PERFORMANCE, CAREER DYNAMICS, AND SPAN OF CONTROL

by

Valerie Smeets
Department of Economics and Business
Aarhus University
vas@asb.dk

Michael Waldman
Johnson Graduate School of Management
Cornell University
mw46@cornell.edu

and

Frederic Warzynski
Department of Economics and Business
Aarhus University
fwa@asb.dk

March 2016

* We thank seminar and conference participants at the Aarhus/Xiamen workshop, Cornell University, the National University of Singapore, Queens University, Stockholm University, Universidad Carlos III, the University of Hong Kong, the University of New South Wales, HECER, Kansai Labor Research Group seminar, the 6th Annual Organizational Economics Workshop at the University of Sydney, the 2012 NBER Organizations Conference, the IOS session at the 2012 ASSA meetings, the 2013 SOLE conference, the 2013 NBER Personnel Conference, the 2014 CEPR Workshop on Incentives, Management, and Organisation, Guido Friebel, Esteban Rossi-Hansberg, and Maria Guadalupe for helpful comments on previous drafts. We also thank Jed DeVaro and Bob Gibbons for conversations helpful for the initial formulation of the paper.

ABSTRACT

There is an extensive theoretical literature based on what is called the scale-of-operations effect, i.e., the idea that the return to managerial ability is higher the more resources the manager influences with his or her decisions. This idea leads to various testable predictions including that higher ability managers should supervise more subordinates, or equivalently, have a larger span of control. And although some of this theory's predictions have been empirically investigated, there has been little systematic investigation of the theory's predictions concerning span of control. In this paper we first extend the theoretical literature on the scale-of-operations effect to allow firms' beliefs concerning a manager's ability to evolve over the manager's career, where much of our focus is the determinants of span of control. We then empirically investigate testable predictions from this theoretical analysis using a unique single firm dataset that contains detailed information concerning the reporting relationships at the firm. Our investigation provides support both for the model's predictions concerning wages, wage changes, and probability of promotion, and also for the model's predictions concerning span of control including predictions derived from the learning component of the model. Overall, our investigation supports the notion that the scale-of-operations effect and additionally learning are important determinants of the internal organization of firms including span of control.

I. INTRODUCTION

A standard building block of hierarchical models of the firm is the scale-of-operations effect. The basic idea is that a manager's ability has a larger effect on firm productivity the more workers that are below the manager in the firm's hierarchy. This standard building block which can be found in various theoretical models of the firm, such as Lucas (1978), Rosen (1982), and Garicano (2000), leads to a number of testable predictions concerning various important aspects of internal organization such as assignment, promotion, and span of control.[1] And although some of these predictions have been investigated previously, there has been little systematic empirical study of the theory's predictions concerning span of control.[2] This paper is a theoretical and empirical investigation concerning the scale-of-operations effect and, in particular, what the scale-of-operations effect tells us about span of control.

The most basic prediction that the scale-of-operations effect makes concerning span of control is that higher ability or more productive managers should have larger spans of control. In the standard model that makes this prediction a subordinate's productivity is positively related to the ability of the subordinate's manager, but there is also a cost of having a manager that supervises a large number of workers such as having the extra productivity due to a higher ability manager be lower the more workers the manager supervises, i.e., the larger is the manager's span of control. In deciding upon a manager's span of control, therefore, a firm must trade-off the benefit of having this extra productivity apply to a larger number of workers with the costs associated with a larger span of control. In the standard set-up, in turn, this trade-off yields that a higher ability manager should have a larger span of control.

This argument which can be found in early papers on the topic such as Lucas (1978) and Rosen (1982) ignores the idea that during a manager's career firms' beliefs concerning the manager's ability will evolve as the manager accumulates human capital and firms observe managerial performance in each period. Gibbons and Waldman (1999a,2006) consider models

---

[1] The scale-of-operations effect was first discussed in Mayer (1960). See Section II for a discussion.

[2] One exception is Lazear, Shaw, and Stanton (2014) which we discuss in Section II.

with this type of learning and show that they match well with various empirical findings in the literature concerning wage and promotion dynamics, including findings in the well known study of Baker, Gibbs, and Holmstrom (1994a,b). But Gibbons and Waldman's models do not incorporate the scale-of-operations effect and also do not consider span of control (in the Gibbons and Waldman models there is a job ladder but there is no sense in which a manager supervises a particular set of workers). In this paper we incorporate the human capital and learning assumptions from the Gibbons and Waldman analyses into a model of the scale-of-operations effect and span of control.

In our basic theoretical analysis workers vary in terms of their innate ability levels, while effective ability is a function of innate ability and general human capital which workers accumulate as they gain labor market experience. When a worker enters the labor market firms initially do not know a worker's innate and effective ability levels but firms update their beliefs concerning these abilities when output is produced, where we assume output realizations are publicly observable. There are identical competitive firms described by a two-level job ladder where each division in a firm has a single worker on an upper level who supervises employees at a lower level. Consistent with the above discussion, we further assume that the return to managerial ability increases with the number of workers the manager supervises. We also consider a number of extensions including what happens when workers vary in terms of publicly observable schooling levels and what happens when there are three-level job ladders.

We first show that this model is consistent with the basic scale-of-operations result that higher ability managers have larger spans of control, where to be precise in this model this positive correlation is between a manager's expected effective ability and the manager's span of control. We also find a number of other theoretical results that have been found elsewhere such as a positive correlation between wage increases and performance, a positive correlation between worker education and probability of subsequent promotion holding worker performance fixed, and a positive correlation between worker performance and probability of subsequent promotion holding the education level fixed.

We then derive a number of results concerning span of control that are not found in previous theoretical analyses. These new results are based on the idea that span of control in our framework is positively correlated with a manager's expected effective ability, so factors that are correlated with expected effective ability will also have the same correlation with span of control. For example, we find that, holding education fixed, a manager's span of control should be positively correlated with previous performance. The logic is that higher performance causes firms to positively update their beliefs concerning the manager's expected effective ability, so the positive correlation between expected effective ability and span of control translates into a positive correlation between prior performance and span of control. Similarly, holding performance fixed, a manager's span of control should be positively correlated with the manager's education. The logic here is similar. Because performance ratings are noisy measures of expected effective ability, higher education even controlling for performance ratings is positively correlated with expected effective ability. So the positive correlation between expected effective ability and span of control now translates into a positive correlation between the manager's education and span of control.

In the second part of the paper we test these predictions using confidential performance and personnel data from a large EU "high tech" manufacturing firm that has production facilities in various counties around the world and whose products are sold globally. What distinguishes this dataset from the dataset investigated in Baker, Gibbs, and Holmstrom (1994a,b) and datasets used in other similar investigations is that, in addition to performance ratings like those found in the Baker, Gibbs, and Holmstrom study, this dataset includes information about the firm's chain of command, i.e., we know who each individual worker and manager reports to.[3] And we are able to use this information to calculate the span of control for each manager and how that span of control varies over time.

[3] Other studies of single firm datasets that find results similar to those found by Baker, Gibbs, and Holmstrom include Lazear (1992), Seltzer and Merrett (2000), Treble et al. (2001), and Dohmen, Kriechel, and Pfann (2004). A similar dataset is also investigated in the well known studies of Medoff and Abraham (1980,1981), although those studies mostly focus on different issues.

In our empirical analysis we test various predictions of our theoretical analysis. We start with predictions that are consistent with earlier related theoretical studies and for which there is already empirical evidence. For example, as predicted by the theory, we find that wage growth is positively related to a worker's performance evaluations. We also find that the probability of promotion is positively related to a worker's education level and to performance ratings. Note that, consistent with our discussion, related empirical results can be found in various empirical studies such as Baker, Gibbs, and Holmstrom (1994a,b), Seltzer and Merrett (2000), and DeVaro and Waldman (2012).

We then turn our attention to predictions concerning span of control many of which are new and for all of which there is little or no previous empirical evidence. As discussed, our theoretical model predicts that span of control should be positively related to factors positively correlated with expected effective ability – performance ratings and education. We find evidence consistent with the prediction concerning performance ratings and mixed evidence concerning the education prediction. Our theoretical model also predicts positive correlations between current span of control and the current wage as well as between current span of control and the probability of subsequent promotion. And our empirical analysis finds clear support for the former prediction at managerial levels above the bottom one and support for the latter prediction at the managerial levels in the middle of the firm's job ladder.

We also conduct an additional test which we believe is a better test of the theory because it is less likely to be explained by alternative theories. If span of control is positively related to expected effective ability as our theory predicts, then changes in span of control should be positively correlated with factors that are positively correlated with changes in expected effective ability. Specifically, performance ratings should be positively correlated with how a manager's span of control varies over time. And indeed we find evidence that changes in span of control are positively correlated with performance ratings.

Overall, our theoretical and empirical analysis supports the idea that the scale-of-operations effect and learning are important determinants of the design of job hierarchies and, in particular, play an important role in span of control and how a manager's span of control varies over the manager's career. The paper thus contributes to the literature referred to above that investigates the role of job assignment, learning, and human capital acquisition in the operation of internal labor markets. The earlier literature has mostly focused on showing that these factors can explain a wide variety of empirical findings concerning wage and promotion dynamics in internal labor markets, while here we show that a related approach is also consistent with empirical evidence concerning span of control.

The outline for the paper is as follows. Section II reviews the related literature. Section III presents and analyzes a theoretical model that combines the scale-of-operations effect with human capital acquisition and learning about worker ability as workers gain labor market experience. Section IV describes the data and presents some basic facts about the firm. Section V focuses on tests of predictions of our theoretical model, where much of the focus is on tests concerning span of control because many of these predictions are new. Section VI discusses what we learn from our empirical findings. Section VII provides concluding remarks.

II. RELATED LITERATURE

As indicated in the Introduction, there is an extensive theoretical literature concerning the operation of hierarchies and much of this literature is consistent with the scale-of-operations effect which was first discussed in Mayer (1960). For example, numerous models that focus on assignment such as Lucas (1978), Rosen (1982), and Waldman (1984a) are characterized by the scale-of-operations effect, while the knowledge-based hierarchy models of Garicano (2000) and Garicano and Rossi-Hansberg (2006) are also consistent with this idea. In addition, there are other models of hierarchical production such as the supervision models of Williamson (1967), Calvo and Wellisz (1978,1979), and Qian (1994) and tournament models such as Lazear and Rosen (1981) and Rosen (1986) where there is no clear scale-of-operations effect either because workers do not vary in terms of ability (the supervision models) or the managerial production function is not modeled (the tournament models).[4] But we believe that it is possible to build extensions of such models that would be consistent with the scale-of-operations effect.[5] In our theory and empirical work we do not try to distinguish between different approaches that might yield the scale-of-operations effect, but rather focus on identifying the empirical implications of the scale-of-operations effect and the extent to which our data is consistent with those implications.

Another literature our paper builds on is the literature on symmetric learning in labor markets studied initially in Harris and Holmstrom (1982) and Holmstrom (1982). The basic idea in this literature is that firms are imperfectly informed about a worker's ability when the worker enters the labor market and learn about ability gradually as worker outputs are publicly observed.[6] Specifically, we build on the more recent contributions found in Gibbons and Waldman (1999a,2006) which consider this type of learning in a setting characterized by workers who accumulate human capital with labor market experience and move up a job ladder as they age. Gibbons and Waldman show that such a model can explain various empirical findings concerning wage and promotion dynamics found in Baker, Gibbs, and Holmstrom (1994a,b) and elsewhere.[7] But, as mentioned earlier, the Gibbons and Waldman models do not capture the scale-of-operations effect because there is no span of control in their model. We

[4] Calvo and Wellisz (1979) do allow workers to vary in terms of ability, but they purposely do not incorporate the scale-of-operations effect because they want to show that in a model of supervision higher ability workers wind up at higher levels of job hierarchies even in the absence of the scale-of-operations effect.

[5] See Sattinger (1993) and Gibbons and Waldman (1999b) for surveys that discuss many of these literatures.

[6] See Farber and Gibbons (1996) and Altonji and Pierret (2001) for empirical studies that find evidence consistent with the symmetric learning approach. The alternative assumption is that learning is asymmetric which means that only a worker's current employer directly gathers information concerning the worker's ability and other firms observe the actions of the current employer in updating beliefs about the worker. This approach was initially modeled in Greenwald (1979,1986) and Waldman (1984b), while Gibbons and Katz (1991), Pinkston (2009), DeVaro and Waldman (2012), and Kahn (2013) find empirical evidence consistent with this assumption. See the Conclusion for a related discussion.

[7] Other important early papers in this literature include Lazear (1992), McLaughlin (1994), McCue (1996), and Podolny and Baron (1997). Also, see Kauhanen and Napari (2012) for empirical results similar to those found by Baker, Gibbs, and Holmstrom but in a study that employs a large linked employer-employee panel dataset, while a number of studies including Lima and Pereira (2003), Lluis (2005), Dias da Silva and van der Klaauw (2011), and Hunnes (2012) empirically estimate the Gibbons and Waldman (1999a) model using large datasets from other countries and in general these studies find evidence that supports the framework. See Waldman (2012) for a survey that discusses this literature.

incorporate span of control into a Gibbons and Waldman type set-up, where much of our focus is the resulting testable implications concerning span of control and whether the data is consistent with those implications.

This paper extends in two ways the previous empirical work in Smeets and Warzynski (2008) that used personnel data from the same firm we study. Specifically, we incorporate performance data into the empirical analysis and also provide a formal theoretical model that we use to guide our empirical work. Smeets and Warzynski analyzed how the firm's hierarchy had evolved over time and showed that the hierarchy had become flatter and that span of control has increased consistent with earlier results in Rajan and Wulf (2006). They concluded that these results are best explained by the knowledge-based theory of hierarchical production and, in particular, by the specific model put forth in Garicano and Rossi-Hansberg (2006) under the assumption that communication costs have fallen over time. They then considered the dynamics of managerial careers including various factors such as promotions, compensation, and span of control. They found a number of results that they argued might be explained by combining elements of Rosen's (1982) static analysis of assignment and hierarchies with the learning and human capital acquisition elements of Gibbons and Waldman (1999), and then suggested that theoretically and empirically investigating such a model would be worthwhile. Our paper follows up this suggestion.

