© *Academy of Management Journal*
2009, Vol. 52, No. 2, 294–313.

SEARCHING FOR THE OPTIMAL LEVEL OF EMPLOYEE TURNOVER: A STUDY OF A LARGE U.K. RETAIL ORGANIZATION

W. STANLEY SIEBERT
University of Birmingham

NIKOLAY ZUBANOV
CPB Netherlands Bureau for Economic Policy Analysis

We study the relationship between sales assistant turnover and labor productivity in 325 stores of a large U.K. clothing retailer tracked over 1995–99. We find that the turnover-productivity relationship is contingent on type of work system. For a large group of part-timers, managed under a "secondary" work system, the relationship clearly has an inverted U-shape, but for the smaller group of full-timers, managed under a "commitment" system, the relationship is the conventional negative one. Implications for the contingency view of the link between turnover and productivity are discussed.

Research on the link between employee turnover and organizational performance has long been a battlefield of conflicting views. Thus, Dess and Shaw (2001) contrasted "human capital" theory, predicting losses in performance as turnover erodes firm-specific human capital, with cost-benefit approaches predicting an optimal level of turnover maximizing the difference between its benefits and costs. A particular benefit of turnover is replacement of poorly performing workers with better job matches (Ableson & Baysinger, 1984)—indeed, the Big Four accountancy firms have recently been reported as wanting to "keep the performance bar high" by maintaining a turnover rate of 10–12 percent (*Economist*, 2007: 74). Some turnover is required under economic theories of job matching (Jovanovic, 1979). Worker productivity can only be known imperfectly at the time of hiring. Therefore, some positive turnover must be beneficial, since it is rational for both employers and workers to continue an employment contract only if the workers' productivity is matched by their pay, while mismatches should separate.

Yet Dess and Shaw (2001: 448) noted that no organization-level studies have yet supported the curvilinear, inverted U-shaped, cost-benefit prediction. Much more common is the finding of a negative link between turnover and performance. The position is little different today, with only Glebbeek and Bax (2004) having found signs, albeit not statistically significant ones, of an inverted U-shaped relationship. Harris, Tang, and Tseng (2006) found a curvilinear relationship, but at an aggregate level in a cross-firm study using Australian data. Studies by Shaw, Gupta, and Delery (2005) and Shaw, Duffy, Johnson, and Lockhart (2005) have confirmed the usual negative relationship. However, the difficulty of testing these theories empirically is that "unobservables," such as management ability, affect performance and confound the true turnover-performance link. Our study brings forward new organization-level data with which to test the cost-benefit prediction.

As Glebbeek and Bax (2004: 279) noted, the methodology of using different units from within the same organization—rather than comparing organizations—offers the advantage of more control. By making an organization-level study, we aimed to control for some of the factors affecting the link between performance and turnover that are not easily observed and thus are difficult to capture in a multiorganization study. Examples include man-

We wish to thank the executives and staff at the study company for providing us with the data and assisting us with management insights. We also wish to thank the editor (Peter Bamberger) and three anonymous reviewers for their helpful comments and their patience during the review process. Arnaud Chevalier and Tarja Viitanen performed early work on the data set, and thanks are due to them as well. Finally, we are grateful for valuable suggestions from participants at the University of Aarhus Low Wage Employment Network (LOWER) Conference, and at Birmingham University Economics Department's seminar, and also Erling Barth, Mary Gregory, Maura Sheehan, Peter Sinclair, and Niels Westergaard-Nielsen.

Because the data used in this study are proprietary, the authors will only be able to release them with prior permission from the company.

agement practices, the nature of operations, and economic conditions facing the industry or region in question. Also, our methodology could be directly used to assist central management calculations in any large organization with multiple subsidiary plants. The single-firm literature is growing (see Lin [2005] for the industrial relations literature), with noteworthy studies by Bartel (2004), of branches of a Canadian bank; by Kacmar, Andrews, Van Rooy, Steilberg, and Cerrone (2006), of restaurants in the Burger King chain; by Morrow and McElroy (2007), of bank branches; and by Shaw, Duffy et al. (2005), also of restaurants. Admittedly, there is the disadvantage that a single organization is not representative. Nevertheless, lessons can be learned, so long as its relevant characteristics are made clear. Basically, we believe that our results will apply to most situations requiring management of many unskilled workers over dispersed sites.

We conducted our analysis in the context of the work systems for sales assistants in a large U.K. retail organization, whose stores we tracked over 1995–99. We leave aside the more skilled store and central management groups, as well as central administration and warehousing. For these groups, the inverted U-shaped turnover-performance association is likely to be ruled out because substantial human capital is lost with turnover; a considerable literature on such losses exists (Arthur, 1994; Batt, 2002; Huselid, 1995; Shaw, Duffy, et al., 2005; Shaw, Gupta et al., 2005). But for sales assistants the position may be different. Their turnover is high, with an average headcount separation rate of about 25 percent a year in our organization. The job might be thought of as low skilled, with low involvement, and so some level of employee turnover, and perhaps quite a lot, may be appropriate. Therefore, here we would expect to find an inverted U-shaped turnover-performance relationship.

But even within a seemingly homogeneous employee group such as sales assistants, different work systems may be necessary. We needed to be alert to this possibility, following Lepak and Snell's (1999: 42) questioning of the "one-size-fits-all" approach to analyzing work systems. In fact, as we will show below, two work systems appear to coexist for sales assistants within our organization: a "commitment" system for the core full-time group (20 percent of the workforce), and a "secondary" system for the many part-timers. For the core workers, we expected a straight negative turnover-performance link, and we expected the inverted U to apply only to the secondary group. Some evidence for this contingency view exists, in that both Arthur (1994: 683) and Guthrie (2001: 186) found that turnover-performance relationships were negative only for organizations using "high-involvement" work systems (the commitment system in our terms) and were positive for organizations with low-involvement work systems (our secondary work system) (but see Shaw, Gupta et al. [2005: 64] for a no-contingency finding). The cited studies were conducted in multiple organizations; ours was within one organization. We thus carry the contingency debate further (see Datta, Guthrie, & Wright, 2005; Lepak, Taylor, Tekleab, Marrone, & Cohen, 2007; Youndt, Snell, Dean, & Lepak, 1996), allowing for heterogeneity within our organization. Following this line of argument, we propose a contingency model in which the turnover-performance relationship is contingent on the nature of work systems within an organization and differentiate between commitment and secondary sales assistant groups.

The possible coexistence of two different work systems for rank-and-file workers has implications both for central management and for the wider strategic human resource management (HRM) research effort concerned with "high-performance" work systems. For management, the problem will be to allow for the work systems to differ between groups within their organization and to try to achieve the correspondingly different optimal levels of turnover. Our findings should assist managers in making such decisions. For strategic HRM research, the problem will be how to characterize the work system of an organization if some of its operatives are optimally managed with high-performance practices (commitment), and others with low-performance practices (secondary). More differentiation is needed, as Becker and Huselid (2006: 905) called for, and as is indeed beginning to emerge with Lepak et al.'s (2007) distinction between "core" and "support" workers. Our findings below further examine differentiation.

Our approach also offers two measurement innovations. The first is more accurate measurement of employee turnover. For each of our organization's 325 stores, we had daily data on when a worker was hired or separated (in fact, most separations are quits, and dismissals are few; see below for management methods for inducing "voluntary" separations), and on his/her contracted hours. These data allowed us to analyze turnover not simply in terms of headcount, as is the norm, but with adjustment for full-time equivalents—that is, hours gained or lost from hirings or separations. In fact, as emerges below, most turnover occurred among the short-hours, part-time workers. There are many part-timers in our studied organization, but not many hours are lost when one leaves. Thus, the adjusted turnover rates are below the headcount measures. We show that the performance-turnover link is much clearer when turnover rates are measured in their full-time equivalents, a reasonable convention given that output primarily depends on worker hours, not headcounts.

The second innovation our approach offers was derived from taking advantage of the panel structure of our data, which allowed every store to be tracked over five years. Having five annual observations for each store meant that we could include in the regression equation a term representing store-specific unobservables, or "omitted variables" (Becker & Huselid, 2006), that affect sales—such as management ability and store location. This approach would constitute a fixed-effects analysis if the store-specific unobservables were correlated with the regression variables, or random-effects analysis if they were not. By allowing for these unobservables we can obtain true "average" (see Capelli & Neumark, 2001: 744) turnover-performance relationships.

THEORETICAL BACKGROUND FOR THE TURNOVER-PERFORMANCE LINK

Turnover in Commitment versus Secondary Work Systems

Employee turnover has benefits and costs for organizational performance (Ableson & Baysinger, 1984), and an inverted U-shaped relationship between employee turnover and performance may be the net result of these, depending upon the work system in question. Following Bamberger and Meshoulam (2000), we contrast two archetypical work systems, the commitment and the control, or secondary, systems. The difference is based on the nature of the production process. On the one hand, the commitment system can be said to arise when managers are unable to monitor or fully evaluate employee behavior. On the other hand, the secondary system arises when operations are routinized and low-cost. In relation to staffing and training, in the commitment context (Bamberger & Meshoulam, 2000: 67) careful selection is paramount, as are career development and dependence on the internal labor market of the organization in question. In contrast, the secondary system calls for little reliance on selection and for little career development, and it depends on the external labor market.

Within this framework, let us look in turn at the benefits and costs of turnover. On the benefits side are two major components: First, the benefits of improved job matching can be seen—revitalizing a workforce by weeding out the bad workers and preventing Dalton and Todor's (1979: 226) "trained incapacity." Turnover here is "involuntary," involving dismissals (though, as Dalton and Todor noted, dismissal rates are a poor guide, since "voluntary separation may be viewed merely as a convenient substitute for termination" [1979: 331]). Although a worker's productivity can only be known imperfectly at the time of her or his hiring, the costs of mistakes are higher in the commitment system, so there must be fewer "false positives," and thus less (involuntary) turnover from this source. In a secondary labor market, on the other hand, where the costs of selection errors are relatively small, there are likely to be more such mistakes, which will require more turnover to clear. Second, there is the benefit of keeping labor costs variable where sales are variable—"greasing the wheels"—which our studied organization does by keeping pay rates low (dismissals are thus less needed, since voluntary turnover is high). Such flexibility benefits also imply marked turnover in the secondary system. In the commitment system, where there is careful selection and training, labor hoarding will dampen employee turnover. It must be emphasized that both these benefits reach a limit beyond which further turnover from this source does not help.

