Methods

Empirical Setting and Data Preparation

The empirical setting for our analyses was a Fortune 500 technology firm (TechCo) and our data spanned 2008 to 2015. TechCo was particularly well suited for studying the concepts of mobility and differential gender effects because they had recently announced a significant effort towards promoting diversity in all levels of their organization, citing diversity and inclusion as among the most important forces driving their organization's continued success. In particular, one of their key focus areas was increasing the representation of and ensuring equal pay for women in their workforce. Given this particular strategical bent, TechCo provided a conservative estimate of our results, as examining mobility and differential gender effects in such a research environment would have been a more stringent test of these effects.

We received personnel data from TechCo that comprised 4,270,115 monthly observations for 68,438 employees. Although TechCo was founded several decades ago, the personnel data started in January, 2008, and therefore represented a more recent snapshot of the company. In 2008, the organization consisted of 46,147 employees, and since then has fluctuated between 42,787 employees and 50,547 employees annually.

In order to prepare the data for analysis, we performed two major processing steps on the raw data. First, we limited the data to the population of workers who were exempt for at least a portion of their career and for whom we had standard pay grade data, which will be described in a later section. Exempt workers are traditionally salaried workers in engineering, management, and administrative positions[1]. This resulted in a condensed data set of 3,246,031 monthly observations for 53,311

[1] For what exactly constitutes exempt status, please refer to the Federal Fair Labor Standards Act

unique employees.

Second, since our data began in 2008, we met the traditional left-censored problem of panel data. Specifically, for the year 2008, there was a mix of employees who had already been employed at TechCo and were just beginning to be tracked, and those who had just joined TechCo. In order to distinguish between the two, we utilized the length of service field. However, for a number of potential reasons, the length of service field is not always close to zero when an employee starts at the company. For example, employees whose first observation was in the years 2009-2015, meaning that they had joined the company in that period or else they would have been seen in 2008, had a median starting length of service ranging from .74 to 1.15, depending on the year. Therefore, to determine which employees whose first observation is in the year 2008 actually started in 2008, we used a starting Length of Service threshold determined by the empirical cumulative distribution of starting length of services after 2008. We found that there was a sharp dropoff around 1.3 years, which encompassed roughly 85% of all starting length of services after 2008, and so use that as our cutoff to derive a full population of employees who started their careers at TechCo during the observation period[2]. We then used this population, which consisted of 18,322 employees (1,357 of which entered in 2008), to present a cleaner and more appropriate view of how employees started their careers at TechCo, and where they went from there.

Variables

Dependent Variable—Pay Grade

Our first main dependent variable was pay grade because pay outcomes are an unambiguously important indicator of success within a firm and evidence of pay differentials are of substantive

[2]We also tried a different threshold, determined by the 1st percentile of the distribution of all the job tenures for employees immediately prior to moving jobs. The intuition here is that an employee would still have starting characteristics for any starting length of service below that threshold. We find substantially similar results.

interest. Furthermore, we understand that many large organizations benchmark pay with one another, and, as a result, findings with respect to pay are inherently generalizable.

The pay grades in our data set ranged from two to twenty and were roughly linear in the sense that a higher pay grade generally made more than a lower pay grade. This was somewhat dependent on job family (i.e. technical versus non-technical)—a lower pay grade technical worker could have made more than a higher pay grade non-technical worker. Additionally, the pay grades were roughly non-overlapping, meaning that each pay grade represented a range of salaries, and the lower salaries in a higher pay grade would generally be higher than the higher salaries in a lower pay grade. However, there was more overlap the higher the pay grade, such that in the higher echelons a higher pay grade even within the same job family and function may not have translated to a higher actual salary. Despite this noise, the data were still quite meaningful as we could control down to the job title, and there were not a significant number of employees in the upper ranges of the scale, meaning that an increase in pay grade conceptually indicated a better outcome.

Finally, since pay grades were also used for other purposes, there were some pay grades outside of the two to twenty range. We were provided a translation chart and were thus able to translate those pay grades into the two to twenty range, but have additionally added an indicator for these non-standard pay grades in our analyses.

Dependent Variable—Performance

Our second main dependent variable was performance ratings. Performance ratings may be less straightforward as consequential outcomes within a firm, but are nonetheless important indicators of success, and are recognized to influence both pay and promotion opportunities, despite potentially unfair effects (Castilla, 2008).

We derived our dependent variable of performance from the performance ratings provided in the data set. We were told there were four meaningful distinctions: B, or below expectations, and I, or improvement required, constituted the lowest category of performance ratings; S, or successful, constituted the next; E, or excellent, constituted the next; and O, or outstanding, constituted the highest category.

Since a significant portion of the performance appraisals were in the successful category, we used stock share levels to further differentiate performance. Stock share levels ranged from one to five and were based on a relative performance decision, in line with but separate from the base performance ratings. Therefore, we constructed a linear scale of performance ratings plus stock share levels, such that a below expectations or improvement required with a stock share level of five made up the lowest category, and an excellent with a stock share level of one made up the highest category[3].

Dependent Variable—Exit

We additionally used exit in our supplementary models as a proxy for aggregated outcomes. To the extent that a series of career outcomes were beneficial or detrimental to an individual employee, we should see those career histories resulting in a certain kind of exit. In order to perform these analyses, we leveraged the granular exit data provided by the organization.

There were three fields in our data that we used to categorize exits: voluntary, involuntary or neutral; desired, undesired, or neutral; and high-level reasons for exit. Using these fields, we defined our first two forms of exit as all involuntary exits and all voluntary exits. We further defined two more granular forms of voluntary exit as voluntary and desired exit, and voluntary and undesired exit, where the desirability of an exit was evaluated from the perspective of the organization. Finally, we defined a form of voluntary exit considering the following specific reasons: dissatisfied with career

[3]Using only the performance rating without the stock share level does not substantially change our results.

prospects, dissatisfied with work conditions, dissatisfied with pay, and dissatisfied with supervisor.