Our paper also contributes to a small but important empirical literature concerning hierarchical production, where most of this literature is based on survey data.[8] Ortin-Angel and Salas-Fumas (2002) use survey data from a large number of Spanish firms for the period 1990 to 1992. Their main findings are that the elasticity of managerial compensation to number of subordinates seems to be less than one and that differences in measurable human capital explain a large fraction of wage differences within firms. Another important paper is Rajan and Wulf

[8] Another related strand of the literature analyzes how information and communication technology and product market competition affect firm organization (see Delmastro (2002) and Bloom et al. (2014) for analyses related to information and communication technology and Guadalupe and Wulf (2010) for an analysis of product market competition). These issues are beyond the scope of our paper. See Smeets (2016) for a discussion.

(2006) referred to above. They focus on 300 large US firms over the period 1986 to 1999 and, as indicated, document a flattening of firms' hierarchies over time, i.e., span of control has increased while number of layers has decreased. They also find growth in pay inequality. Garicano and Hubbard (2007, 2009) employ the 1992 Census of Services to look at hierarchical production in law firms where much of their focus is on specialization. They find a number of results including that, as predicted by the theory they develop, span of control depends on the extent of the market. In a recent paper, Caliendo, Monte, and Rossi-Hansberg (2015) focus on the relationship between changes in the number of hierarchical layers and wage and employment changes.

A couple of other papers in this literature are closer to our paper in that they are related to the scale-of-operations effect. Fox (2009) uses a Swedish linked employer-employee dataset and US data from the 1996 Survey of Income and Program Participation to show that the wage gap between large and small firms increases with "job responsibility" which is consistent with predictions of hierarchical models characterized by the scale-of-operations effect. Finally, the closest paper to ours is the recent paper by Lazear, Shaw, and Stanton (2015) that focuses on the extent to which high ability managers affect worker productivity. Although it is not their main focus, they do test for a scale-of-operations effect concerning span of control but in their dataset they find no evidence that higher ability managers have larger spans of control.[9]

III. MODEL AND THEORETICAL ANALYSIS

In this section we first present and analyze a model with identical competitive firms described by a two-level job ladder and multiple divisions, where a division consists of a single manager at an upper level who supervises workers on a lower level. There are overlapping generations of workers, where workers are in the labor market for exactly two periods but enter

[9] One possibility for why our empirical analysis supports the scale-of-operations effect and their analysis does not is that we employ the actual performance evaluations employed by the firm while they construct performance data from data on worker productivity. We would like to thank Ed Lazear for suggesting this possibility to us.

the labor market with heterogeneous schooling levels. The model captures both the scale-of-operations effect as in, for example, Lucas (1978) and Rosen (1982), and symmetric learning about worker abilities as in, for example, Harris and Holmstrom (1982) and Gibbons and Waldman (1999a,2006). We then discuss what happens when the analysis is extended to workers with longer labor market lives and to job ladders with more levels.

A) The Model

There is free entry into production, where firms are identical and the only input is labor. A worker's career lasts two periods and labor supply in each period is fixed at one unit for each worker. A worker is referred to as young in the worker's first period in the labor market and old in the second. Worker i enters the labor market with schooling level s_i, where s_i can take on any integer value between 1 and S. We also assume that each cohort of workers has exactly z_s workers of schooling level s, where $z_s>0$ for all s, s=1,…S.

Let η_{it} denote worker i's effective ability in period t, where

(1) $$\eta_{it}=\theta_i f(x_{it}).$$

In equation (1), θ_i is the worker's innate ability and x_{it} is the worker's labor market experience prior to period t, i.e., $x_{it}=0$ for young workers and $x_{it}=1$ for old workers. Also, we assume $f(1)>f(0)>0$ which captures that workers acquire general human capital as they age. We assume that innate ability can be either high or low: $\theta_i \in [\theta_L, \theta_H]$. The ex ante probability that a worker with schooling level s has high innate ability equals p(s), where p(s)>p(s-1) for all s, s=2,…,S. That is, a higher schooling level translates into a higher probability that the worker has high innate ability. This can be the case either because innate ability and schooling are positively correlated which could occur because of education signaling like in Spence (1973), or because schooling enhances human capital. In the latter case it might be more appropriate to refer to θ_i as worker i's starting ability rather than worker i's innate ability.

A firm consists of two job levels and m divisions, where in each division a single worker is assigned to job level 2 while the number of workers assigned to job level 1 in a division is a choice variable for the firm. Note that a richer and more realistic specification would make the number of divisions in each firm a choice variable, but that enrichment would not change any of the testable predictions we derive in the next subsection. So for tractability reasons the number of divisions is exogenously determined.

If worker i is assigned to job 1 in period t, then the worker produces

(2) $$y_{i1t}= (1+v_{it})[c_1+c_2(\eta_{it}+\varepsilon_{it})],$$

where c_1 and c_2 are constants known to all labor market participants and ε_{it} is a noise term drawn from a normal distribution with mean 0 and variance σ^2. The term v_{it} equals v, v>0, if the worker was employed at the firm in the previous period and zero otherwise (so all young workers assigned to job 1 in period t are characterized by $v_{it}=0$). The term v_{it} thus captures that workers acquire firm specific human capital as they gain experience at a firm and this human capital increases productivity at job level 1.

It is assumed that firm specific human capital is required to produce at the high level job, so only an old worker with previous experience at the firm is ever assigned to job 2. If old worker i with previous experience at firm k is assigned to job 2 at division j in firm k in period t, then the worker produces

(3) $$y_{ijkt}=g(n_{jkt})(\eta_{it}+\varepsilon_{it}),$$

where n_{jkt} is the number of level 1 workers employed in division j in firm k in period t. We assume g(0)=G, g'>0, and g''<0. The assumption g'>0 captures the scale-of-operations effect – formally, the partial derivative of y_{ijkt} with respect to η_{it} is increasing in n_{jkt}. We also assume that g''<0 and that the concavity of the g(·) function is such that firms are "small" relative to the population of workers, or equivalently "many" firms operate in equilibrium. We come back to this assumption below.[10] We further assume that $G>(1+v)c_2$ which ensures that each firm finds

[10] We also assume that the g(·) function is such that in each period a firm hires at least m total young workers so that it is able to fully staff its managerial positions by promoting from within in the following period.

it profitable to assign its retained old workers with the highest expected effective abilities to the job 2 positions (see Waldman (1984a) for a related discussion).[11]

At the beginning of a worker's career, a worker with schooling level s is known to be of innate ability θ_H with probability p(s) and of innate ability θ_L with probability 1-p(s). Learning takes place at the end of the worker's first period in the labor market when the realization of the worker's output for that period becomes common knowledge. The presence of the noise term in equation (2) means that learning at the end of this first period is incomplete, i.e., there is updating after the worker's first period output is realized but after this updating there is still uncertainty concerning the worker's true innate ability.

Workers and firms are assumed to be risk neutral and there is no discounting, while workers face no mobility costs and firms face no hiring or firing costs. To make the model consistent with standard wage determination at most firms we assume spot-market contracting. Also, we focus on wages paid in advance of production rather than piece-rate contracts.

At the beginning of each period, all firms simultaneously offer each old worker a wage for that period and the worker then chooses to work for the firm that offers the highest wage. Further, if multiple firms are tied for this highest wage, the worker chooses randomly among the firms unless one is the worker's employer from the previous period in which case the worker remains with that firm. After this stage, firms hire young workers. Because within a schooling group all young workers look ex ante identical, in any period the young worker wage will vary with the schooling level but within a schooling group all young workers receive the same wage.

---

[11] Frequently the scale-of-operations effect, as in the Introduction, is described in terms of managerial ability positively affecting the outputs of workers below the manager in the hierarchy. An alternative specification to the one we employ that would be more similar to this description would be to assume that a manager directly produces zero, but each worker in division j in firm k in period t has an extra term in his or her production function equal to $g(n_{jkt})(\eta_{jkt}+\varepsilon_{j2t})/n_{jkt}$, where η_{jkt} is the effective ability of the manager in division j in firm k in period t. Although more consistent with typical descriptions of the scale-of-operations effect, in the type of learning environment that we consider in which firms observe workers' outputs each period such a specification would be analytically very challenging. The reason is that when a worker's output is realized in a period there would be learning both about the worker's ability level and the ability level of the manager the worker reports to. We are not familiar with any paper that successfully analyzes learning in such an environment. Our alternative specification simplifies the learning process because when a worker's (or manager's) output is realized the only learning that takes place involves that worker's (or manager's) ability level.

Specifically, we assume there are sufficiently many firms such that firms are price takers in the young worker labor market, so for any schooling level the wage equates supply and demand for young workers with that schooling level.[12]

Finally, by prior assumption we know that no old worker switches employers to work at job 2 in a new firm. We also assume that the magnitude of firm specific human capital in the production function for job level 1, v, is sufficiently large that no old worker switches firms to work at job level 1 at a new firm and no old workers are unemployed. See Ghosh (2004) and DeVaro and Morita (2012) for related analyses which allow for turnover.

Note that our specification includes a number of simplifying assumptions such as that firms have identical production functions and there is sufficient firm specific human capital so there is no turnover in equilibrium. Most of the results we focus on should be robust to relaxing these assumptions, so we are imposing these assumptions both to make the model more easily tractable and to make the arguments more transparent.[13]

B) Equilibrium and Testable Implications

In this subsection we describe equilibrium behavior in this model and derive testable implications. In the next subsection we discuss extensions. Also, we focus on behavior in equilibria in which there is no entry or exit after the first period.[14]

We start by describing the nature of equilibrium when S=1, i.e., workers do not vary in terms of schooling levels. This case is a bit simpler and analyzing this case first helps build intuition. Consider firm k in period t. The firm starts the period with some number of old workers it employed in job 1 in period t-1 when the workers were young, where these workers

---

[12] This specification for the young worker wage determination process is consistent with young workers not being able to observe prior wage and promotion practices of each firm so firms are not able to establish reputations concerning these practices.

[13] A number of empirical papers such as Neal (1995), Parent (2000), and Kambourov and Manovskii (2009) find a limited role for firm specific human capital in most real world settings.

[14] If firm specific human capital is sufficiently high there will never be exit. And if the managerial job is sufficiently important in terms of a division's total production then there will never be entry after the beginning of the game.

vary in terms of their period t-1 outputs. Let η_{it}^e be the market's expectation of worker i's effective ability at the beginning of period t. Given there is a single schooling level so all workers are ex ante identical, for every young worker i the firm employed in period t-1 we have $\eta_{it-1}^e=[p(1)\theta_H+(1-p(1))\theta_L]f(0)$. In turn, after observing period t-1 outputs the market updates its beliefs based on these outputs. Specifically, let θ_{it}^e denote the expected innate ability of old worker i in period t, i.e., $\theta_{it}^e=E(\theta_{it}|y_{i1t-1})$. Then the expected effective ability of old worker i in period t is given by $\eta_{it}^e=\theta_{it}^e f(1)$, where η_{it}^e is an increasing function of y_{i1t-1}. That is, because of Bayesian updating, the expectation concerning expected effective ability when a worker is old is positively related to the worker's output when the worker was young.

Now consider how firm k assigns these old workers to jobs in period t (remember we are focused on equilibria in which there is no turnover). Firm k has m divisions and a single job level 2 or managerial job in each division and effective ability is more valuable in job level 2 than in job level 1. So the firm promotes the m old workers with the highest values for η_{it}^e, i.e., the firm promotes the m old workers with the highest period t-1 outputs and the remaining old workers are assigned to job level 1. In turn, taking the market clearing wage for young workers as given, the firm hires the number of young workers such that each division has the efficient number of level-1 workers. The result is that, given the scale-of-operations effect captured by the g(n) term in equation (3), the number of workers in a division is positively related to the division manager's expected effective ability.

Proposition 1 formalizes results from the above discussion. Note that all proofs are in the Appendix.

Proposition 1: Consider any firm k and period t. If S=1, then the following describe firm k's behavior in period t.

i) The firm promotes the m old workers it employed in the previous period who produced the highest outputs, while the other workers are assigned to job 1.

ii) The wage increase from period t-1 to t for old workers at the firm is a strictly positive function of the period t-1 output.

iii) Promoted workers receive larger pay increases from t-1 to t than the workers who were not promoted.

iv) All young workers hired by the firm are assigned to job 1 and are paid the same wage.

v) Managerial span of control (weakly) increases with the manager's t-1 output.

We now consider how the nature of equilibrium changes when S>1, i.e., there are multiple education groups. Consider firm k in period t that starts the period with some number of workers it employed in job 1 in period t-1, where these workers vary in terms of their education levels and their period t-1 outputs.[15] Because the schooling level determines the ex ante probability a worker has high ability, for every young worker i with schooling level s_i the firm employed in period t-1 we now have $\eta_{it-1}^e=[p(s_i)\theta_H+(1-p(s_i))\theta_L]f(0)$. Further, we also again have that after observing period t-1 outputs the market updates its beliefs concerning the workers' innate ability levels. So the expected innate ability of old worker i in period t is given by $\theta_{it}^e=E(\theta_{it}|s_i,y_{i1t-1})$. In turn, the expected effective ability of old worker i in period t is still given by $\eta_{it}^e=\theta_{it}^e f(1)$. The difference is that now the expectation concerning expected effective ability when a worker is old increases with y_{i1t-1} holding s_i fixed and also increases with s_i holding y_{i1t-1} fixed.

Now consider firm k's assignment decisions in period t. As in the single education group case, the firm promotes the m old workers with the highest values for expected effective ability to the level-2 or managerial jobs and the remaining old workers are assigned to level-1 jobs. There is a difference, however, which is that this decison rule does not translate into promoting

[15] Although we do not show it formally, in our model the return to hiring a young worker with a high education level falls with the number of other young workers with high education levels employed. The reason is that the probability the worker will be assigned to a managerial job in the following period is lower when the firm has a large number of other young workers with high education levels. The result is that, rather than a firm hiring a homogeneous group of young workers, in our model the typical case is that a firm hires a heterogenous group of young workers.

the m old workers who produced the highest outputs in period t-1. Because expected effective ability is a function of both the period t-1 output and the schooling level, there can be a pair of workers only one of whom is promoted where the promoted worker had a lower t-1 output but a higher education level. Further, because of the scale-of-operations effect, a manager's span of control is again a positive function of the manager's expected effective ability. In this case this means that span of control will be positively related to output in t-1 holding the schooling level fixed, and will also be positively related to the schooling level holding output in period t-1 fixed. And since span of control is a positive function of the manager's expected effective ability, there is also a positive correlation between span of control and managerial wages.

