Which Workers Earn More at Productive Firms? Position Specific Skills and Individual Worker Hold-up Power*

Justin Bloesch
Columbia Business School

Birthe Larsen
Copenhagen Business School

Bledi Taska
Lightcast

October 7, 2022

Abstract

We argue that productive firms share rents with workers only in occupations where workers have individual hold-up power. Workers have this power if the output of positions is individually complementary and workers acquire position-specific skills on the job. We estimate individual worker hold-up power by occupation using the effect of worker deaths on firm profits in Denmark and a measure of task differentiation from US job postings. High hold-up occupations exhibit higher wage levels and higher long-run passthrough of permanent firm productivity innovations to wages. We examine inequality implications for the gender wage gap and the effect of superstar firms.

*Contact: Justin Bloesch (jb4768@columbia.edu), Birthe Larsen (bl.eco@cbs.dk), and Bledi Taska (bledi.taska@lightcast.io). We thank Antoine Bertheau, Adrien Bilal, Gabriel Chodorow-Reich, John Coglianese, Harris Eppsteiner, Jason Furman, Xavier Gabaix, Andrew Garin, Ed Glaeser, Fane Groes, Lawrence Katz, Omeed Maghzian, Namrata Narain, Anna Stansbury, Ludwig Straub, James Stock, and Jacob Weber and seminar participants at Harvard University, the 2021 Search and Matching Conference, the University of Copenhagen, and the Copenhagen Business School for thoughtful discussions and comments. This research was supported by the James M. and Kathleen D. Stone PhD Scholarship in Inequality and Wealth Concentration.

1 Introduction

As wage inequality has grown in a number of advanced economies, researchers have taken a renewed interest in understanding the role that firms play in determining wages. Recent research has explored whether firms have heterogeneous wage effects for different types of workers across numerous countries, typically finding that firms have larger effects on the wages of high-income or highly educated workers.[1] Figure 1 shows data from Denmark, ploting the cross-sectional relationship between individual worker hourly wages and firm output per worker by occupation between 2008 and 2016. The cross-sectional productivity-wage elasticity is 0.2 for managers, a little under 0.1 for most middle-wage occupations, and nearly zero for low-wage service occupations.

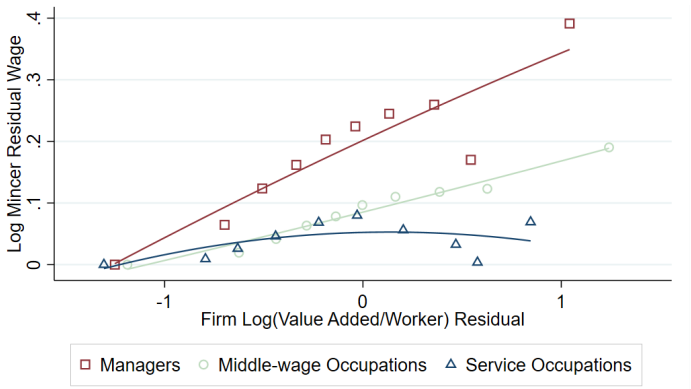
In this paper, we argue that at least half of the cross-sectional elasticity of wages to productivity in Figure 1 reflects rents. The underlying source of different productivity-wage elasticities is heterogeneity in individual worker hold-up power across occupation groups and exists when (i) labor in production is organized into distinct, differentiated positions, (ii) the output of positions is individually complementary or "critical" in the production process, and (iii) skills are position-specific, i.e., skills are acquired on the job and are not transferable across positions or firms. If output losses from an unfilled position are larger at higher productivity firms, then a worker in a given occupation can hold up more output at a high productivity firm than at a low productivity firm. With on-the-job search, incomplete contracts incentivize productive firms to pay higher wages to decrease differentially costly turnover.

Consider a simple example of a restaurant with two positions, using only a cook and a waiter. The production function is Leontief: if the tasks in either position are left undone, then the restaurant sells nothing. Suppose also that training a new cook takes longer than training a waiter, as even an experienced cook must take time to learn the new recipes. In this setting, the cook has hold-up power, but the waiter does not. Both positions are essential in production, but only the cook has position specificity, while the waiter is replaceable. If the restaurant becomes more popular and can charge higher prices, the amount of profit that the cook can hold up increases, and the owner will be incentivized to raise the cook's pay, while the waiter's pay will be set to the going market rate. Our more general model formalizes this example to a continuum of positions and diminishing returns to labor.