Independent Variable—Erraticism

Our main independent variable was erraticism, which we operationalized in such a way as to capture the intricacies of atypical career paths. A common way to measure the typical career path and deviations from it is to use sequence analysis (Blair-Loy, 1999), but we found sequence analysis inappropriate for our theoretical conceptualization for a number of reasons. First, the sheer number of states we considered was not amenable to sequence analysis. Since the idea behind sequence analysis for careers in general is to cluster movement between states, it becomes more difficult and less meaningful when the number of states grows large; the number of distinct job titles in our data set was 363. Second, related to but distinct from the number of states, selecting the number of meaningful paths, or clusters, is a subjective decision. Although there are more analytical and data-guided ways to derive cleaner paths, determining the number of distinct paths in our setting would have been difficult. If we had selected too few clusters, the defined paths might have been too broad and uninformative, and the smaller, unique paths, which we may have been particularly interested in, may have been grouped together. On the other hand, if we had selected too many clusters, the more prototypical paths may not have been weighted correctly, and rarer paths might have been seen as equally possible or likely as those prototypical paths. Third, career sequence analysis requires a realized sequence of states. This makes it more difficult to assess effects throughout one's entire career, as one might be clustered differently after x months of observation than after y months of observation.

In light of these considerations, we instead operationalized erraticism on the basis of realized transitions. Specifically, we collected all the transitions from one job to a different job in our data set and used those occurrences to construct a transition probability matrix. We used probabilities to normalize across the starting jobs so that the total probability of transitioning to any other job sums to one. As a simple example, we construct the transition probabilities for the starting job of chemical engineer. As seen in Figure 1 below, there were nine instances of job title changes from chemical engineer in our data set. However, for the nine instances, there were only five distinct job title changes, with four being to construction project manager and one to each of process engineer, facilities engineer, and materials TD engineer. Therefore, the probabilities from left to right would be .44, .22, .11, .11, and .11, respectively.

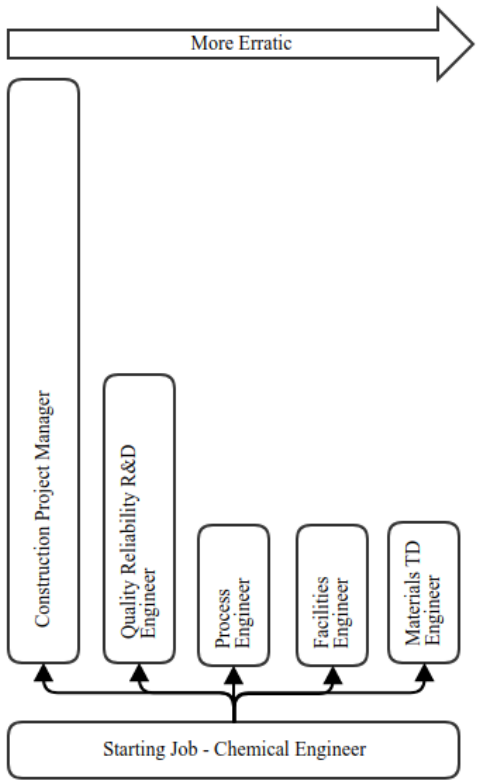


Figure 1: Conceptual Diagram of Erraticism

Then, the erraticism of each transition is calculated by subtracting the transition probabilities from one, which formally is

$$E_{ij} = 1 - (n_{ij}/\Sigma_{j \in J} n_{ij}) \tag{1}$$

where n_{ij} is the number of realized transitions from job title i to job title j. Using this definition of the erraticism of a single transition, we then calculated an employee's time-varying erraticism by averaging the erraticisms of that employee's moves through his or her career at TechCo. Under this formulation, employees who have never changed job titles have an erraticism of zero, those who changed job titles closer to more probabilistic paths (e.g. those that other employees have also taken) have medium levels of erraticism, and those who changed job titles closer to less probabilistic paths (e.g. they are perhaps the only realizations of those job title changes) would have the highest erraticism.

The distribution of the erraticism of job title changes in our data set is shown in Figure 2 below. It is left skewed, such that if a job title change was made, it tended to be more erratic than not. As an example of a career that was more erratic (erraticism~.918), the career path was technical marketing engineer to platform architect to platform manager to field applications engineer to field sales engineer. As an example of a career that was less erratic (erraticism~.578), the career path was accountant to financial analyst to finance specialist to finance business TechCo specialist to senior finance business analyst.

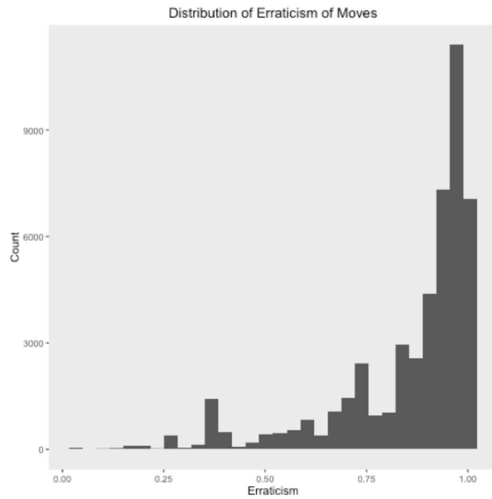


Figure 2: Distribution of Erraticism of Job Title Changes

Controls

We had a rich set of controls in our data set, which allowed us to rule out a host of alternative explanations. First, the data set contained demographic characteristics. There were indicators for the gender, ethnicity, and age group of employees, as well as fields for their highest degree status and college major. Second, the data set contained geographic details, starting from the level of the country down to the level of the city of one's local office. Third, the data set contained organizational tenure attributes, including the time the employee had been at TechCo as a whole, the time the employee had been in his or her particular job function, the time the employee had been in his or her organizational unit, the time the employee had been in his or her pay grade, and whether or not the employee had come from an acquired company. Fourth, the data set contained job characteristic data, including different levels of categorization for an employee's job, nine levels of organizational structure, a unique identifier for the employee's supervisor, whether or not the employee was himself or herself a supervisor, and the type of an employee's shift and status.

In addition to the above measures, we derived two additional sets of controls—hierarchy variables and movement type—and an indicator if an employee's manager was the same gender as the employee. For the hierarchy variables, we used the unique supervisor identifier to calculate the number of ranks above and number of ranks below an employee at a given time. We first determined which supervisors in our data set either have no supervisor, or whose supervisor is not in the data set, and mark these as the top of our hierarchy. Then, we followed the chain down, such that each link in the chain represented a rank. For example, if employee x had a supervisor in time period t, and that supervisor also had a supervisor, then employee x would have had two ranks above him or her.

For the movement type variables, vertical and horizontal moves, we used a set of change characteristics to capture the directionality of movement. We defined a vertical move as a move in which one changed hierarchical structure, where a vertical up move would be either having fewer ranks above or greater ranks below, and a vertical down move would be either having greater ranks above or fewer ranks below. Because organizational structure is not always rigid, we allowed for independent effects regarding ranks above and ranks below. Then, we coded an any change variable that captured other potential events, such as a gain or loss in managerial status; a change in pay grade or pay grade group; a change in organizational unit; or a change in job family, type, or coarse job family. By capturing what we believed to be vertical moves as well as any other type of change, the remainder is controlled for as horizontal moves.