Proposition 2 formalizes results from the above discussion.

Proposition 2: Consider any firm k and period t. If S>1, then the following describe firm k's behavior in period t.

i) The firm promotes the m old workers it employed in the previous period with the highest values for expected effective ability, while the other workers are assigned to job 1.

ii) Within a schooling group the old workers promoted are the ones who produced the highest outputs, but there can be pairs of old workers where only one is promoted and this worker produced less in t-1 but has a higher education level.

iii) The period-t wage for old workers at the firm and the wage increase from t-1 to t for old workers at the firm both increase with the t-1 output holding education fixed, while the wage for old workers increases with education holding the t-1 output fixed.

iv) Within a schooling group, promoted workers receive larger pay increases from t-1 to t than the workers who were not promoted.

v) All young workers hired by the firm are assigned to job 1, where the young worker wage increases with the worker's education level.

vi) Managerial span of control (weakly) increases with the manager's t-1 output holding education fixed, and also (weakly) increases with the manager's education level holding the t-1 output fixed.

vii) Within a schooling group, managerial span of control is positively correlated with the managerial wage.

The above discussion and Proposition 2 tell us that our model characterized by the scale-of-operations effect and learning yields a number of testable implications. First, wages for both young and old workers should be increasing functions of worker schooling levels. Second, within a schooling group, promoted workers should be those who performed better prior to promotion. Third, within a schooling group, performance should be positively related to subsequent wage increases and there is a corresponding prediction that promoted workers receive larger wage increases than those that are not promoted. Fourth, a manager's span of control should be positively related to prior performance holding the schooling level constant. Fifth, span of control should also be positively related to the schooling level holding performance constant. Sixth, managerial span of control should be positively related to the wage paid.[16]

In summary, in this subsection we have incorporated the scale-of-operations effect into a Gibbons and Waldman (1999a,2006) type model of job assignment, human capital acquisition, and symmetric learning. The model captures results concerning wage and promotion dynamics similar to those found in the earlier Gibbons and Waldman analyses and also captures new results concerning the determinants of span of control. The basic idea is that wages, wage changes, probability of promotion, and span of control are all related to a worker's expected effective ability, so the determinants of expected effective ability – in this model schooling and performance – are also determinants of wages, wage changes, promotions, and span of control.

---

[16] All of these implications would also hold in the absence of learning, i.e., if each worker's innate ability was known with certainty when the worker entered the labor market. We come back to this issue in the next subsection where we discuss extensions.

C) Extensions

In the previous subsection we formally analyzed a model with two job levels and workers whose labor market lives are two periods. In this subsection we discuss three extensions of the model. We begin with a discussion of what happens when workers are in the labor market for more than two periods. We then consider what happens when there are more than two job levels in addition to workers being in the labor market for more than two periods. We end this subsection with a discussion of what happens when firm size changes over time.[17]

Suppose that everything is the same as in the model considered in the previous subsection except that labor market careers last T periods rather than two and the equation for expected effective ability is changed to accommodate the longer labor market lives. Specifically, we now have the following. First, equation (1) still determines expected effective ability, where now we assume $f(T-1)>f(T-2)>\ldots>f(1)>f(0)>0$. Second, worker i's output in job 1 in period t is still given by equation (2) and output in job 2 in period t for workers with at least one prior period of experience at the firm is still given by equation (3).

This model yields results similar to those found in the previous subsection. For example, focusing on workers assigned to job 1 in period t-1 of a given schooling level, labor market experience, and performance history, the workers promoted to job 2 in period t will be those with higher period t-1 outputs. Further, for those promoted span of control will also be positively related to performance in t-1. Similarly, education will also affect probability of promotion and span of control in basically the same way it did in the previous subsection.

One interesting aspect of this extension is that it allows us to distinguish between our approach with symmetric learning and a model where firms have full information about each worker's effective ability. As mentioned in footnote 16, a model characterized by full information can capture most or all of the testable implications derived in the previous

[17] In a mathematical supplement available from the authors upon request we provide formal analyses that match these extensions: i) a three-period version of the model; ii) a version with three job levels; and iii) a version of the model in which one of the firms grows from t-1 to t.

subsection. For example, in such a model wages and probability of promotion would vary positively with effective ability. So if effective ability positively depends on schooling, then in this type of alternative model wages and probability of promotion will depend positively on schooling just like they did in our model with symmetric learning.

But now consider the model just analyzed in which careers last T periods rather than two. In this model wage changes and changes in span of control are determined by changes in expected effective ability. So wage changes and changes in span of control are correlated with the most recent performance even after controlling for education, labor market experience, firm tenure, and the starting value for the variable, because the most recent performance is correlated with changes in expected effective ability. But in a world characterized by full information this is not the case because firms know each worker's innate ability, so effective ability changes with labor market experience and firm tenure but after controlling for these factors, education, and the starting value for the variable, changes in expected effective ability should be uncorrelated with the most recent performance.

Note that another prediction one could derive from this T-period extension of our model is that predictions derived from the learning aspect of the model should become less important as workers get old. The logic is that as firms observe a worker's outputs early in worker's careers they learn about the worker's ability so predictions derived from the learning component of the model become unimportant late in careers. However, if instead, consistent with the empirical findings in Kahn and Lange (2014), we more realistically assumed that worker ability moved in a stochastic fashion as workers age, then learning would remain important even late in workers' careers.

Now suppose that we extend this model further so that there is a third job level in each firm's hierarchy. Specifically, each firm continues to have m divisions with a single manager at the top of each division who is now defined to be in the division's single level-3 job. What is new is that each division now has an endogenously determined number of subdivisions, where each subdivision is characterized by a single manager defined to be in the subdivision's single level-2 job and an endogenously determined number of workers reporting to the subdivision's manager. Further, assume that the production functions are analogous to those in the two-level model above so that the scale-of-operations effect applies both to a level-2 manager in terms of the number of workers who report to that manager and to a level-3 manager in terms of the number of level-2 managers who report to this level-3 manager.

Analysis of this model would yield results similar to those discussed above for firms with two job levels, except these results would now apply to multiple levels of the hierarchy. For example, holding job level, schooling level, labor market experience, and previous performance fixed, it will be the workers with high current performance who will be promoted in the next period for workers both in job level 1 and job level 2 (since there are only three job levels in this model workers in job level 3 cannot be promoted). Similarly, holding the same set of factors fixed and focusing on individuals who are on levels 2 or 3 both in this period and the next period, we get the prediction that current performance will be positively related to the change from this period to the next period in the managerial span of control. We also now get the additional prediction that current span of control is positively related to the probability of subsequent promotion. The idea is that span of control and expected effective ability are positively related, so a manager with a higher current span of control has a higher expected effective ability and thus a higher probability of subsequent promotion.

Note that an alternative specification that would make similar predictions is that, instead of having the scale-of-operations effect apply to the level-3 manager in terms of the number of level-2 managers who report to this level-3 manager, the scale-of-operations effect applies in terms of the total number of workers below this level-3 manager in the firm's hierarchy. With this in mind, in the empirical analysis we conduct the span of control tests both in terms of the number of workers directly reporting to a manager and in terms of the total number of workers below the manager in the firm's job hierarchy.

Let us also briefly discuss what would happen if firm size changes over time. In the main model analyzed in the previous subsection all firms were similar in size and that size did not change over time (at each date every firm had two job levels and m divisions). And we only moved away from this a little in the first two extensions – in particular, in the second extension the number of subdivisions reporting to a level 2 manager was assumed to be endogenous but the number of job levels and the number of divisons was assumed to be fixed. In our empirical analysis, however, the firm we focus on is growing during the time period of our study.

But we do not believe this is a concern. All the main predictions that we focus on generalize to the case of a growing firm. For example, consider the main model analyzed in the previous subsection with the single change that from t-1 to t one of firms increases the number of divisions from m to m+Δ, Δ>0. All of the main predictions would continue to hold even focusing on that firm during the period in which the growth occurred. That is, wage increases would rise with performance, promotion probabilities would depend on both performance and education, and span of control would rise with performance and education. So although our main theoretical analysis does not perfectly match the firm we look at in the sense we do not focus on a growing firm, we do not believe this creates any problems in terms of testing the theory.

In summary, in this subsection we have have argued that extending the model by allowing for longer work lives, more jobs, and a growing firm does not change the basic predictions of the model where workers have two-period careers, there are two jobs, and firms are not growing. But it does result in the new prediction that changes in span of control should be positively correlated with performance.

IV. DATA

In this section we describe our data and present some basic facts about the firm. In the next section we provide tests of our model.

We received confidential performance data from one large European Union “high tech” manufacturing firm that produces in various countries around the globe and sells its products in almost every country. The data covers the calendar years 2006 to 2011, where each year’s data were collected in the spring of the following year. The firm employs a 1 to 5 scale for its evaluation ratings, where 1 denotes the lowest performance and 5 the highest.[18] Performance evaluation is relatively recent in the company. While the firm initially focused its attention on the top managers of the firm, in later years the ratings were extended to include almost all of the firm's white-collar workers with wider global coverage in later years (in particular the US, China and Japan). Our focus will be on Danish workers since that is the location where the data is most complete and the distribution of performance ratings varied significantly across countries.[19]

Individuals are assessed by their direct supervisor. An assessment is made in December at the end of each year which covers that year's performance. The assessment is entered into the HR system and also communicated to the worker in a face-to-face interview with the supervisor. The supervisor and the worker also discuss an action plan for the following year which includes various performance targets for the worker. At this firm HR managers spend significant time and effort ensuring that supervisors understand the importance of the evaluation process and provide fair ratings.

We combine the data on performance ratings with data from confidential monthly personnel records we received for all workers from January 2003 (January 1997 for Danish workers) to December 2011. In this way we create a panel dataset for the years 2006 to 2011 that includes for each observation the worker's firm tenure, age, salary, bonus, cost center category, job level, nationality, gender, schooling level, a promotion indicator variable (1 if the worker was promoted that year and 0 if not), and performance evaluation.[20]

---

[18] To be precise, 1 means the worker does not meet expectations, 2 means the worker's performance approaches expectations and goals, 3 means performance meets expectations, 4 is for performance that exceeds expectations and goals, and 5 is for outstanding performance.

[19] For example, the percentage of workers who received a 4 or 5 was roughly twice as high in Denmark than in Japan, while this percentage for each of China and the US is between the values for Denmark and Japan. Also, the percentage of workers who received a 1 or 2 was much higher in Japan than in any of the other countries. There are various possible explanations for these differences. They could be due to differences in the quality of the firm's labor force across the countries, differences in organizational structures, or cultural differences that result in differences in how workers are evaluated across the countries.

[20] Cost center categories refer to the functional divisions of the firm. These are administration, administration in production, production, sales and marketing, and research and development.

What is distinctive about this dataset which sets it apart from other similar datasets based on firms' personnel records is that we were also provided information about the firm's chain of command. Before describing this part of the dataset and how we used it we need to describe the hierarchical structure of the firm. Until relatively recently, the firm was organized into five hierarchical layers. Workers at job level 0 are the non-managerial employees. Workers at job level 1 are lower management and these workers are referred to as managers. Job level 2 refers to middle management and these workers are referred to as VPs. Job level 3 refers to upper management and workers at this level are referred to as SVPs. And workers at job level 4 are the top management and these workers are referred to as EVPs. Also, to cope with significant growth during the last decade the firm recently introduced two new levels. Team leaders or assistant managers are between the old job levels 0 and 1 (above non-managerial employees but below managers), while corporate vice presidents, CVPs, are between the old job levels 2 and 3 (above VPs but below SVPs).

As mentioned above, we were provided information about the firm's chain of command. That is, for each individual and each year we were given the names of all the individuals directly above the worker in the firm's hierarchy. For example, in the latter part of the dataset, for each worker at level 0 we know the names of the team leader, manager, VP, CVP, SVP, and EVP, as well as the name of the department, the name of the subsidiary, and the geographic area where it operates. We use this information to construct span of control measures for workers above level 0 (see Smeets and Warzynski (2008) for more detail). In our empirical analysis of span of control, we focus on team leaders, managers, VPs, and CVPs.

We define span of control in two ways. An individual's direct span of control is the number of subordinates who report directly to the individual, while an individual's full span of control is the total number of subordinates underneath the individual in the firm's hierarchy. For example, consider a three level job ladder with a single individual on level 3, two individuals on level 2 who report to the level 3 manager, and four individuals on level 1 where two of these four report to one level two manager and the other two report to the other. For each manager on level

2 direct span of control and full span of control are the same and each equals two. For the single level 3 manager direct span of control equals two and full span of control equals six.

Table 1 provides summary statistics by job level for our main variables of interest, including our two span of control measures, for each of the levels in our dataset. We can see that firm tenure, experience, education, and performance increase with the job level, especially when we focus on the managerial levels. Workers tend to have relatively high firm tenure and experience because many of them are production workers with little probability of promotion. The promotion rate varies by job level and is particularly high for VPs. This is at least partly explained by the fact that the CVP layer was introduced during our period of analysis and most CVPs were promoted from the VP level. The bottom panel shows more detailed summary statistics concerning span of control for individuals with managerial responsibility. The direct span of control tends to decrease along the hierarchy, declining from around 25 subordinates at the manager level to less than 2 at the CVP level. On the other hand, the full span of control increases along the hierarchy, as expected. We also see that variation in span of control within level increases at higher levels of the job ladder which could be due to learning causing variation in perceived managerial ability to be larger at higher managerial levels.

V. EMPIRICAL TESTS

In this section we present various empirical tests many of which are tests of the theory developed in Section III. The theory makes predictions about three sets of variables: i) wages and wage changes; ii) probability of promotion; and iii) span of control and changes in span of control. In the first subsection we consider wages, wage changes, and probability of promotion. In the second subsection we consider span of control and changes in span of control.