The main contributions of this paper are four-fold. First, we establish conditions

[1] See Song et al. (2019) and Kline et al. (2019) in the United States, Garin and Silverio (2018) in Portugal, Friedrich et al. (2019) in Sweden, and Maibom and Vejlin (2021) and Chan et al. (2020) in Denmark.

Figure 1: Cross-Sectional Productivity-Wage Elasticity by Occupation



The wage and labor productivity measures are residuals of a regression of log wages or firm value added per worker on standard Mincer regression covariates: sex, years of education, potential experience, potential experience squared (with interactions on years of education), and industry, year, and 6-digit occupation code fixed effects. The sample is full-time workers in their "main" November job between 2008-2016, in private sector firms that are at least 5 years old.

under which an individual worker has hold-up power over inframarginal firm output in a multi-worker firm and develop a novel and tractable production function to generate these conditions. Second, we measure hold-up power by occupation group in two ways, both by estimating the effect of worker deaths on firm profits in Denmark and by constructing a separate measure of within-firm, across-position task differentiation in US job posting data. Third, we show that hold-up power predicts long-run passthrough of permanent productivity innovations to wages. Fourth, we demonstrate that heterogeneity in hold-up power is useful for analyzing dimensions of wage inequality such as the gender wage gap and the effect of superstar firms on wage dispersion.

We begin the paper by introducing a production function that generalizes the O-Ring production in Kremer (1993), allowing for intermediate cases of complementarity in which a failed task or vacant position decreases output in proportion to average product, rather than total product as in the traditional O-Ring case. We introduce a key parameter that varies by occupation called the *degree of complementarity*, which governs how disruptive it is for a position to suddenly go unfilled. Mathematically, the key feature is that turnover generates output losses that are multiplicative in firm average labor productivity. Thus, the output loss from sudden turnover is not just "at the margin" but instead interacts with the firm's inframarginal productivity. This distinction allows us to introduce two separate concepts of marginal product, where the *extensive marginal product* is the marginal output of a firm being one position smaller or larger, and the *intensive marginal product* is the

lost output due to an incumbent worker leaving after the production arrangement has been set. These notions of marginal product differ due to the multiplicative production and the requirement that the combination of positions needed to produce is set each period before turnover and production occur. This is a particular form of lower short-run substitutability of factors, incorporating the idea that labor inputs are indivisible after the firm has assigned tasks.

Next, in a simple two-period model with worker turnover, we show that both individual position production complementarities and position-specific skills are necessary for individual workers to have hold-up power, which generates wage premia at productive firms.[2] We derive a simple closed-form elasticity of wages to firm productivity as a function of worker hold-up power.[3] High labor productivity arises from greater concavity in the revenue function with respect to the number of positions, and productive firms share rents with workers to discourage turnover that is especially costly at productive firms.

To obtain an empirical measure of individual hold-up power for each occupation, we use two sources of information. First, we estimate the effects of worker deaths on a measure of firm profits in Danish administrative data, measured as value added less wages and salaries. Profit losses are highest when the deceased worker is a manager, even relative to the worker's prior wages. When the deceased worker was employed in an occupation in the middle of the wage distribution, such as professionals, technicians, and skilled blue collar jobs, firm profit losses are moderate. Firm profit losses are the smallest relative to prior wages for low-wage service, administrative, and manual occupations. We further show that profit losses tend to be higher at high productivity firms, especially for managers, providing a direct test of our production function.

Second, we estimate occupation-level measures of how differentiated a job tends to be from other jobs within the firm based on the skill requirements in online job postings from Burning Glass Technologies. Following Lazear (2009)'s "skill-weights approach", skills are specific to a position because positions require unique combinations of general skills. Therefore we construct an index that measures, for a given occupation, how different are the task requirements in a given job from the task requirements of other jobs within the firm. This measure should proxy well for individual hold-up power, as task uniqueness

[2]We show that firm-specific skills raise the level of wages but not the slope of wages with respect to firm productivity.

[3]This setup implies that firms are worse off when workers have hold up power and may attempt to "despecify" workers or make production less fragile. For most of this paper, we assume that the complementarities and position specificity associated with employing a particular occupation is an exogenous feature of producing a given type of product. A more general model could allow firms to pay a cost to despecify positions, for example by training workers to learn each others' roles or to change production such that output is less fragile. We discuss an example in Appendix A.5.

should be an indicator both that the output of positions is complementary and that workers must learn position-specific skills to be productive. While our task differentiation measure is correlated with traditional measures of skill such as years of education, there is significant variation in the degree of task differentiation among occupations of similar levels of education. We show that this measure of task differentiation is predictive of profit losses in the event of a worker death, as well as higher losses at high-productivity firms, providing evidence that this measure captures hold-up power well at the detailed occupation level.

