from Superiors	--	--	--	0.0139** (0.0069)
Existing Worker Cross-Counter Peer-based learning from Inferiors	--	--	--	0.0108** (0.0055)
Overall Experience Learning	--	0.0006*** (0.0002)	0.0006*** (0.0002)	--
New Worker Experience Learning	--	--	--	0.0010*** (0.0003)
Existing Worker Experience Learning	--	--	--	0.0003*** (0.0001)

Next, we investigate how the knowledge transfer between coworkers depends on the difficulty of job tasks. We examine the two cosmetic sub-categories, skin care and makeup, as discussed in previous sections. Interviews with cosmetic salespeople and managers in China and the United States consistently reveal that skin care products are much more difficult to sell than makeup products. We identify all the cosmetics products sold during our sample period as belonging to the skin care or makeup categories, and re-run our asymmetric Model 4 using a salesperson’s average hourly sales of skincare and makeup products as dependent variables. We present these as Model 7 and Model 8 in Table 4, respectively. Peer-based learning effects remain the most important predictors of monthly productivity gains for both types of products, but the differences in the coefficients suggest that the learning process is not identical for each product type. Within-counter learning from superiors is much stronger for skin care, while cross-counter learning from superiors is stronger for makeup.

Table 4: Estimation Results of Product Category and Compensation Models

	(1)		(2)	(3)
	Model 6		Model 7	Model 8
	IC Counters	TC Counters	Skin Care Products	Makeup
Carry-Over	0.9834 (0.1659)		0.9941 (0.1815)	0.9936 (0.1789)
Within-Counter Peer-Based Learning from Superiors	0.2713*** (0.0627)	0.2689*** (0.0603)	0.3044*** (0.0637)	0.1672*** (0.0422)
Within-Counter Peer-Based Learning from Inferiors	0.0098 (0.0081)	0.0302*** (0.0095)	0.0335*** (0.0133)	0.0228*** (0.0107)
Cross-Counter Peer-Based Learning from Superiors	0.0971*** (0.0289)	0.0336*** (0.0124)	0.0166*** (0.0076)	0.1588*** (0.0564)
Cross-Counter Peer-Based Learning from Inferiors	0.0052 (0.0062)	0.0136*** (0.0053)	0.0327*** (0.0133)	-0.0077 (0.0081)
New Worker Experience Learning	0.0013*** (0.0004)	0.0008*** (0.0003)	0.0004*** (0.0002)	0.0006*** (0.0002)
Existing Worker Experience Learning	0.0002** (0.0001)	0.0003*** (0.0001)	0.0001 (0.0001)	0.0003*** (0.0001)

These results are consistent with the difference between observing the practice of peers and teaching from peers. Knowledge transfer across firm boundaries is unlikely to involve active teaching given the competition between the firms,¹¹ and is mostly likely to occur through pure observation. In contrast, learning within firms will likely involve both active teaching and observation. For skin care, the much stronger within-counter peer effect from superiors than from cross-counter peers suggests that for difficult sales tasks new workers need active help from peers to learn how to pitch these products to consumers. Since makeup sales are less complex and easier to observe, learning to sell these products from observing high-ability peers at adjacent counters is much more feasible than it is for learning skin care sales. The magnitude similarity of learning from inside and outside superiors suggests that learning to sell makeup may occur mainly through observation. Our results therefore provide some evidence that knowledge transfer between coworkers involves both observing and teaching from peers.

4.2 Alternative Explanations to Peer-Based Learning

We have attributed the positive relationship between peers' productivity and a salesperson's long-term productivity growth to peer-based learning. In this section, we will examine several explanations of our findings that are alternative to the learning hypothesis and test them from data. The first obvious alternative explanation is related to the endogenous selection issue – firms with good workers may be more likely to hire new workers with good learning capability, or new workers with good learning capability may select firms with good salespeople to work. If salespeople with higher learning ability are consistently matched with high productivity peers, we will observe the positive relationship in our results. To test whether or not our findings are driven by this reason, we ran a regression using sales growth of new

¹¹ There is a possibility that social preferences may lead workers to actively help their competitors, as in the tournament-based compensation of fruit-pickers in Bandiera, Barankay, and Rasul (2005). Still, the magnitude of the active help across counters should be far smaller than within counters.

salespeople in the first three months of employment as the dependent variable in the regression. This growth rate is measured by the log of the average hourly sales in the second month relative to the first month, $\ln(Y_{i2}/Y_{i1})$, and in the third month relative to the second month, $\ln(Y_{i3}/Y_{i2})$. By constructing this variable we have controlled for individual fixed effects representing the worker heterogeneity in the level of sales. The remaining endogeneity issue is that firms select workers based on learning capability.