These sets of controls[4] allowed us to rule out some common alternative explanations. For example, it could be argued that the longer one is in the organization or works in one's specific job function, the better pay and performance reviews one would get, regardless of how one moves. For such a reason, we controlled for organizational tenure, both in the organization as a whole as well as in the organizational unit, and job tenure. A second argument could be made from the literature that managers who share similar features as their subordinates are susceptible to homophily effects in

[4]For sake of brevity, we do not list the exact set of controls used for each model. While those used for each model are not exactly the same, they generally capture the wide breath of obvious alternative explanations.

deciding pay or performance outcomes (Castilla, 2011). To mitigate such concerns, we controlled for whether one's supervisor is of the same gender or not. A third argument could be made that there was some inherent quality difference that resulted in differential effects for pay and performance. Not only would that difference need to be consistent and correlated with our variables of interest, but also we attempted to control for it by using proxies such as college major, highest degree earned, and exit type (for models not examining exit as a dependent variable).

Analytical Approach

In order to best present our results and findings, we felt it necessary to leverage a mix of analytical methods. For our descriptive analyses, our descriptive statistics were constructed by subsetting the data in several different ways. Often, as in the case of displaying the mean starting pay grades for the top five male and top five female populated coarse job families, we limited the illustration to a characteristic subset of the data to make a specific point. For other descriptive statistics, we presented a view across the majority of the data set (with some extreme outliers removed) to provide a sense of our context. We additionally included sets of formal models, including OLS regressions for starting and ending pay grade analysis, and slope estimates from OLS regressions with employee fixed effects when appropriate.

For our main results, we condensed our data down to employee-year observations, as performance reviews were conducted on a yearly basis, by using the characteristics a month prior to the next performance review to predict the next period performance review and pay grade outcomes. We used OLS regressions pooling together the employee-year observations with controls to estimate the main gender effect. Additionally, for these models we cluster the standard errors on unique employees. However, as a more stringent examination of the differential gender effects on next period outcomes, we used OLS regressions with employee fixed effects to account for time-invariant individual differences. These latter models do not allow us to estimate the main gender effects but do enable us to confirm our next period effects.

For our supplementary analyses, we used a wider variety of models. Specifically, we used a mix of exact matching, OLS regressions, and piecewise exponential proportional hazard rate models. In our exact matching models, since paths that are equally erratic may be substantially different, we match females to males on the exact career sequence taken as well as the starting coarse job family and highest organizational level in order to ensure that the paths are as similar as possible. Matching models operate under the conditional independence assumption, or the assumption that, conditional on the covariates matched between the treatment and control groups, the independent variable is independent of the dependent variable. The intuition here, then, is that there may be something unobserved about the actual path that employees take, regardless of the erraticism of that path, that may affect that employee's next period outcomes. With respect to this omitted variable, our matching models help us determine if the effect occurs as a result of the actual path or some perception of the same path.

Finally, we used our piecewise exponential proportional hazard rate models to determine which covariates affected exit from TechCo. These models are an extension to the traditional proportional hazard rate models (Cox et al., 1972), in which the hazard of exit is estimated as:

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t} \tag{2}$$

and represents the instantaneous probability of experiencing a hazard event (i.e. exit) at time T between t and Δt, given that the hazard event has not yet happened at time t. The traditional model estimates a hazard function of the form

$$\log h(t) = \lambda_0(t) + \beta X \tag{3}$$

where $h(t)$ is the hazard rate at time t, $\lambda_0(t)$ is the baseline hazard rate at time t, and β is a vector of coefficients for covariates X, while the piecewise model allows for a separate constant baseline hazard within each interval, such that

$$\lambda_0(t) = \lambda_j \forall\, t \in [\tau_{j-1}, \tau_j) \tag{4}$$

where the intervals are defined by cutpoints $\tau_0 < \tau_1 < \cdots < \tau_J$. Intuitively, the piecewise element allows us to estimate a different baseline hazard rate by month, and the exponential proportional hazard rate formulation allows us to focus on how the hazard rate is affected by covariates, as coefficients can be interpreted as hazard ratios, or relative increases or decreases to the baseline hazard rate.

Results

Descriptive Analyses—Examining the Context

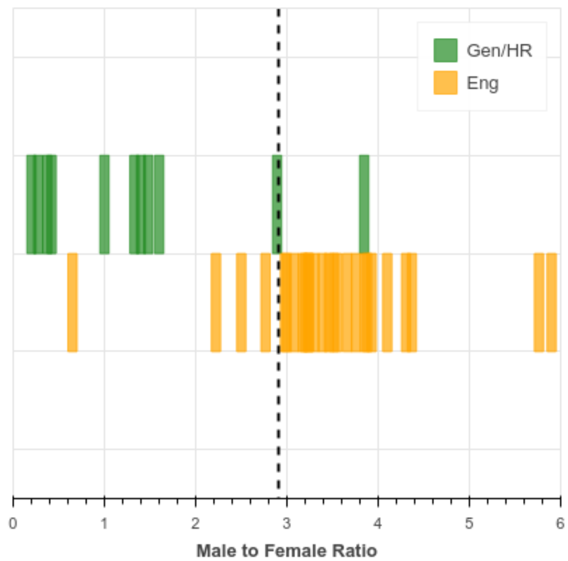
We begin our analyses with a description of our context, as understanding our context is necessary not only to establish credibility for our theoretical claims under which intra-firm mobility would be a prominent mechanism for overcoming initial disparities, but also to substantiate the impact of our results and provide a grounded basis for the later discussion of mechanisms. Specifically, we have theorized that, in accordance with prior literature, women generally start in organizational settings in disadvantaged positions compared to similar men, and therefore mobility may act as a key to unlock resources that are generally limited by organizational positions. Without establishing the former, differential mobility effects may perhaps be more an idiosyncrasy of a uniquely gendered environment than a viable mechanism of advancement.

We consider first the patterns of entry into TechCo. Job titles are nested in one of 75 coarse job families, which represent a level of detail that is usefully distinctive, yet not too unwieldy to analyze. The coarse job families are themselves divided into larger categories, such as general, human resources, engineering, finance, marketing, etc, and for the sake of our analyses, we remove those coarse job families that do not have at least ten employees.

