Learning from Peers

Tat Y. Chan, Jia Li, and Lamar Pierce

September 20, 2010

Abstract. In this paper we study the impact of peers on coworker learning in a retail setting. Using four years of individual cosmetic sales data from a Chinese department store, we incorporate peer-based learning into a model with learning-by-doing and forgetting. We observe peer-based learning both within and across firm boundaries, identifying that the relative ability of coworkers impacts salesperson learning rate more than own experience. Using two sales tasks of different difficulty and two compensation systems, we observe evidence consistent with the presence of both active teaching and learning through observation. Our paper shows that the inter-organizational knowledge spillovers observed in past studies have micro-foundations in individual employees.

1 Tat Y. Chan is Associate Professor of Marketing and Lamar Pierce is Associate Professor of Strategy at Olin Business School, Washington University in St. Louis. Jia Li is Assistant Professor of Marketing at Krannert School of Management, Purdue University. Authors can be reached at chan@wustl.edu, jial@purdue.edu, or pierce@wustl.edu.
1. Introduction

An extensive literature in economics and management explains organizational learning through three primary processes. The main focus of this literature has been the process of learning-by-doing (Arrow 1962), with empirical work identifying production costs decreasing with cumulative production in aircraft (Alchian 1963; Benkard 2000), ships (Rapping 1965; Thompson 2001), trucks (Argote and Epple 1990), chemicals (Lieberman 1984), and semiconductors (Hatch and Mowery 1998).\(^2\) Related work has identified a second process in productivity declines through forgetting or knowledge depreciation (Argote et al. 1990; Darr et al. 1995; Benkard 2000, 2004; Thompson 2007). Yet a growing literature in economics and management argues that a critical third source for learning may be knowledge spillovers from other parties, where knowledge is transferred either through active teaching or observation. This work has identified the importance of knowledge spillovers across firm boundaries (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 2006) within the same firm.

Yet while much of the literature discusses the role of individuals in facilitating these knowledge spillovers, there exists a surprising paucity of evidence on the interactive learning of individual workers and their peers. While studying learning at the group and the firm level is important, the investigation of individual worker learning can help tease out the foundational processes that drive the heterogeneity of learning observed at more macro levels (Adler and Clark 1991; Argote 1999; Lapre and Tsikriktsis 2006). For example, the observation of learning-by-doing at the firm level could be explained by either the learning-by-doing of individual workers or alternatively by knowledge spillovers between workers that occur with experience.

---

\(^2\) See Thompson (2010) for a more extensive discussion of this literature.
working together (Pisano, Bohmer, and Edmondson 2001; Reagans, Argote, and Brooks 2005; Huckman, Staats, and Upton 2009). Similarly, the knowledge depreciation of a firm may be due to individual worker forgetting, or due to the lack of knowledge spillovers among workers when turnover rate is high (Ton and Huckman 2008). Furthermore, knowledge spillovers across firms may be only possible if there are sufficient (physical or non-physical) interactions between workers across firm boundaries.

In this paper we complement the literature on organizational learning by empirically examining how individual worker learning dynamically changes worker productivity. Our primary interest is to understand knowledge spillovers, or peer-based learning, among workers either within or across firm boundaries. This focus is consistent with an arising stream of literature on temporary peer effects on productivity. While previous studies (e.g. Bandiera et al. 2005; Mas and Moretti 2009; Chan et al. 2010) have shown how the ability of contemporaneously working peers impacts the performance of individual workers, peer-based learning may have a more important impact on the future performance of the worker. For complicated job tasks in environments with intense worker interactions, individual learning-by-doing may be insufficient. Instead, growth in worker productivity may require learning from peers, and the ability and effort of such peers may considerably impact dynamic performance improvements at both the individual and organizational level.

Our approach to studying peer-based learning is similar to the peer effects literature in that we identify the impact of the ability of coworkers on a worker’s performance. The key difference in our approach, however, is that we study how peer effects change the capital of knowledge and hence impact the future productivity of the worker.\(^3\) To study this long-term

\(^3\) It is important to note that empirical work in the economics of education demonstrates peer effects on academic performance among classmates (Hoxby 2000) or college roommates (Sacerdote 2001) that may relate to learning.
impact we build a learning model that allows an individual’s work ability to depend on the accumulation of her past interactions with other coworkers, learning-by-doing, and forgetting. We allow a rich learning mechanism that allows for whether the job tenure of the worker and whether the ability of peers is above or below the worker. A unique feature in our study is that we identify peer-based learning from both within and outside the firm at the individual worker level, something that to our knowledge has never before been identified in a market setting.

Our empirical context is cosmetic sales in a Chinese department store, where multiple manufacturers employ salespeople at co-located counters on the same retail floor. Salespeople at brand-based counters compete for customers with a constantly changing mix of peers assigned by the department store manager across shifts. Sales ability varies considerably across salespeople and changes dynamically with experience as new employees replace those who depart. A salesperson can learn sales techniques from other salespeople at the same and adjacent counters through two different channels: she may simply learn from observing how other salespeople serve customers, or she may learn in a more active way by seeking the advice and comments of peers. Either way provides the opportunity for knowledge spillovers between workers, but the direction and magnitude of such spillovers may depend on the difference in ability between each pair of workers and whether or not they work at the same counter.