We then estimate the effect of innovations to firm productivity on wages across occupation groups using a common instrumental variables strategy that regresses long-run changes of individual workers' wages on long-run changes in productivity, conditioning on workers who stay in the firm. We find the highest passthrough among managerial occupations, with an elasticity of wages in response to permanent shocks of approximately 0.14. Workers in craft and assembly occupations, as well as professionals and technicians, show passthrough elasticities of around 0.06-.07. Workers in administrative, sales, or low-wage service and manuals jobs exhibit passthrough elasticities of 0.04. We find that an occupation's degree of task differentiation is a much better predictor of the effect of firm productivity on wages than an occupation's average educational attainment. Together, these results support the main prediction that individual hold-up power generates wage premia at productive firms, while also showing that hold-up power is not simply a function of traditional measures of skill.

Turning to implications, we show that heterogeneity in individual hold-up power across occupations has implications for both the gender wage gap and the effect of superstar firms on wage inequality. We show that hold-up power, and higher turnover costs in general, raise the level of an occupation's wage. Because men tend to be employed in higher hold-up occupations, rents from hold-up power raise mens' wages by 3 percentage points more than for women, accounting for approximately one fifth of the gender wage gap in Denmark. We also show that in recent years as productivity dispersion increased between small and large firms, wages rose relatively rose at the large, productive firms, but only for occupations that have individual hold-up power. This provides time series evidence that increases in firm productivity dispersion raises wage inequality if occupations differ in their degree of hold-up power.

1.1 Related Literature

A large literature has sought to explain heterogeneous wages for observably identical workers, based on the employer, industry, and institutional status of the worker. Blanchard and

Summers (1986) and Lindbeck et al. (1989) explore insider-outsider theories in which insiders use market power or institutional advantages to exploit high turnover costs. Gibbons et al. (2005) and Gibbons and Katz (1992) explore whether inter-industry wage premia are explained by unobserved worker quality or rent-sharing. Krueger and Summers (1988) argue that the inter-industry wage differentials cannot be explained by competitive wage setting alone. That high wages and queuing may result from training costs and lost output is discussed in Stiglitz (1974). Katz (1986) explores various efficiency wage mechanisms, and Montgomery (1991) offers a search-theoretic explanation in which productive firms offer higher wages to find workers faster.

This paper differs substantially from the much of the literature that relies on upward sloping labor supply in order to generate passthrough of productivity to wages. Card et al. (2018) provide the canonical model of rent sharing in multi-worker firms. Lamadon et al. (2019) use a related framework, where rents arise due to finite long-run extensive margin labor supply elasticities. Card et al. (2016) find heterogeneous labor supply elasticities by gender. Manning (2011) provides a comprehensive overview of imperfect competition in the labor market. Our production function shares a feature with that in Manning (1994), in which firms operate with a 'blueprint,' and deviations in the firm's labor force away from the optimal level specified in the blueprint generate larger output losses than would be lost by operating with a smaller blueprint.

This paper is related to research that explores how differentiation of tasks within a firm makes skills specific to the firm. Lazear (2009) argues that skills are firm specific because firms require particular combinations of general skills. Leping (2009) measures how task differentiation differs across occupations as a predictor of on-the-job training. Yamaguchi (2012) conceives of occupations as bundles of tasks to measure the similarity of and distance between occupations. Edmond and Mongey (2019) argue that low-education occupations have homogenized and that firms' inability to 'unbundle' labor gives workers in highly differentiated occupations rents due to comparative advantage. Lise and Postel-Vinay (2020) describe occupations as bundles of tasks and argue that cognitive skills take longer to acquire than manual skills. Haanwinckel (2018) models workers as assigned to tasks in different degrees, where a firm is a bundle of tasks and firms face upward sloping labor supply due to idiosyncratic preferences across firms. Goldin and Katz (2016) argue that increasing substitutability of workers within an occupation decreases the part-time hours premium and the gender wage gap.

A range of papers suggest that workers have heterogeneous ability to bargain with employers. Hall and Krueger (2012) show that employers of high-wage occupations tend to bargain, while employers of low-wage workers tend to post wages. Caldwell and Harmon

(2019) show that when a worker's knowledge of outside offer increases due to growth at firms of former co-workers, higher wage occupation see the largest percent gains in wages. Lachowska et al. (2021) show that when dual job-holders receive a raise in their secondary job, high-wage workers tend to get raises in their primary job, while low-wage workers tend to switch jobs. Our framework provides a microfoundation for why match surplus to the firm is larger for high-wage occupations.