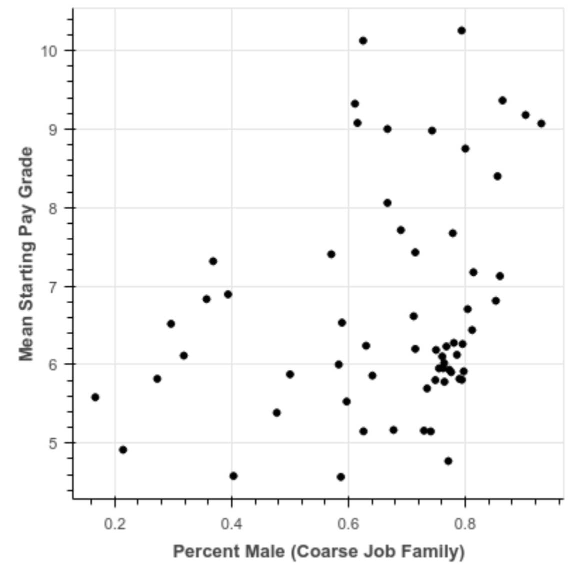
When we examine the entry of employees into TechCo by coarse job family, we find that there seems to be a general gender sorting taking place. In Figure 3(a) below, the dashed line at 2.9 represents the entry rate into TechCo as a whole; on average, 2.9 males enter for every female. Therefore, those coarse job families on the right of the dashed line represent those in which males enter at a disproportionate rate, and those coarse job families on the left of the dashed line represent those in which females enter at a disproportionate rate. We find that males tend to disproportionately enter engineering positions, while females tend to disproportionately enter general or human resources positions.

However, not only is there evidence of sex segregation with respect to types of jobs, but also there is evidence of disproportionate pay for those sex segregated jobs. In Figure 3(b) below, we see that there is a 0.273 correlation between percent male of a coarse job family, and the mean starting pay grade, such that those coarse job families that are more male dominated are generally paid higher[5].

[5]Percent male is a more intuitive measure, but if we consider the correlation between male to female ratio and mean starting pay grade it is 0.376.



(a) Select Coarse Job Families by Male to Female Ratio



(b) Mean Starting Pay Grade for Coarse Job Families by Percent Male

Figure 3: Entry into TechCo

As an illustrative example, we examine the top five male dominated and female dominated coarse job families in Figure 4 below. Notice that the top five male dominated coarse job families are all in engineering or sales, and are usually in management. Conversely, the top five female dominated coarse job families are in human resources or finance. For these ten coarse job families, we see very clearly the pay grade disparity, such that the female dominated coarse job families start at lower pay grades on average.

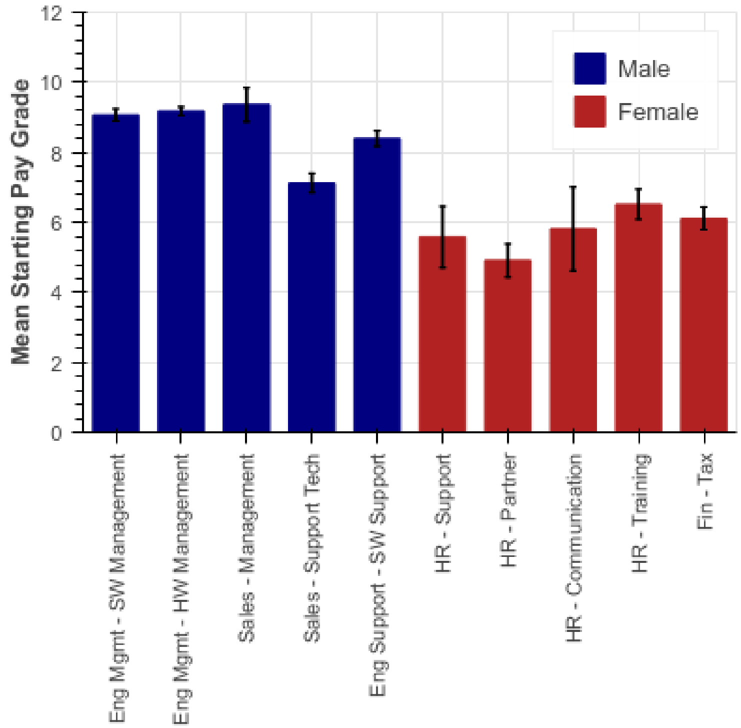


Figure 4: Mean Starting Pay Grade for Top Five Male and Female Coarse Job Families

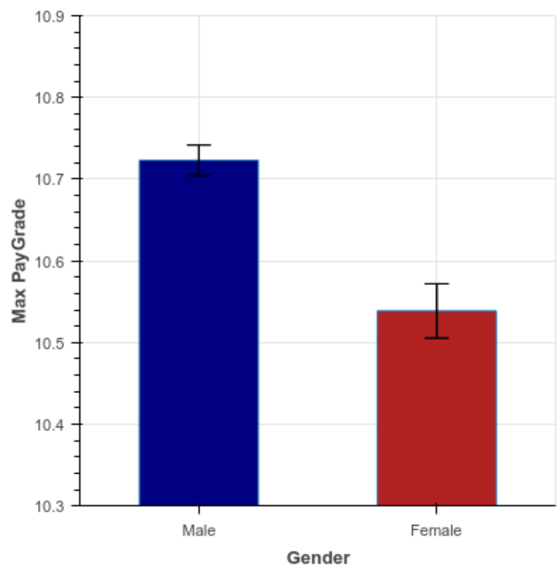
However, there are a number of known reasons for why these disparities may exist from the sex segregation literature. For example, it is not unreasonable that those coarse job families that start at higher pay grades require more responsibilities or different skills. Therefore, we further examined if there was evidence of women and men being paid differentially within jobs. Through a simple OLS regression on starting pay grade by gender, we find a significant difference of women being paid less than men, even after controlling for common confounds (Table 1).