We employ a nested non-linear least squares algorithm to simultaneously estimate changes in the monthly ability of workers and the impact of these changes on other workers both within and across counters. A major challenge in the identification of peer effects is that people

---

Carrell and West (2010) provide evidence on the relationship between teacher’s evaluation and student’s academic performance over years. Similar work in development economics shows social learning from peers in agricultural and health technology adoption (Bandiera and Rasul 2006; Duflo et al. 2006; Kremer and Miguel 2007; Conley and Udry 2010). 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 2001; Uzzi 1997; Reagans and McKeVily 2003).
may select peer groups based on unobserved characteristics. To control for this selection issue, we include worker fixed effects in the model; hence, peer-based learning is estimated from the within-worker variation in the pool of peers over time. Compared with previous studies that identify peer effects from cross-sectional data, we believe that the individual panel data used in this study help to mitigate the selection problem. We further exploit other variations in data that can be reasonably treated as exogenous shocks to worker selection. First, frequent leaves of absence by workers for vacation and personal reasons (e.g. sick or maternity leave) provide additional fluctuations in the ability level of peers. Furthermore, due to the construction surrounding the department store in the middle of our sample period, cosmetic counters are relocated on the retail floor, immediately changing the stock of peers at adjacent counters for every worker.

Using four years of individual sales data for 92 salespeople working for 11 collocated brand counters, we find evidence of both learning-by-doing and peer-based learning. Our results show that for a newly-hired salesperson, working 100 hours in a month on average improves her ability by about 10%. Yet we observe an even larger source of learning in the new worker’s peers, finding that superior peers produce learning benefits for coworkers while inferior peers appear to produce minor degradation in coworker ability. In our data, the magnitude of learning from peers, particularly for new employees, could be even larger than that of learning-by-doing. For instance, a new employee’s working for 100 hours with a peer with twice her ability enjoys triple the learning rate of an employee working alone. These results are consistent with findings in previous empirical studies on peer effects that suggest that heterogeneity in worker ability helps to enhance overall team productivity (Hamilton et al. 2003; Mas and Moretti 2009; Chan et al. 2010).
Moreover, relative peer ability influences knowledge spillovers not only within each counters but also across firm boundaries. Although the magnitude of peer-based learning is much stronger within the firm than across firm boundaries, this suggests that individual workers are an important source for knowledge spillovers across firms, locations, and product categories. In addition, we compare the magnitude of these peer-based learning effects in two sub-categories of cosmetics – skin care and makeup. Results suggest that task difficulty influences the relative importance of peer-based learning. For the more difficult task of selling skin care products, cross-counter learning is far less effective than within-counter learning, suggesting that close-distance observation alone is not sufficient. Where workers can glean knowledge from observation on simpler tasks such as selling makeup products, propinquity may be a sufficient condition for peer-based learning. Cross-counter learning hence is as efficient as within-counter learning. Finally, we find that peer-based learning effects at firms adopting individual-based compensation system are different from those adopting team-based compensation, suggesting that financial incentives may induce workers to invest different levels of time and effort to learn from peers.

2. 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 sells a wide range of products including apparel, jewelry, watches, home furnishings, appliances, electronics, toys, and food. The store 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.

We observe cosmetic sales for 11 of the 15 counters over a four-year period (January 1, 2003 – December 31, 2006). Descriptive statistics for those brands are summarized in Table 1. The counters vary both in average price and total revenue. While the average transaction size and price are within a fairly tight range, the number of sales varies considerably. The largest counter (Brand 1) is over seven times the size of the smallest (Brand 8), although their price range and transaction size are approximately equivalent.

<table>
<thead>
<tr>
<th>Table 1: Descriptive Statistics of Cosmetic Brands</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Annual Sales Revenue (US$)</strong></td>
</tr>
<tr>
<td>Min</td>
</tr>
<tr>
<td>Brand 1</td>
</tr>
<tr>
<td>Brand 2</td>
</tr>
<tr>
<td>Brand 3</td>
</tr>
<tr>
<td>Brand 4</td>
</tr>
<tr>
<td>Brand 5</td>
</tr>
<tr>
<td>Brand 6</td>
</tr>
<tr>
<td>Brand 7</td>
</tr>
<tr>
<td>Brand 8</td>
</tr>
<tr>
<td>Brand 9</td>
</tr>
<tr>
<td>Brand 10</td>
</tr>
<tr>
<td>Brand 11</td>
</tr>
</tbody>
</table>

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. The department store takes responsibility for the staffing of salespeople at all counters, with workers typically rotating shifts every day. For example, if a salesperson works the first shift on Monday, she will typically work the second or third shift on Tuesday. This scheduling process, while not completely random, ensures that each salesperson will rotate workdays and times, and thereby share their shifts with a variety of peers.
While workers work an average of six hours per day, they often exceed this amount, particularly on peak weekends and holidays. During our sample period, salespeople work alone at their counter roughly 30% percent of the time. For the smaller counters we occasionally observe all workers staffed simultaneously, which is likely due to workers staffing multiple shifts on high volume days.

The law requires that workers average two days off per week. However, since the department store is open 7 days a week, and weekends are usually its busiest time, it cannot adopt the standard way to accommodate the law (i.e., let workers off on every Saturday and Sunday). For the same reason, salespeople in the store typically cannot take off public holidays such as New Year, Chinese New Year, Labor Day, and National Day, nor would they 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. This data feature will provide a nice exogenous instrument to facilitate the identification of peer-based learning. When an individual salesperson completes a sale, the cashier records the identity of that salesperson, the product identity, quantities, prices, and other information. The time of transaction is also recorded in database. This careful sales tracking provides the store with detailed information about every cosmetic sale for each of its brands at the individual level. Each salesperson is assigned a unique ID. This allows us to infer the exit of an existing worker and the entry of a new worker. Ninety-two female salespeople worked for the 11 brands, with a 44 percent turnover among workers during the entire sample period. We observe a maximum of 49 salespeople employed in the department store compared with a minimum of 38. 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 many salespeople hired during the sample period leave the job before the end of the sample period.