This paper shares a feature with Stole and Zwiebel (1996) that inframarginal productivity affects wage setting. However, our model does so without relying on off-equilibria breakdowns in negotiations to affect wage determination. Additionally, our framework allows for estimating heterogeneity in how inframarginal productivity affects wages across occupations. Cahuc et al. (2006) estimates heterogeneous bargaining power accounting for on-the-job search.

Empirically, we build on Bertheau et al. (2021) and Jäger (2016) in the worker deaths literature. We find in common with Jäger (2016) that managers in particular exhibit high complementarities with other workers in the firm.[4] Bennedsen et al. (2020) finds that hospitalization of CEOs also affects firm performance. Our study relates to the estimation of employee replacement costs as in Dube et al. (2010) and Muehlemann and Leiser (2018). These studies generally find that replacement costs are larger relative to wages in high-skill occupations, and that training costs relatively outweigh recruiting costs as a share of all turnover costs. Bertheau et al. (2021) estimate that turnover costs are much higher than in previous studies.

Cobb and Lin (2017), Bloom et al. (2018), and Mueller et al. (2017) point out that large firms continue to offer a premium to high-wage college educated workers, but non-college workers no longer earn a premium at large firms. While our study does not approach this question directly, we hypothesize that (i) some low-education occupations may have 'de-specified', and (ii) the occupational composition of low-education work has shifted toward low-specificity occupations. In the absence of institutions that impose wage compression within the firm, large, productive firms may show growing within-firm wage inequality. Gregory et al. (2020) demonstrates that firms have heterogeneous effects on workers' wage growth, arguing that firms provide heterogeneous learning environments for acquiring human capital.

Various authors have also studied turnover costs and complementarities to understand the durability of employment relationships. Nagypál (2001) develops a related two-worker firm model, in which two types of labor are complements, but the quality of the match in the skilled job is learned over time. Oi (1962) argues that firms are less likely to layoff

[4]We explicitly reconcile our findings with Jäger (2016)'s findings in Appendix A.9.

high "fixity" workers in recessions. In companion papers Jovanovic (1979a) and Jovanovic (1979b) argues that workers learn on the job about match quality over time and that the probability of turnover depends on the current beliefs of the quality of the match.

That workers may be complementary in production is explored in the organizations literature, such as Baldwin (2018). Eisfeldt and Papanikolaou (2013) argue that key talent that is specific to the firm is a key component of organization capital, and that better outside options of key workers renders firms with high organizational capital to be riskier, further implying that key talent has claims to firm's cash flows.

The rest of the paper is organized as follows. Section 2 develops the theory of individual hold-up power and derives a closed form elasticity of wages to average product that depends on workers' individual hold-up power. Section 3 describes the data used in this paper, and Section 4 estimates the degree of hold-up power by occupation. Section 5 estimates passthrough of changes in firm productivity to wages by occupation. Section 6 discusses implications for wage inequality. Section 7 concludes.

2 Theory

In this section, we derive a production function with production complementarities between individual positions within firm. We then embed firms that use this production technology into a simple frictional labor market and show that workers can extract higher wages from high productivity firms only if production exhibits complementarities *and* if skills are position-specific.

Note that we argue that individual hold-up power is an occupation property. In Sections 2.1 and 2.2, we will assume that any given firm produces using positions of only one type of occupation. In Section 2.3, we will generalize the model so that firms produce using multiple occupations, each with their own degree of complementarity and degree of hold-up power, and we show that the predictions are nearly identical as under one-occupation firms. We denote occupations with subscript j.

2.1 Firms with Positions and Individual Complementarities

We begin by introducing a general production function that takes as arguments the number of positions in the firm and the share of positions that are filled with productive workers. Firms produce in only one period and is exogenously endowed with a production function and define two production inputs, the number of positions N, and the share of positions filled with fully specific, trained workers X. An unfilled position indicates that either a worker is unproductive or that the position is vacant. We take a stand on this

distinction later in this section. We also define the degree of complementarity $m \geq 0$, which is a parameter that governs how disruptive an unfilled position is to the production process. In the simplest version of the model, all positions will have the same degree of complementarity.[5] Let output take the form:

$$Y = F(N)g(m, X).$$

The general form of production has two parts: a familiar outer structure that determines the potential revenue product of the firm as a function of the number of positions $F(N)$, with $F'(N) > 0$ and $F''(N) \leq 0$.[6] The second term $g(m, X)$ is the share of potential output that is achieved due to the extent that the firm is producing with specific workers and has three properties:

$$(1): \ \ g(m,1) = 1 \qquad (2): \ \ g_X(m,X) > 0 \qquad (3): \ \ \frac{\partial^2 \log(g(m,X))}{\partial X \partial m} > 0.$$

The first condition states that if all positions are filled with trained workers, then the firm achieves its potential output given N, regardless of degree of complementarity m in production. The second states that the share of potential output that is achieved is increasing in the share of positions filled X. The third states that the percent response of g to changes in the share of filled positions X is higher when m is higher. That is, when complementarities are high, filling positions with productive workers generates more output in percent terms, and losing productive workers is more costly to output in percent terms.