To control for this issue, we include fixed effects for each firm in the regression. Compared with the significant fluctuation in within- and cross-counter peers due to departure or long leaves, the ability of firms to attract workers with different learning capability should be relatively time-invariant; hence, the worker selection issue will be mitigated by estimating these fixed effects. We estimate two regression models. The first uses four covariates of change in peers: (i) an indicator of a high productivity within-counter peer departing or taking a long leave in the previous month; (ii) an indicator of a low productivity within-counter peer departing or taking a long leave in the previous month; (iii) an indicator of a high productivity cross-counter peers departing or taking long leave in the previous month; and (iv) an indicator of a low productivity cross-counter peer departing or taking a long leave in the previous month. The second model further separates departing peers from those taking a long leave. We use the median of average hourly sales for all salespeople at the same counter to differentiate workers with high productivity from those with low productivity. We also incorporate year and month dummies to control for seasonality. Finally, we calculate the robust standard errors clustered at the worker level.

Regression results are presented in Table 5. Effects from within-counter peers are consistent with our previous findings – leaves of high-productivity peers in the previous month

in the two models significantly lower the productivity of a new salesperson, and *vice versa* for the leaves of low-productivity peers. Relative magnitudes between the two types of peers are also consistent. The significance and consistency of the within-counter peer effects provides support that our peer-based learning findings are at least not entirely driven by the selection issue. We can also use these results to rule out another alternative explanation again based on the selection story: knowing its star salesperson is leaving, a firm will select workers with good learning capability to hire as replacement. Such argument does not explain the effect of peers taking leave-of-absence; more importantly, if the argument is valid, we should expect positive instead of the negative coefficients for the departure of high-productivity peers in our regressions.

Regarding the cross-counter effects, the only significant result (at 10% significance level) is the negative effect from the departure of a high-productivity peer, which is again consistent with our previous results. The lack of significance for other effects is probably due to the few observations we use in the regression. It may also be because of the crude measurement for peer productivity (e.g. a low-productivity cross-counter peer may still be more productive than the new salesperson).

Another alternative explanation for our findings is that a salesperson gains long-term customers when working with productive peers, and consequently she will experience higher sales in the next period from these returning customers. This explanation, if true, implies that the underlying mechanism of our findings is irrelevant to knowledge transfer. However, this story does not explain the positive association between the salesperson's sales and the productivity of cross-counter peers, since counters always compete for instead of share customers. Further, results in Table 5 show that when a high-productivity within-counter peer departs or takes a leave-of-absence, productivity of a new salesperson declines in the following month. If customer

transfer is the underlying mechanism, the new salesperson should gain more customers from the departure of the peer, which contradicts the negative effect in the regressions.

Table 5: Effects of Peer Leaves on Productivity Change					
		Model 1		Model 2	
		<i>Estimate</i>	<i>Std. Error</i>	<i>Estimate</i>	<i>Std. Error</i>
Within-Counter	High Productivity Peers Departed Last Month	-0.403*	-0.237	-0.321**	-0.157
	High Productivity Peers Took Leave Last Month			-0.489**	-0.179
	Low Productivity Peers Departed Last Month	0.049*	0.028	0.008**	0.003
	Low Productivity Peers Took Leave Last Month			0.263**	0.121
Cross-Counter	High Productivity Peers Departed Last Month	-0.247	-0.172	-0.084*	-0.050
	High Productivity Peers Took Leave Last Month			-0.334	-0.254
	Low Productivity Peers Departed Last Month	-0.010	-0.007	-0.214	-0.145
	Low Productivity Peers Took Leave Last Month			0.077	0.056
R-Square		0.39		0.44	
Observations		90		90	

*: significant at 10% significance level; **: significant at 5% significance level.

The third argument is that a new salesperson may experience lower sales when working with superior peers in the previous month, but sales will recover the next month when she works with other salespeople with average productivity. This reflection problem would also explain the positive correlation between a salesperson’s current sales and the productivity of peers either within or across counters in the previous month. However, this explanation does not apply here since we model the level and not the growth or change of sales. The fact that working with superior peers may lower sales in the previous month should have no bearing on the salesperson’s sales level in the current month.