The costs side also has two main components. The first is the loss of specific human capital (or unique human capital, in Lepak and Snell's [1999] terminology), which will become increasingly onerous as turnover increases and is a problem likely to be associated particularly with voluntary turnover (quits). This factor, above all, will make high turnover detrimental. Specific human capital includes both formal training and "tacit knowledge" (Kacmar et al., 2006: 135) acquired through practice. It is clear that high turnover leads to loss of formal training because new hires have to be trained up. Indeed, reduced worker productivity during the learning period has even been categorized as the highest cost of turnover (Cascio, 1998). Turnover also reduces tacit knowledge of, for example, service values and norms, the transmission of which to new employees is disrupted (for a survey, see Morrow and McElroy [2007: 830]). Again, there is likely to be a distinction between the two systems simply because specific human capital is so much more important in the commitment work system.

The second component of turnover's costs are the costs of losing particularly productive, "star" workers through dysfunctional turnover. As Shaw, Delery, Jenkins, and Gupta (1998) showed in their study of voluntary and involuntary turnover in trucking, this type of dysfunctional voluntary turnover depends more upon pay-for-performance practices rewarding good workers than on careful selection practices (see Harrison, Virick, and William [1996] for the effects of pay for performance). In the commitment system, where pay is more tailored to the individual, dysfunctional turnover is less likely than in the secondary system, where it is simply derived from surveys of the external labor market (Bamberger & Meshoulam, 2000: 129). This said, dysfunctional turnover is likely to be more costly per quit in the commitment system, where able workers are more valuable, and these costs will also work to reduce turnover.

Summing the benefits and costs gives the overall turnover-performance relationship, which is likely to be contingent on the type of work system in force. Under the commitment work system, selection is careful, so not much involuntary turnover (dismissal) is required to weed out the bad workers. At the same time, the costs of dysfunctional turnover in the form of voluntary turnover (quits) and the losses of human capital owing to turnover rise quickly. The net effect is that the range of turnover that is beneficial for performance is likely to be very small, if it exists. In general, under this scenario, we then expect to find a negative turnover-performance relationship.

Under the secondary work system, careful selection is not as profitable, so accordingly more turnover is required to weed out bad workers. In fact, within the secondary model there is "the presumption of considerable interfirm movement" (Osterman, 1987: 51). The costs of dysfunctional turnover per separation are lower, and human capital losses are clearly less important. In the secondary system, therefore, all these factors point to a wider range of turnover values over which the effect of turnover on performance is positive than that for the commitment system. This context was applicable to most of the sales assistants in our case, and we therefore expected an inverted U-shaped relationship here. A particular further point in our case, as noted above, is variability of demand, which was expected to drive the optimum point of the inverted U further outward.

Coexistence of Work Systems

Although the commitment and control (secondary) work systems have different implications for the turnover-performance link, the two systems may coexist in the same organization. Indeed, Lepak and Snell (1999: 32) questioned the appropriateness of taking a single HR architecture as optimal for managing all employees. Clearly, different work systems will apply to managers and rank-and-file employees, requiring a distinction between "core" and "support" employees (Lepak et al., 2007), or concentration simply on nonmanagerial employees (Batt, 2002). But even within a seemingly homogeneous rank-and-file group, such as the sales assistants in our organization, different work systems may be necessary. Sales assistants may seem to be all alike, but they are not. Although all of them perform some basic, easy-to-monitor functions, such as assisting customers in their choices, shelving, stock taking, and cleaning, their levels of commitment and responsibility differ, and with these differences come extra tasks that are less easily monitored and require more managerial insight.

In the organization studied here, at one end of the commitment spectrum are part-timers, some of whom are extremely part-time, filling in particular shifts for example on busy Saturdays; 17 percent worked fewer than five hours a week. Their career aspirations lie elsewhere, which is why their head-count turnover rate is much higher than that of the sales force as a whole (21 percent per year). Naturally, this group receives less responsibility, less specialist training, and fewer promotion opportunities. Their pay is flat and is determined by salary surveys of similar occupations in the county (that is, by the external labor market). Although not formally on temporary contracts—only 5 percent of this organization's employees are on such contracts—they are temporary workers in all but name. Thus, following Bamberger and Meshoulam's (2000) categorization, we identify this group as managed under a secondary work system. Given the part-time nature of employment in our organization, we propose that sales assistants working fewer than 30 hours per week belong to this group. These part-timers constitute 80 percent of all the firm's sales assistants.

At the other end of the spectrum, our organization contains sales assistants working regular full-time hours (30 or more per week). These workers form the core of every store. They are the most responsible and committed and are less likely to separate than the part-timers. In addition to performing the basic functions outlined above, they open and close the stores in which they work, reconcile cash with receipts ("cash up"), help induct and train newcomers, and generally act as communication nodes, for example sometimes conducting team briefings instead of the manager. They receive specialist training in administration, wage control, and health and safety. The process through which they are selected is tougher: the yearly intake of full-timers is lower than that of part-timers and, of part-timers with a given tenure, some, but not all, become full-timers. It is this group that forms the pool from which future store managers are generally drawn, particularly for the smaller stores (65 percent of store managers derive from internal promotions, according to our interviews). This possibility of promotion provides an implicit link of pay to performance. Therefore, this group fits the description of being managed under a commitment work system.

Our allocation of sales assistants into two different work systems—the basis of our contingency view of the turnover-performance link—is testable. It may be objected that sales assistants, even the most committed ones, are not sufficiently skilled to be managed under a true commitment system (they are far from Lepak and Snell's [1999: 36] Intel engineers). Yet differentiation makes sense if the full-timers' job is more strategic, thereby requiring disproportionately greater investment (Becker & Huselid, 2006: 919). Our description above provides good grounds on which to propose such differentiation for turnover-performance links, and below we present results of regression analyses both with and without the differentiation. Thus, readers can see for themselves whether the differentiation that accords with our theoretical priors also achieves a better fit with the data.

Hypotheses

We have suggested that some level of sales assistant turnover is beneficial for performance in a secondary work system, while accepting that there can be too much turnover. We have assumed that the part-timers working fewer than 30 hours per week are managed under a secondary work system. Therefore, with respect to these workers we propose our first hypothesis:

> *Hypothesis 1. Under a secondary work system, there is an inverted U-shaped relationship between employee turnover and performance.*

Given their job description, the full-timers (30+ hours per week) are managed under a commitment work system. Accordingly, for this group we propose:

> *Hypothesis 2. Under a commitment work system, there will be less of an inverted U-shaped relationship between employee turnover and performance than under a secondary system, and more of a straight, negative turnover-performance relationship.*

Although the two systems have different outcomes, we must also allow for their interdependence where they coexist in one workplace. Increased turnover among one group might exacerbate the negative effects of turnover among the other (see also Broschak and Davis-Blake's [2006] study of the way work systems for part-timers affect full-timers' behavior, and conversely). In fact, the social capital theory of turnover and performance (Shaw, Duffy et al., 2005: 595–596) provides good grounds for expecting interdependence. Thus, the costs of secondary turnover will depend on how quickly new hires are trained, which requires input from the full-timers who are supposed to communicate tacit knowledge to them. But the higher the full-time turnover, the rougher the transfer of this knowledge, increasing the costs of secondary turnover without affecting its benefits. A lower inverted U-shaped curve with an optimum closer to zero will be the result.

Equally, higher part-time turnover may exacerbate the negative effects of full-time turnover on productivity. This effect in turn may be attributed to an argument of human capital theory that turnover depletes "organizational skill banks" (Shaw, Duffy et al., 2005: 594). Thus, when part-time turnover is high, less will be known about the abilities of the part-timers in an organization's pool of workers, and there will be fewer to choose from for promotion to the core full-timer group. Since it will be more difficult to find replacements, the costs of full-time turnover will rise. Although we do not propose a separate hypothesis for the interaction effect between part- and full-time turnover, we do control for it by allowing for a cross-product term of the two turnover rates, as is outlined in our next section, on measures and regression analysis specification.

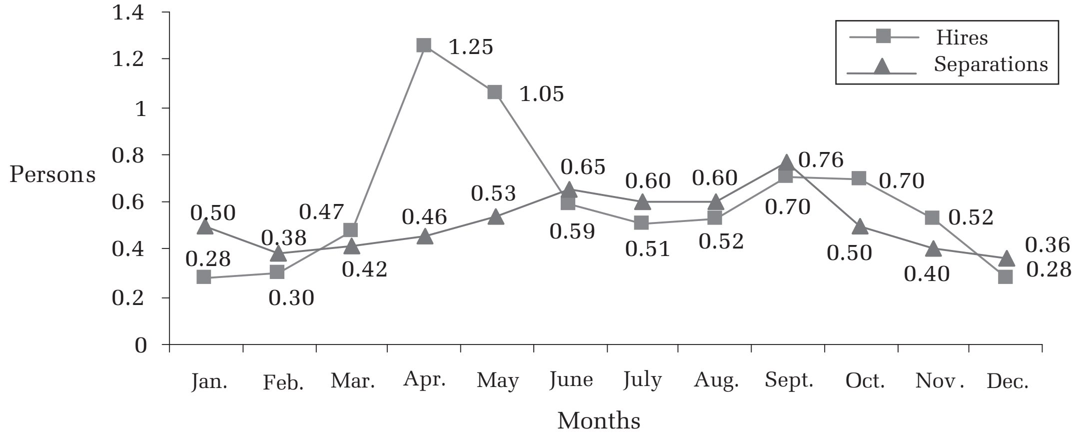
METHODS

Organization

The organization from which we derive our data offered an opportunity for testing both of our hypotheses. As discussed above, most of its employees are managed according to the principles of the secondary work system, a situation that was appropriate for testing Hypothesis 1. At the same time, the management of a sizeable minority of employees according to elements of a commitment work system provided a set-up for testing Hypothesis 2. We now describe the features of our organization that were relevant to the hypotheses in a greater detail.