In Appendix A.1 we introduce a production function that generalizes the O-ring production function of Kremer (1993). We show that this generalized O-ring is approximated by the production function in the next section, which is a special case of the general production function $Y = F(N)g(m, X)$ as above.

Generalized O-Ring Production For the rest of the paper, we will use a large firm approximation of generalized O-ring:

$$Y = AN^{\alpha} e^{-(\alpha+m)(1-X)}, \tag{1}$$

where N is a continuous measure of the number of positions, m is the degree of complementarity for position i, and the function that determines the share of potential output

[5] In Section 2.3, we will generalize the function to allow firms to utilize multiple occupations, each with their own degree of complementarity.

[6] We do not distinguish whether diminishing returns of $F()$ come from declining marginal physical product or downward sloping demand.

is $g(m, X) = e^{-(\alpha+m)(1-X)}$.[7] Recall that X is the share of positions that are filled with trained workers with position specific skills, and we assume that the remaining $1-X$ share of positions are filled with unproductive workers.[8] Intuitively, the exponential structure indicates that marginal unfilled positions decrease output by a constant proportion, and this proportion increases with the degree of complementarity m.

We define two new terms that will be central to the following analysis: the *extensive marginal product* and *intensive marginal product*, denoted MPL^E and MPL^I, respectively. In words, the extensive marginal product is the additional output from the firm becoming larger, holding the share of positions filled constant:

$$MPL^E = \frac{\partial Y}{\partial N}\Big|_X = \alpha A N^{\alpha-1} e^{-(\alpha+m)(1-X)},$$

with $X \in [0, 1]$.[9] To understand this expression, consider the case where all positions are filled and the firm adjusts its size: $MPL^E|_{X=1} = \alpha A N^{\alpha-1}$, which is the expression for marginal product under neoclassical production with Cobb-Douglas. If the share of positions filled X is less than 1, then the marginal output from the firm growing the number of positions, but holding the share of positions fixed, will be lower than if all positions were filled: $e^{-(\alpha+m)(1-X)} < 1$ if $X < 1$.

The intensive marginal product is the additional output gained by filling positions with productive workers, holding the number of positions N fixed:

$$\begin{aligned} MPL^I = \frac{\partial Y}{\partial X}\Big|_N \times \frac{1}{N} &= (\alpha + m) A N^{\alpha-1} e^{-(\alpha+m)(1-X)} \quad \text{if } X \in [0, 1] \\ &= \underbrace{\alpha A N^{\alpha-1}}_{MPL^E|_{X=1}} + \underbrace{m A N^{\alpha-1}}_{\substack{\text{Multiplicative} \\ \text{losses}}} \qquad \text{if } X = 1. \end{aligned}$$

Evaluating the intensive marginal product at all positions being filled ($X = 1$), we see that the intensive marginal product is the sum of two terms: the extensive marginal product and an additional term that is equal to m multiplying the firm's average product

[7] We discuss this discrete case in more detail in Appendix A.1 and show that the continuous version is a large firm approximation of the discrete generalized O-ring production.

[8] This will be important when defining firm average product.

[9] We assume that the remaining $N(1 - X)$ positions are occupied by workers who are unproductive in their roles. While in output terms there is no difference between a vacant position and one filled with an unproductive worker, this assumption allows for a more convenient derivation of model properties and is consistent with the remainder of the paper. We relax the assumption that untrained workers are completely unproductive in the next section, Section 2.2.

$mAN^{\alpha-1} = m \times \frac{Y}{N}$. Lastly, the expression for average product is:

$$\begin{aligned} APL = \frac{Y}{N} &= AN^{\alpha-1}e^{-(\alpha+m)(1-X)} && \text{if } X \in [0,1] \\ &= AN^{\alpha-1} && \text{if } X = 1. \end{aligned}$$

Together, these expressions generate two convenient properties. First, similar to standard Cobb-Douglas, the ratio of average product to extensive marginal product is constant. Secondly, relating the two marginal product terms, the intensive marginal product MPL^I is weakly greater than the extensive marginal product MPL^E and is strictly greater if $m > 0$:

$$\frac{APL}{MPL^E} = \frac{1}{\alpha}, \qquad MPL^I = \big(1 + \frac{m}{\alpha}\big)MPL^E = (\alpha + m) \times APL.$$

The key features is that the output losses from losing productive workers interacts with inframarginal productivity only if $m > 0$.