The possibility of this explanation can be further examined by comparing the peer effects in our model based on the competitive relationship between salespeople. Unlike IC counters,

there is no financial incentive for salespeople at TC counters to compete for customers. Chan, Li and Pierce (2011) used the same dataset and demonstrated that, within TC counters, working with superior peers will increase the current sales of a salesperson. Columns 1 and 2 in Table 4 show positive effects from within-counter superiors at both TC and IC counters. If the competition between salespeople drives our previous findings, we should not observe the positive effect from TC counters, therefore we can exclude this alternative explanation.

We have offered evidence that salespeople learn from peers, but we have not shown what has been learned. The final alternative explanation for our findings is that the improvement in a salesperson's productivity comes from her adopting the work ethic of superior peers instead of learning the knowledge or ability of selling. Evidence in Carrell et al (2011) suggests that the asymmetry of peer effects in work ethic is of the opposite direction of our findings. Their peer effects are primarily from inferior peers, while ours are from superior peers. Even if this alternative hypothesis were true, we believe that adopting work ethic could be classified as a type of peer-based learning and therefore does not contradict our primary claims. Still, given our hypotheses of knowledge transfer, it is important for us to explore if there is evidence of learning the parameter values in the production function from peers.

To support our claims, we again compare the peer effects from the skin care and makeup categories. Model 7 and Model 8 in Table 4 show that within-counter learning and cross-counter learning from superiors are different between skin care and makeup products. If the peer-based learning mechanism only involves adopting the work ethic from superiors, there is no reason why the learning effects in the two categories should be different. Since the impact of observing work ethic within-counter and cross-counter is similar for makeup sales, one cannot explain the large difference between within-counter and cross-counter peer effects for skin care. Suppose

observing work ethics from cross-counter peers is more difficult than from within-counter peers, and hence the cross-counter effect is negligible for the skin care category, we should also find such asymmetry in the makeup category. The inconsistency in the results for the two product categories suggests that the identified peer effects cannot be solely explained by observing work ethic from superior peers.

4.3 A “What-If” Experiment: The Impact of Employee Turnover on New Worker Learning

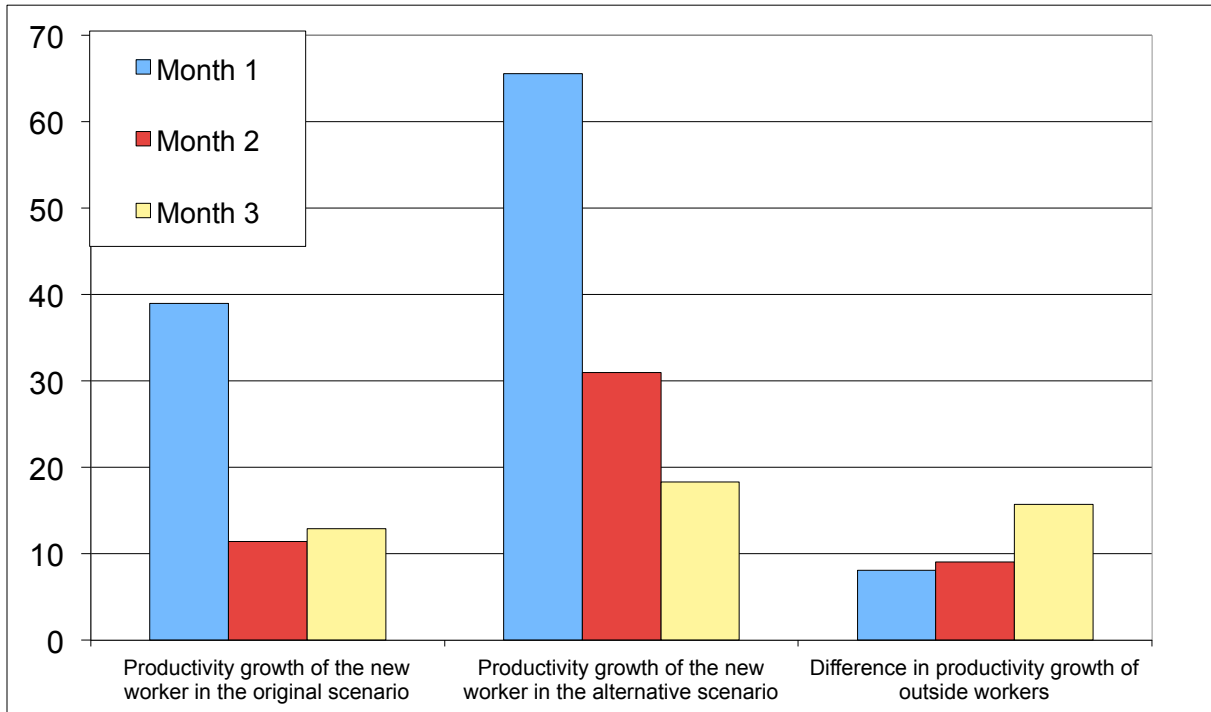
Theoretical (Becker 1962) and empirical (Ton and Huckman 2009) work suggests that one major cost of turnover may be lost opportunities for knowledge transfer, and thus productivity growth or maintenance, within the firm. To explore this, we compare a hypothetical scenario of new worker learning with the real scenario from data to illustrate the implications of our findings of peer-based learning for a firm’s employee turnover policies. We specifically examine a new worker A from counter 3 who replaces an exiting worker B in April 2005 (See Figure 3(b)). Worker A has a low estimated first month productivity at CNY111.8, relative to the average initial productivity CNY165.76 of all new workers in our data. The exiting worker B is a star salesperson whose productivity is at CNY392.2 in her last month; she is the third most productive salesperson based on our estimates. The left bars under “Productivity growth of the new worker in the original scenario” in Figure 4 illustrates the growth of A’s productivity from the second to fourth month through peer-based learning (learning-by-doing is negligible). Her productivity after the fourth month is CNY175.1 and stabilizes after that.