Given that our units of observation were the stores of a large company, it is necessary to understand the centralized management system under which these stores operate. Our sample is 325 of

FIGURE 1
Hires and Separations in an Average Store by Month and Headcount



the company's more than 500 stores chosen because they had continuous data available for our period, 1995–99, and were wholly owned rather than franchised. The stores are spread all over the United Kingdom; they are quite small (containing on average about seven full-time-equivalent workers); and they sell a similar range of goods distributed from the company's warehouses. In these circumstances, the company has naturally adopted centralized management.

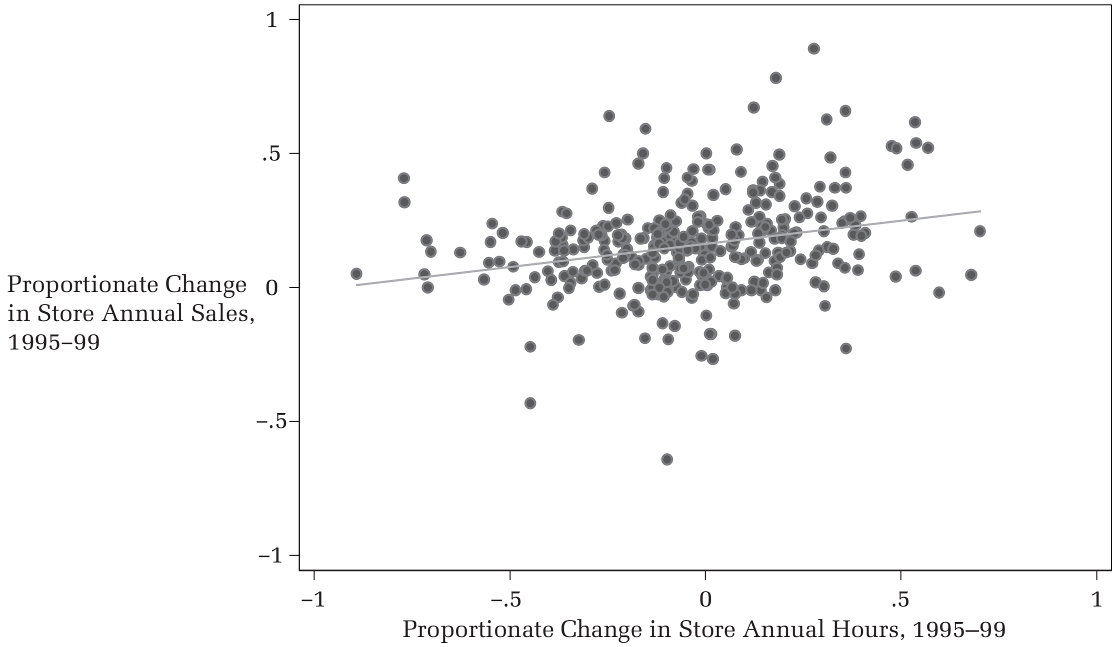
The wage policy is quite simple. Central management sets out a standard hourly wage schedule for sales assistants, determined every year, for all stores. During our period of study, this schedule set wages that differed by region and age. Four region groups (rural, town, city, and central London) and two age groups (18 or younger, and older than 18) are specified. There are no wage increments for seniority and no immediate pecuniary rewards for performance for sales assistants. Training varies by group (see above) but is basic for the majority of sales assistants.

The company's management rule, "right people, right time," is a good summary of its personnel policies with respect to both part- and full-timers. Within the constraints of central policies, store managers select workers, negotiate hours, and regulate turnover. By hiring workers to make up for separations, store managers match the workforce to the seasonal pattern in demand and also to yearly fluctuations, which are quite marked (see below). In addition, the managers negotiate contract hours, maintaining a pool of part-timers to be employed at the "right time," and can if necessary pressure a worker to leave by reducing hours—a useful alternative to terminating contracts, which is difficult under U.K. employment law. Most separations are "voluntary," therefore. In these circumstances, high employee turnover among part-timers is to be expected (Batt, Colvin, & Keefe, 2002; Shaw et al., 1998), particularly if wages are set relatively low, as is the case. This strategy is required by the turbulence of the retail environment, which we delineate in the next subsection. On the other hand, core workers whom managers particularly want to keep are given longer hours and more training. A commitment approach is required for managing this group of "right people" to ensure coherent communications and a supply of management talent.

The retail environment leads to both seasonal and year-to-year fluctuations in sales, employment, hirings, and separation that are worth illustrating. Figure 1 illustrates the seasonal turbulence affecting the average store (though we emphasise that we used only annual data for our empirical work). We see a relatively steady headcount separation rate of around half a person per month. However, peaks in the hiring rate occur at the beginning of summer, and then in September-October, which is the season for children's clothing sales. Store managers have to try to even out labor productivity over the year, so as to avoid having idle staff on the one hand, and overlong customer queues on the other.

Store managers also have to meet year-to-year fluctuations in demand. Figure 2 shows how difficult this problem can be, graphing large changes in labor input and in sales for many stores during 1995–99. In fact, around one-fifth of the stores expanded or contracted hours worked by more than 50 percent over our study period. Sales were almost as changeable. Further, the link between changes in labor inputs 1995–99 and changes in sales is weak, correlated at only .26. Therefore, labor productivity was difficult to predict, and store managers had to work hard to provide the appropriate labor inputs to service their customers. In such a turbulent en-

FIGURE 2
Changes in Store Employment and Sales, 1995–99[a]



[a] $r = .26$, $p < .001$, 325 stores.

vironment, it is as important to downsize as upsize, so a certain amount of employee turnover, at least among part-timers, has clear advantages. The question is, What is the optimum level of such turnover?

Data

Our data set contains detailed information on employee turnover and personnel characteristics. It consists of three parts. The first part is based on personnel records of all the employees who worked at any time between 1995 and 1999. Altogether, there are 72,669 records of 32,546 individuals. Individual records include age, gender, date hired, date left employment, and weekly contract hours, our measure of hours worked. We thus could observe job matches as short as one day and accurately adjust our measures of turnover to their full-time equivalents. Since individual productivity data were not available, we derived employee average data for each store in order to compute store annual average productivity.

Then, to the employee information we matched the second part of our data, which contains store information, including the output variable, real annual sales per store, store square footage (our capital variable), and other potentially relevant characteristics (see below). The third part of the data consists of areawide wage and unemployment data for the county in which each store was situated. These data were obtained from the UK Labour Force Survey, which is based on samples of individuals 16 to 65 years of age in 63 U.K. counties. To avoid dealing with store entry and exit events, which might present special effects of downsizing on performance (see McElroy, Morrow, & Rude, 2001), we excluded stores that opened or closed down during 1995–99 and, for the remaining stores, controlled for those having more than a 20 percent hours reduction in any year. The resulting data set includes 1,625 observations from 325 stores for 1995–99.

Measures

Dependent variable. Our measure of store performance is labor productivity. We measured this variable as annual sales per store, adjusted for inflation, divided by total annual hours worked in the store. An alternative measure of performance would be store profitability, but the company could not reveal this information. Nevertheless, labor productivity is an important variable in its own right, and it has been widely used in studies of organizational performance (e.g., Glebbeek & Bax, 2004; Shaw, Gupta et al., 2005).

Turnover variables. Our measure of turnover is the separation rate, as is common (see, for example, Glebbeek and Bax [2004] and Shaw, Duffy et al.

TABLE 1
Turnover Calculations: Full-Time Equivalents versus Headcount Measures[a]

Scenario	Changes in Workforce due to Separations: Reduction in Hours	Changes in Workforce due to Separations: Reduction in Headcount	Separation Rate per Year: Full-Time Equivalents	Separation Rate per Year: Headcount
Full-timer leaves, and another hired December 1, so 160 hours gained and lost in the year:	160 hours	One person	$\frac{160}{14{,}000} = 0.01$	1/8 = 0.125
Half-timer leaves and another hired December 1, so 80 hours gained and lost:	80 hours	One person	$\frac{80}{14{,}000} = 0.005$	1/8 = 0.125
Full-timer leaves and another hired January 1, so 2,000 hours gained and lost:	2,000 hours	One person	$\frac{2{,}000}{14{,}000} = 0.145$	1/8 = 0.125
Half-timer leaves and another hired January 1, so 1,000 hours gained and lost:	1,000 hours	One person	$\frac{1{,}000}{14{,}000} = 0.07$	1/8 = 0.125

[a] Several assumptions underline our example. (1) There are eight people in the store initially, six full-time (40 hrs/week), and two half-time (20 hrs/week), which makes seven full-time equivalents (FTEs): 6 staff × 40 hrs/week × 50 weeks + 2 staff × 20 hrs/week × 50 weeks = 6 × 2000 hrs/year + 2 × 1,000 hrs/year = 14,000 hrs/year. (2) Separations are instantly replaced by equivalent hires, so the headcount and FTE employment stays the same throughout the year. (3) The FTE-adjusted separation rate is calculated as the number of hours that leavers would have worked had they not left, divided by the total hours worked during the year. (4) The headcount separation rate is calculated as the number of leavers divided by the average number of workers employed in the store during the year (eight).

[2005b]; see Morrow and McElroy [2007] and Shaw, Gupta et al. [2005] for studies using the quit rate alone). We calculated separations from employee records for the sales assistants working fewer than 30 hours per week (defined as part-time separations) and 30 or more hours per week (defined as full-time separations), as described in the Theory section above. (We also conducted a sensitivity test, using 20 hours a week as the cut-off; see below.) Although the records do not reveal the reason for worker separation, our measure of separations included few terminations because it is company policy not to terminate.

We calculated the full-time equivalent separation rate as the number of hours that leavers would have worked had they not left, relative to annual total hours worked. (Here we used our daily data on when workers left and were hired; but we must emphasize that all our variables are in annual terms, so seasonality is not an issue.) Hence, if someone separated at the very end of a year, the impact of his/her leaving on labor productivity in that year would be zero, and the converse would apply if he or she left at the beginning of the year. Table 1 gives an example of our calculation. Thus, our measure of separations allows both for their real impact in terms of hours, and for their timing. The squares of the part-time and full-time separation rates, as well as their cross-product, were included in the regression equations testing our hypotheses.