Table 2: Descriptive Statistics of Cosmetics Sales Teams

<table>
<thead>
<tr>
<th>Brand</th>
<th>Total # of Salespeople during the Sample Period</th>
<th># of Entering Salespeople</th>
<th># of Exiting Salespeople</th>
<th>Observed Working Days</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Min</td>
<td>Max</td>
<td>Mean</td>
<td>Min</td>
</tr>
<tr>
<td>Brand 1</td>
<td>13</td>
<td>8</td>
<td>5</td>
<td>34</td>
</tr>
<tr>
<td>Brand 2</td>
<td>14</td>
<td>9</td>
<td>7</td>
<td>45</td>
</tr>
<tr>
<td>Brand 3</td>
<td>7</td>
<td>3</td>
<td>3</td>
<td>37</td>
</tr>
<tr>
<td>Brand 4</td>
<td>6</td>
<td>3</td>
<td>2</td>
<td>51</td>
</tr>
<tr>
<td>Brand 5</td>
<td>9</td>
<td>6</td>
<td>4</td>
<td>14</td>
</tr>
<tr>
<td>Brand 6</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>39</td>
</tr>
<tr>
<td>Brand 7</td>
<td>9</td>
<td>5</td>
<td>4</td>
<td>22</td>
</tr>
<tr>
<td>Brand 8</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>72</td>
</tr>
<tr>
<td>Brand 9</td>
<td>11</td>
<td>8</td>
<td>6</td>
<td>32</td>
</tr>
<tr>
<td>Brand 10</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>133</td>
</tr>
<tr>
<td>Brand 11</td>
<td>8</td>
<td>5</td>
<td>4</td>
<td>30</td>
</tr>
<tr>
<td>Mean</td>
<td>8.36</td>
<td>4.82</td>
<td>3.73</td>
<td>46.27</td>
</tr>
</tbody>
</table>

The 11 brands in our data use two different compensation systems: team-based commissions (TC) and individual-based commissions (IC). For of the brands use team-based commissions and pay each worker a monthly salary of $900-$1000 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 $800-$900 plus 2% of personal monthly sales. Under both compensation systems, a salesperson’s monthly salary depends on their performance in the month, giving them strong personal incentives to improve their ability to sell, and in TC counters, strong incentives to improve the ability of coworkers. Our data exhibit dramatic changes in worker ability over time, especially for new hired ones. Figure 1 plots the average learning curve of all new workers in our data across their months of service. Figure 1b is a smoothed prediction version of Figure 1a.
Figure 1: The Learning Curve of New Workers

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 2(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. The relocation decision was made by the store’s management because the construction surrounding the store caused the west gate to replace the east gate as the store’s main entry point; the store wanted to keep its cosmetics department in the area with the heaviest customer traffic. Figure 2(b) presents the floor plan and location of the cosmetics counters after the relocation. Such a relocation offers another important exogenous shock for us to identify peer-based learning effects – each salesperson after relocation faces a new set of peers from competing counters. For example, a salesperson at counter 2 was facing counters 1, 3, 4 and 5 before the change, but afterwards had a new opportunity for learning from her new neighbors at counters 8, 9 and 11.
Figure 2a: Cosmetics Floor Layout before the Relocation

Figure 2b: Cosmetics Floor Layout after the Relocation

Fed Counters: IC; Blue Counters: TC; Grey Counters: not in dataset
3 Identification of Learning from Peers

We first discuss those features in the data that allow us to separately identify learning-by-doing, forgetting, and learning from peers. The identification of learning-by-doing and forgetting in each month comes from variation in a salesperson’s work experience over time. If learning-by-doing is the key force in the change in worker productivity, by working fewer hours (e.g. for vacation) in a month, a person will have a lower growth in her average hourly sales in the next month, holding other factors constant. Note that empirical model, learning-by-doing will never reduce work ability. If we observe from data that salespeople who worked fewer hours in a month experience lower productivity in the next month, this will be inferred as forgetting in our model.

The identification of peer-based learning comes from variation in the pool of peers working at the same time in each month. Variation of co-workers in the same counter will identify the within-counter peer-based learning, while variation in other counters will identify cross-counter learning. During the entire sample period, we consistently observe exit and entry from workers for all 11 counters (See Column 3 and 4 of Table 2), with the overall turnover rate across all the counters approximately 44%. Suppose we observe from data that, after a star salesperson quits her job in a month, the productivity growth of other salespeople in the same counter in future months becomes lower. This implies a positive within-counter peer-based learning effect. If the productivity growth of salespeople in other adjacent counters also slows, we can infer that a positive cross-counter peer-based learning effect also exists. To distinguish peer-based learning from the peer effects that have been identified in the previous literature (e.g. Bandiera et al 2005, Mas and Moretti 2009, Chan et al 2010), note that the former increases the stock of knowledge or ability of peer workers, providing a long term impact on productivity,
with the latter provides a temporary shock only in that time period. In our example of peer-based learning, if the sales revenues of other workers do not decrease after a star salesperson quits her job, but instead produce slower productivity growth rates, this effect cannot be explained by the previously identified temporal peer effects in the existing literature.

It is widely-known that peer effects are difficult to identify from data because of endogenous selection issues. That is, people may select peer groups based on characteristics unobservable by researchers, which in turn leads to correlations in peer productivity observed in data. To control for this selection issue, we will exploit the panel nature of our data by including a fixed effect, i.e., \( \gamma_i \), for every salesperson in our model. Peer-based learning effects are therefore estimated not from the difference in peer groups across workers, but from the within-worker changes in the pool of peers through exit and entry of other workers over time. We believe that compared with other peer effects studies that use cross-sectional data, the time-series aspect of our data helps to mitigate the selection issue in this study.