In the alternative hypothetical scenario, we consider how the new worker’s learning might have benefited from a change in employee turnover. We assume that worker B had been retained, and instead a lower productivity worker C (at CNY162.4) had exited. This scenario holds the total number of peers constant while changing the average productivity of inside peers

for the new worker. We assume that the worker B works in the same hours as worker C. The new productivity growth of A because of this personnel change from month 2 to 4 is presented in the middle bars under “Productivity growth of the new worker in the alternative scenario”. The total productivity growth in the three months would be CNY114.87, representing CNY51.57 or 81 percent higher growth than the original scenario. These simulated results suggest that retaining the best workers not only improves firm performance through their own productivity, but also through the indirect effect they have on the learning of new peers.

This peer effect also extends to competitors. In our example, counter 3 is adjacent to counters 1, 2, and 9 (see Figure 3(b)). In the second month after worker B left, a new salesperson at counter 2 replaces another worker, and in the third month another was employed at counter 1. By retaining worker B, the productivity of these new workers is also enhanced. Although it does not directly benefit counter 3, total sales of the cosmetic category are likely to increase and hence this is an important consideration from the profit perspective of the department store. We calculate the expected sales revenue of all other outside workers from the second to the fourth month based on our estimates. The right bars in Figure 4 show the results of the differences between the original scenario and the hypothetical scenario. Total external benefits of retaining worker B for the competing counters in the first three months is CNY32.88, a non-negligible impact, although it is smaller than the within-counter effect on worker A’s learning. This exercise highlights that when productivity growth is driven by peer-based learning, the economic consequence of losing or retaining existing high-ability workers can be significant.

Figure 4: Impacts of Retaining a Star Salesperson on the Productivity Growth of Peers



V. Discussion and Conclusion

This study identifies that, in the empirical setting of cosmetic sales, changes in worker productivity over time are mainly influenced by the presence of peers from which workers can learn. While knowledge transfers have been documented at the organization or firm level, we have identified individual-level peer-based learning that dramatically outweighs learning-by-doing in driving worker productivity growth. These occur both within and across firm boundaries, and their magnitude depends on the ability of both the source peer and the recipient of the knowledge. It is crucial for a new worker to work with more productive workers when she starts. In contrast, learning-by-doing and forgetting have little impact on worker's productivity

growth in our study. We believe this paper makes important contributions to the literature on learning and knowledge transfer in economics. To the best of our knowledge, this paper is the first to decompose the underlying mechanisms of individual worker learning in an industrial setting. The simultaneous and separate identification of individual peer-based learning and learning-by-doing has broad implications on firm's dynamic competitive strategies (Cabral et al 1994; Besanko et al 2010), productivity growth (Ghemawat and Spence 1985) and agglomeration effects (Zucker et al. 1998).