Table 1 shows how our full-time equivalent measure of separations differs from the usual headcount measure of separations—that is, the number of leavers during the year divided by the average number employed (e.g., Glebbeek & Bax, 2004; Shaw, Gupta et al., 2005). For example, if a full-timer leaves at the end of the year on the December 1, only 160 hours are lost, which represents a turnover rate of about 1 percent for the average store using 14,000 hours (seven full-time equivalent workers) of labor. In headcount terms, however, the turnover rate would be one-eighth, or 12.5 percent. The main point is that the same headcount separation rate will be observed under quite different scenarios of full-time-equivalent turnover. This error in measuring turnover by headcount separation rate may result in a large standard error of its estimate in a regression equation, which will lower statistical significance. We show below that the headcount measure of separations indeed proves to be insignificant.

In calculating both full- and part-time separation rates, we preferred to use total annual hours worked in the store as the denominator. This procedure ensures that separation rates are bounded between zero and one. An alternative would be to express full-time separations as a percentage of full-time hours worked, and equivalently for part-time, but then separation rates are much more variable. We discuss results using these alternative measures below and show that they are consistent with our preferred specification.

We also had information on hires, which we used to check the robustness of our results. In particular, we calculated the rate of full-time-equivalent hires as the ratio of the hours worked by employees hired in a given year to the total hours worked. Although little research has addressed the effect of hires on performance (see Bingley and Westergaard-Nielsen [2004] for an exception), it is worth including hires as a possible confound for the effect of separations. Because hires are normally closely connected with separations, separations might be picking up the possible effect of hires—that is, the benefits from invigorating the workforce and the costs of training and integrating new workers. However, we found that the inclusion of hires along with separations in equations had little consequence for the regression results for separations, as is shown by the robustness checks we report below.

Controls. Labor productivity is fundamentally determined by capital and labor inputs. We measured capital input as store size in square feet, and labor input, as the sum of hours worked by every sales assistant employed in a given store in a given year. In addition, we included a number of variables relating to store environment and employee characteristics that might also affect productivity. The store environment variables were store location (eight dummies, including "city center" and "retail park"), type of product (three dummies indicating more or less expensive goods), share of children's goods, number of floors, and area wealth and unemployment. By including these variables, we aimed to control for the fact that it is easier to sell in prime locations, and sales volumes may vary with type of product, store configuration, and customer target group.

As regards employee characteristics, we include sales assistants' weekly hours (shares of employees working 0–4, 5–14, 15–29, and 30+ hours per week), which determine labor flexibility, important in the retail environment. We also included the relative wage (sales assistants' pay relative to county average), and sales assistants' average age and tenure, which helped control for workforce quality. We used full-time equivalents for average age and tenure as well as employee turnover. (This adjustment is important since it is younger workers with shorter tenures who are more likely to separate.) Finally, we included 20 regional manager dummy variables to control for possible effects of regional management on store productivity.

Estimation Issues

We tested our hypotheses by regressing our measure of store performance (labor productivity) on part- and full-time separation rates, their squares and cross-product, and the control variables outlined above. At the same time, the consistency of our estimation results, and thus our judgment of Hypotheses 1 and 2, could be compromised by the presence of store-specific unmeasured factors, such as management ability and store location. These factors are hard to measure and, in any case, the data were not available to us. Such store-specific "unobservables" would cause omitted variable bias if correlated with the regression variables (see Halaby [2004] on these "problematic" error structures). For example, simple economic reasoning implies (e.g., Lucas, 1978) that better managers are likely to run bigger stores and employ more labor. Thus, not controlling for management ability will give an overestimate of the effect of labor input on sales. Furthermore, if labor input is correlated with turnover, not controlling for store management would bias the estimates for turnover as well. Indeed, as we show in the next subsection, even simple correlations between pairs of our regression variables differ depending on whether we calculate them within-store (thus excluding the unobservables) or between-store; this variation underlines the need to control for the unobservables.

Fortunately, the problem of unobservables, common to all studies of employee turnover using cross-sectional data, can largely be overcome by using within-store variation in productivity and turnover available from our panel data, as we now explain. Our regression equation could be written as follows:

$$\begin{aligned} &\textit{Labor productivity}_{it} = \textit{regression variables}_{it} \\ &\qquad \times \textit{regression coefficients} + \\ &\qquad \underbrace{\underline{\textit{store-specific error}_{i} + \textit{ordinary error}_{it}.}}_{\text{the total error term}} \end{aligned}$$

Here, i is the index for the store and t for the year. In our data, i ranges from 1 to 325, and t from 1995 to 1999. Note that the store-specific error (caused by the unobservables) has only index i, because we assumed that the unobservables vary by store but not by year. Both the total and the ordinary error terms are meant to follow the standard assumption of no correlation with the regression variables, which is required for regression identification.

Estimating the productivity equation correctly required separate identification of the unobservables and the ordinary error term. The ordinary least squares (OLS) estimator treats the two separate error terms as one, and thus it might produce inconsistent estimates if the store-specific unobservable were correlated with any of the regressors (see above). A practical solution to this potential problem was to apply a fixed-effects estimator, which works by taking store averages from all variables in the regression equation and applying OLS to the transformed equation. As a result of this transformation, the regression equation loses the store-specific error term (as well as other time-invariant variables) while giving the same estimates to the remaining (time-varying) variables as they would have in the true regression model. (For an example, see Capelli and Neumark's [2001: 256] comparison of OLS and fixed-effects estimates; also relevant is Guest, Michie, Conway, and Sheehan's [2003] finding that the link between high-performance HRM practices and firm productivity disappears when longitudinal rather than cross-sectional regression techniques are applied.) Another way of looking at the fixed-effects procedure is that it produces the "within-store" effects, relying only on within-store variation in the regression variables over time. Thus, fixed-effects results should be interpreted as pertaining to the "average" store.

The fixed-effects estimates were suitable for testing our hypotheses regardless of whether there were unobservables or not. However, fixed effects might not always be the most efficient estimator. Lacking unobservables, OLS would be more efficient, since it would not require the estimation of 325 store averages. Moreover, even given unobservables, so long as they were not correlated with the regressors, the random-effects estimator was preferred. Random-effects estimation, which treats unobservables and the ordinary error term separately but requires that neither be correlated with the regressors, would allow use of the between-store regression, through the 325 store averages, to estimate the variance of store-specific error (Wooldridge, 2002: 261). In fact, the random-effects estimates are a type of average of the fixed-effects and between-store estimates. Thus, random-effects had an advantage over OLS when unobservables were not correlated with regressors. Furthermore, random effects had advantages over fixed effects since random effects did not require the estimation of 325 store averages, and it also returned estimates of the time-invariant variables, such as store space.

The question therefore arose: How do we choose between OLS, random-effects, and fixed-effects estimation? A sequence of statistical tests clarified this choice. The first test compared the goodness of fit of OLS and random-effects regressions. If OLS was preferred, it was the best estimator of the three. However, if (as was the case), OLS was rejected in favor of random effects, the next test was the standard Hausman (1978) test between random effects and fixed effects. The acceptance of the former would then indicate that the unobservables, although present, were not correlated with the regressors. However, if—as was the case—fixed-effects estimation was preferred, the conclusion is that the unobservables are correlated with the regressors, and therefore fixed effects is the only reliable estimator of the turnover-performance relationship.

RESULTS

Basic Statistics

Table 2 reports key descriptive statistics for the main variables. We present simple correlations both between and within stores. Between-store correlations are the correlations between store averages. Within-store correlations, like the fixed-effects estimator, are derived from the "demeaned" variables so that they show correlations within an average store. Although our study required a multivariate analysis to produce conclusive results, the simple between- and within-store correlations are instructive as a first step.

Figure 3 illustrates why the within-store and between-store correlations can differ, using an example of productivity and full-time separations for two stores. The scatter of points for stores 1 and 2 are illustrated as both sloping upward. A within-store regression is designed to go through both scatters and is consequently positive, corresponding to a positive within-store correlation, as is actually the case, .08. (The shift in the regression line intercept between stores 1 and 2 is the fixed effect.) However, a between-store regression would simply go through the means for each store (the heavy circles), and is negative as drawn—corresponding to a negative correlation, again as is actually the case, −.24. It seems that some unobservable such as poor management is causing store 2 to have both persistently low productivity and high full-time separations. An OLS regression is a mixture of within and between effects, and ends up quite flat (heavy dashed line), corresponding to the approximately zero overall correlation between productivity and full-time separations. This flat line is mis-

TABLE 2
Means, Standard Deviations, and Pairwise Correlations[a]

	Mean	s.d.	1	2	3	4	5	6	7	8	9
Between store (n = 325)											
1. Sales per hour worked	52.8	14.2									
2. Total hours worked	14,642	11,160	−.30								
3. Separation rate, full-timers[b]	0.05	0.08	−.24	.37							
4. Separation rate, part-timers[b]	0.08	0.07	−.02	.07	.09						
5. Average age in years	31.0	7.1	.22	−.55	−.50	−.28					
6. Average tenure in years	5.30	2.70	.15	−.30	−.44	−.34	.63				
7. Store wage relative to county average wage for sales assistants	0.82	0.08	.18	.01	−.17	−.08	.15	.21			
8. County unemployment rate (%)	6.8	2.5	−.02	.17	.04	−.13	−.09	.05	.31		
9. Store employees working less than five hours weekly (%)	17.0	16.0	.06	.05	−.14	.12	.00	−.01	.04	−.17	
10. Space in square feet	1,600.0	878.0	−.12	.80	.21	.10	−.38	−.20	.01	.07	.16
Within store (n = 1,625)											
2. Total hours worked			−.73								
3. Separation rate, full-timers[b]			.08	−.08							
4. Separation rate, part-timers[b]			.30	−.28	.03						
5. Average age in years			.30	−.24	.01	.01					
6. Average tenure in years			.34	−.14	−.01	.05	.44				
7. Store wage relative to county average wage for sales assistants			.01	.01	−.01	.01	.01	.02			
8. County unemployment rate (%)			−.39	.04	−.08	−.15	−.19	−.39	−.04		
9. Store employees working more than five hours weekly (%)			.27	−.06	−.02	.11	.16	.20	.06	−.33	

[a] Means and standard deviations (s.d.) are for the overall sample (all store-years, n = 1,625).