A) Wages, Wage Changes, and Probability of Promotion

We start with tests concerning wage determination. The theory predicts that wages should be related to job level (since wages rise with promotions), human capital variables, and the worker's performance history. In considering the determinants of the wage in our dataset we employ an ordinary least squares regression where the dependent variable is the log of the worker's wage in period t. Our independent variables include the worker's job level in period t which we capture with a set of indicator variables (the lowest job level is the omitted category), while our human capital variables are experience, experience squared, firm tenure, firm tenure squared, and education indicator variables.[21] We capture the worker's performance history by including the worker's performance rating six months prior to t and in one test we also include an earlier performance measure. We also include year and cost center indicator variables.

The first column of Table 2 reports results not including the performance variables. The results concerning job level and education are consistent with the theory. Specifically, the coefficients on the job level indicator variables increase with job level and in many cases differences between coefficients are statistically significant, while the coefficients on the education variables are also consistent with higher levels of education being correlated with higher wages.

The coefficient on the experience variable is positive and statistically significant while the coefficient on the experience squared variable is negative and statistically significant. This result is consistent with human capital theory if there is depreciation over time and/or investment levels fall with age (see Ben-Porath (1967) for an analysis). The coefficient on the tenure variable is positive, although not statistically significant, while the coefficient on the tenure squared variable is positive and statistically significant. So the effect of tenure on the wage does not have the same concave shape as does experience.

In columns 2 and 3 we add performance evaluations, where column 2 adds the most recent performance evaluation and column 3 adds to the column 2 regression the prior performance measure. Coefficients on the other variables are mostly qualitatively unchanged, while the coefficients on the performance measures are positive and statistically significant.[22,23]

---

[21] Specifically, one education indicator variable is for workers with bachelor's degrees, a second is for workers with advanced degrees (master's degree or PhD), while other education outcomes are grouped into the omitted category.
[22] The only noticeable change concerning the coefficients on the other variables is that the coefficient on the tenure variable is now negative in each of columns 2 and 3, although it is not statistically significant in either regression.

For example, as reported in column 3 of Table 2, when the most recent performance measure and the prior performance measure are both included, the coefficient on each is positive and statistically significant.

Our next set of tests concerns wage changes. The theory predicts that wage changes should be positively related to the most recent performance evaluation and whether the worker was promoted in the most recent period. To investigate the determinants of wage changes we employ an ordinary least squares regression where the dependent variable is the change in the log wage from t-1 to t. Our independent variables are an indicator for whether the worker was promoted in period t, the worker's most recent performance evaluation, and the log wage in t-1. We include the log wage in t-1 to capture the possibility that average percentage wage changes vary with the t-1 wage. We also include year and cost center indicator variables.

The first column of Table 3 does not include performance variables, the second column does not include a promotion variable, while the third and fourth columns include both performance and promotion variables. In each column we find a statistically significant negative coefficient on the log wage t-1 variable which means that percentage wage increases are on average smaller for workers who are already being paid a high wage. We also find that being promoted is positively related to the change in the log wage, that performance is also positively related to the change in the log wage, and that the positive coefficient for the most recent performance variable is basically double the coefficient on the prior performance variable.[24]

---

[23] In all of the tests we report that employ a performance measure we use the raw performance score. We have also investigated alternative approaches such as measuring performance as the difference between an individual's raw performance score and an average based on job level and period. In general our results are robust to employing alternative performance measures.

[24] Note that one could argue that since in the theory promotions are associated with large wage increases because they are associated with high performance, in columns 3 and 4 there should be little or no correlation between the promotion variable and wage increases because we control for performance. But given that the performance measure has only five categories which is likely too coarse to fully capture true performance, promotions should be positively correlated with true performance even when measured performance is controlled for, i.e., the theory is in fact consistent with a positive coefficient on the promotion variable in columns 3 and 4. Alternative theories concerning why promotions are associated with large wage increases are that these wage increases are prizes in a tournament such as in Lazear and Rosen (1981) or promotions serve as signals such as in Waldman (1984b). See the conclusion for a related discussion.

In our next set of tests, we consider the probability of promotion. In the theory, promotion is determined by both the education level and performance. That is, holding performance fixed the probability of promotion rises with the education level, while holding education fixed the probability of promotion rises with performance. To investigate the probability of promotion we conduct probit tests where the dependent variable is an indicator variable that takes on a value of one if the worker was promoted that period and zero if not, where independent variables include performance measures, indicators for the worker's education, experience and tenure variables, and year and cost center indicator variables. Our sample includes all observations of individuals who were workers, team leaders, managers, and VPs in the prior period, where we include indicator variables which capture whether the individual was a team leader, manager, or VP prior to the promotion decision (being a worker prior to the promotion decision is the omitted category).

Results are reported in Table 4. In column 1, we report marginal effects from a probit regression that includes only the most recent performance measure, while in column 2, we add as an explanatory variable the prior performance measure. One result found in both regressions is that the coefficient on the manager indicator variable is negative, although statistically significant only in column 2, while the coefficient on the team leader indicator variable is positive and statistically significant in both columns. Also, consistent with the theory, we find that both the most recent performance and performance from one period earlier are positively related to the probability of promotion. In terms of how education is related to probability of promotion, in column 1 we find that both coefficients on the education indicator variables are positive and statistically significant, although inconsistent with the theory the two coefficients are roughly equal. In column 2 both of the coefficients are positive and statistically significant, but now consistent with the theory the coefficient on the advanced degree indicator variable is larger than the coefficient on the bachelor's degree indicator variable (although the difference between the coefficients is not statistically significant).[25]

So overall our results in this subsection are consistent with the theory presented in Section III. First, wages are positively related to performance, education, and job level. Second, wage increases are positively related to performance and receiving a promotion. Third, performance and education are positively related to the probability of promotion, although the evidence is mixed concerning whether having an advanced degree improves the promotion probability relative to having a bachelor's degree. Note, however, although the results are in general consistent with Section III's model, they can also be explained by Gibbons and Waldman (1999a,2006). In the next subsection we focus on predictions from our model that are not found in the Gibbons and Waldman papers.

B) Span of Control

In this subsection we consider tests related to span of control, where we begin with tests concerning the determinants of span of control. The scale-of-operations effect which is a driving force in our theoretical model predicts that a manager's span of control should be positively related to the manager's ability level (expected ability level in the model). In turn, translating this into a testable prediction yields that an individual's span of control should be positively related to the manager's performance history and the manager's education level.

To investigate our predictions concerning the determinants of span of control we conduct ordinary least squares regressions using a sample consisting of all observations of individuals who were team leaders, managers, VPs, and CVPs. The dependent variable in our regressions is the log of either the individual's direct span of control or full span of control, while independent

[25] To further investigate the difference between columns 1 and 2 concerning the coefficients on the education variables, in column 3 we rerun the column 1 test but use only the 4,499 observations from column 2. Here we find that, as in the column 2 regression, the coefficient for advanced degrees is larger than the bachelor's coefficient (and all the other coefficients are also qualitatively similar to the column 2 coefficients). So it seems that the differences between the column 1 and column 2 regression results concerning the education variables are mostly driven by differences in the sample rather than differences in the specification.

variables include performance measures, indicator variables for the various job levels, education indicator variables, year indicator variables, and cost center indicator variables.

Results are reported in Tables 5 and 6, where Table 5 reports results for our direct span measure and Table 6 for our full span measure. In column 1 of each table we report results for ordinary least squares regressions where only the individual's most recent performance measure is included. The coefficients on the job level indicator variables match how average span varies across levels as found in Table 1. In terms of the theoretical predictions, in each regression the coefficient on the performance variable is positive and statistically significant at the one percent level, while the coefficient on the bachelor's indicator variable is positive but not statistically significant. Additionally, in column 1 of Table 5 the coefficient on the advanced degree indicator variable is negative but not statistically significant, while in Table 6 this coefficient is negative and statistically significant at the five percent level. We can translate the coefficient on the performance variable in the first column of each of Tables 5 and 6 into a prediction concerning how positive performance affects span of control. For example, the coefficient in the first column of Table 5, 0.077, indicates that managers who are rated as exceeding rather than meeting expectations (a rating of 5 rather than 3) have on average 3.8 more workers reporting to them which is approximately a 20 percent increase over the average.

In column 2 of each table, we add as an explanatory variable the lagged value for the individual's performance. In Table 5, the main result is that the coefficient on each performance variable is positive, where the coefficient on the most recent performance variable is significant at the one percent level and the coefficient on the prior performance variable is significant at the five percent level. Also, the coefficient on the most recent performance variable is larger. In Table 6 the pattern is similar. The coefficient on the most recent performance variable is positive and statistically significant at the one percent level, while the coefficient on the prior performance measure is positive but smaller than the coefficient on the most recent performance variable and is statistically significant at only the ten percent level. One additional result worth pointing out is that in column 2 of Table 6 the coefficient on the advanced degree indicator variable is negative like in column 1 but it is not statistically significant.

To further investigate why the education results in the first two columns of each of Tables 5 and 6 do not support the theory, in column 3 of each table we add to the column 2 specification interactions of the team leader indicator variable with the two education variables. The results indicate that having a bachelor's degree increases span of control for individuals in managerial positions above the team leader level, but for the team leader level having a bachelor's degree has a negative effect on span of control. For those with advanced degrees, we find no statistically significant effect on span of control for individuals in managerial positions above the team leader level or for those at the team leader level.[26] So overall our tests concerning span of control support the predictions concerning performance, but the results concerning education are mixed.

In our next set of tests we consider changes in span of control. To investigate changes in span of control we begin with an ordinary least squares regression where changes in the logs of direct span of control and the full span of control are the dependent variables, while independent variables are the starting values for the logs of direct span of control and full span of control, indicator variables for job level, performance measures, year indicator variables, and cost center indicator variables. Results are reported in Tables 7 and 8, where Table 7 focuses on direct span of control and Table 8 on full span of control.[27] In the first column in each table we only include the most recent performance measure. One result in each regression is that the coefficient on the starting value for the log of the span of control measure is negative and statistically significant which means that the percentage increase in the log of the span of control measure is on average smaller when the starting value for the measure is larger. The other main result is that the

---

[26] We have also looked at tests with interactions between indicator variables for other job levels and the education variables. These further tests indicate no statistically significant differences concerning how education affects span of control across levels other than the results concerning the team leader level found in column 3 of Tables 5 and 6.
[27] Because of outliers in this set of tests, the distribution is trimmed below the first percentile and above the ninety ninth percentile.

coefficient on the performance variable is positive and statistically significant in each regression which is exactly what one would expect if changes in span of control are driven by changes in expected effective ability as predicted by the theory.

In column 2 of each table, we add the lagged performance measure to the regression. The main result is that in each regression the coefficient on the most recent performance measure remains positive and statistically significant while the coefficient on the lagged performance variable is statistically insignificant. This is consistent with span of control changing quickly in response to changes in expected effective ability.

In columns 3 and 4 of Tables 7 and 8, we add the education indicator variables. In each regression the coefficients on the education indicator variables are statistically insignificant, the size or sizes of the coefficients on the performance variables are basically unchanged, and the statistical significance of the coefficient on the most recent performance measure is also unchanged. Columns 3 and 4 in each table tell us that the positive correlation between change in the log of our span of control measure and the most recent performance measure is not driven by a positive correlation between performance and education.[28]

We now consider the relationship between wages and span of control. The theory predicts a positive relationship between the two because both are positively related to a worker's expected effective ability. To test this prediction we conduct the same tests as in Table 2 but add variables that capture interactions between a log span measure and the various managerial levels (the team leader level is the omitted category), where Table 9 focuses on the log of the direct span of control and Table 10 on the log of the full span of control. The two tables yield similar results. That is, each table provides clear support for the prediction at the manager, VP, and CVP levels in that for each of these levels the relevant coefficients are positive and statistically

[28] Another alternative explanation for our findings in columns 1 and 2 of each table is that division level profitability, controlling for true managerial performance, is positively correlated with the managerial performance rating and future division level growth. Unfortunately, we are unable to test for this possibility because we do not have division level profitability data.

significant at the one percent level. But at the team leader level there is no support for the prediction since the coefficients are negative and statistically insignificant.

The last set of tests we consider concerns the relationship between the current span of control of a worker who is currently in a managerial position and the probability the worker is subsequently promoted. Similar to the previous set of tests, the theory predicts that the two should be positively correlated because both are positively related to expected effective ability. To test this prediction we conduct the same tests as in Table 4 but add the interaction variables employed in Tables 9 and 10, where Table 11 focuses on the log of the direct span of control and Table 12 focuses on the log of the full span of control.

The two tables are similar. The manager and VP levels both exhibit results consistent with the theoretical prediction. That is, in both tables there is evidence of a positive and statistically significant relationship for each of these managerial levels between current span of control and probability of subsequent promotion, where the correlation is stronger for the VP level than for the manager level. On the other hand, for the CVP level there seems to be no correlation while at the team leader level there is some evidence of a negative correlation.

That there is no correlation for the CVP level, at least in Table 11, is not surprising. As reported in Table 1, the average direct span of control at the CVP level is only 1.69 so there may not be enough variability in the direct span of control at the CVP level to accurately identify the relationship. The more surprising finding is the evidence for a negative relationship found at the team leader level between current span of control and the probability of subsequent promotion. But note that this result is in a sense consistent with results in the third column of Table 5 and results in Tables 9 and 10 in that it suggests that the theoretical predictions hold less well at the team leader level.

VI. DISCUSSION

As briefly mentioned at the end of Subsection V.A, the results in that subsection can be explained by the models (or a slight variant of the models) investigated in Gibbons and Waldman

(1999a,2006). Like the model investigated here, those models combine job assignment, human capital acquisition, and symmetric learning in an exploration of wage and promotion dynamics in internal labor markets. They are consistent with our findings that wages are positively related to performance, education, and job level, wage changes are positively related to performance and having received a promotion, and that the probability of promotion is positively correlated with performance and education.

But the Gibbons and Waldman (1999a, 2006) models do not say anything about span of control. In those models there are two or three job levels and like the model in Section III the return to ability is higher at higher job levels. But in those models there is no sense in which a worker assigned to a higher level job supervises workers at a lower level, so those models do not capture the idea of span of control. So it is not the case that any of our empirical findings in Subsection V.B are inconsistent with testable predictions of the Gibbons and Waldman models, but rather those models are simply silent concerning how span of control should work in real world firms. Thus, to the extent our findings in Subsection V.B concerning span of control are consistent with testable predictions of our model, this is evidence that our extension of the Gibbons and Waldman framework provides a more complete picture of the operation of internal labor markets than the Gibbons and Waldman models.