While we cannot directly observe the mechanism through which learning occurs, the difference in the magnitude of learning from inside peers and from outside peers suggests that workers learn from both observation and the active teaching of their peers, with the latter mechanism unlikely to occur across firm boundaries. The differential importance of learning from inside and outside peers in skin care and makeup categories suggests that active teaching from peers is critical to the learning of difficult tasks. For simpler tasks with easily observable knowledge, the observation of high-ability peers may be sufficient. We have also argued that the identified peer-based learning is likely to involve knowledge transfers and not entirely from adopting the work ethic of superior workers.

Our study has only focused on the empirical setting of cosmetic sales in a department store. The role of peer-based learning and learning-by-doing on worker productivity growth can be very different in other workplaces. We hypothesize that their importance relies on the nature of job tasks. For jobs requiring the knowledge of high-dimensional parameters in a stochastic production function (e.g. how to match unique customer needs with various cosmetic products), it is more effective for a new salesperson to learn the knowledge from experienced peers. In contrast, learning-by-doing may play a bigger role in more standardized job tasks with fewer

parameters, and in those where speed and precision of each work procedure are the primary concerns (e.g. farm and factory production). It is important for future research to explore the determinants of the job nature and workplace environment on the roles of learning-by-doing vs. peer-based learning. Furthermore, sources of the learning curve of a firm may come from the improvement of the production process or the adoption of better technologies. The exploration of such technology growth is out of the scope of this study, since in our data the retail technology has remained constant throughout the sample period. Yet this can be an important source of productivity growth for technology-intensive industries such as aircraft (Alchian 1963; Benkard 2000), ship building (Rapping 1965; Thompson 2001) and semiconductors (Hatch and Mowery 1998).

Finally, we have not been able to fully investigate the conditions that facilitate knowledge transfers between workers within and across firms. Our finding of the asymmetric peer-based learning between IC and TC counters indicates that financial incentives may be crucial for workers to invest effort in learning and teaching. Other factors that may impact worker interaction include the physical proximity and group identity of workers in the same or different organizations, individual worker personality and organizational culture. To what extent these factors facilitate the peer-based learning and how they affect firms' productivity growth are to be addressed by future research.

References

- Adler, Paul and Kim Clark. "Behind the Learning Curve: A Sketch of the Learning Process." *Management Science* 37 (1991): 267-281.
- Alchian, Armen. "Reliability of Progress Curves in Airframe Production." *Econometrica* 31 (1963): 679-693.
- Argote, Linda. *Organizational Learning: Creating, Retaining, and Transferring Knowledge*. Boston, Kluwer Academic.
- Argote, Linda, Sara Beckman, and Dennis Epple. "The Persistence and Transfer of Learning in Industrial Settings." *Management Science* 36 (1990): 140-154.
- Argote, Linda and Dennis Epple. "Learning Curves in Manufacturing." *Science* 247 (1990): 920-924.
- Arrow, Kenneth. "The Economic Implications of Learning by Doing." *Review of Economic Studies* 29 (1962): 155-173.
- Arrow, Kenneth. "Methodological Individualism and Social Knowledge." *The American Economic Review* 84 (1994): 1-9.
- Audretsch, David, Maryann Feldman. "R&D spillovers and the geography of innovation and production." *American Economic Review* 86 (1996) 630-640.
- Bala, Venkatesh and Sanjeev Goyal. "Learning from Neighbours." *Review of Economic Studies* 65 (1998): 595-621.
- Bandiera, Oriana, Iwan Barankay, and Imran Rasul. "Social Preferences and the Response to Incentives: Evidence from Personnel Data." *Quarterly Journal of Economics* 120 (2005): 917-962.
- Bandiera, Oriana and Imran Rasul. "Social Networks and the Adoption of New Technology in Northern Mozambique." *Economic Journal* 116 (2006): 862-902.
- Bayer, Patrick, Randi Hjalmarsen, and David Pozen. "Building Criminal Capital Behind Bars: Peer Effects in Juvenile Corrections." *Quarterly Journal of Economics* 124 (2009): 105-147.
- Becker, Gary. "Investment in Human Capital: A Theoretical Analysis." *Journal of Political Economy* 70 (1962): 9-49.
- Benkard, Lanier. "Learning and Forgetting: The Dynamics of Aircraft Production." *American Economic Review* 90 (2000): 1034-1054.
- Benkard, Lanier. "A Dynamic Analysis of the Market for Wide-Bodied Commercial Aircraft." *Review of Economic Studies* 77 (2004): 581-611.
- Bercovitz, Janet and Maryann Feldman. "Academic Entrepreneurs: Organizational Change at the Individual Level." *Organization Science* 19 (2008): 69-89.
- Besanko, David, Ulirch Doraszelski, Yaroslav Kryukov, and Mark Satterthwaite. "Learning-by-Doing, Organizational Forgetting, and Industry Dynamics." *Econometrica* 78 (2010): 453-508.
- Bucheli, Marcelo, Joseph Mahoney, and Paul Vaaler. "Chandler's Living History: The Visible Hand of Vertical Integration in Nineteenth Century America Viewed Under a Twenty-First Century Transaction Costs Economics Lens." *Journal of Management Studies* 47 (2010): 859-883.
- Buttrick, John. "The Inside Contracting System." *Journal of Economic History* 12 (1952): 205-221.
- Cabral, Luis and Michael Riordan. "The Learning Curve, Market Dominance, and Predatory Pricing." *Econometrica* 62 (1994): 1115-1140.