Our study further exploits two other sources of data variation that can be reasonably assumed to be exogenous shocks in our model of peer-based learning. First, the counter relocation in the middle of our sample period brings an exogenous 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 worked in the store since 2001 and had exhausted the potential knowledge spillovers from neighboring counters. Second, we find from the data that the total working hours for each salesperson varies significantly across months. The store manager explained to us that this is because the department store requires that workers average two days off per week. Many
salespeople choose to work six or seven days a week in order to save for a longer vacation. While we do not directly observe the absence of any workers, the sales data shed some light on worker vacation and leave. A salesperson on average works 181 hours per month with a standard deviation of 61 hours. The standard deviation of the average monthly hours across all workers is 21, suggesting that the significantly variation in monthly working hours is within rather than across workers. 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 absence length is 63 days. These absences serve as shocks to the learning of new workers. For example, a new salesperson who has to work alone for more hours because her more experienced peers are on vacation will have a completely different learning process compared with others who can spend their first months observing and learning from peers.

4. An Empirical Model of Worker Learning

Our primary interest in this paper is to identify peer-based learning among workers. We therefore aim to build a model that identifies how co-located salespeople in our data learn from one another through repeated interactions, but also must account for baseline learning-by-doing and forgetting. Combining these sources for learning presents a more complete picture about how workers increase their skill and knowledge through accumulated work experience and knowledge spillovers from peers, and which sources of learning might be most important in new worker development.

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 selling ability in the month, \( \hat{Y}_{it} \), and a (row) vector of covariates that may affect sales, \( Z_{it} \).
including year (Year 2 through Year 4), month (February – December), day of week (Monday through Saturday), and brand indicators as follows:

\[
Y_{it} = \hat{Y}_{it} \cdot e^{\beta \cdot \eta_{it}}
\]  

(1)

where \( \eta_{it} \) is a (positive) error term representing demand shocks, such as store promotion of cosmetics or other product categories, that may affect the salesperson’s revenue. 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

\[
y_{it} = \hat{y}_{it} + Z_{it} \beta + \eta_{it}
\]  

(2)

The focus of our model is the evolution of the salesperson’s ability as a result of her learning process.

### 4.1 Modeling Peer-based learning

To identify peer effects in learning, we model how the abilities of individual salespeople change over time through repeated interactions with their peers. In our analysis, a salesperson’s peers are defined as all the department store’s other cosmetics salespeople who either work at the same counter (within-counter peers or inside peers) or work in any adjacent counters (cross-counter peers or outside peers).\(^5\) For example, in Figure 2a, counter 7 is adjacent to counter 4 while counter 9 is not.

We start with a baseline model in which a salesperson’s ability is affected by her learning from various peers in the previous month:

\[
\hat{y}_{it} = \hat{y}_{i,t-1} + \theta_1 \sum_{h} \left[ \frac{\sum_{k \in N_{j}, k \neq j} (\hat{y}_{k,t-1} - \hat{y}_{i,t-1})}{N_{j,h} - 1} \right] + \theta_2 \sum_{h} \left[ \frac{\sum_{k \in N_{j}} (\hat{y}_{k,t-1} - \hat{y}_{i,t-1})}{N_{j,h}} \right] + \tau_{it}
\]  

(3)

\(^5\) 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) also showed that cross-counter peer effects diminish with distance between counters.
where $N_{jh}$ and $N_{j}'$ denote the total number of salespeople working in $i$’s own counter $j$ and in any adjacent counters $j'$ in any hour $h$ the salesperson works, respectively. In the specification, $\hat{y}_it$ is first assumed to be dependent on her ability in the previous month, $\hat{y}_i,t-1$. We assume that it is also affected by the ability of all peers of the salesperson within and across counters. As we have discussed, salespeople in our data could work with different peers in different working hours throughout month $t-1$. Taking this coworker variation into account, we begin by specifying how the peer interaction at the hour level impacts worker $i$'s ability -- we assume that $\hat{y}_it$ is affected by the average ability difference in the previous month between $i$ and all her peers within and across counters who also work in any hour $h$ the salesperson works, represented by

$$\frac{\sum_{k\in N_{jh} \land k \neq j} (\hat{y}_{k,t-1} - \hat{y}_{i,t-1})}{N_{jh} - 1} \quad \text{and} \quad \frac{\sum_{k \in N_{jh}} (\hat{y}_{k',t-1} - \hat{y}_{i,t-1})}{N_{j'} - 1}$$

respectively. We then aggregate these hourly interactions to the month level in equation (2). Parameters $\theta_1$ and $\theta_2$ represent the within-counter and cross-counter peer-based learning effects, respectively, to be estimated from the data.

Finally, $\tau_{it}$ 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 ($\eta_t$) 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 $\hat{y}_i = (\hat{y}_{i1}, \hat{y}_{i2}, ..., \hat{y}_{it})'$ and $\hat{y}_{i,t-1} = (\hat{y}_{i1,t-1}, \hat{y}_{i2,t-1}, ..., \hat{y}_{it-1})'$ 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:

$$\hat{y}_i = \Theta_{t-1} \hat{y}_{i,t-1} + \tau_i \quad (4)$$
where \( \tau_t = (\tau_{t1}, \tau_{t2}, \ldots, \tau_{th})' \), and \( \Theta_{t-1} \) is a \( I \) by \( I \) square matrix with the \([i,s] \) element (assume worker \( i \) works in counter \( j \)) as

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

(5)

where the variable \( H_{i,t-1} \) is the natural log of total working hours for the salesperson in month \( t-1 \), and \( 1\{s \in N_{jh}\} \) and \( 1\{s \in N_{j'h}\} \) are indicators that worker \( s \) worked during hour \( h \) in month \( t-1 \) (for counter \( j \) or any of \( j \)'s adjacent counters, respectively).