Production Function Intuition While the production function describes firms with a continuum of positions, it is useful to consider a discrete case. Consider two firms that share identical production functions, but firm A has 100 positions and firm B has 99 positions. If all positions at both firms are filled with fully productive workers, then firm A will produce more than firm B. However, suppose that firm A loses a trained worker, and so has 100 positions but only 99 filled. If $m > 0$, then output will be *lower* at firm A than at firm B. This feature can be seen as a form of lower ex-post substitutability of labor inputs once production has been set, i.e. production follows a "blueprint".[10]

2.2 Position Specific Skills and Individual Hold-Up Power

In this section, we use a simple two period model to show that production complementarities alone are insufficient to generate productivity-wage premia in multi-worker firms with diminishing revenue returns to labor, and it is necessary that both the production process exhibits complementarities *and* that workers' skills in performing complementary

[10]The short-run rigidity of production, and its fragility to disruption from individual workers, raises the question of whether firms can adjust production to be less dependent on a single worker. Appendix A.3 outlines a dynamic model in which newly hired workers acquire position-specific skills each period with probability p. This model can also be reinterpreted as the firm being able to rearrange production to fit the skills of the newly hired worker with probability p each period. Another extension includes a choice margin where firms can expend resources to decrease the degree of complementarity m, either by paying a fixed or a flow cost. In Appendix A.5, we derive one such scenario where firms can expend flow resources to decrease m.

tasks are position specific. We additionally show that when firms choose the number of positions with which to produce, measured firm productivity in terms of output per worker is a function of the concavity of the revenue function with respect to the number of positions α. We discuss how our model differs from other common models of imperfect competition and wage premia in Appendix A.7.

Two Period Model There are two periods. At the beginning of period 1, the firm chooses the number of positions N with which it will produce. The firm sets a common wage policy for all workers w, which it must commit to pay in the later period to all workers in the firm. At the end of period 1, N workers are hired and trained in their specific positions. In period 2, workers draw from a distribution of outside wages $F(w')$. Workers care only about wages and change jobs if the competing firm offers a higher wage. Prior to production at the end of period 2, the firm can hire untrained workers to fill the vacant positions. New hires are less productive than trained incumbents, and let $d \in [0, 1]$ denote the gap in productivity in complementary tasks between fully trained and new workers; d is therefore the measure of position specificity. New hires are paid the same wage as incumbent, trained workers.[11] The firm's problem becomes:

$$\max_{N,w} AN^{\alpha} e^{-(\alpha+dm)(1-X)} - wN \tag{2}$$

subject to:

$$X = F(w),$$

where $F()$ is the CDF of the outside offer distribution. Let $F(w') = 1 - \left(\frac{w'}{\underline{w}}\right)^{-\gamma}$, for $w' \geq \underline{w}$. In this functional form, γ is the elasticity of the quit probability with respect to the wage. Solving for the firm's optimal choice of wage gives:

$$w^* = \left(\gamma\Big(1 + \frac{dm}{\alpha}\Big)\right)^{\frac{1}{\gamma}} \underline{w}. \tag{3}$$

There are a few important aspects to note about the firm's choice of wage. First, as the market minimum wage $\underline{w}$ increases, the firm's optimal wage increases as well: $\frac{\partial w^*}{\partial \underline{w}} > 0$. Second, as the degree of complementarity m increases, the optimal wage increases, as the losses of a worker leaving grow as m increases. $\frac{\partial w^*}{\partial m} > 0$. Third, the wage is increasing in

[11] This assumptions allows us to derive a closed-form solution for the elasticity of wages to average product. In Appendix A.4, we relax this assumption and allow the untrained workers to be paid a lower wage, specifically the minimum outside wage $\underline{w}$.

position specificity (the gap of productivity between trained and untrained replacement workers): $\frac{\partial w^*}{\partial d} > 0$.[12]