	Dependent variable:			
	Starting Pay Grade (2-20)			
	(1)	(2)	(3)	(4)
Female	-0.560^{***}	-0.177^{***}	-0.093^{***}	-0.097^{***}
	(0.030)	(0.020)	(0.021)	(0.028)
Time Controls	YES	YES	YES	YES
Demographic Controls		YES	YES	YES
Job Characteristic Controls			YES	YES
Exit Type Controls				YES
Performance Rating Controls				YES
Observations	18,311	18,311	18,311	18,177
R^2	0.047	0.600	0.922	0.966
Adjusted R^2	0.045	0.599	0.823	0.846
Note:	*p<0.05; **p<0.01; ***p<0.001			

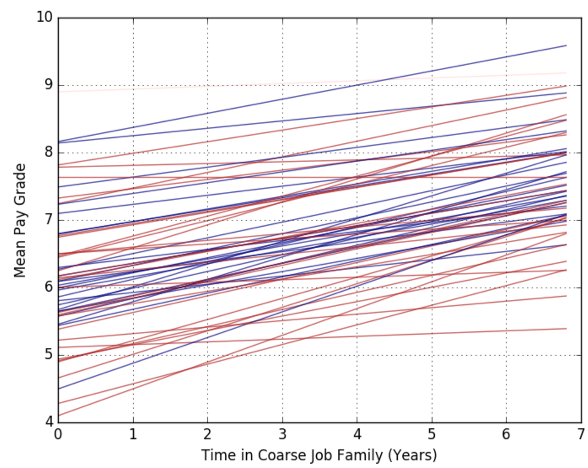
Table 1: Starting Pay Grade by Gender

While these descriptive results suggest that women do start in disadvantaged positions, there is ample opportunity for movement once within the company. Informants at TechCo have told us that movement is “encouraged” and that a common belief within the firm is that “changing jobs is the only way to get ahead in the firm.” To ease this process, TechCo has an internal job interface, in which descriptions for open jobs are posted and current employees can apply. Mobility is generally encouraged, and to this effect 27,339 of the 53,311 (51.3%) employees have at least one job title change throughout their career at TechCo.

Along these lines, we see that not only is there ample opportunity for movement, but for women, it may disproportionately advantageous. In Figure 5(a) below, we see that the max paygrade for males and females differs significantly ($p < 0.001$), so it may benefit women to access job titles and paths that are more male dominated. Using projected within person estimates of mean pay grades over time in a coarse job family, we see that coarse job families can have significantly different paths (blue represents male dominated paths and red represents female dominated paths). And, often, it may make more of a difference to shift from one coarse job family path to another rather than to continue in a single coarse job family. This is illustrative evidence that mobility may be the path to organizational rewards.



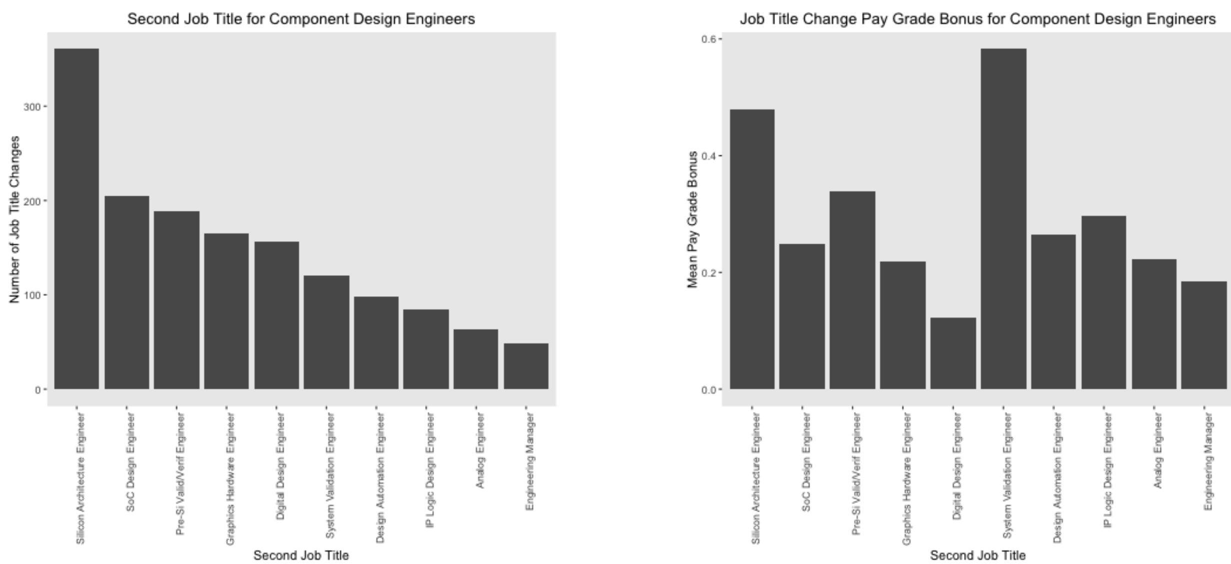
(a) Mean Pay Grade Increase by Job Title



(b) Distribution of Second Job Titles

Figure 5: Differential Benefits by Starting Position and Trajectories

However, given the encouragement of movement in the company, it is reasonable for there to be more and less well-defined internal ladders, and those that are more well-defined may not always provide the greatest benefits. We find that a large number of employees start in one of three job titles—2,591 as component design engineers, 2,573 as software engineers, and 1,249 as process engineers—accounting for 31.8% of all starting employees, but the job titles that these employees move to after these initial ones are diverse and broad. As an example, for employees who start as component design engineers, the most popular second job titles are silicon architecture engineer, SoC design engineer, and pre-si valid/verif engineer. There are a total of 53 second job titles, and the distribution of moves into the top ten are found in Figure 6(a) below. Notice that the distribution of changes in pay grade is varied for the different job title moves (Figure 6(b)), suggesting that the most common moves may not always afford the greatest benefits.



(a) Distribution of Second Job Titles

(b) Mean Pay Grade Increase by Job Title

Figure 6: Job Title Transitions from Component Design Engineer

Therefore, it may benefit employees, especially female employees, to find less typical career paths that offer better benefits. In line with this reasoning, we find a differential effect for erraticism by gender, such that females who are more erratic end up with a higher ending pay grade than those who are not erratic (Table 2), after controlling for censoring and other interim processes.

	Dependent variable:			
	Ending Pay Grade (2-20)			
	(1)	(2)	(3)	(4)
Female	−0.016		−0.017	−0.073***
	(0.009)		(0.009)	(0.011)
Erraticism		0.054***	0.054***	0.026**
		(0.008)	(0.008)	(0.009)
Female:Erraticism				0.117***
				(0.015)
Observations	52,599	52,599	52,599	52,599
R^2	0.931	0.931	0.931	0.931
Adjusted R^2	0.901	0.901	0.901	0.901
Note:		*p<0.05; **p<0.01; ***p<0.001		

Table 2: Ending Pay Grade by Gender and Erraticism

In summary, we see that while women do seem to start off in disadvantaged positions with respect to pay and advancement opportunities, intra-firm mobility serves as an understood, encouraged, and realistic pathway for career advancement. However, while mobility seems a viable way to correct the original imbalance, what is less understood is the impact of such mobility itself, which we turn to in the next section.