Let \( y_t = (y_{t1}, y_{t2}, \ldots, y_{th})' \) be the vector of the (log) average hourly sales revenues of the \( I \) salespeople in month \( t \), and similarly let \( Z_t \) be a matrix with the row vectors \( Z_{jt} \) for all salespeople combined. From equations (1) and (4) we establish a relationship between workers’ sales in month \( t \) and their ability in month \( t-1 \) as:

\[
y_t = \Theta_{t-1} \hat{y}_{t-1} + Z_t \beta + \tau_t + \eta_t \tag{6}
\]

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 of \( T \) months. For any given month \( t=2, 3, \ldots, T \), we can repeat the iteration \( t-1 \) times to obtain

\[
y_t = \left( \prod_{g=1}^{t-1} \Theta_{t-g} \right) \hat{y}_1 + Z_t \beta + \epsilon_t \tag{7}
\]

where \( \hat{y}_1 \) represents the initial ability of all salespeople, which is a vector of parameters we will estimate. This also represents the fixed effects for workers. 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 ability in the first month.
of our sample. The error term $\epsilon_t$ is a combination of demand shocks and past learning shocks as the following:

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

It is easy to see that under such specification $\epsilon_t$ for every salesperson is serially correlated. If $\eta_t$ consists of macro demand shocks (e.g. store-wide promotions that lift sales for the whole cosmetic category), $\epsilon_t$ will also be correlated across salespeople in any month. This implies that we have to allow for a general structure of heteroskedasticity for $\epsilon_t$ in our econometric implementation.

### 4.2 Modeling Learning-by-Doing and Forgetting

Equation (7) represents our baseline model (Model 1), by which we aim to identify how salespeople learn from one another. Previous economics studies have documented that an organization’s knowledge could significantly improve through learning-by-doing or experience learning (Thompson 2010) but deteriorate with forgetting (Argote et al. 1990; Benkard 2000), making it critical to include both of these processes in any model of peer-based learning. In a selling environment like ours, a salesperson could improve her skills to make her customers more beautiful through continuous practice, even though this sometimes does not lead to purchases. On the other hand, without continuous practice her skills may be forgotten over time. With both of these components and peer-based learning in one model, we are able to compare their magnitudes to provide insight on which factors are most important in skill development by salespeople.

To model learning-by-doing, we assume that a worker’s ability in month $t$ is affected by her “working experience” in month $t-1$, which we proxy using the total number of hours she
worked in month $t-1$. The use of hours instead of sales assumes that salespeople learn equally well from failed sales as from successful ones. We redefine equation (3) in our baseline model to include peer-based learning, learning-by-doing and forgetting (Model 2):

$$
\hat{y}_{it} = \gamma \hat{y}_{it-1} + \theta_1 \sum_{h} \left[ \frac{\sum_{j \in N_{it}:j \neq i} (\hat{y}_{j,t-1} - \hat{y}_{ij,t-1})}{N_{jh}-1} \right] + \theta_2 \sum_{h} \left[ \frac{\sum_{j \in N_{ij}} (\hat{y}_{j,t-1} - \hat{y}_{ij,t-1})}{N_{jh}-1} \right] + \lambda H_{it-1} + \epsilon_{it}
$$

(8)

where the parameter $\gamma$ captures the carry-over of knowledge from last month of the current month. An estimate being significantly smaller than one implies the salesperson’s ability depreciation or forgetting. The parameter $\lambda$ captures the effect of learning-by-doing from the work experience in the previous month, and $H_{it-1}$ is the log of total working hours for the salesperson in month $t-1$, which proxies work experience.

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

$$
y_{i} = \Theta_{t-1} \hat{y}_{t-1} + \lambda H_{t-1} + \beta Y_{i} + \epsilon_{i}
$$

(9)

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

$$
y_{i} = \left( \prod_{g=1}^{t-1} \Theta_{t-g}^h \right) \hat{y}_{t-1} + \lambda \sum_{h=1}^{t-1} \left[ \left( \prod_{g=1}^{h-1} \Theta_{t-g} \right) H_{i,g} \right] + \beta Y_{i} + \epsilon_{i}
$$

(10)

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

4.3 Model Estimation

In equations (7) and (10), the $\Theta$’s of different months are functions of unknown parameters $\gamma$, $\theta_1$, and $\theta_2$. They multiply themselves over months and interact with the vector of salespeople’s initial ability $\hat{y}_i$ that is also unknown. Thus, both models 1 and 2 are non-linear. If we estimate equation (7) or (10) for all parameters (including parameters $\lambda$ and $\beta$)
simultaneously via the 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 (2010) to solve this problem. This procedure recognizes that the dimensionality problem mainly comes from \( \hat{y}_1 \), since there are 92 salespeople in our data. It exploits the observation that, conditional on the three parameters \( \gamma \), \( \theta_1 \), and \( \theta_2 \), the \( \Omega \)'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 \( \gamma \), \( \theta_1 \), and \( \theta_2 \), computing \( \Omega \)'s from equation (5), and then estimating \( \hat{y}_1 \) (and other parameters including \( \lambda \) and \( \beta \)) via standard linear least-square methods (inner procedure). The outer procedure is to search for the optimal \( \gamma \), \( \theta_1 \), and \( \theta_2 \) 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 \( \gamma \), \( \theta_1 \), and \( \theta_2 \). Since given \( \Omega \)'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 \( \Omega \)'s the estimate \( \hat{y}_1 \) 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 where the initial values for \( \gamma \), \( \theta_1 \), and \( \theta_2 \) are, the procedure always converge 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 $\varepsilon$ in equations (7) and (10) are serially correlated and probably correlated across salespeople working in the store in any month, we need to account for the general heteroskedastic structure of the error term when computing the standard errors. We compute the robust standard errors as