The first finding in Subsection V.B is that span of control is positively correlated with both the most recent performance measure and a lagged performance measure. We also found evidence of a positive correlation between education and span of control for managerial levels above the team leader level, although having a masters or higher degree does not seem to increase span relative to that of managers with bachelor's degrees. These findings are generally consistent with the scale-of-operations effect being important at this firm. That is, the scale-of-operations effect predicts that span of control should be higher for higher ability managers, so span of control being clearly positively correlated with performance and to some extent positively correlated with education is consistent with the scale-of-operations effect since performance and education should both be positively correlated with managerial ability. But these findings tell us nothing about whether learning is an important determinant of span of control at our firm. Or another way to put this is that these findings can be explained by static analyses such as in Lucas (1978) and Rosen (1982) and do not require the learning component of our theoretical approach.

Our second finding in Subsection V.B is that change in span of control is positively correlated with the most recent performance measure and this is the case even after controlling for education. This finding is consistent with learning being an important determinant of span of control since, if learning was unimportant, then perceptions about ability would not change with performance so changes in span of control which should depend on changes in such perceptions would be uncorrelated with performance. That is, dynamic versions of Lucas (1978) or Rosen (1982) that do not incorporate learning would not be consistent with our second finding. In the absence of learning, high performance in a period does not translate into changes in beliefs about the worker's ability. So there would be no prediction that high performance should be followed by a change in the span of control.

So overall we find results consistent with the Gibbons and Waldman (1999a,2006) approach that combines job assignment, human capital acquisition, and symmetric learning, results consistent with the scale-of-operations effect and span of control as in the early papers of Lucas (1978) and Rosen (1982), and results that can be explained by combining the two approaches like we do in Section III's model. We also do not know of any alternative theoretical approach capable of explaining our findings, although as we discuss in the Conclusion, the pure symmetric learning assumption we employ in our theoretical model may not fully capture how learning works in our firm.

VII. CONCLUSION

A standard idea in the theoretical literature concerning hierarchies is the scale-of-operations effect, i.e., the idea that the return to managerial ability is higher the more resources the manager influences with his or her decisions. This theoretical idea leads to various testable implications including that higher ability managers should have a larger span of control. And although the empirical literature on organizations finds evidence consistent with some predictions that follow from the scale-of-operations effect, predictions concerning span of control have not been investigated.

In this paper we first extended the theory by looking at the scale-of-operations effect in a model in which there is uncertainty concerning workers' abilities when they enter the labor market and firms learn about these abilities as careers progress. This model yields a large number of testable predictions some of which are consistent with models focused solely on the scale-of-operations effect, some of which are consistent with models focused solely on learning, and some of which follow from combining the scale-of-operations effect with learning. These various testable predictions concern wages, wage changes, promotion probabilities, and span of control.

We then empirically investigated these predictions using a unique single firm dataset that contains detailed information concerning the reporting relationships at the firm. Our results support most of our theoretical predictions. We find results concerning wages, wage changes, promotion probabilities, and span of control all consistent with our theoretical model. Most interestingly, we find results that follow from combining the scale-of-operations effect with learning about worker ability. For example, in a world characterized by the scale-of-operations effect but no learning, performance should have little or no effect on changes to a manager's span of control because with no learning performance does not change beliefs concerning the manager's ability. But with learning there should be a positive correlation because performance is positively correlated with subsequent changes in expected effective ability and changes in expected effective ability are correlated with changes in span of control. Our empirical analysis finds a positive correlation between performance and subsequent changes in span of control as predicted by our analysis that combines the scale-of-operations effect with learning. Our conclusion, therefore, is that both the scale-of-operations effect and learning are important factors in the operation of real-world hierarchies and, further, that the interaction of these two factors which has not previously been explored is in its own right an important factor.

There are a number of directions in which the analysis in this paper could be extended. One that we think has particular promise is to introduce an element of asymmetric learning into the analysis. In this paper we have focused on symmetric learning similar to the approach taken, for example, in Harris and Holmstrom (1982), Holmstrom (1982), and Gibbons and Waldman (1999a,2006). In a world of symmetric learning all firms are equally informed about each worker's ability at any point in time. An alternative approach first explored in Greenwald (1979,1986) and Waldman (1984b) is that learning is asymmetric, i.e., only a worker's current employer directly receives information about the worker's ability, while other firms learn about the worker's ability by observing the actions of the current employer such as promotion decisions. There are numerous empirical studies that test for asymmetric learning in labor markets such as Gibbons and Katz (1991), Schoenberg (2007), Pinkston (2009), DeVaro and Waldman (2012), and Kahn (2013) and, on net, we believe the evidence supports asymmetric learning being important. We thus believe it would be interesting to incorporate an element of asymmetric learning into our theoretical analysis and then empirically investigate the new testable predictions that result.

APPENDIX

Proof of Proposition 1: As indicated in the set-up of the model, v is assumed to be sufficiently large that an old worker always stays at the firm that employed the worker when he or she was young. Given this, consider young workers at firm k in periods t-1 and t. As also indicated in the set-up, each firm is sufficiently small that it is a price taker in the labor market for young workers where the young worker wage equates supply and demand for young workers. Call the equilibrium young worker wage in periods t-1 and t, W_{t-1}^{Y} and W_t^{Y}. Since a young worker assigned to job 2 produces zero and adding a worker to a division always increases the aggregate output produced by a division's level 1 and level 2 workers, all young workers hired by firm k in each of periods t-1 and t are assigned to job 1. This proves iv).

Now consider period t. Let $z_{it-1}=(y_{i1t-1}-c_1)/c_2=\eta_{it-1}+\varepsilon_{it-1}$, where z_{it-1} is the signal about ability that can be extracted from observing young worker i's output in period t-1. By Bayes' rule we have

(A1) $$prob(\theta = \theta_H \mid z_{it-1}) = \frac{p(1)h(z_{it-1} - \theta_H f(0))}{p(1)h(z_{it-1} - \theta_H f(0)) + (1-p(1))h(z_{it-1} - \theta_L f(0))},$$

where $h(\cdot)$ is the density of ε_{it} in equation (2) which is normal with mean 0 and variance σ^2. This yields:

(A2) $$\frac{h(z_{it-1} - \theta_L f(0))}{h(z_{it-1} - \theta_H f(0))} = \exp[-\frac{1}{2\sigma_1^{\,2}}\{[z_{it-1} - \theta_L f(0)]^2 - [z_{it-1} - \theta_H f(0)]\}]$$

which is monotonically decreasing in z_{it-1}. Combining this with (A1) yields θ_{it}^{e} is strictly increasing in y_{i1t-1} which in turn yields that η_{it}^{e} is strictly increasing in y_{i1t-1}.

Given this, suppose that in period t the firm does not promote the m old workers who when young in period t-1 produced the highest outputs, i.e., there exists a pair of workers a and b such that $y_{a1t-1}>y_{b1t-1}$ where b is promoted but not a. Given $G>(1+v)c_2$, the firm could reverse the assignment, leave everything else the same, and because $\eta_{at}^{e}>\eta_{bt}^{e}$ it must be the case that expected profit would increase. So firm k in period t must promote the m old workers it employed in t-1 who produced the highest outputs. This proves i).

Given the wage determination process for old workers, in period t an old worker's period t-1 employer will simply match the highest alternative offer the worker receives. Since an old worker's productivity at an alternative employer is a strictly increasing function of the worker's value for η_{it}^e, this means the worker's wage is an increasing function of η_{it}^e which, since from above we know η_{it}^e is itself an increasing function of y_{i1t-1}, means the period t wage is a strictly increasing function of y_{i1t-1}. This proves ii). Since from above we also know that it is the old workers with the highest values for y_{i1t-1} who are promoted, this also means that promoted workers receive larger wage increases than the old workers who are not promoted. This proves iii).

Finally, consider two old workers in period t, a and b, assigned to job 2 by firm k where $\eta_{at}^e>\eta_{bt}^e$ and suppose that the number of level-1 workers in a's division is strictly smaller than the number in b's division. The firm could reverse the assignments of a and b, leave everything else the same, and because $\eta_{at}^e>\eta_{bt}^e$ and $g'>0$ it must be the case that expected profit would increase. So the number of workers in a's division must be at least as large as the number in b's division. In turn, since η_{it}^e is an increasing function of y_{i1t}, we now have that the number of level-1 workers in a division is weakly increasing in the value for y_{i1t-1} of the division's level-2 manager. This proves v).

<u>Proof of Proposition 2</u>: It is again the case that v is assumed to be sufficiently large that an old worker always stays at the firm that employed the worker when he or she was young. Given this, consider young workers at firm k in periods t-1 and t. Using an argument similar to that in the proof of Proposition 1 yields that in each period there are wage functions $W_{t-1}^Y(s)$ and $W_t^Y(s)$ which give young worker wages in t-1 and t as functions of the worker's schooling level, while all young workers hired by firm k in each of periods t-1 and t are assigned to job 1. Further, both $W_{t-1}^Y(s)$ and $W_t^Y(s)$ are strictly increasing with s because $p'>0$ means expected productivity when a worker is young increases with s and, based on results below, the expected rents a firm earns in

the following period from employing a young worker in either t-1 or t (which will be reflected in the current wage due to competition) also increases with s. This proves v).

Now consider the firm's decision concerning who to promote in period t. Suppose that the firm does not promote the m old workers with the highest values for η_{it}^e, i.e., there exists a pair of workers a and b such that $\eta_{at}^e>\eta_{bt}^e$ where b is promoted but not a. Given $G>(1+v)c_2$, the firm could reverse the assignment, leave everything else the same, and because $\eta_{at}^e>\eta_{bt}^e$ it must be the case that expected profit would increase. So firm k in period t must promote the m old workers it employed in t-1 with the highest values for η_{it}^e. This proves i).

Using the same argument as used to prove i), ii), and iii) of Proposition 1, one can show that within a schooling group promoted workers are the ones who produced the higher outputs in the previous period, the period t wage increases with the period t-1 output (this also implies that the wage increase goes up with the period t-1 output since the period t-1 wage is independent of t-1 output), promoted workers receive larger wage increases, managerial span of control (weakly) increases with period t-1 output, and span of control is positively correlated with the managerial wage.

So the only parts of Proposition 2 left to prove are the second parts of ii), iii), and vi). Generalizing (A1) for the case of multiple schooling groups yields (A3).

$$\text{(A3)} \qquad prob(\theta=\theta_H \mid z_{it-1}) = \frac{p(s)h(z_{it-1}-\theta_H f(0))}{p(s)h(z_{it-1}-\theta_H f(0))+(1-p(s))h(z_{it-1}-\theta_L f(0))}$$

Given the definition of z_{it-1}, (A3) immediately yields that, holding y_{i1t-1} fixed, the probability $\theta=\theta_H$ increases with s which in turn means that η_{it}^e increases with s. Given this, arguments above now yield that the wage for old workers increases with education holding the t-1 output fixed and the managerial span of control (weakly) increases with the manager's education level holding the t-1 output fixed.

So the only part of Proposition 2 left to prove is the second half of ii). From i) we know that the promoted workers are the ones with the highest values for η_{it}^e. But (A3) tells us that η_{it}^e increases with the schooling level holding the period t-1 output fixed, while an earlier argument showed that η_{it}^{e} increases with period t-1 output holding the schooling level fixed. So therefore there can be a worker promoted who has higher schooling, lower period t-1 output, and a higher value for η_{it}^{e} than a worker who is not promoted.

REFERENCES

Altonji, J.G. and C.R. Pierret (2001), "Employer Learning and Statistical Discrimination," *Quarterly Journal of Economics*, 116, pp. 313-350.

Baker, G., M. Gibbs, and B. Holmstrom (1994a), "The Internal Economics of the Firm: Evidence from Personnel Data," *Quarterly Journal of Economics*, 109, pp. 881-919.

Baker, G., M. Gibbs, and B. Holmstrom (1994b), "The Wage Policy at a Firm," *Quarterly Journal of Economics*, 109, pp. 921-955.

Ben-Porath, Y. (1967), "The Production of Human Capital and the Life Cycle of Earnings," *Journal of Political Economy*, 75, pp. 352-365.

Bloom, N., Garicano, L., Sadun, R. and Van Reenen, J. (2014), "The Distinct Effects of Information Technology and Communication Technology on Firm Organization," *Management Science*, 60, pp. 2859-2885.

Caliendo, L., F. Monte, and E. Rossi-Hansberg (2015), "The Anatomy of French Production Hierarchies," *Journal of Political Economy*, 123, pp. 809-852.

Calvo, G. and S. Wellisz (1978), "Supervision, Loss of Control and the Optimal Size of the Firm," *Journal of Political Economy*, 86, pp. 943-952.

Calvo, G. and S. Wellisz (1979), "Hierarchy, Ability, and Income Distribution," *Journal of Political Economy*, 87, pp. 991-1010.

Delmastro M. (2002), "The Determinants of the Management Hierarchy: Evidence from Italian Plants," *International Journal of Industrial Organization*, 20, pp. 119-137.

DeVaro, J. and H. Morita (2012), "Internal Promotion and External Recruitment: A Theoretical and Empirical Analysis," *Journal of Labor Economics*, 31, pp. 227-269.

DeVaro, J. and M. Waldman (2012), "The Signaling Role of Promotions: Further Theory and Empirical Evidence," *Journal of Labor Economics*, 30, pp. 91-147.

Dias da Silva, A. and B. van der Klaauw (2010), "Wage Dynamics and Promotions Inside and Between Firms," Jo*urnal of Population Economics*, 24, pp. 1513-1548.

Dohmen, T.J., B. Kriechel, and G.A. Pfann (2004), "Monkey Bars and Ladders: The Importance of Lateral and Vertical Job Mobility in Internal Labor Market Careers," *Journal of Population Economics*, 17, pp. 193-228.

Farber, H. and R. Gibbons (1996), "Learning and Wage Dynamics," *Quarterly Journal of Economics*, 111, pp. 1007-1047.

Fox, J.T. (2009), "Firm-Size Wage Gaps, Job Responsibility and Hierarchical Matching," *Journal of Labor Economics*, 27, pp. 83-126.

Garicano, L. (2000), "Hierarchies and the Organization of Knowledge in Production," *Journal of Political Economy*, 108, pp. 874-904.