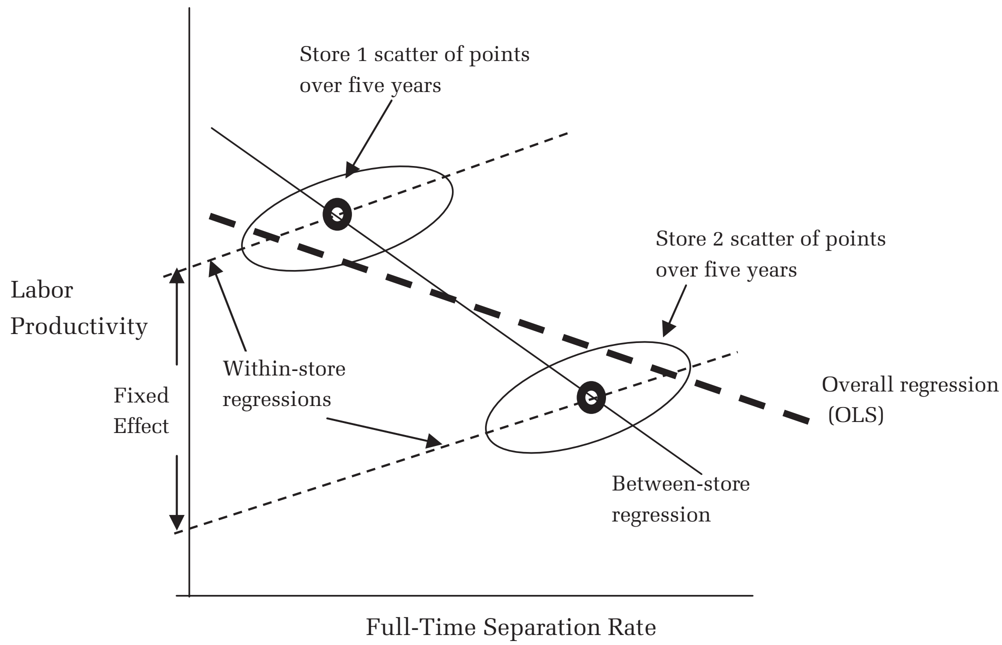
[b] The separation rate was calculated as the full-time-equivalent value. Full-timers, defined as working 30+ hours per week, comprised 20 percent of the workforce. The average headcount separation rate for full-timers was 0.03 (s.d. = 0.05), and it was 0.21 for part-timers (s.d. = 0.12). The total headcount separation rate is the weighted sum of the two: $0.03 \times 0.2 + 0.21 \times 0.8 = 0.174$. For comparison, the average headcount hiring rate for full-timers is 0.03 (s.d. = 0.05), and 0.23 for part-timers (s.d. = 0.14).

leading. Thus, we argue that it was within-store statistics, which are not affected by unobserved factors such as management ability, that we should best use for testing our hypotheses.

Table 2 shows differences for several variables in the within- and between-store correlations. A further example is the correlation between productivity and part-time separations, which is close to zero in the between-store matrix, yet positive (.30) in the within-store array. We do not intend to stress these particular values, since they are only simple correlations. But they show the conflicting signals with which central management planners in a large organization have to deal. Again, the lesson is that care needs to be taken when comparing stores differing in many unmeasured ways, and hence it is safest to build on what is happening within stores, which requires controlling for the unobservables.

Noteworthy features that emerge from the averages in Table 2 are, first, that the full-time-equivalent turnover rates are quite low: the average store loses only 13 percent of hours yearly as a result of separations (5% full-timers + 8% part-timers = 13%). The headcount separation rate is higher, at 24 percent (3% full-timers + 21% part-timers). Second, the company's pay schedule gives a relatively low wage on average, only 82 percent of the prevailing wage for sales assistants in the county. Also, as a result of the company's simple pay structure, the wage is quite uniform across the country, and not very well matched to local rates, causing considerable variability in relative wages (the relative wage in some stores goes down to 60 percent). As noted above, relatively low pay rates push the separation rate upward, which is to some extent therefore a company choice. Finally, the proportion of very part-time workers, those working fewer than five hours a week, is high, 17 percent. Such extreme part-time work shows the unskilled nature of the job, and the turbulence of the retail environment, since alert managers will rely on such part-timers to meet the "right people, right time" company management rule. On the other hand, this flexible workforce is combined with a sizeable core

FIGURE 3
Illustration of Between-Store and Within-Store Correlations for Labor Productivity and Separation Rate



of full-timers, 20 percent. This full-time core is quite stable, with an only 5 percent a year separation rate. We therefore expected turnover among the full-timers to have a qualitatively different effect on store performance.

Regression Analyses

Table 3 shows the results of our regression analysis for labor productivity. The first column gives the simple OLS specification and provides a starting point. The immediate point to note is the overall lack of significance of the coefficients on our measures of turnover. Only the separation rate for part-timers is significant at the 10 percent level (0.26), and this counterintuitive estimate simply suggests that higher turnover brings higher performance. If we were to stop here, Hypothesis 1, concerning the inverted U-shaped effect of turnover for secondary workers (as measured by part-timers) would not be confirmed. Nor would Hypothesis 2, concerning core workers (as measured by full-timers) receive any backing.

However, the random-effects specification in the second column tells quite a different story. It shows a definite inverted U-shaped turnover-performance relationship for the part-timers. As can be seen, the part-timer separation rate is now significantly positive (0.23), and the squared term significantly negative (−0.59). The results also show a negative effect for the full-timers operating through the cross-product term, −0.81. Therefore, both our hypotheses begin to receive support as we account for the store-specific unobservables by applying the random-effects method. Indeed, as the first of the diagnostic tests in Table 3 shows, we have to allow for store-specific unobservables. With the share of variance of unobservables accounting for 0.79 of total error variance in the random-effects specification (last row of Table 3), controlling for unobservables improves regression fit dramatically.

We still have to further investigate the role of these important store-specific unobservables. For this we turn to the fixed-effects results in the third column of Table 3. These estimates for the turnover rates are similar to the random-effects estimates. The positive (and significant) coefficients on part-time separations (0.19) and negative part-time separations squared (−0.65) give rise to a clear inverted U-shape for part-timers. At the same time, insignificant full-time separations and the same squared plus a significant negative cross-product term (−0.71) give the conventional negative turnover-performance link for full-timers. Moreover, as noted in the estimation issues subsection, the fixed-effects (within-store) technique is not biased by any correlation between unobservables and the regressors. As the second of our diagnostic tests (Table 3, test 2, reported in the footnote) showed, unobservables were certainly correlated with some of the regressors. Hence, we were compelled to reject random effects in favor of fixed effects, de-

TABLE 3
Results of Regression Analyses for Labor Productivity

Variables[b]	Estimator[a]: Ordinary Least Squares	Random Effects	Fixed Effects
Time-varying			
Total hours worked[b]	−0.34**	−0.70**	−0.82**
Separation rate, full-timers	0.06	0.06	0.05
Separation rate, full-timers, squared	0.37	0.13	0.08
Separation rate, part-timers	0.26†	0.23**	0.19**
Separation rate, part-timers, squared	−0.43	−0.59**	−0.65**
Separation rate, full-timers, × separation rate, part-timers	−1.11	−0.81†	−0.71†
Average age	0.00	0.00	0.00
Average tenure	0.01	0.02**	0.02**
Average tenure squared	0.00	−0.001**	−0.001**
Share of male employees	0.01	0.04	0.05
Share of employees working less than 5 hours weekly	0.30**	0.05	0.04
Share of employees working less than 15 hours weekly	0.30**	0.08	0.07
Share of employees working less than 30 hours weekly (compared to 30+)	0.13*	−0.01	0.03
Hours reduction 20% in year over	0.03†	0.03**	0.02*
Relative wage	0.09†	−0.01	−0.02
County average pay	−0.15*	−0.14	−0.19*
County unemployment rate	−0.78**	−0.18	0.01
Year dummies (compared to 1999)	Up to −0.16**	Up to −0.17**	Up to −0.17**
Time-invariant			
Space	0.23**	0.45**	
Type of product (3 dummies)	Insignificant	Up to −0.12**	
Location (8 dummies)	−0.18** to 0.22**	−0.32** to 0.52**	
Share of children's goods	−0.39**	−0.47**	
Number of floors (compared to one)	Up to −0.17**	Up to −0.24**	
Regional manager (20 dummies)	−0.09* to 0.25**	−0.13† to 0.24**	
R^2 (within)		0.77	0.77
R^2 (between)		0.26	0.09
R^2 (overall)	0.43	0.39	0.13
Fraction of error variance due to store-specific unobservables	Not applicable	0.79	0.98

[a] (1) Comparing random effects with OLS, we obtained a likelihood-ratio test statistic equal to 1,444.1 distributed as chi-square with 1 degree of freedom, and the 1% significance level is 6.63, much smaller than 1,444.1, implying that random effects has significantly more explanatory power than OLS. In fact, unobservables have a large variance equal to 0.79 of the total variance. (2) Comparing random effects with fixed effects, we obtained a Hausman (1978) specification test result of 178.7 distributed as chi-square with 21 degrees of freedom, and the 1% significance level is 38.9, much smaller than 178.68, implying that the unobservables are correlated with the regressors. Thus, random effects was inconsistent and we based our conclusions on fixed effects. Stata was used to generate estimates.

[b] Total hours worked, relative wage, county average pay, and space were logged. Separation rates were calculated as full-time-equivalent values.

† $p < .10$
* $p < .05$
** $P < .01$

spite the convenience of the random-effects estimates of the time-invariant variables. Note that unobservables account for 0.98 of total error variance in the fixed-effects specification (last row of Table 3), an increase over the random-effects model. This increase was expected, since the variance explained by the time-invariant variables (e.g., store space) no longer contributed to our equation's explanatory power but instead was subsumed in the fixed effects.