$$Var_{rob} = (X'X)^{-1} \left[ \sum_{i=1}^{N} (e_i \star x_i) \star (e_i \star x_i) \right] \ast (X'X)^{-1},$$

in which $N$ denotes the total number of observations, $e_i$ is the residual for the $i$th observation, and $x_i$ is a row vector of predictors.

### 4.4 Asymmetric Learning Models

Our models thus far have several limitations. First, they are “symmetric” in the sense that the $\theta$’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 $100 more per day is cancelled by the negative effect of learning from another peer who sells $100 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 starting ability 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. The magnitude of the learning effects for a new worker hence can be very different from an incumbent worker. 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 of this asymmetric model from the symmetric model in
equation (7) is that for the peer-based learning effects \( \theta_g, g = 1, 2 \) we now estimate two separate effects, \( \theta^a_g \) and \( \theta^b_g \). The former (latter) represents the within- or cross-counter peer-based learning effects from coworkers with higher (lower) ability. That is, for a focal worker \( i \) and her peer worker \( k \), we estimate \( \theta^a_g \) if \( \hat{y}_i \leq \hat{y}_k \) and \( \theta^b_g \) otherwise. Now we have four \( \theta \)'s in our Model 3 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 \( \hat{y}_i \)'s are linear in the \( y_i \)'s conditional on \( \gamma \) and \( \theta \)'s. With asymmetric effects, however, \( \hat{y}_i \)'s now interact with indicator functions \( \{ \hat{y}_i \leq \hat{y}_k \} \) or \( \{ \hat{y}_i > \hat{y}_k \} \). We employ a trick in order to avoid estimating them non-linearly. Notice that to construct the indicators \( \{ \hat{y}_i \leq \hat{y}_k \} \) or \( \{ \hat{y}_i > \hat{y}_k \} \), all we need is the ability 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 \( t \), \( \bar{y}_i \), to construct proxies \( \{ \bar{y}_i \leq \bar{y}_k \} \) or \( \{ \bar{y}_i > \bar{y}_k \} \) for indicators \( \{ \hat{y}_i \leq \hat{y}_k \} \) or \( \{ \hat{y}_i > \hat{y}_k \} \). 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 \( \theta \)'s and other parameters.

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.\(^6\)

---

\(^6\) 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 $\lambda$ and the peer-based learning parameters $\theta_1^a$, $\theta_1^b$, $\theta_2^a$, and $\theta_2^b$ to be different between new and incumbent workers in our model. This becomes our final model – Model 4.

4.5 Results

The results from our models are presented in Table 3. Column 1 presents our symmetric model with peer-based learning only. The ability of peers both within and across counters clearly impacts monthly sales growth. Workers learn at a faster rate when working with high-ability peers both within their own counter and at adjacent counters, with knowledge spillovers from within-counter peers being 3-4 times as large. This is likely because salespeople can learn more from within-counter peers through closer observation and through soliciting active teaching and advice. Still, the significant cross-counter learning effect suggests that it may be beneficial for a counter with only average workers to be located close to a competing counter with star salespeople. While sales may be hurt by stronger competition in the short run, the long term benefit for the counter is substantial since its salespeople can directly observe and learn from those star peers. Column 2 adds learning-by-doing and forgetting to the symmetric model. Working longer hours in the prior month appears to increase productivity, consistent with the literature on learning-by-doing at the organizational level. The forgetting parameter is less than but not statistically different from one, suggesting no forgetting or knowledge depreciation in our
Column 3 presents estimation results from the asymmetric model. While the effects from learning-by-doing and forgetting remain unchanged, we observe substantial asymmetry in knowledge spillovers 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.\(^8\) This may result from the focal salesperson learning bad sales techniques or practices, or from some spillover in adopting a poor work ethic that has long term effects. The substantive implication from these results is that, compared to teams consisting of homogeneous salespeople with average ability, 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 (2010), but it suggests that the benefit from team heterogeneity can be even stronger in long run, as new salespeople will learn from superior peers early in their careers in ways that will improve their performance long into the future.

Model 4 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 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

---

\(^7\) 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 absences off of our data support structure.

\(^8\) Given that the difference in ability between an inferior peer worker and the focal worker is negative (see equation 3), a positive coefficient implies the lower the ability of her peer the lower the ability of the worker.
from within-counter superior peers on new workers is about 10 times as large as effect on incumbents. In contrast, learning-by-doing for new workers is only about 3 times as large as for incumbents.