Garicano, L. and T.N. Hubbard (2007), "Managerial Leverage is Limited by the Extent of the Market: Hierarchies, Specialization, and the Utilization of Lawyers' Human Capital," *Journal of Law and Economics*, 50, pp. 1-45.

Garicano, L. and T.N. Hubbard (2009), "Specialization, Firms and Markets: The Division of Labor within and Between Law Firms," *Journal of Law, Economics, & Organization*, 25, pp. 339-371.

Garicano, L. and E. Rossi-Hansberg (2006), "Organization and Inequality in a Knowledge Economy," *Quarterly Journal of Economics*, 121, pp. 1383-1435.

Ghosh, S. (2004), "Job Mobility and Careers in Firms," *Labour Economics*, 14, pp. 603-621.

Gibbons, R. and L. Katz (1991), "Layoffs and Lemons," *Journal of Labor Economics*, 9, pp. 351-380.

Gibbons, R. and M. Waldman (1999a), "A Theory of Wage and Promotion Dynamics Inside Firms," *Quarterly Journal of Economics*, 114, pp. 1321-1358.

Gibbons, R. and M. Waldman (1999b), "Careers in Organizations: Theory and Evidence," in O. Ashenfelter and D. Card, eds., *Handbook of Labor Economics, Volume 3* (North Holland: Amsterdam), pp. 2373-2437.

Gibbons, R. and M. Waldman (2006), "Enriching a Theory of Wage and Promotion Dynamics Inside Firms," *Journal of Labor Economics*, 24, pp. 59-107.

Greenwald, B. (1979), *Adverse Selection in the Labor Market* (Garland Press: New York).

Greenwald, B. (1986), "Adverse Selection in the Labour Market," *Review of Economic Studies*, 53, pp. 325-347.

Guadalupe, M. and J. Wulf (2010), "The Flattening Firm and Product Market Competition: The Effect of Trade Liberalization on Corporate Hierarchies," *American Economic Journal: Applied Economics*, 4, pp. 105-127.

Harris, M. and B. Holmstrom (1982), "A Theory of Wage Dynamics," *Review of Economic Studies*, 49, pp. 315-333.

Holmstrom, B. (1982), "Managerial Incentive Schemes – A Dynamic Perspective," in *Essays in Economics and Management in Honour of Lars Wahlbeck* (Swenska Handelshogkolan: Helsinki, Finland).

Hunnes, A. (2012), "Testing the Role of Comparative Advantage and Learning in Wage and Promotion Dynamics," *International Journal of Manpower*, 33, pp. 556-582.

Kahn, L.B. (2013), "Asymmetric Information Between Employers," *American Economic Journal: Applied Economics*, 5, pp. 165-205.

Kahn, L.B. and F. Lange (2014), "Employer Learning, Productivity and the Earnings Distribution: Evidence from Performance Measures," *Review of Economic Studies*, 81, pp. 1575-1613.

Kambourov, G. and I. Manovskii (2009), "Occupational Specificity of Human Capital," *International Economic Review*, 50, pp. 63-115.

Kauhanen, A. and S. Napari (2012), "Career and Wage Dynamics: Evidence from Linked Employer-Employee Data," *Research in Labor Economics*, 36, pp. 35-76.

Lazear, E. (1992), "The Job as a Concept," in W. Bruns, ed., *Performance Measurement, Evaluations, and Incentives* (Harvard University Press: Boston, MA), pp. 183-215.

Lazear, E. and S. Rosen (1981), "Rank-Order Tournaments as Optimum Labor Contracts," *Journal of Political Economy*, 89, pp. 841-864.

Lazear, E., K. Shaw, and C. Stanton (2015), "The Value of Bosses," *Journal of Labor Economics*, 33, pp. 823-861.

Lima, F. and P.T. Pereira (2003), "Careers and Wages within Large Firms: Evidence from a Matched Employer-Employee Data Set," *International Journal of Manpower*, 24, pp. 812-835.

Lluis, S. (2005), "The Role of Comparative Advantage and Learning in Wage Dynamics and Intrafirm Mobility: Evidence from Germany," *Journal of Labor Economics*, 23, pp. 725-768.

Lucas, R.E. (1978), "On the Size Distribution of Business Firms," *Bell Journal of Economics*, 9, pp. 508-523.

Mayer, T. (1960), "The Distribution of Ability and Earnings," *Review of Economics and Statistics*, 42, pp. 189-195.

McCue, K.J. (1996), "Promotions and Wage Growth," *Journal of Labor Economics* 14, pp. 175-209.

McLaughlin, K.J. (1994), "Rigid Wages," *Journal of Monetary Economics*, 34, pp. 383-414.

Medoff, J. and K. Abraham (1980), "Experience, Performance, and Earnings," *Quarterly Journal of Economics*, 95, pp. 703-736.

Medoff, J. and K. Abraham (1981), "Are Those Paid More Really More Productive?," *Journal of Human Resources*, 16, pp. 186-216.

Neal, D. (1995), "Industry-Specific Human Capital: Evidence from Displaced Workers," *Journal of Labor Economics*, 13, pp. 653-677.

Ortin-Angel, P. and V. Salas-Fumas (2002), "Compensation and Span of Control in Hierarchical Organizations," *Journal of Labor Economics*, 20, pp. 848-876.

Parent, D. (2000), "Industry Specific Capital and the Wage Profile: Evidence from the National Survey of Youth and the Panel Study of Income Dynamics," *Journal of Labor Economics*, 18, pp. 306-323.

Pinkston, J. (2009), "A Model of Asymmetric Employer Learning with Testable Implications," *Review of Economic Studies*, 76, pp. 367-394.

Podolny, J.M. and J.N. Baron (1997), "Resources and Relationships: Social Networks and Mobility in the Workplace," *American Sociological Review*, 62, pp. 673-693.

Qian, Y. (1994), "Incentives and Loss of Control in an Optimal Hierarchy," *Review of Economic Studies*, 61, pp. 527-544.

Rajan, R. and J. Wulf (2006), "The Flattening Firm: Evidence on the Changing Nature of Firm Hierarchies from Panel Data," *Review of Economics and Statistics*, 88, pp. 759-773.

Rosen, S. (1982), "Authority, Control, and the Distribution of Earnings," *Bell Journal of Economics*, 13, pp. 311-323.

Rosen, S. (1986), "Prizes and Incentives in Elimination Tournaments," *American Economic Review*, 76, pp. 701-715.

Sattinger, M. (1993), "Assignment Models and the Distribution of Earnings," *Journal of Economic Literature*, 31, pp. 831-880.

Schoenberg, U. (2007), "Testing for Asymmetric Employer Learning," *Journal of Labor Economics*, 25, pp. 651-691.

Seltzer, A. and D.T. Merrett (2000), "Personnel Policies at the Union Bank of Australia: Evidence from the 1888-1900 Entry Cohorts," *Journal of Labor Economics*, 18, pp. 573-613.

Smeets, V. (2016), "Why Hierarchies, Span of Control and Talent Allocation Matter for Firm Performance," Mimeo, Aarhus University.

Smeets, V. and F. Warzynski (2008), "Too Many Theories, Too Few Facts? What the Data Tell Us About the Link Between Span of Control, Compensation and Career Dynamics," *Labour Economics*, 15, pp. 688-704.

Spence, M. (1973), "Job Market Signaling," *Quarterly Journal of Economics*, 87, pp. 355-374.

Treble, J., E. van Gameren, S. Bridges, and T. Barmby (2001), "The Internal Economics of the Firm: Futher Evidence from Personnel Data," *Labour Economics*, 8, pp. 531-552.

Waldman, M. (1984a), "Worker Allocation, Hierarchies, and the Wage Distribution," *Review of Economic Studies*, 51, pp. 95-109.

Waldman, M. (1984b), "Job Assignments, Signalling, and Efficiency," *Rand Journal of Economics*, 15, pp. 255-267.

Waldman, M. (2012), "Theory and Evidence in Internal Labor Markets," in R. Gibbons and J. Roberts, eds., *The Handbook of Organizational Economics* (Princeton University Press: Princeton, NJ), pp. 520-574.

Williamson, O. (1967), "Hierarchical Control and Optimal Firm Size," *Journal of Political Economy*, 75, pp. 123-138.

Table 1: Summary statistics by rank

Panel A: summary statistics on human capital and performance variables

Variable	Worker	Team leader	Manager	Vice president	Corporate vice president
Average firm tenure	9.56	9.35	10.67	13.15	15.03
Average experience	20.99	19.48	20.44	23.12	24.90
Performance	3.42	3.48	3.70	3.84	4.06
- perf=3 (%)	*56.59%*	*50.32%*	*37.12%*	*27.57%*	*20.05%*
- perf=4 (%)	*34.85%*	*40.58%*	*51.42%*	*56.46%*	*52.81%*
- perf=5 (%)	*5.20%*	*5.75%*	*10.02%*	*14.64%*	*26.65%*
Promotion probability	1.39%	5.16%	2.77%	16.84%	0.98%
Education					
Bachelor	18.08%	22.95%	21.58%	16.43%	12.86%
Master and above	38.44%	45.18%	72.49%	82.34%	87.14%
# worker-year data	18,817	1,703	2,365	526	409

Panel B: summary statistics on span of control 6 months after performance evaluation

Variable	Team leader	Manager	Vice president	Corporate vice president
Average direct span	17.50	25.84	2.84	1.69
- 25% percentile	*12*	*8*	*1*	*1*
- median	*16*	*14*	*2*	*1*
- 75% percentile	*21*	*30*	*4*	*2*
Average full span			41.41	136.63
- 25% percentile			*10*	*24.5*
- median			*22*	*60.5*
- 75% percentile			*52*	*177*
# worker-year data	1,350	1,953	375	356

Note: Tenure is the number of years working for the firm. Experience is defined as age minus the number of years of education minus 6. Performance of year t-1 is assessed by the direct supervisor in December of year t-1. Promotion probability is the probability that the individual was promoted from December in year t-1 to December in year t. Bachelor is a dummy equal to 1 if the individual has a bachelor degree and 0 otherwise. Master and above is a dummy equal to 1 if the individual has a MSc or PhD degree, and 0 otherwise. The average span of control is the average number of individuals reporting (directly or indirectly) to their supervisor as observed in our dataset.

Table 2: Wages, Human Capital and Performance

Dep. var.: log (wage)	(1)	(2)	(3)
Experience	0.018***	0.018***	0.015***
	(0.001)	(0.001)	(0.001)
Experience²/100	-0.028***	-0.028***	-0.023***
	(0.002)	(0.002)	(0.002)
Tenure	0.0001	-0.0004	-0.001
	(0.0006)	(0.0006)	(0.0007)
Tenure²/100	0.004**	0.005***	0.005***
	(0.002)	(0.002)	(0.002)
Bachelor	0.214***	0.213***	0.204***
	(0.005)	(0.005)	(0.005)
Master and above	0.296***	0.295***	0.281***
	(0.004)	(0.004)	(0.005)
Team Leader	0.219***	0.215***	0.211***
	(0.005)	(0.005)	(0.005)
Manager	0.417***	0.409***	0.415***
	(0.004)	(0.004)	(0.005)
Vice President	0.641***	0.628***	0.620***
	(0.008)	(0.008)	(0.009)
Corporate Vice President	0.838***	0.816***	0.819***
	(0.010)	(0.009)	(0.009)
Performance in t-1	-	0.034***	0.022***
		(0.002)	(0.002)
Performance in t-2	-	-	0.019***
			(0.002)
Constant	10.161***	10.045***	10.052***
	(0.008)	(0.009)	(0.014)
Year dummies	YES	YES	YES
Cost center dummies	YES	YES	YES
R2	0.82	0.83	0.85
N	21,537	21,537	12,697

Note: the dependent variable is the log of the wage (not including bonuses) in June, 6 months after the latest performance evaluation. Experience is defined as age minus the number of years of education minus 6. Tenure is the number of years working for the firm. Performance of year t-1 (t-2) is assessed by the direct supervisor in December of year t-1 (t-2). The other explanatory variables are job level dummies, education dummies for college education, cost center dummies and year dummies. Standard errors are clustered by worker.

Table 3: Wage Growth and Performance

Dep. var.: Δ log (wage)	(1)	(2)	(3)	(4)
log (wage)	-0.005***	-0.007***	-0.009***	-0.010***
	(0.001)	(0.001)	(0.001)	(0.001)
Performance in t-1	-	0.009***	0.008***	0.006***
		(0.0003)	(0.0003)	(0.0004)
Performance in t-2	-	-	-	0.003***
				(0.0004)
Promotion	0.028***	-	0.024***	0.023***
	(0.003)		(0.003)	(0.003)
Constant	0.093***	0.091***	0.107***	0.115***
	(0.007)	(0.007)	(0.007)	(0.009)
Year dummies	YES	YES	YES	YES
Cost center dummies	YES	YES	YES	YES
R2	0.08	0.11	0.12	0.11
N	23,266	23,266	23,266	13,553

Note: the dependent variable is the difference in log between the wage 6 months after performance evaluation (in June of year t) and the wage in December of year t-1. Promotion is a dummy variable equal to 1 if the individual is promoted during the same period. The performance variables are defined as in table 2. Standard errors are clustered by worker.

Table 4: Promotion and Performance, probit regression (marginal effects reported)

Dep. var.: Promotion	(1)	(2)	(3)
Performance in t-1	0.013***	0.011***	0.015***
	(0.001)	(0.001)	(0.001)
Performance in t-2	-	0.007***	-
		(0.001)	
Team Leader	0.028***	0.026***	0.028***
	(0.005)	(0.006)	(0.006)
Manager	-0.002	-0.004**	-0.004*
	(0.002)	(0.002)	(0.002)
Vice president	0.005	-0.003	-0.001
	(0.005)	(0.004)	(0.004)
Bachelor	0.027***	0.022***	0.025***
	(0.005)	(0.006)	(0.006)
Master and above	0.026***	0.027***	0.029***
	(0.003)	(0.004)	(0.004)
Experience	-0.0006	-0.001**	-0.001**
	(0.0004)	(0.0005)	(0.0005)
$Experience^2/100$	0.0006	0.0019	0.0019
	(0.001)	(0.0012)	(0.0012)
Tenure	0.0001	0.0003	0.0004
	(0.0003)	(0.0004)	(0.0004)
$Tenure^2/100$	0.0005	-0.0002	-0.0004
	(0.001)	(0.001)	(0.001)
Year dummies	YES	YES	YES
Cost center dummies	YES	YES	YES
Pseudo R2	0.13	0.15	0.14
Log Likelihood	-2063.41	-1274.85	-1292.53
N	22,015	13,028	13,028

Note: the dependent variable is a dichotomic variable equal to 1 if the individual is promoted: 1) from worker to team leader; 2) from worker to manager; 3) from team leader to manager; 4) from manager to vice president; 4) from vice president to corporate vice president; within a 12-month period following the performance evaluation (i.e. from January to December of year t). The independent variables are as described in table 2. Standard errors are clustered by worker.