We also have that $\frac{\partial w^*}{\partial \alpha} < 0$ if $dm > 0$. While in simple terms α is just the curvature of the production function, α^{-1} is also the ratio of average to extensive marginal product. Therefore, as α falls and output becomes more concave with respect to number positions, average product endogenously rises, and the product $dm > 0$ gives rise to a positive relationship between average product and wages. Under the assumption that heterogeneity in α the only source of heterogeneity in average product across firms (i.e., complementarities m and position specificity d are constant across firms), then the elasticity of wages to average product is:

$$\varepsilon_{w,apl} = \frac{\mathbf{h}}{\gamma\alpha + (1+\gamma)\mathbf{h}}, \tag{4}$$

where $\mathbf{h} = dm$ is the measure of individual hold-up power over firm average product, equal to the product of the degree of complementarity m and position specificity d.[13] This expressions shows that $\varepsilon_{w,apl} \to 0$ as $\mathbf{h} \to 0$ (i.e., $m \to 0$ or $d \to 0$), and $\varepsilon_{w,apl} \to \frac{1}{\gamma+1}$ as $\mathbf{h} \to \infty$. This says that if positions exhibit no complementarities *or* if skills are not position specific, i.e. $\mathbf{h} = 0$, then higher productivity firms in terms of observed average productivity pay no wage premium. As the product of complementarities and position specificity $\mathbf{h} = dm$ grows, the elasticity of wages to average product increases.

Another result is that total factor productivity, A, has no effect on a firm's optimal choice of wages. Instead, A only affects firms size: as A increases, the firm's optimal number of positions increases. Facing the same retention function, two firms with the same α, m, and d but different values of A will each choose to expand until their extensive marginal products are equalized.[14] Since the ratio of average to extensive marginal product is just $1/\alpha$, the two firms will have identical average products and wages. This result ultimately comes from the assumption of linear extensive margin wage costs: the marginal cost of another position is constant and independent of firm size, unlike in standard monopsony

[12]This problem is subject to a technical condition that replacement workers must be sufficiently productive that the firm finds it optimal to hire the replacement workers: $(1-d)m > 1$. That is, untrained workers must be at least as productive as w^* to justify hiring them. This requires both that (i) complementarities m are sufficiently large and replacement productivity $(1-d)$ is sufficiently high. However, this technical restriction is a function of the two-period setup, since in a dynamic setting, the value of untrained workers to the firm is their future productivity. In Appendix A.3, we derive a dynamic version of the firm's problem that yields an identical wage policy but has no technical restrictions on d and m.

[13]We provide a full derivation in the Appendix A.2.

[14]Bloesch and Larsen (2022) explore the conditions under which TFP results in long-run changes in firm size rather than wages. We show suggestive evidence in Appendix B.7 that the firm size-wage premium is very small in Denmark.

models. Instead, the choice of the wage reflects the firm's tradeoff between per-position wage and turnover costs, and the firm's size is chosen independently given that optimal wage.

Lastly, a key endogenous result is that the wage is equal to the extensive marginal product:

$$w^* = MPL^E.$$

This condition simply results from firm optimization. The extensive marginal product is the additional output the firm gets from being one position larger, and the wage is the cost.[15] Firm optimizing implies that these two terms will be equalized if MPL^E is decreasing in the number of positions. This result will helpful for identifying hold-up power $\mathbf{h}$ when we turn to estimation in Section 4.

Complementarities without Position Specificity - If output exhibits production complementarities ($m > 0$), but skills are not specific to a position ($d = 0$), then workers will not have individual hold-up power over the rest of the firm's output. In this scenario, a firm can swap in workers as soon as workers leave and so never faces multiplicative losses. Firms may even over-hire if positions require on-the-job learning but workers are replaceable across positions ex-post. As an example, consider a catering firm that requires twenty staff to be present to staff an event, and the services are dysfunctional if the team is understaffed. The firm may train more than twenty workers and compensate workers to be on call, so that if a workers suddenly leaves, the firm never falls below the critical threshold of twenty. While workers in this setting may have significant collective hold-up power over firm output, they have little hold-up power individually.

Position Specificity without Complementarities - If workers have position specific skills, but output is not individually complementary, then workers will have no hold-up power over the rest of the firm's output. Some production environments may include highly differentiated labor, but ultimately the outputs for the production unit are not interdependent. For example, consider a newspaper that loses an international reporter. This reporter's work may be highly specialized, but it is unlikely to affect the work of the newspaper's opinion columnists.

2.3 Bridging the Model to Data

In Sections 2.1 and 2.2, we developed a stylized model to derive predictions relating individual hold-up power $\mathbf{h}$ to wages. In order take the model to the data, we must make

[15] This also implies that if there are production complementarities, i.e., $m > 0$, the wage will be less than the intensive marginal product MPL^I.

two modifications: (i) properly account for turnover costs that are not due to production complementarities (i.e., training costs, search costs, etc.), and (ii) generalize to allow firms to produce using multiple occupations, each with their own degree of hold-up power.