- Carrell, Scott, Richard Fullerton, and James West. "Does Your Cohort Matter? Measuring Peer Effects in College Achievement." *Journal of Labor Economics* 27 (2009): 439-464.
- Carrell, Scott, Mark Hoekstra, and James West. "Is Poor Fitness Contagious? Evidence from Randomly Assigned Friends." Forthcoming in *Journal of Public Economics*. (2011).
- Carrell, Scott and James West. "Does Professor Quality Matter? Evidence from Random Assignment of Students to Professors." *Journal of Political Economy* 188 (2010): 409-432.
- Chan, Tat, Jia Li, and Lamar Pierce. "Compensation and Peer Effects in Competing Sales Teams." *Unpublished Working Paper* (2011)
- Conley, Timothy and Christopher Udry. "Learning about a New Technology: Pineapple in Ghana." *American Economic Review* 100: 35-69.
- Darby, Michael and Edi Karni. "Free Competition and the Optimal Amount of Fraud." *Journal of Law and Economics* 16 (1973): 67-88.
- Darr, Eric, Linda Argote, and Dennis Epple. "The Acquisition, Transfer and Depreciation of Knowledge in Service Organizations: Productivity in Franchises." *Management Science* 41 (1995): 1750-1762.
- Duflo, Esther, Pascaline Dupas, and Michael Kremer. "Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from Randomized Evaluation in Kenya." Forthcoming in *American Economic Review*. (2011).
- Easley, David and Nicholas Kiefer. "Controlling a Stochastic Process with Unknown Parameters." *Econometrica* 56 (1988): 1045-1064.
- Edmondson, Amy, R.M. Bohmer, and Gary Pisano. "Disrupted Routines: Team Learning and New Technology Implementation in Hospitals." *Administrative Science Quarterly* 46 (2001): 685-716.
- Ellison, Glenn and Drew Fudenberg. "Rules of Thumb for Social Learning." *Journal of Political Economy* 101 (1993): 612-643.
- Ellison, Glenn and Drew Fudenberg. "Word-of-Mouth Communication and Social Learning." *The Quarterly Journal of Economics* 110 (1995): 93-125.
- Emons, Winand. "Credence Goods and Fraudulent Experts." *RAND Journal of Economics* 28 (1997): 107-119.
- Epple, Dennis, Linda Argote, and Kenneth Murphy. "An Empirical Investigation of the Microstructure of Knowledge Acquisition and Transfer Through Learning by Doing." *Operations Research* 44 (1996): 77-86.
- Ghemawat, Pankaj and Michael Spence. "Learning Curve Spillovers and Market Performance." *Quarterly Journal of Economics* 100 (1984): 839-852.
- Gruber, Harald. "Learning by Doing and Spillovers: Further Evidence for the Semiconductor Industry." *Review of Industrial Organization* 13 (1998): 679-711.
- Hatch, Nile and David Mowery. "Process Innovation and Learning by Doing in Semiconductor Manufacturing." *Management Science* 44 (1998): 1461-1477.
- Hoxby, Caroline M. "Peer Effects in the Classroom: Learning from Gender and Race Variation." *Unpublished Working Paper*. (2000).
- Huang, Chong, Hongbin Li, Pak Wai Liu, and Junsen Zhang. "Why Does Spousal Education Matter for Earnings? Assortative Mating and Cross-Productivity." *Journal of Labor Economics* 27 (2009): 633-652.
- Huckman, Robert, Bradley Staats, and David Upton. "Team Familiarity, Role Experience, and Performance: Evidence from Indian Software Services." *Management Science* 55 (2009): 85-100.