The Unobservables

Turning back to the OLS estimates, why are they so different from both the random- and fixed-effects ones? The reason is that the effects of store-specific unobservables on performance are caught up by the store-specific error term and swept out by the fixed-effects technique, whereas the OLS method does no such sweeping out. Let us briefly compare the OLS and fixed-effects estimates to explain better how

TABLE 4
Results of Regression Analyses for Full-Time-Equivalent and Headcount Measures of Turnover Explaining Labor Productivity

Variables	Full-Time Equivalent Measures: Separations, Full-Timers and Part-Timers	Full-Time Equivalent Measures: Total Separations	Full-Time Equivalent Measures: Total Separations and Hires	Headcount Measures: Separations, Full-Timers and Part-Timers	Headcount Measures: Total Separations	Headcount Measures: Total Separations and Hires
Separations, full-timers	0.05			−0.12		
Separations, part-timers	0.19**			−0.01		
Separations, full-timers, squared	0.08			0.98†		
Separations, part-timers, squared	−0.65**			0.01		
Separations, full-timers, × separations, part-timers	−0.71†			−0.15		
Separations, total		0.14**	0.11†		−0.02	−0.03
Separations, total, squared		−0.36**	−0.35**		0.01	−0.11
Hires, total			0.10			−0.02
Hires total, squared			−0.47*			−0.16
Separations, total, × hires, total			0.22			0.33*
Optimum levels						
Separations, full-timers	0.00			Not defined		
Separations, part-timers	0.15			Not defined		
Total separations		0.20	0.21		Not defined	Not defined
Total hires			0.16			Not defined

[a] All these equations use the same fixed-effects specification as in the last column of Table 3, only varying the measure of separations and hires. Optimum values of separations and hires are the values that maximixe labor productivity.

† $p < .10$

* $p < .05$

** $p < .01$

unobservables can confound inferences from regression results.

In Table 3, we see that labor input has a coefficient of −0.34 in the OLS equation, but a much larger negative coefficient in the fixed-effects equation, −0.82. The latter shows the diminishing returns that an average store manager faces when employing more labor. The OLS coefficient, on the other hand, picks up the between-store estimate for labor input, which shows returns diminishing less because some unobservable factor, such as abler management, is boosting sales in bigger stores.

Several variables are significant in the OLS equation but not in the fixed-effects one. Thus, the hours composition variables are significant in the OLS equation (0.30 for the share of employees working fewer than 15 hours per week, 0.13 for those working 15–29 hours), but are small and insignificant in the fixed-effects equation. The OLS coefficient is picking up the fact that only stores with a high unobservable (presumably those having a good manager) can maintain an effective pool of part-timers and fine-tune their working hours to achieve the best returns to labor. Another example is the county unemployment rate: stores located in areas with high unemployment (OLS estimate, −0.78) find it harder to sell. Here, the fixed-effects estimate is again insignificant, this time probably because we did not have enough within-store variation in unemployment over time to establish its relationship with labor productivity.

To return to our focus on the turnover-performance link, the insignificance of the OLS coefficients on separations indicates that the store-specific unobservables are working against the inverted U-shaped relationship. It might be that stores with good unobserved management or location have both higher turnover and higher labor productivity than stores with bad.

Sensitivity Tests

We now investigate the sensitivity of our results to alternative definitions of turnover, and also the causality of the turnover-performance relationship. Table 4 reports the results for various measures of turnover. The first column repeats the fixed-effects results from Table 3 for convenience and gives our estimated optimum separation rates for full-timers (0), and part-timers (0.15). The next two columns use full-time-equivalent total separations (part-time and full-time together), and then both total separations and hires. In the final three columns, we give parallel results using headcount measures of turnover.

The first column gives results for the turnover-productivity relationship contingent on type of work system, which we can compare with the results in the second column obtained under a no-contingency assumption. The contingency model fares better. If we restrict the separation rate and its square to be equal for the part- and full-timers, the test result (not shown) allows rejection of the restriction at the 5 percent significance level, thus empirically justifying our contingency view of the turnover-performance link. Moreover, the inverted U-shaped relation shown in the second (no-contingency) column has smaller coefficients (0.14 for total separations and −0.36 for the squared term): it is flattened. This result is also reasonable, since the second column runs together full-timers, for whom there is no inverted U, with part-timers for whom there is, thus blurring the results. A similar blurring (not shown) emerges when we identify the core group as workers working more than 20 hours, rather than 30 hours. Thus, our preferred 30-hour definition for the core group of full-timers, in conjunction with treating those working fewer hours as part-time/secondary, provides good differentiation, and the data support our contingency model.

Turning to the third column, which includes hires, we see a negative and significant effect of the full-time-equivalent hiring rate squared in the third column (−0.47). This result is consistent with the earlier finding on the negative effect of hires reported in Bingley and Westergaard-Nielsen (2004: 561), and it shows how recruitment and training costs rise as hiring rates increase. We experimented (results not shown) with making the hires-performance relationship also contingent on the core/secondary distinction, and the coefficients on separations did not change much.

The next columns show how the situation with the inverted U-shaped association and the contingency model changes dramatically if we use headcount turnover. The estimates lose significance, and we are no longer able to establish either the inverted U or the contingency model, since the only significant results suggest a spurious positive, and increasing, effect of full-timer separations on productivity (0.98). These difficulties persist for all three specifications, a consequence of measurement error in the headcount rates. Our results thus suggest that the turnover-performance relationship requires careful measurement of turnover.

Another measure of separations is also possible, as we noted above—that is, hours lost because of separations of part-timers relative to total part-time hours (rather than total full-time plus part-time hours), with the corresponding definition for full-timers. We tested for changes using this measure and found results to be largely unaffected. Specifically, we estimated the optimum separation rate for part-timers to be 0.25 as a proportion of part-time hours worked, which when measured in terms of total hours worked is 0.20 (= 0.25 × 0.80, where 0.8 is the proportion of part-timers). This optimum is statistically indistinguishable from the optimum of 0.15 in our preferred model shown in Table 4.

Finally, there is the issue of causality. We used a procedure resembling a Granger causality test (the proper test could not be applied because of the short panel time series of only five years). In this test, we compared the explanatory power of lagged separations on productivity (i.e., separations cause productivity being our preferred concept), with a reverse set-up, in which lagged productivity is on the right-hand side of the equation, explaining separations. Glebbeek and Bax (2004) used a similar approach. We found that our preferred set-up, with separations causing productivity, performed much better than the reverse. Thus, lags of part- and full-time separations, their squares, and the cross-product term affect productivity in a significant way: their estimates are 0.26, 0.21, −0.94, −0.25, and 1.26, respectively. On the other hand, lagged productivity, although significant statistically, had a very minor economic impact on separations. Our estimates imply that a large, 10 percent change in the last year's productivity is correlated with only 0.3 and 0.4 percent reductions in the current year's part- and full-time separations, respectively. Thus, like Glebbeek and Bax (2004), we concluded that turnover indeed influences labor productivity, and not the other way around.

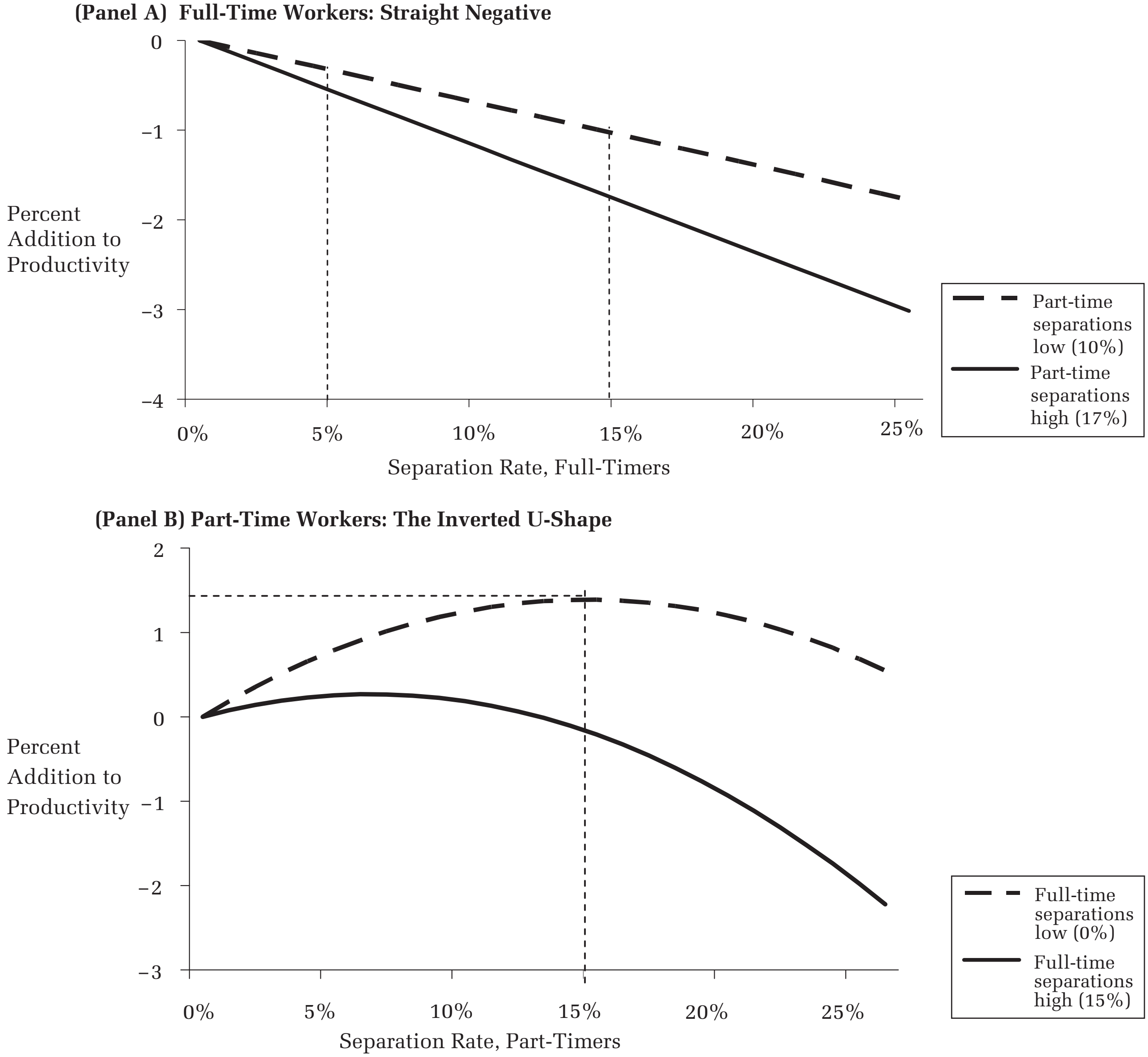
DISCUSSION

The Straight Negative and the Inverted U-Shape

We have obtained two sets of regression results for the turnover-performance link: one—negative—for full-timers, who we maintain are managed under a commitment work system, and the other—an inverted U-shape—for part-timers, managed under a secondary system. The two work systems appear to coexist for rank-and-file workers in our organization. Let us discuss the consequences of employee turnover for labor productivity under each system in turn. We will then look at the way the two systems interact. Finally, we provide some calculations of the magnitude of the effects.