Table 5: Direct Span of Control and Performance

Dep. var.: log (direct span)	(1)	(2)	(3)
Performance in t-1	0.077***	0.063***	0.058**
	(0.022)	(0.023)	(0.023)
Performance in t-2	-	0.045**	0.041*
		(0.022)	(0.022)
Team leader	2.034***	2.055***	2.345***
	(0.076)	(0.080)	(0.154)
Manager	2.350***	2.366***	2.361***
	(0.069)	(0.072)	(0.072)
Vice President	0.521***	0.554***	0.546***
	(0.084)	(0.088)	(0.087)
Bachelor	0.058	0.088	0.405***
	(0.067)	(0.073)	(0.144)
Master and above	-0.096	-0.061	0.154
	(0.066)	(0.073)	(0.132)
Team leader*BA/BSc	-	-	-0.556***
			(0.157)
Team leader*(Master and above)	-	-	-0.244*
			(0.144)
Experience	0.003	-0.003	-0.002
	(0.012)	(0.014)	(0.014)
$Experience^2/100$	-0.026	-0.016	-0.017
	(0.026)	(0.029)	(0.028)
Tenure	0.029***	0.032***	0.032***
	(0.008)	(0.009)	(0.009)
$Tenure^2/100$	-0.058**	-0.062**	-0.066**
	(0.027)	(0.031)	(0.030)
Constant	0.107	-0.006	-0.193
	(0.201)	(0.247)	(0.266)
Year dummies	YES	YES	YES
Cost center dummies	YES	YES	YES
R2	0.61	0.63	0.63
N	3,776	2,929	2,929

Note: the dependent variable is the log of the direct span of control as team leader, manager, vice president or corporate vice president 6 months after the latest performance evaluation, i.e. in June of year t. The independent variables are as defined in table 3. Standard errors are clustered by worker.

Table 6: Full Span of Control and Performance

Dep. var.: log (full span)	(1)	(2)	(3)
Performance in t-1	0.094***	0.097***	0.093***
	(0.027)	(0.034)	(0.034)
Performance in t-2	-	0.063*	0.056*
		(0.033)	(0.032)
Team leader	-1.831***	-1.815***	-1.531***
	(0.154)	(0.152)	(0.213)
Manager	-1.430***	-1.427***	-1.434***
	(0.144)	(0.141)	(0.142)
Vice President	-0.973***	-0.935***	-0.944***
	(0.170)	(0.173)	(0.172)
Bachelor	0.019	0.043	0.361**
	(0.080)	(0.088)	(0.162)
Master and above	-0.172**	-0.139	0.067
	(0.078)	(0.087)	(0.145)
Team leader*BA/BSc	-	-	-0.597***
			(0.179)
Team leader*(Master and above)	-	-	-0.220
			(0.164)
Experience	0.002	-0.001	-0.002
	(0.019)	(0.022)	(0.022)
$Experience^2/100$	-0.046	-0.044	-0.043
	(0.043)	(0.049)	(0.049)
Tenure	0.042***	0.046***	0.047***
	(0.010)	(0.012)	(0.012)
$Tenure^2/100$	-0.078**	-0.084**	-0.090**
	(0.034)	(0.038)	(0.038)
Constant	4.012***	3.757***	3.593***
	(0.303)	(0.371)	(0.381)
Year dummies	YES	YES	YES
Cost center dummies	YES	YES	YES
R2	0.31	0.32	0.33
N	3,760	2,914	2,914

Note: the dependent variable is the log of the full span of control as team leader, manager, vice president or corporate vice president 6 months after the latest performance evaluation, i.e. in June of year t. The independent variables are as defined in table 3. Standard errors are clustered by worker.

Table 7: Change in Direct Span and Performance

Dep. var.: Δ log (direct span)	(1)	(2)	(3)	(4)
log (direct span)	-0.107***	-0.105***	-0.108***	-0.105***
	(0.009)	(0.010)	(0.009)	(0.010)
Performance in t-1	0.024**	0.026**	0.024***	0.026**
	(0.009)	(0.011)	(0.009)	(0.011)
Performance in t-2	-	-0.00002	-	0.000
		(0.010)		(0.011)
Team leader	0.246***	0.234***	0.237***	0.227***
	(0.031)	(0.034)	(0.032)	(0.035)
Manager	0.269***	0.259***	0.266***	0.257***
	(0.029)	(0.034)	(0.030)	(0.034)
Vice president	0.108***	0.111***	0.107***	0.110***
	(0.026)	(0.027)	(0.026)	(0.027)
Bachelor	-	-	-0.007	-0.009
			(0.020)	(0.021)
Master and above	-	-	-0.023	-0.018
			(0.021)	(0.023)
Experience	-0.002	-0.003	-0.003	-0.004
	(0.003)	(0.004)	(0.004)	(0.004)
Experience^2	0.002	0.004	0.003	0.004
	(0.008)	(0.009)	(0.008)	(0.009)
Tenure	0.003	0.004	0.002	0.004
	(0.003)	(0.003)	(0.003)	(0.003)
Tenure^2	-0.008	-0.014	-0.008	-0.013
	(0.010)	(0.012)	(0.010)	(0.012)
Constant	0.009	-0.002	0.043	0.024
	(0.060)	(0.077)	(0.069)	(0.086)
Year dummies	YES	YES	YES	YES
Cost center dummies	YES	YES	YES	YES
R2	0.06	0.06	0.06	0.06
N	3,612	2,805	3,612	2,805

Note: the dependent variable is the difference between the log of the direct span of control 6 months after the latest performance evaluation (i.e. in June of year t) as in table 5, and the log of the direct span of control in December of year t-1. The explanatory variables are defined as in tables 3 and 5. The distribution is trimmed below the first percentile and above the 99th percentile. Standard errors are clustered by worker.

Table 8: Change in Full Span and Performance

Dep. var.: Δ log (full span)	(1)	(2)	(3)	(4)
log (full span)	-0.085***	-0.078***	-0.085***	-0.078***
	(0.009)	(0.011)	(0.009)	(0.011)
Performance in t-1	0.022**	0.028**	0.022**	0.028**
	(0.009)	(0.011)	(0.009)	(0.011)
Performance in t-2	-	-0.002	-	-0.002
		(0.012)		(0.012)
Team leader	-0.140***	-0.126***	-0.149***	-0.132***
	(0.032)	(0.036)	(0.034)	(0.037)
Manager	-0.112***	-0.098***	-0.116***	-0.100***
	(0.029)	(0.032)	(0.030)	(0.032)
Vice president	-0.054*	-0.039	-0.056*	-0.040
	(0.030)	(0.032)	(0.030)	(0.032)
Bachelor	-	-	-0.015	-0.014
			(0.020)	(0.022)
Master and above	-	-	-0.021	-0.014
			(0.021)	(0.023)
Experience	-0.006*	-0.009**	-0.007*	-0.009**
	(0.004)	(0.004)	(0.004)	(0.005)
Experience^2	0.010	0.015	0.011	0.015
	(0.008)	(0.010)	(0.008)	(0.010)
Tenure	0.004	0.0066*	0.004	0.0064*
	(0.003)	(0.0035)	(0.003)	(0.0035)
Tenure^2	-0.013	-0.019	-0.012	-0.019
	(0.010)	(0.012)	(0.010)	(0.012)
Constant	0.374***	0.321***	0.407***	0.343***
	(0.075)	(0.095)	(0.083)	(0.104)
Year dummies	YES	YES	YES	YES
Cost center dummies	YES	YES	YES	YES
R2	0.05	0.05	0.05	0.05
N	3,578	2,774	3,578	2,774

Note: the dependent variable is the difference between the log of the full span of control 6 months after the latest performance evaluation (i.e. in June of year t) as in table 6, and the log of the full span of control in December of year t-1. The explanatory variables are defined as in tables 3 and 6. The distribution is trimmed below the first percentile and above the 99th percentile. Standard errors are clustered by worker.

Table 9: Wages, Human Capital and Performance
(controling for direct span)

Dep. var.: log (wage)	(1)	(2)	(3)
Experience	0.004***	0.004***	0.003**
	(0.001)	(0.001)	(0.001)
Experience²/100	-0.008***	-0.008***	-0.006**
	(0.002)	(0.002)	(0.003)
Tenure	0.0012	0.001	0.001
	(0.0008)	(0.001)	(0.001)
Tenure²/100	-0.001	-0.001	-0.0003
	(0.002)	(0.002)	(0.003)
Bachelor	0.062***	0.061***	0.064***
	(0.008)	(0.008)	(0.009)
Master and above	0.080***	0.079***	0.081***
	(0.007)	(0.007)	(0.008)
Manager	0.201***	0.196***	0.198***
	(0.018)	(0.018)	(0.020)
Vice President	0.468***	0.461***	0.454***
	(0.018)	(0.018)	(0.020)
Corporate Vice President	0.691***	0.677***	0.674***
	(0.017)	(0.017)	(0.019)
Performance in t-1	-	0.021***	0.018***
		(0.002)	(0.002)
Performance in t-2	-	-	0.007***
			(0.002)
log (direct span)	-0.006	-0.008	-0.007
	(0.006)	(0.006)	(0.006)
log (direct span) * Manager	0.022***	0.023***	0.023***
	(0.006)	(0.006)	(0.007)
log (direct span) * VP	0.044***	0.043***	0.032***
	(0.010)	(0.010)	(0.010)
log (direct span) * CVP	0.063***	0.059***	0.061***
	(0.016)	(0.016)	(0.015)
Constant	10.686***	10.617***	10.605***
	(0.022)	(0.023)	(0.026)
Year dummies	YES	YES	YES
Cost center dummies	YES	YES	YES
R2	0.91	0.91	0.92
N	4,084	4,084	3,128

Note: variables are defined as in tables 2 and 5. Standard errors are clustered by worker.

Table 10: Wages, Human Capital and Performance
(controling for full span)

Dep. var.: log (wage)	(1)	(2)	(3)
Experience	0.004***	0.004***	0.003**
	(0.001)	(0.001)	(0.001)
Experience²/100	-0.008***	-0.008***	-0.006**
	(0.002)	(0.002)	(0.003)
Tenure	0.001	0.001	0.0005
	(0.0008)	(0.001)	(0.001)
Tenure²/100	-0.001	-0.0002	0.0003
	(0.003)	(0.002)	(0.003)
Bachelor	0.062***	0.061***	0.064***
	(0.008)	(0.008)	(0.009)
Master and above	0.080***	0.078***	0.081***
	(0.008)	(0.007)	(0.008)
Manager	0.199***	0.195***	0.196***
	(0.018)	(0.018)	(0.020)
Vice President	0.422***	0.418***	0.429***
	(0.022)	(0.022)	(0.024)
Corporate Vice President	0.643***	0.633***	0.628***
	(0.027)	(0.026)	(0.027)
Performance in t-1	-	0.021***	0.018***
		(0.002)	(0.002)
Performance in t-2	-	-	0.008***
			(0.002)
log (full span)	-0.006	-0.007	-0.007
	(0.006)	(0.006)	(0.006)
log (full span) * Manager	0.022***	0.023***	0.023***
	(0.006)	(0.006)	(0.007)
log (full span) * VP	0.030***	0.030***	0.021***
	(0.007)	(0.007)	(0.008)
log (full span) * CVP	0.022***	0.023***	0.022***
	(0.008)	(0.008)	(0.008)
Constant	10.685***	10.616***	10.601***
	(0.022)	(0.023)	(0.027)
Year dummies	YES	YES	YES
Cost center dummies	YES	YES	YES
R2	0.91	0.91	0.92
N	4,084	4,084	3,128

Note: variables are defined as in tables 2 and 6. Standard errors are clustered by worker.

Table 11: Promotion and performance, probit regression (marginal effects reported) (controling for direct span)

Dep. var.: Promotion	(1)	(2)	(3)
Performance in t-1	-	0.033***	0.025***
		(0.004)	(0.004)
Performance in t-2	-	-	0.016***
		.	(0.004)
log (direct span)	-0.014	-0.014*	-0.007
	(0.009)	(0.008)	(0.008)
log (direct span) * Manager	0.019*	0.018*	0.015
	(0.010)	(0.009)	(0.009)
log (direct span) * VP	0.062***	0.054***	0.047***
	(0.015)	(0.013)	(0.014)
log (direct span) * CVP	0.004	-0.005	-0.017
	(0.017)	(0.015)	(0.014)
Manager	-0.107***	-0.109***	-0.107***
	(0.034)	(0.034)	(0.039)
Vice president	-0.052***	-0.047***	-0.038***
	(0.006)	(0.005)	(0.005)
Corporate Vice President	-0.010	-0.021	-0.015
	(0.021)	(0.013)	(0.014)
Experience	-0.002	-0.002	-0.003
	(0.002)	(0.002)	(0.002)
Experience²/100	0.005	0.006	0.006
	(0.005)	(0.004)	(0.004)
Tenure	-0.001	-0.001	0.0002
	(0.001)	(0.001)	(0.001)
Tenure²/100	0.002	0.002	-0.0003
	(0.004)	(0.004)	(0.004)
Bachelor	0.040**	0.037**	0.033***
	(0.020)	(0.019)	(0.010)
Master and above	0.037***	0.033***	0.030*
	(0.011)	(0.010)	(0.021)
Year dummies	YES	YES	YES
Cost center dummies	YES	YES	YES
Pseudo R2	0.10	0.14	0.18
Log Likelihood	-787.13	-755.35	-539.92
N	4,100	4,100	3,137

Note: variables are defined as in tables 2, 4 and 5. Standard errors are clustered by worker.