- Ingram, Paul. and Peter Roberts. "Friendship among Competitors in the Sydney Hotel Industry." *American Journal of Sociology* 106 (2000): 387-423..
- Jaffe, Adam, Manuel Trajtenberg, and Rebecca Henderson. "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations." *Quarterly Journal of Economics* 108 (1993): 577-598.
- Kremer, Michael and Edward Miguel. "The Illusion of Sustainability." *Quarterly Journal of Economics* 122 (2007): 1007-1065.
- Lapr , Michael, Amit Mukherjee, and Luk Wassenhove. "Behind the Learning Curve: Linking Learning Activities to Waste Reduction." *Management Science* 46 (2000): 597-611.
- Lerner, Josh, and Jean Tirole. Some Simple Economics of Open Source. *The Journal of Industrial Economics* 50 (2002): 197-234
- Lieberman, Marvin. "The Learning Curve and Pricing in the Chemical Processing Industries." *RAND Journal of Economics* 15 (1984): 213-228.
- Mas, Alexandre, and Enrico Moretti. "Peers at Work." *The American Economic Review*, 99 (2009): 112-145.
- Mostafa, Romel and Steven Klepper. "Industrial Development through Tacit Knowledge Seeding: Evidence from the Bangladesh Garment Industry." *Unpublished Working Paper*. (2011).
- Pisano, Gary, Richard Boehmer, Amy Edmondson. "Organizational Differences in Rates of Learning: Evidence from the Adoption of Minimally Invasive Cardiac Surgery." *Management Science* 47 (2001): 752-768.
- Rapping, Leonard. "Learning and World War II Production Functions." *The Review of Economics and Statistics* 47 (1965): 81-86.
- Reagans, Ray and Bill McKeivily. "Network Structure and Knowledge Transfer: The Effects of Cohesion and Range." *Administrative Science Quarterly* 48 (2003): 240-267..
- Reagans, Ray. and Ezra Zuckerman. "Networks, Diversity, and Productivity: The Social Capital of Corporate R&D Teams." *Organization Science* 12 (2001): 502-517.
- Sacerdote, Bruce "Peer Effects with Random Assignment: Results for Dartmouth Roommates." *Quarterly Journal of Economics* 116 (2001): 681-704.
- Thompson, Peter. "How Much Did the Liberty Shipbuilders Learn? New Evidence for an Old Case Study." *Journal of Political Economy* 109 (2001): 103-137.
- Thompson, Peter. "How Much Did the Liberty Shipbuilders Forget?" *Management Science*, 53 (2007): 908-918.
- Thompson, Peter. "Learning by Doing." In Bronwyn Hall and Nathan Rosenberg (eds.) *Handbook of Economics of Technical Change*, Elsevier/North-Holland. Pp. 429-476.
- Thornton, Rebecca and Peter Thompson. "Learning from Experience and Learning from Others: An Exploration of Learning and Spillovers in Wartime Shipbuilding." *American Economic Review* 91 (2001):1350-1368.
- Ton, Zeynep and Robert Huckman. "Managing the Impact of Employee Turnover on Performance: The Role of Process Conformance." *Organization Science* 19 (2008): 56-68.
- Williamson, Oliver. The organization of work: A comparative institutional assessment. *Journal of Economic Behavior and Organization* 1 (1980): 5-38.
- Young, Peyton. "Innovation Diffusion in Heterogeneous Populations: Contagion, Social Influence, and Social Learning." *The American Economic Review* 99 (2009): 1899-1924,
- Zucker, Lynne, Michael Darby, and Marilyn Brewer. "Intellectual Human Capital and the Birth of U.S. Biotechnology Enterprises." *American Economic Review* 88 (1998): 290-306.