Main Results

In our main models, we examine the consequences of erratic moves for men and women with respect to pay and performance. Specifically, we examine how an employee's erraticism in one period affects their pay and performance in the next period, where each period is defined by the periodicity of performance reviews. Each time a new performance review is performed, that marks a new period. Therefore, these models allow us to examine how accumulated erraticism affects downstream outcomes.

In Table 3 below, we find that there is a significant and positive main effect of erraticism ($\beta = 0.175$, $p < 0.001$), meaning that individuals who are more erratic on average get higher pay grades in the next period. We additionally see that this effect is contingent on gender, such that for women, there is a significant and negative main effect ($\beta = -0.030$, $p < 0.001$), which is only overcome if one is relatively erratic ($\beta = 0.040$, $p < 0.001$). This indicates that women in general tend to fare worse in terms of pay grades, that both men and women benefit from being erratic, and that women tend to receive a differential bonus for being erratic with respect to their next period pay grades. We find that the same general set of results holds when we include individual fixed effects; in models 5 and 6 we actually find that the differential effect is larger than the main effect of erraticism.

	Dependent variable:					
	Next Period Pay Grade (2-20)					
	Pooled				Within	
	(1)	(2)	(3)	(4)	(5)	(6)
Female	−0.017**		−0.018**	−0.030***		
	(0.006)		(0.006)	(0.006)		
Erraticism		0.186***	0.186***	0.175***	0.049***	0.034***
		(0.007)	(0.007)	(0.007)	(0.005)	(0.006)
Female:Erraticism				0.040***		0.050***
				(0.009)		(0.006)
Observations	257,904	257,750	257,750	257,750	257,686	257,686
R^2	0.918	0.919	0.919	0.919	0.971	0.971
Adjusted R^2	0.908	0.908	0.908	0.908	0.958	0.959
Note:				*p<0.05; **p<0.01; ***p<0.001		

Table 3: Next Period Pay Grade by Gender and Erraticism

In Table 4 below, we find that there is a significant and positive main effect of erraticism ($\beta = 0.100$, $p < 0.01$), meaning that individuals who are more erratic also on average get higher performance ratings in the next period. Again, we see that this effect is contingent on gender, but for women, there is a significant and positive main effect ($\beta = 0.064$, $p < 0.05$), and a significant and negative differential effect ($\beta = -0.145$, $p < 0.01$). Notice that in both the pooled models and the within models, the main effect of erraticism is overcome by the differential effect (in the within models they essentially cancel out), suggesting that more erratic women generally fare worse than their similarly erratic male counterparts with respect to next period performance reviews.

	Dependent variable:					
	Next Period Performance					
	Pooled				Within	
	(1)	(2)	(3)	(4)	(5)	(6)
Female	0.020		0.019	0.064*		
	(0.028)		(0.028)	(0.032)		
Erraticism		0.062	0.062	0.100**	0.154**	0.216***
		(0.033)	(0.033)	(0.036)	(0.055)	(0.058)
Female:Erraticism				−0.145**		−0.217***
				(0.049)		(0.061)
Observations	257,868	257,715	257,715	257,715	257,651	257,651
R^2	0.210	0.210	0.210	0.210	0.477	0.477
Adjusted R^2	0.111	0.111	0.111	0.111	0.262	0.262
Note:				*p<0.05; **p<0.01; ***p<0.001		

Table 4: Next Period Performance by Gender and Erraticism

These results indicate that there exists a dilemma of mobility: erratic women tend to benefit from their erratic moves with respect to pay, which may help in equalizing the initial pay differential. However, erratic women also suffer a performance detriment as compared to similarly erratic men, suggesting that erratic moves may act as a double-edged sword.

Supplementary Analyses—Examining the Mechanism

While our main results support the claim that intra-firm mobility events lead to differential outcomes for men and women, they still leave important questions about the mechanisms of these effects unanswered. Although we must be clear that it is nearly impossible to tease apart these deeper mechanisms given our empirical setting and data, we have elicited and can partially examine two potential mechanisms driving the dilemma, one in which female employees may be under-prepared as compared to male counterparts to handle the new responsibilities of their new job titles, and another in which female employees may be merely perceived more negatively. Our supplementary analyses therefore help us answer some critical questions and provide nuance to our mechanistic pathways.

First, the actual paths employees take may affect pay and performance outcomes. Two employees may have the same level of erraticism but can have taken drastically different paths. So, it remains an open question whether the differential effect we see is a result merely of women taking different paths. For example, if men are moving from one engineering position to another, while women are moving from human resources or general to engineer or from non-managerial positions to managerial positions, then we might imagine that the change in responsibilities for the latter moves might be greater, resulting in a lower performance rating in the next period. Moreover, if men and women are moving through the same job titles, then the argument has been made that each job title prepares the individual for the next (Doeringer and Piore, 1971; Dokko et al., 2009), and therefore any effect we see would be more perceptual. We use a matching analysis to answer these two questions, whereby we match males and females on starting positions and exact job title changes, and we find that both differential effects remain significant (Tables 5 and 6).

	Dependent variable:			
	Next Period Pay Grade (2-20)			
	(1)	(2)	(3)	(4)
Female	-0.181^{***}		-0.181^{***}	-0.187^{***}
	(0.014)		(0.014)	(0.015)
Erraticism		0.092^{**}	0.084^{**}	0.053
		(0.030)	(0.030)	(0.032)
Female:Erraticism				0.062^{*}
				(0.030)
Observations	76,607	76,566	76,566	76,566
R^2	0.826	0.825	0.826	0.826
Adjusted R^2	0.761	0.759	0.761	0.761
Note:		*p<0.05; **p<0.01; ***p<0.001		

Table 5: Next Period Pay Grade (Matched)

	Dependent variable:			
	Next Period Performance			
	(1)	(2)	(3)	(4)
Female	0.057		0.058	0.087
	(0.048)		(0.048)	(0.051)
Erraticism		0.232*	0.235*	0.370**
		(0.105)	(0.105)	(0.121)
Female:Erraticism				−0.268*
				(0.120)
Observations	76,417	76,376	76,376	76,376
R^2	0.355	0.355	0.355	0.355
Adjusted R^2	0.112	0.113	0.113	0.113
Note:				*p<0.05; **p<0.01; ***p<0.001

Table 6: Next Period Performance (Matched)

The fact that both effects remain significant even after matching on exact job title transitions suggests that the effect is not purely driven by men and women making different types of moves, and also that it may be more of a perceptual effect.