Our four models show consistent patterns in learning by cosmetic salespeople. We see considerable evidence that experience, in the form of hours worked, increases future productivity, consistent with past evidence on learning-by-doing at the firm level. However, learning-by-doing alone does not support the observation from store managers that it only takes a new salesperson three months to obtain full knowledge of product and selling techniques. Given that the worker exit rate is 44% and an average salesperson only works for two years in the data, learning-by-doing alone may 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 knowledge spillovers or peer-based learning. For a new salesperson, working together with an experienced peer of superior ability greatly enhances her future sales productivity. Furthermore, working closely with superior workers in adjacent counters also enhances her learning.
Table 3: Worker Learning Models

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

4.6 Peer-Based Learning in Different Product Categories

In the previous section we demonstrated that salespeople actively learn both from themselves and from peers, but we are unable to separate the different ways in which workers might learn. As we discussed earlier, workers might learn through simple observation, with or without the consent of their peers. But they might also learn through the cooperative teaching of
their peers when observation is insufficient for learning more complex or difficult sales tasks. In our context, worker learning involves selling cosmetics products, yet not all cosmetic products require the same knowledge and skills to sell. There are many categories of cosmetics products, including mascara, lipstick, face cleanser, body lotion, and sunscreen, all of which may have different difficulty levels in learning. Salespeople may be able to learn how to sell certain products purely by observing others, while others may require a combination of observation and coworker help through active teaching. In this section, we separately examine how salespeople learn to sell two main types of cosmetics products: makeup and skincare.

Interviews with cosmetic salespeople and managers in both China and the United States consistently reveal that skin care products are much more difficult to sell than makeup. Skin care products require the worker to have intimate knowledge of the chemical and mineral composition of the product, and are also much more expensive than makeup. Furthermore, while the benefits of makeup in improving appearance are immediately observable and verifiable by the customer, skin care products are experience goods, where the value of the good can only be observed through usage. Many skin products can be further classified as credence goods, where the quality of services might only be verifiable through long-term usage (Darby and Karni 1973). The information asymmetry of credence goods such as skin care products provides considerable incentives for opportunistic seller behavior (Emons 1997), a problem that makes the selling process much more difficult. Salespeople must convince buyers that a skin care product is worth its considerable price despite the absence of immediate verification.

To examine how learning differs across task difficulty, we identify and separate all the cosmetics products sold during our sample period into these two 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 models in columns 1 and 2 of Table 4, respectively. While the parameter on forgetting remains unidentified, learning-by-doing and peer-based learning remain significant predictors of monthly productivity gains for both types of products. Furthermore, the differences in the parameters suggest that the learning process is not identical for each product type. Makeup products have slightly stronger learning-by-doing parameters for both new and existing workers. Yet within-counter peer-based learning is much stronger for skin care, while across-counter knowledge spillovers are stronger for makeup.

Table 4: Product Category and Compensation Models

<table>
<thead>
<tr>
<th></th>
<th>Model 6</th>
<th>Model 7</th>
<th>Model 8</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Skin Care Products</td>
<td>Makeup</td>
<td>IC Counters</td>
</tr>
<tr>
<td><strong>Forgetting</strong></td>
<td>0.9941 (0.1815)</td>
<td>0.9936 (0.1789)</td>
<td>0.9834 (0.1659)</td>
</tr>
<tr>
<td><strong>Within-Counter Peer-based learning from Superiors</strong></td>
<td>0.3044*** (0.0637)</td>
<td>0.1672*** (0.0422)</td>
<td>0.2713*** (0.0627)</td>
</tr>
<tr>
<td><strong>Within-Counter Peer-based learning from Inferiors</strong></td>
<td>0.0333*** (0.0133)</td>
<td>0.0228*** (0.0107)</td>
<td>0.0098 (0.0081)</td>
</tr>
<tr>
<td><strong>Cross-Counter Peer-based learning from Superiors</strong></td>
<td>0.0166*** (0.0076)</td>
<td>0.1588*** (0.0564)</td>
<td>0.0971*** (0.0289)</td>
</tr>
<tr>
<td><strong>Cross-Counter Peer-based learning from Inferiors</strong></td>
<td>0.0327*** (0.0133)</td>
<td>-0.0077 (0.0081)</td>
<td>0.0052 (0.0062)</td>
</tr>
<tr>
<td><strong>New Worker Experience Learning</strong></td>
<td>0.0004*** (0.0002)</td>
<td>0.0006*** (0.0002)</td>
<td>0.0013*** (0.0004)</td>
</tr>
<tr>
<td><strong>Existing Worker Experience Learning</strong></td>
<td>0.0001 (0.0001)</td>
<td>0.0003*** (0.0001)</td>
<td>0.0007** (0.0001)</td>
</tr>
</tbody>
</table>

While we cannot positively identify the reasons for these differences, we can speculate based on the differences in how knowledge spillovers might occur differently within vs. across firm boundaries. Knowledge spillovers across firm boundaries are unlikely to involve active teaching given the competition between the firms, and are mostly likely to occur through observing techniques and sales tactics. In contrast, learning within the firm will likely involve both active teaching and learning through observation. The much stronger impact of inside peers  

---

9 We must allow for the 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).
on the learning of skin care sales is consistent with the higher difficulty of selling these products. New workers need active help from peers to learn how to pitch these products to consumers.

Since makeup sales is less complex and easier to observe, learning to sell these products from high-ability peers as adjacent counters is much more feasible than it is for learning skin care sales. The magnitude similarity of learning from inside superiors and outside superiors suggests that learning to sell makeup occurs mainly through observation. The higher rate of learning from experience for new workers selling makeup likely reflects the greater ease of learning as well, although the continued learning of existing workers is difficult to explain.