For workers managed under a secondary work system, we find a clear inverted U-shaped relation-

FIGURE 4
Linked Effects of Core and Secondary Worker Turnover on Performance[a]



[a] Separation rates were calculated as full-time-equivalent values.

ship between turnover and performance. Figure 4, panel B, which is based on the part-timer's coefficients from Table 3's fixed-effects column, illustrates this result. The inverted-U finding supports Abelson and Baysinger's (1984) theory that turnover has benefits and costs to firms and that initially the benefits of turnover outweigh the costs. This theory has also received some support in Arthur (1994: 683), Guthrie (2001: 186), and Glebbeek and Bax (2004: 285), but we now provide a more solid empirical foundation. The initial positive impact of part-time separations on productivity implies that less productive workers are more likely to separate (consistently with findings reported in McEvoy and Cascio [1987] and Williams and Livingstone [1994]). Another benefit of turnover is particularly important for our organization: labor flexibility in a turbulent retail environment (Figure 2). However, there comes a point at which the benefits of improved job-worker match and workforce flexibility become offset by dysfunctional turnover and the loss of firm-specific human capital. So, our Hypothesis 1 receives empirical support.

As for employees managed under a commitment work system, the link is purely negative. Figure 4's panel A illustrates. Although according to our theory, there must be an optimum nonzero level of turnover for each work system, this level is likely to be small here, so taking zero as optimum is an acceptable simplification. A number of the earlier across-organization studies cited in the introduction have found this type of link, although they have not allowed for differentiation in work systems within an organization. The economywide studies of turnover (Bingley & Westergaard-Nielsen, 2004; Harris et al., 2006) face even more difficulty in providing for contingency. As Table 3 shows, although the estimates for the full-time separation term and its quadratic term are insignificant, the negative effect (−0.71) is manifested through a cross-product term. Thus, as we theorized earlier, the costs of core worker turnover appear to be higher than the costs of secondary worker turnover, with few benefits. Hence, the negative turnover-performance link fits, bearing out our Hypothesis 2.

Interactions

The way in which full-time separations affect labor productivity merits further discussion. It appears that the two work systems interact, one affecting the outcomes of the other. The significant cross-product term implies that the effect of full-time separations is exacerbated by secondary turnover, and conversely. We have given the theoretical grounds for expecting this type of interaction above; for example, losing full-timers at a high rate means that the new part-time hires are trained less effectively, raising the costs of part-time separations.

Figure 4 shows these effects. For example, in panel A, holding part-time separations at 10 percent, which is about average, the effect of the average full-time separation rate of 5 percent is to reduce productivity by about 0.35 percent ($= -0.71 \times 0.1 \times 0.05$) on the dashed line. Increasing full-time separations to 15 percent, which is about a one standard deviation increase, would move us down the dashed line to give a loss of around 1 percent. However, losses from full-time separations would be worse, about 1.5 percent, if part-time separations themselves increased by one standard deviation to 17 percent, as shown by the solid line.

Analogously, panel B of Figure 4 illustrates how the inverted U-shaped relationship depends on the level of full-time turnover, giving two examples. If full-time separations are zero (their optimum level), we trace out the dashed top curve, with an optimum secondary worker turnover of about 15 percent (more than twice this figure in headcount, given that an average worker works less than half time). As can be seen, such a level of turnover increases productivity by about 1.4 percent. However, if full-time separations are at a high level (15 percent is illustrated, giving the solid line) the beneficial effects of secondary turnover are almost lost, though the curve still has quite a long flat region, to about 10 percent, where secondary worker turnover does not seriously damage performance. Still, panel B illustrates that for stores where core worker turnover is high, managers must be careful not to allow secondary worker turnover to rise much as well.

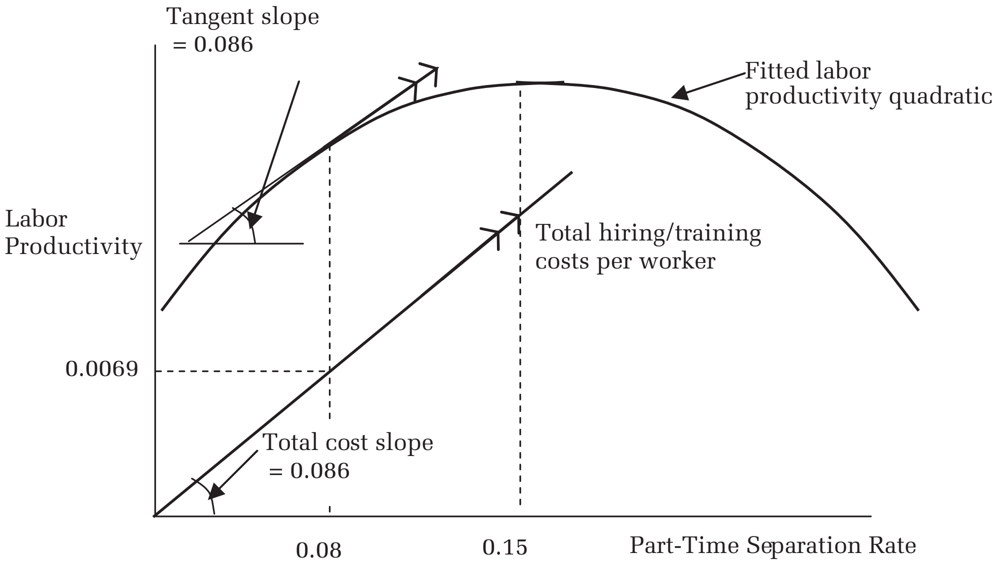
Magnitude of Effects

Let us make some back-of-the-envelope calculations about the effects of turnover on productivity. For full-time turnover, we find no benefits, only losses to productivity—so that the optimum level of full-time turnover is zero. We calculate that if all the stores operated at the optimum level of full-time turnover instead of at their observed levels, the organization would gain 0.3 percent of total sales over 1995–99, or £4.2 million.

As for part-time turnover, Figure 4's panel B shows that our organization can gain up to 1.4 percent in productivity from choosing the optimum levels of full- and part-time turnover. This 1.4 percent gain translates into £0.73 per each hour worked per year, on average. For some stores, much greater gains are possible from reducing part-time turnover. For example, taking the store with the highest secondary turnover rate (0.80), we calculate that reducing it to optimal would result in a 27.8 percent gain in productivity. Finally, summing up individual productivity gains from moving to optimal part-time turnover for all stores and years, we find that the overall gain for the organization would be about 0.6 percent of total sales over the period 1995–99, or £7.5 million.

It is true that our estimated optimum rate of part-time turnover of 15 percent is higher than its observed average of 8 percent (Table 2), but this difference can be explained by hiring and training costs for the new hires who replace the separated workers. Figure 5 illustrates this position. If profit maximization is assumed, the marginal revenue product and marginal costs of part-time separations are equal at their observed average of 0.08. Using our regression estimates, the fitted labor productivity quadratic curve has a slope of 0.086 ($= 0.19 - 2 \times 0.65 \times 0.08$) at a part-time separation (= hire) rate of 0.08. Then, assuming that the costs of hiring and training are linear, the slopes of the hiring-training costs and labor productivity curves will be the same at this point. Hence, the slope of the hiring-training costs curve will equal 0.086. There-

FIGURE 5
Calculation of Hiring and Training Costs for an Average Store[a]



[a] Labor productivity and total hiring and training costs were logged. We observed a part-time separation rate of 0.08. Our fitted labor productivity quadratic curve has a slope at this point of 0.086. If the part-time separation rate (= hiring rate) of 0.08 maximizes profits, then the slopes of the (log) total hiring/training costs and labor productivity curves will be the same at this point. Hence the slope of the total hiring/training costs curve will equal 0.086. Therefore, when the part-time separation rate is 0.08, total hiring/training costs per worker will equal 0.0069 (= 0.08 × 0.086), or 0.69% of annual output per worker.

fore, when the part-time separation rate is 0.08, total hiring-training costs per worker will equal 0.08 times 0.086, or 0.69 percent of output per worker.

The conclusion of this reasoning gives the plausible result that for an average store with sales of £773,098 (= 52.8 £/hour × 14,642 hours) per year and hiring seven people, hiring-training costs per newly hired sales assistant are £762 (= £773098 × 0.0069/7). Contemporary practitioner estimates (Institute of Personnel and Development, 1997: Table 6) of hiring/training costs for unskilled labor are similar, £840. Thus, our estimate is reasonably close to an alternative, independently derived estimate, which gives confidence in the statistical results.

Conclusions

Our findings suggest that the turnover-performance relationship is contingent on the nature of an organization's work system, and indeed that commitment and secondary work systems can coexist within the rank-and-file workforce of an organization. Thus, our study provides empirical evidence for a number of claims made by Lepak and Snell: that not all employee groups are equally unique and valuable to a company and that current strategic HRM research may be "overly simplistic" (1999: 45). Although researchers have generally accepted a distinction between work systems appropriate for management and for rank-and-file workers, our results make it clear that in organizations with many nonstandard (part-time or contract) workers, there may well be considerable differences in work systems within the rank-and-file workforce itself.