Second, understanding how these effects may persist or not may give us insight into the nature of the effect. In order to examine this, we limited the sample to those individuals with only one job title change in their career, and looked at the next *n* periods later outcomes. For example, if an employee had a job title change in one period that influenced his or her erraticism, we would then look at outcomes in the next period, in the period after that, and so on. This gives us a sense for how far-reaching erraticism in one period is.

If the effect were truly due to an over-promotion mechanism, then we might expect the performance effect to persist over time, as women would have a more difficult time adjusting, and, provided that men and women may learn at similar rates in their new job positions, would consistently lag behind their similarly erratic male counterparts with respect to performance. On the other hand, if the effect were more perceptual, we might see an initial discount, but then the effect should diminish or disappear in later years.

In Tables 7 and 8 below, we see a pattern of effects consistent with the latter, such that the pay grade bonus is immediate, but then goes away after the next period, while the performance detriment persists for a little while, then goes to non-significance. We were told that often, if the timing of the performance review is soon after a move, both the new and old supervisors will have input into the performance review process, which suggests that the performance review hit in the second period after a move would be slightly stronger as we find. After 2 periods, however, we see that the effect diminishes to non-significance.

	Dependent variable:				
	Pay Grade (2-20): Number of Periods Later				
	1	2	3	4	5
	(1)	(2)	(3)	(4)	(5)
Erraticism	0.009	0.025	0.035*	−0.017	0.026
	(0.012)	(0.015)	(0.016)	(0.020)	(0.029)
Female:Erraticism	0.042***	0.023	0.001	0.014	0.0004
	(0.011)	(0.013)	(0.015)	(0.020)	(0.030)
Observations	86,265	71,547	57,936	45,608	34,592
R^2	0.978	0.977	0.979	0.982	0.985
Adjusted R^2	0.962	0.959	0.962	0.966	0.969
Note:	*p<0.05; **p<0.01; ***p<0.001				

Table 7: Lag Pay

	Dependent variable:				
	Performance: Number of Periods Later				
	1	2	3	4	5
	(1)	(2)	(3)	(4)	(5)
Erraticism	0.425***	0.551***	−0.044	0.032	0.771*
	(0.125)	(0.154)	(0.188)	(0.243)	(0.372)
Female:Erraticism	−0.282*	−0.413**	−0.191	0.124	0.402
	(0.113)	(0.142)	(0.177)	(0.242)	(0.390)
Observations	86,256	71,302	57,495	45,184	34,190
R^2	0.585	0.596	0.583	0.610	0.657
Adjusted R^2	0.290	0.285	0.259	0.256	0.281
Note:					*p<0.05; **p<0.01; ***p<0.001

Table 8: Lag Performance

Third, it may be that the effect is driven at least partially by erraticism being related to some aspect of talent or ability above and beyond our controls, such that those who are more erratic are less skilled, thus accounting for the differential performance detriment. One possibility is that women who are erratic are simply being pushed around into different jobs in the organization as a means of retaining women in the workforce. If this is the case, then we should expect to see that women who are more erratic are less likely to leave involuntary, as the organization would prefer to move them around rather than have them leave the organization.

However, this is the opposite of the pattern we see from the exit results. We fail to find evidence that erratic women are more likely to leave involuntarily than their similarly erratic male counterparts (Table 9) and actually find that erratic women are less likely to leave voluntarily (Table 10). When we break down the voluntary leave by desired and undesired (from the perspective of the organization) in Table 11 below, we further corroborate the story that erratic women are less likely to leave for reasons against the organization's desires, which would suggest that these women are not low performers; if they were low performers, then their voluntary leave would most likely be coded as desired. Finally, we see that when we examine specific types of voluntary exit, specifically those due to some kind of dissatisfaction, erratic women are significantly and highly less likely to leave (Table 12), suggesting that erratic moves do not have the negative connotation they would if it were merely a means to retain women.

These patterns of findings, if anything, would suggest that erratic women are actually higher performing than their less erratic female counterparts, meaning that the performance hit after moving job titles would be more surprising. We find evidence that erraticism is less likely to be a signal of poor performance as it is to be of high performance.

	Dependent variable:			
	Involuntary Exit			
	(1)	(2)	(3)	(4)
Female	1.021		0.985	0.978
	(0.706)		(0.789)	(0.739)
Erraticism		0.692^{***}	0.692^{***}	0.688^{***}
		(0.000)	(0.000)	(0.000)
Female:Erraticism				1.026
				(0.842)
Observations	93,302	91,109	91,109	91,109
LR Test Statistic	594.72	566.02	566.09	566.13
Max. Log-Likelihood	-7986.5	-7630.2	-7630.2	-7630.2
Note: p-value in parentheses	*p<0.05; **p<0.01; ***p<0.001			

Table 9: All Involuntary Turnover

	Dependent variable:			
	Voluntary Exit			
	(1)	(2)	(3)	(4)
Female	0.988		0.988	1.027
	(0.611)		(0.612)	(0.352)
Erraticism		0.812^{***}	0.812^{***}	0.840^{***}
		(0.000)	(0.000)	(0.000)
Female:Erraticism				0.879^{*}
				(0.019)
Observations	378,576	368,359	368,359	368,359
LR Test Statistic	2829.51	2724.03	2724.29	2729.87
Max. Log-Likelihood	-42502	-40712	-40712	-40709
Note: p-value in parentheses		*p<0.05; **p<0.01; ***p<0.001		

Table 10: All Voluntary Turnover

	Dependent variable:							
	Desired Voluntary Exit				Undesired Voluntary Exit			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Female	0.901^{*}		0.912^{*}	0.926	1.052		1.046	1.092^{*}
	(0.013)		(0.031)	(0.160)	(0.079)		(0.131)	(0.012)
Erraticism		0.883^{*}	0.883^{*}	0.892		0.758^{***}	0.758^{***}	0.792^{***}
		(0.032)	(0.031)	(0.068)		(0.000)	(0.000)	(0.000)
Female:Erraticism				0.960				0.855^{*}
				(0.667)				(0.022)
Observations	149,341	145,938	145,938	145,938	226,580	219,840	219,840	219,840
LR Test Statistic	1987.56	1959.28	1961.55	1966.82	989.02	909.11	913.81	913.99
Max. Log-Likelihood	-27972	-26873	-26872	-26869	-13297	-12615	-12612	-12612
Note: p-value in parentheses						*p<0.05; **p<0.01; ***p<0.001		