4.7 Peer-based learning under Different Compensation Systems

Chan, Li, and Pierce (2010) demonstrate that short-term peer effects are significantly influenced by the incentive structures that firms use to compensate their salesforces. The compensation system could also impact the long-term peer-based learning effects studied in this paper. Workers under team-based incentives may be willing to spend more time teaching new workers at their counters, while those with individual-based compensation have stronger incentives to improve their own ability. To investigate this possibility, we separately estimate the asymmetric peer-based learning effects for team-based commissions (TC) and individual-based commissions (IC) counters. As we discussed earlier, four brands use team-based commissions while seven use individual-based commissions. In Figure 2, the 4 TC counters are represented in red color and the 7 IC counters are represented in blue color.

Specifically, this can be done by extending equation (1) to the following:10

---

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.
\[
\hat{y}_{it} = \gamma \hat{y}_{i,t-1} + (\theta_1^{IC} \cdot 1\{j \in IC\} + \theta_1^{TC} \cdot 1\{j \in TC\}) \sum_h \frac{\sum_{k \in N_{jth}} (\hat{y}_{k,j-1} - \hat{y}_{i,t-1})}{N_{jh}} - 1 \\
+ (\theta_2^{IC} \cdot 1\{j \in IC\} + \theta_2^{TC} \cdot 1\{j \in TC\}) \sum_h \frac{\sum_{k \in N_{jth}} (\hat{y}_{k,j-1} - \hat{y}_{i,t-1})}{N_{jh}} + \tau_{it}
\]

in which the variables \(1\{i \in IC\}\) and \(1\{i \in TC\}\) are indicators that brand \(j\) is an IC counter or a TC counter. Parameters \(\theta_1^{IC}\) and \(\theta_1^{TC}\) represent the within-counter peer-based learning effects for IC and TC counters, respectively. \(\theta_2^{IC}\) and \(\theta_2^{TC}\) measure the peer-based learning effects from workers at competing counters on salespeople at IC and TC counters, respectively.

Columns 3 and 4 of Table 4 report the estimated peer-based learning effects of our Model 8. We see very little difference in the within-counter peer-based learning across the two compensations systems. This lack of difference may reflect two countervailing impacts of compensation system. While team-based compensation provides a stronger incentive for workers to actively teach their peers, those peers have weaker incentives to devote effort to learning. In fact, we observe that learning-by-doing is higher in IC counters than in TC counters, consistent with stronger financial incentives to increase individual performance. Furthermore, we 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. While team-based compensation may increase incentives for active teaching, these benefits appear to be overwhelmed by the decreased individual incentive to learn provided by this system. We must be cautious in our interpretation of these results, however, due to our inability to observe the endogenous selection process of workers into compensation systems.

5. Discussion and Conclusion
In this paper we identify that changes in worker ability over time are not purely functions of learning-by-doing and forgetting. They are also influenced by the presence of peers from which workers can learn through observation or active teaching. While knowledge spillovers are typically observed at the organization or firm level, individual-level spillovers are clearly important as well. These individual knowledge spillovers 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. Higher-ability peers increase learning, while lower-ability workers have a small negative impact on the permanent ability of their peers.

While we cannot directly observe the mechanism through which learning occurs, our parameter estimates support stories told to us by workers and managers and game theoretical models of learning. Workers learn from peers through both observation and the active teaching or their peers, with the latter mechanism unlikely to occur across firm boundaries. The differential importance of learning from inside and outside peers in our two major product categories suggests that active teaching is critical to the learning process on difficult tasks with tacit knowledge. For simpler tasks with easily observable knowledge, the observation of high-ability peers may be sufficient.

We believe this paper makes important contributions to both the economics literature on learning and the extensive literature on knowledge transfer and learning in management (Argote and Todorova 2010), where the need to “get behind” the learning curve has been extensively noted (Adler and Clark 1991; Argote 1999; Lapré et al. 2000). To the best of our knowledge, this paper is the first to directly observe individual workers impacting their peers’ learning in an industrial setting. While studies have looked at individual peer-based learning effects in
educational, consumer, and development settings, this work typically infers learning from limited observations of future performance, as opposed to the continuous time-series nature of our study. Prior papers on individual knowledge spillovers in industrial settings have instead relied on changes in team performance to infer knowledge spillovers between workers (Edmondson, Bohmer, and Pisano 2001; Pisano, Bohmer, and Edmondson 2001; Huckman, Staats, and Upton 2009; Huckman and Staats 2009). Furthermore, this paper is the first to break down knowledge spillovers across firms to the individual worker level. We are able to observe firms learning from competitors at the dyad level, showing more detailed mechanisms about how these widely-studied spillovers occur, including some suggestion about the balance of learning through observation vs. active teaching.

Furthermore, this paper adds a dynamic element to the literature on peer effects in personnel economics (Ichniowski and Shaw, 2010). We implement the methodological advances of Chan, Li, and Pierce (2010) into a dynamic setting, allowing worker ability, and its resulting impact on peers, to change across time. We show that peer effects in industrial settings are not only important in the cross-section, but also have a more permanent impact in the learning of workers. Whom you work with today matters not only now, but also long into the future.

While we believe our results are robust and novel, we must acknowledge the potential of endogeneity in hiring. While our cosmetic brands could potentially hire workers based on their potential to learn from the specific peers around them, we believe this is unlikely to account for our results. The exogenous shock to the workers (and the brands) of counter relocation means that brands were unlikely to know which peers would be adjacent. Furthermore we must also acknowledge the potential of a reflection problem (Manski 1993), where two workers

---

11 The literature on peer-based learning in development economics involves agricultural settings with self-employed farmers. See Conley and Udry (2010) for a discussion.
simultaneously learn from each other. Yet this is unlikely to be a major problem in a model of peer-based learning, since the strongest effects are likely to occur for new workers. If an experienced worker impacts the learning of a new worker, reflection is unlikely.
References