Table 12: Promotion and performance, probit regression (marginal effects reported) (controling for full span)

Dep. var.: Promotion	(1)	(2)	(3)
Performance in t-1	-	0.033***	0.025***
		(0.004)	(0.004)
Performance in t-2	-	-	0.015***
		.	(0.004)
log (full span)	-0.0078*	-0.009**	-0.009**
	(0.0045)	(0.004)	(0.004)
log (full span) * Manager	0.013*	0.012**	0.016***
	(0.007)	(0.006)	(0.006)
log (full span) * VP	0.057***	0.050***	0.047***
	(0.013)	(0.012)	(0.013)
log (full span) * CVP	-0.002	-0.010	-0.016
	(0.015)	(0.013)	(0.012)
Manager	-0.086***	-0.089***	-0.107***
	(0.023)	(0.023)	(0.028)
Vice president	-0.045***	-0.040***	-0.035***
	(0.005)	(0.005)	(0.005)
Corporate Vice President	0.046***	0.025*	0.015
	(0.022)	(0.017)	(0.015)
Experience	-0.002	-0.002	-0.003
	(0.002)	(0.002)	(0.002)
Experience²/100	0.004	0.005	0.005
	(0.005)	(0.004)	(0.004)
Tenure	-0.001	-0.001	0.0004
	(0.001)	(0.001)	(0.0013)
Tenure²/100	0.002	0.002	-0.001
	(0.004)	(0.004)	(0.004)
Bachelor	0.040**	0.036**	0.028*
	(0.020)	(0.019)	(0.020)
Master and above	0.036***	0.032***	0.030***
	(0.011)	(0.010)	(0.010)
Year dummies	YES	YES	YES
Cost center dummies	YES	YES	YES
Pseudo R2	0.10	0.14	0.18
Log Likelihood	-781.24	-749.74	-534.15
N	4,098	4,098	3,135

Note: variables are defined as in tables 2, 4 and 6. Standard errors are clustered by worker.

MATHEMATICAL SUPPLEMENT FOR "PERFORMANCE, CAREER DYNAMICS, AND SPAN OF CONTROL"

By Valerie Smeets, Michael Waldman, and Frederic Warzynski

In this mathematical supplement we consider three extensions of the two-period careers, two job level, identical firm model presented and analyzed in Subsections III.A and III.B of the paper. In the first extension we consider what happens when there are two job levels but careers last three periods rather than two. In the second extension we consider what happens when a third job level is added. In the third extension we return to the case in which careers are two periods and there are two job levels, but we assume that one of the firms is a growing firm.

Extension 1: Careers Last Three Periods

In this extension we consider what happens when careers last three periods rather than two. Also, below a young worker refers to a worker for whom the current period is the worker's first period in the labor market, an old worker refers to a worker for whom the current period is the worker's second period in the labor market, and a very old worker refers to a worker for whom the current period is the worker's third period in the labor market.

Proposition 1A: Consider any firm k and period t. If careers last three periods and $S>1$, then the following describe firm k's behavior in period t.

i) The firm assigns to job 2 the m old and very old workers with the highest values for expected effective ability, while the other old and very old workers are assigned to job 1.

ii) Within a schooling group the old workers promoted are the ones who produced the highest outputs, but there can be pairs of old workers where only one is promoted and this worker produced less in t-1 but has a higher education level.

iii) The period-t wage for old workers at the firm and the wage increase from t-1 to t for old workers at the firm both increase with the t-1 output holding education fixed, while the wage for old workers increases with education holding the t-1 output fixed.

iv) Within a schooling group, promoted old workers receive larger pay increases from t-1 to t than the old workers who were not promoted.

v) Within a schooling group and holding period t-2 output fixed, very old workers assigned to job 1 in period t-1 and to job 2 in period t are the very old workers assigned to job 1 in period t-1 who produced the highest outputs in period t-1. But across schooling groups and holding period t-2 output fixed, there can be pairs of very old workers where only one is promoted and this worker produced less in t-1 but has a higher education level.

vi) Within a schooling group and holding period t-2 output fixed, very old workers assigned to job 2 in period t-1 and to job 2 in period t are the very old workers assigned to job 2 in period t-1 who produced the highest outputs in period t-1.

vii) Holding period t-2 output fixed, the period t wage for very old workers and the wage increase from t-1 to t for very old workers both increase with the t-1 output holding education fixed, while the wage for very old workers increases with education holding the t-1 output fixed given the t-1 job assignment is unchanged.

viii) Holding the period t-2 output fixed, very old workers promoted from job 1 to job 2 receive larger pay increases than very old workers in job 1 in period t-1 who were not promoted.

ix) All young workers hired by the firm are assigned to job 1, where the young worker wage increases with the worker's education level.

x) Within a schooling group, holding the worker's t-2 output fixed, the change in span of control for a very old worker assigned to job 2 in both period t-1 and t (weakly) increases with the worker's output in t-1.

xi) Managerial span of control for old workers assigned to job 2 in period t (weakly) increases with the manager's t-1 output holding education fixed, and also (weakly) increases with the manager's education level holding the t-1 output fixed.

xii) Managerial span of control for a very old worker assigned to job 2 in period t, holding t-2 output fixed, (weakly) increases with the manager's t-1 output holding education fixed, and also (weakly) increases with the manager's education level holding the t-1 output fixed.

xiii) Within a schooling group and holding age fixed, managerial span of control is positively correlated with the managerial wage.

Proof of Proposition 1A: The expression for z_{it-1} for old workers in period t and the expression for z_{it-2} for very old workers in period t are both given by (A1) in the Appendix since young workers must be assigned to job 1. Now consider z_{it-1} for very old workers in t-1. There are two separate expressions because the expression for the signal about ability that can be extracted from observing the worker's t-1 output depends on the worker's t-1 job assignment. If the worker was assigned to job 1, then $z_{it-1}=[y_{i1t-1}/c_2(1+v_{it-1})]-(c_1/c_2)=\eta_{it-1}+\varepsilon_{it-1}$. If the worker was assigned to job 2, then $z_{it-1}=(y_{i2kt-1}/g(n_{2kt-1}))=\eta_{it-1}+\varepsilon_{it-1}$. That is, given that the noise terms are drawn from the same normal distribution whether the worker was assigned to job 1 or job 2 in t-1, the rate of learning when workers are old is independent of the job assignment.

Since the rate of learning is independent of the job assignment for old workers and there is no turnover given our assumption concerning v, the firm must maximize expected profits from period t onward by assigning the highest expected effective ability old and very old workers to job 2. This proves i). ii), iii), and iv) then follow from arguments in the proof of Proposition 2. And with a slight modification these arguments also yield v), vi), vii), and viii), while ix) also follows from an argument in the proof of Proposition 2.

Consider two very old workers, a and b, who produced the same output in t-2, were each assigned to job 2 in t-1, and are both assigned to job 2 in period t. Using logic like above the two must have had the same span of control in t-1. Given this, suppose a produced more in t-1 but has the smaller span of control in t. We know that η_{it}^{e} is higher for a rather than b, but this means the firm could switch the span of controls for the two managers and increase profitability. So a's span of control must be weakly higher than b's. This proves x). Finally, xi), xii), and xiii) follow from arguments like arguments in the proof of Proposition 2.

Extension 2: Three Job Levels

In this extension we consider what happens when a third job level is added to the model considered in extension 1, where we now assume a single job level 3 position at the top of each of the m divisions. Further, there is an endogenously determined number of subdivisions in each division, where each subdivision has a single level 2 position and an endogenously determined number of level 1 workers reporting to this manager. The productivity of a level 1 workers is still given by equation (2). The productivity of level 2 worker i in period t with n level 1 workers reporting to this level 2 manager is given by $y_{i2t}=g_2(n)(\eta_{it}+\varepsilon_{it})$, where $g_2'>0$ and $g_2''<0$. Further, the productivity of level 3 manager i in period t with n level 2 managers reporting to this level 3 manager is given by $y_{i3t}=g_3(n)(\eta_{it}+\varepsilon_{it})$, where $g_3'>0$ and $g_3''<0$. We further assume that $g_2(.)$ and $g_3(.)$ are such that it is efficient to staff the level 3 positions with the highest expected effective ability old and very old workers and the level 2 positions with the next highest expected effective ability old and very old workers. Also, let D_{kt} be the total number of subdivisions at firm k in period t.

Proposition 2A: Consider any firm k and period t. If careers last three periods, there are three job levels, and S>1, then the following describe firm k's behavior in period t.

i) The firm assigns to job 3 the m old and very old workers with the highest values for expected effective ability, the D_{kt} old and very old workers with the next highest values for expected effective ability are assigned to job 2, while the other old and very old workers are assigned to job 1.

ii) Within a schooling group the old workers promoted are the ones who produced the highest outputs in the previous period, but there can be pairs of old workers where only one is promoted and this worker produced less in t-1 but has a higher education level.

iii) The period-t wage for old workers at the firm and the wage increase from t-1 to t for old workers at the firm both increase with the t-1 output holding education fixed, while the wage for old workers increases with education holding the t-1 output fixed.

iv) Within a schooling group, promoted old workers receive larger pay increases from t-1 to t then the old workers who were not promoted.

v) Within a schooling group and holding t-2 output fixed, very old workers assigned to job 1 in period t-1 and to job 3 in period 2 are the very old workers assigned to job 1 in period t-1 who produced the highest outputs in t-1 while the ones assigned to job 2 produced the next highest outputs. But across schooling groups and holding period t-2 output fixed, there can be pairs of very old workers both assigned to job 1 in period t-1 where only one is promoted and this worker produced less in t-1but has a higher education level.

vi) Within a schooling group and holding t-2 output fixed, very old workers assigned to job 2 in period t-1 and to job 3 in period 2 are the very old workers assigned to job 2 in period t-1 who produced the highest outputs in t-1 while the ones assigned to job 2 produced the next highest outputs. But across schooling groups and holding period t-2 output fixed, there can be pairs of very old workers both assigned to job 2 in period t-1 where only one is promoted and this worker produced less in t-1but has a higher education level.

vii) Within a schooling group and holding t-2 output fixed, very old workers assigned to job 3 in period t-1 and to job 3 in period 2 are the very old workers assigned to job 3 in period t-1 who produced the highest outputs in t-1 while the ones assigned to job 2 produced the next highest outputs.

viii) Holding period t-2 output fixed, the period t wage for very old workers and the wage increase from t-1 to t for very old workers both increase with the t-1 output holding education fixed, while the wage for very old workers increases with education holding the t-1 output fixed given the t-1 job assignment is unchanged.

ix) Within a schooling group and holding t-2 output fixed, very old workers who are promoted in period t receive larger pay increases than the very old workers who were not promoted.

x) All young workers hired by the firm are assigned to job 1, where the young worker wage increases with the worker's education level.

xi) Within a schooling group, holding the worker's t-2 output fixed, the change in span of control for a very old worker assigned to job 2 in both period t-1 and t (weakly) increased with the worker's output in t-1, while the same condition holds for very old workers assigned to job 3 in both period t-1 and t.

xii) Managerial span of control for old workers assigned to job 2 in period t (weakly) increases with the manager's t-1 output holding education fixed, and also (weakly) increases with the manager's education level holding the t-1 output fixed. Also, the same relationships hold for old workers assigned to job 3 in period t.

xiii) Managerial span of control for a very old worker assigned to job 2 in period t, holding t-2 output fixed, (weakly) increased with the manager's t-1 output holding education fixed, and also (weakly) increases with the manager's education level holding the t-1 output fixed. Also, the same relationships hold for very old workers assigned to job 3 in period t.

xiv) Within a schooling group and holding both age and job level fixed, managerial span of control is positively correlated with the managerial wage.

Proof of Proposition 2A: Using arguments like those in the proof of Proposition 1A, it can be shown that the rate of learning is independent of the job assignment. In turn, since the rae of learning is independentof the job assignment for old workers and there is no turnover given our assumption concerning v, the firm must maximize profits from period t onward by assigning the highest expected effective ability old and very old workers to job 2. This proves i). ii), iii), and iv) then follow from arguments in the proof of Proposition 2. And with a slight modification these arguments also yield v), vi), vii), viii), and ix), while x) also follows from an argument in the proof of Proposition 2. The first part of xi) follows from an argument in the proof of Proposition 1A, while the second part of xi) follows from a slight modification of that argument. Finally, xii), xiii), and xiv) follow from arguments like arguments in the proof of Proposition 2.

Extension 3: A Growing Firm

In this extension we consider what happens when everything is the same as in the model presented and analyzed in Subsections III.A and III.B except that one firm grows from period t-1 to t.

Proposition 3A: Suppose that everything is the same as in the model analyzed in Proposition 2, except that from t-1 to t firm k′ increases the number of divisions from m to m+Δ, Δ>0. i) through vi) below describes the behavior of firm k′ while Proposition 2 describes the behavior of the other firms in period t.[1]

i) The firm promotes the m+Δ old workers it employed in the previous period with the highest values for expected effective ability, while the other workers are assigned to job 1.

ii) Within a schooling group the old workers promoted are the ones who produced the highest outputs, but there can be pairs of old workers where only one is promoted and this worker produced less in t-1 but has a higher education level.

[1] For this extension we assume that the g(.) function is such that in each period a firm hires at least m+Δ total young workers so that firm k′ is able to fully staff its managerial positions in every period t′, t′≥t, by promoting from within.

iii) The period-t wage for old workers at the firm and the wage increase from t-1 to t for old workers at the firm both increase with the t-1 output holding education fixed, while the wage for old workers increases with education holding the t-1 output fixed.

iv) Within a schooling group, promoted workers receive larger pay increases from t-1 to t than the workers who were not promoted.

v) All young workers hired by the firm are assigned to job 1, where the young worker wage increases with the worker's education level.

vi) Managerial span of control (weakly) increases with the manager's t-1 output holding education fixed, and also (weakly) increases with the manager's education level holding the t-1 output fixed.

vii) Within a schooling group, managerial span of control is positively correlated with the managerial wage.

<u>Proof of Proposition 3A</u>: The claim concerning the behavior of firms other than firm k′ follows from the proof of Proposition 2 found in the paper's Appendix. i) follows from the same argument in the proof of Proposition 2 as does i) of Proposition 2 except substitute m+Δ for m in that part of the proof. And then the rest of the results, i.e., ii) through vii), follow from the same arguments as in the proof of Proposition 2.