Table 11: All Voluntary Turnover with Respect to Organizational Desires

	Dependent variable:			
	Voluntary Exit (Dissatisfaction)			
	(1)	(2)	(3)	(4)
Female	0.989		1.018	1.281
	(0.937)		(0.901)	(0.152)
Erraticism		0.609^{**}	0.608^{**}	0.724
		(0.005)	(0.005)	(0.097)
Female:Erraticism				0.500^{*}
				(0.033)
Observations	12,837	12,467	12,467	12,467
LR Test Statistic	155.14	156.66	156.67	161.39
Max. Log-Likelihood	-1480.8	-1416.3	-1416.3	-1414
Note: p-value in parentheses		*p<0.05; **p<0.01; ***p<0.001		

Table 12: Turnover for Reasons of Dissatisfaction

Fourth, it may be that the effect is isolated to a specific population of women and men. Specifically, the literature would suggest that any perceptual backlash should affect women lower in the pay hierarchy, while those higher in the pay hierarchy are shielded from such effects. To this extent, we split our population into two samples, one that contained males and females that started below manager level (generally pay grade 8 and below), and those that started at manager level (generally pay grade 9 and above), and examined our next period effects for each subsample.

We find that for males and females that started below manager level, the effects generally remain the same. There is a significant differential effect for both next period pay grade ($\beta = 0.031$, $p < 0.01$) and next period performance ($\beta = -0.164$, $p < 0.01$) in Tables 13 and 14 below.

	Dependent variable:					
	Next Period Pay Grade (2-20)					
	Pooled				Within	
	(1)	(2)	(3)	(4)	(5)	(6)
Female	-0.012^{*}		-0.013^{*}	-0.023^{***}		
	(0.006)		(0.006)	(0.007)		
Erraticism		0.207^{***}	0.207^{***}	0.198^{***}	0.065^{***}	0.055^{***}
		(0.007)	(0.007)	(0.008)	(0.006)	(0.007)
Female:Erraticism				0.031^{**}		0.034^{***}
				(0.010)		(0.007)
Observations	210,545	210,418	210,418	210,418	210,354	210,354
R^2	0.863	0.864	0.864	0.864	0.952	0.952
Adjusted R^2	0.843	0.844	0.844	0.844	0.930	0.930
Note:						$^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

Table 13: Next Period Pay Grade (Subset 1 - Starting as Non-Manager)

	Dependent variable:					
	Next Period Performance					
	Pooled				Within	
	(1)	(2)	(3)	(4)	(5)	(6)
Female	−0.037		−0.039	0.012		
	(0.030)		(0.031)	(0.035)		
Erraticism		0.058	0.058	0.106^{**}	0.161^{*}	0.237^{***}
		(0.038)	(0.038)	(0.041)	(0.064)	(0.067)
Female:Erraticism				-0.164^{**}		-0.247^{***}
				(0.053)		(0.068)
Observations	210,517	210,391	210,391	210,391	210,327	210,327
R^2	0.216	0.216	0.216	0.216	0.482	0.482
Adjusted R^2	0.101	0.101	0.101	0.101	0.247	0.247
Note:	$^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$					

Table 14: Next Period Performance (Subset 1 - Starting as Non-Manager)

However, for males and females that started at manager level, the effects are significantly weakened. For next period pay grade, we see a significant differential effect only for the within models, which suggests that there may be uncontrolled individual-level heterogeneity. For the next period performance models, we fail to find a significant effect in either the pooled models or the within models. So, while we can not preclude the fact that females who start at the manager level may still experience a differential effect, the results suggest that the effect is much stronger for those starting at lower levels in the pay hierarchy.

	Dependent variable:					
	Next Period Pay Grade (2-20)					
	Pooled				Within	
	(1)	(2)	(3)	(4)	(5)	(6)
Female	−0.026*		−0.027*	−0.025		
	(0.013)		(0.013)	(0.014)		
Erraticism		0.087***	0.087***	0.088***	−0.009	−0.016
		(0.013)	(0.013)	(0.014)	(0.011)	(0.011)
Female:Erraticism				−0.005		0.040**
				(0.024)		(0.015)
Observations	47,359	47,332	47,332	47,332	47,332	47,332
R^2	0.894	0.895	0.895	0.895	0.956	0.956
Adjusted R^2	0.864	0.865	0.865	0.865	0.930	0.930
Note:				*p<0.05; **p<0.01; ***p<0.001		

Table 15: Next Period Pay Grade (Subset 2 - Starting as Manager)

	Dependent variable:					
	Next Period Performance					
	Pooled				Within	
	(1)	(2)	(3)	(4)	(5)	(6)
Female	0.400***		0.400***	0.409***		
	(0.090)		(0.090)	(0.105)		
Erraticism		−0.087	−0.091	−0.086	−0.015	−0.065
		(0.091)	(0.091)	(0.094)	(0.141)	(0.145)
Female:Erraticism				−0.031		0.275
				(0.164)		(0.189)
Observations	47,351	47,324	47,324	47,324	47,324	47,324
R^2	0.361	0.360	0.361	0.361	0.593	0.593
Adjusted R^2	0.181	0.180	0.181	0.181	0.352	0.352
Note:	*p<0.05; **p<0.01; ***p<0.001					

Table 16: Next Period Performance (Subset 2 - Starting as Manager)

This last analysis also brings up an important point about subpopulations that experience these effects. It is quite possible, as in the case of the females who start as managers and those that do not, that some females experience the benefits without the detriments, and potentially some experience the detriments without the benefits. Our models and analyses do not suggest that whenever females are erratic, they always experience both a greater pay grade increase and a performance rating decrease than similarly erratic males simultaneously. They do, however, seem to suggest that on average, there is a potential downside to highly erratic moves which has previously not been considered.

Although each of our supplementary analyses has its weaknesses, we believe that put together, they suggest that the dilemma of mobility is more a perceptual effect, although we do not go so far as to say that over-promotion is not present. Despite the performance detriment, though, it seems that the benefits outweigh the detriments, as erratic women are less likely to leave for reasons of dissatisfaction. Furthermore, it seems the performance detriment is a momentary drawback, while the pay grade increase would be a perpetual benefit.