Our organization-level study in fact calls into question the notion that an organization's labor management systems can be characterized in any simple way as high-performance, because they show that different sections of the workforce may be best managed in different ways, with low-performance strategies used simultaneously. Hence, although some across-organization studies (Arthur, 1994; Guthrie, 2001) have supported the contingency model, and others have not (Shaw, Gupta et al., 2005), there are problems in attributing just one HR system to an organization. (They have also used simple cross-sections.) More differentiation is needed, such as is indeed forthcoming in new lines of strategic HRM research, for example in Lepak et al.'s (2007) contingency approach to examining the impact of high-investment HRM systems on core versus support workers.

We identified core and secondary groups within our organization's sales assistant workforce to which different turnover-performance links apply. For the core group of full-timers, and only for this group, we found the negative turnover-performance link that is conventionally accepted

(Huselid, 1995: 651). For the remainder, we found a robust inverted U-shaped relationship, with an optimum turnover rate of 15 percent a year in full-time-equivalent terms (the rate is more than twice this size in headcount terms). We believe we are the first study to uncover the inverted U-shaped association in retail, and certainly replication using our accurate measurement methods—including panel data—is needed in different sectors and different time periods. Nevertheless, the inverted U-shape finding is reasonable, given the unskilled nature of most sales assistant jobs, and the concomitant low hiring-training costs (as suggested by our £762 per hire estimate) and low losses in productivity via separations, at least over most of the observed separation range. Further, when we consider the large seasonal variation in required retail labor, and also the impact of the business cycle, which can cause variations on the order of 50 percent in a store's labor input over a few years (Figure 2), we can appreciate the advantage of a mobile workforce.

These are good reasons to allow for an inverted U-shaped relationship. We therefore agree with Glebbeek and Bax (2004: 285) that researchers and textbook writers can no longer assume that a lower rate of turnover is always preferable to a higher rate. It is also necessary to allow for work systems to differ between groups in an organization and to have correspondingly different optimal levels of turnover. We have calculated these levels for the two major groups here (Figure 4) and shown that they are in fact interdependent. Thus, as Arthur (1994) concluded, management of the employee turnover process must require careful practitioner attention.

We have reached some quite clear-cut results, but careful measurement has been needed to reveal the true relationships. First, we had to compute turnover in full-time equivalents. The regression results reported in Table 4 suggest that this adjustment is likely to be important for any organization that employs many part-timers or contract workers. Secondly, we have shown that the usual practice of simply comparing production units within an organization runs into difficulties, because of the unobservables problem. It is essential to track each unit over time, build up a panel, and work out the result for the average unit, by allowing each unit to have its own term, fixed or random as the case may be. Of course, in principle central management in a large organization such as the one studied here has such data at its fingertips, but it is tempting to cut corners by making simple comparisons between stores. Confusion will then result. A fundamental lesson for management decision making from our research is: Do not cut statistical corners, but methodically use all the data available from production operations.

REFERENCES

Abelson, M., & Baysinger, B. 1984. Optimal and dysfunctional turnover: Towards an organizational level model. ***Academy of Management Review,*** 9: 331–341.

Arthur, J. B. 1994. Effects of human resource systems on manufacturing performance and turnover. ***Academy of Management Journal,*** 37: 670–687.

Bamberger, P., & Meshoulam, I. 2000. ***Human resource strategy.*** Thousand Oaks, CA: Sage.

Bartel, A. P. 2004. Human resource management and organizational performance: Evidence from retail banking. ***Industrial and Labor Relations Review,*** 57: 181–203.

Batt, R. 2002. Managing customer services: Human resource practices, quit rates, and sales growth. ***Academy of Management Journal,*** 45: 587–597.

Batt, R., Colvin, A. J. S., & Keefe, J. 2002. Employee voice, human resource practices, and quit rates: Evidence from the telecommunications industry. ***Industrial and Labor Relations Review,*** 55: 573–594.

Becker, B., & Huselid, M. 2006. Strategic human resource management: Where do we go from here? ***Journal of Management,*** 32: 898–925.

Bingley, P., & Westergaard-Nielsen, N. 2004. Personnel policy and profit. ***Journal of Business Research,*** 57: 557–563.

Broschak, J., & Davis-Blake, A. 2006. Mixing standard work and nonstandard deals: The consequences of heterogeneity in employment arrangements. ***Academy of Management Journal,*** 49: 371–393.

Capelli, P., & Neumark, D. 2001. Do "high-performance" work practices improve establishment-level outcomes? ***Industrial and Labor Relations Review,*** 54: 737–775.

Cascio, W. 1998. ***Managing human resources: Productivity, quality of work life, profits.*** Boston: Irwin McGraw-Hill.

Dalton, D., & Todor, W. 1979. Turnover turned over: An expanded and positive perspective. ***Academy of Management Review,*** 4: 225–235.

Datta, D., Guthrie, J., & Wright, P. 2005. Human resource management and productivity: Does industry matter? ***Academy of Management Journal,*** 48: 135–145.

Dess, G. G., & Shaw, J. D. 2001. Voluntary turnover, social capital, and organizational performance. ***Academy of Management Review,*** 26: 446–456.

Economist. 2007. Accounting for good people. July 21: 73–75.

Glebbeek, A., & Bax, E. 2004. Is high employee turnover really harmful? An empirical test using company records. ***Academy of Management Journal,*** 47: 277–286.

Guest, D., Michie, J., Conway, N., & Sheehan, M. 2003. Human resource management and corporate performance in the UK. ***British Journal of Industrial Relations,*** 41: 291–314.

Guthrie, J. P. 2001. High-involvement work practices, turnover, and productivity: Evidence from New Zealand. ***Academy of Management Journal,*** 44: 180–190.

Halaby, C. 2004. Panel models in sociological research. In K. Cook & J. Hagan (Eds.), ***Annual review of sociology,*** vol. 30: 507–544. Palo Alto, CA: Annual Reviews.

Harris, M., Tang, K.-K., & Tseng, Y.-P. 2006. Employee turnover: Less is not necessarily more? In B. H. Baltagi (Ed.), ***Panel data econometrics. Theoretical contributions and empirical applications*** (Contributions to economic analysis, vol. 274): 327–350. Amsterdam: Elsevier Science.

Harrison, D., Virick, M., & William, S. 1996. Working without a net: Time, performance, and turnover under maximally contingent rewards. ***Journal of Applied Psychology,*** 81: 331–345.

Hausman, J. A. 1978. Specification tests in econometrics. ***Econometrica,*** 46: 1251–1271.

Huselid, M. A. 1995. The impact of human resource management practices on turnover, productivity, and corporate financial performance. ***Academy of Management Journal,*** 38: 635–672.

Institute of Personnel and Development. 1997. ***IPD Labour Turnover 1996 Survey results.*** London: IPD.

Jovanovic, B. 1979. Job matching and the theory of turnover. ***Journal of Political Economy,*** 87: 972–990.

Kacmar, M., Andrews, M., Van Rooy, D., Steilberg, C., & Cerrone, S. 2006. Sure everyone can be replaced . . . but at what cost? Turnover as a predictor of unit-level performance. ***Academy of Management Journal,*** 49: 133–144.

Lepak, D. P., & Snell, S. A. 1999. The strategic management of human capital: Determinants and implications of different relationships. ***Academy of Management Review,*** 24: 1–18.

Lepak, D., Taylor, M., Tekleab, A., Marrone, J., & Cohen, D. 2007. An examination of the use of high-investment human resource systems for core and support employees. ***Human Resource Management,*** 48: 223–246.

Lin, M.-J. 2005. Opening the black box: The internal labour markets of company X. ***Industrial Relations,*** 44: 659–706.

Lucas, R. E. 1978. On the size distribution of business firms. ***Bell Journal of Economics,*** 9: 508–523.

McElroy, J. C., Morrow, P. C., & Rude, S. N. 2001. Turnover and organizational performance: A comparative analysis of the effects of voluntary, involuntary and reduction-in-force turnover. ***Journal of Applied Psychology,*** 86: 1294–1299.

McEvoy, G., & Cascio, W. 1987. Do good or poor performers leave? A meta-analysis of the relationship between performance and turnover. ***Academy of Management Journal,*** 30: 744–762.

Morrow, P., & McElroy, J. 2007. Efficiency as a mediator in turnover-organizational performance. ***Human Relations,*** 60: 827–849.

Osterman, P. 1987. Choice of employment systems in internal labour markets. ***Industrial Relations,*** 26: 46–67.

Shaw, J. D., Delery, J. E., Jenkins, G. D., & Gupta, N. 1998. An organization-level analysis of voluntary and involuntary turnover. ***Academy of Management Journal,*** 41: 511–525.

Shaw, J. D., Duffy, M., Johnson, J., & Lockhart, D. 2005. Turnover, social capital losses and performance. ***Academy of Management Journal,*** 48: 594–606.

Shaw, J. D., Gupta, N., & Delery, J. E. 2005. Alternative conceptualizations of the relationship between voluntary turnover and organizational performance. ***Academy of Management Journal,*** 48: 50–68.

Williams, C. R., & Livingstone, L. P. 1994. Another look at the relationship between performance and voluntary turnover. ***Academy of Management Journal,*** 37: 269–298.

Wooldridge, J. M. 2002. ***Econometric analysis of cross-section and panel data.*** Cambridge, MA: MIT Press.

Youndt, M., Snell, S., Dean, J., & Lepak, D. 1996. Human resource management, manufacturing strategy and firm performance. ***Academy of Management Journal,*** 39: 836–866.

W. Stanley Siebert (w.s.siebert@bham.ac.uk) is a professor of labor economics in the School of Business at the University of Birmingham. He received his Ph.D. from the London School of Economics. His research interests include the microeconomics of labor management, as well as wider issues, particularly labor regulation.

Nikolay Zubanov (n.zubanov@cpb.nl) is a researcher at CPB Netherlands Bureau for Economic Policy Analysis. He received his Ph.D. from the University of Birmingham for empirical work in labor economics. His current research interests are management practices, human capital, innovation, and productivity.