"Marginal" Turnover Costs What differentiates our model is that position specific skills allow workers to disrupt firm inframarginal productivity. However, not all turnover costs originate from multiplicative output losses. In many models, including those with search costs, training costs, or firm specific skills, turnover costs are paid in wages or affect output at the margin. We will show that these marginal or "non-multiplicative" turnover costs increase the *level* of the wage, but only individual hold-up power increases the *slope* of wages to average productivity.

Define δ as the per-period cost of unfilled positions as a fraction of per-period wages of trained workers. Then the firm's output is:

$$Y = AN^{\alpha}e^{-(\delta\alpha+\mathbf{h})(1-X)}.$$

In the case of $\delta = 0$, all turnover costs come from multiplicative output losses. As δ increases above 0, the non-multiplicative turnover costs increase. For example, the model in the previous section is a special case with $\delta = 1$, resulting from the firms paying full wages to the new workers while new workers are unproductive in non-complementary tasks. Allowing for a flexible value of δ in the firm's problem from the prior section, the optimal wage is:

$$w^* = \left(\gamma\left(\delta + \frac{\mathbf{h}}{\alpha}\right)\right)^{\frac{1}{\gamma}} \underline{w},$$

The elasticity of wages to average product when varying α is then:

$$\varepsilon_{w,APL} = \frac{\mathbf{h}/\delta}{\gamma\alpha + (\gamma+1)\mathbf{h}/\delta}, \tag{5}$$

where the elasticity depends on $\mathbf{h}/\delta$. In practice, the passthrough elasticity will depend on what share of turnover costs come from disruptions to inframarginal product versus non-multiplicative "marginal" turnover costs. Because marginal turnover costs raise the level of the wage, the increase in wages that a worker gets from a lower α (and hence higher average product) when $h > 0$ is proportionally less.

Now that we have fully defined the replacement productivity of untrained workers, we can define the flow surplus to the firm of a match with a fully trained, specific worker, net of the productivity of an untrained worker in occupation j:

$$Surplus_j^{firm} = \Big(\frac{\partial Y}{\partial X}\frac{1}{N}\Big)^{net} = \delta_j MPL_j^E + \mathbf{h}_j APL$$
$$= w_j^*\big(\delta_j + \frac{\mathbf{h}_j}{\alpha}\big). \tag{6}$$

Thus when firms are optimizing, the flow surplus of a match to the firm will be equal to the (optimal) wage times the sum of δ and $\frac{\mathbf{h}}{\alpha}$, which respectively correspond to the non-multiplicative or "marginal" turnover costs and the inframarginal turnover costs from individual hold-up power. This formulation will be useful when we empirically identify $\mathbf{h}$ using exogenous worker separations in Section 4.

The total match surplus to the firm will be the present discounted value of flow surpluses, which would require extending the model into a dynamic setting. In Appendix A.3, we derive a dynamic version of the model in which untrained workers become fully specific, productive workers with probability p each period, with an expected training time of $1/p$.[16] The total surplus of the match will then depend on the speed of learning and the rate at which firms discount.

Multiple Occupations If firms produce with two occupations a and b, we can write the firm's production function and maximization problem as:

$$\max_{N_a,N_b,w_a,w_b} = \Big(N_a^{\sigma_a} N_b^{\sigma_b}\Big)^{\alpha} e^{-\sum_j \sigma_j(\delta_j\alpha+\mathbf{h}_j)(1-X_j)}, \quad j \in \{a,b\}$$

subject to:

$$X_j = F(w_j).$$

We show in Appendix A.6 that the optimal wage is identical to equation (6) for each occupation j and that the passthrough elasticities are nearly identical to equation (5).

3 Data

In the empirical part of this paper, we use two different sources of data: administrative data on workers and firms from Denmark from 2008 to 2016, and job posting data in the US from Burning Glass Technologies from 2010 to 2018.

[16] We show that that the optimal wages is $w^* \approx \left(\frac{\gamma}{p}\left(\delta + \frac{\mathbf{h}}{\alpha}\right)\right)^{\frac{1}{\gamma}} \underline{w}$

3.1 Danish Administrative Data

We use administrative from Denmark on workers and firms from 2008-2016 to generate a merged employer-employee panel data set.[17] For the worker data, we use the IDAN registry (Danish Statistics (2020d)), which reports workers' earnings, hours, occupation, firm, and establishment.[18] Beginning in 2008, the reported wage data is drawn from the e-indkomst monthly online reporting database, which is regarded as highly reliable. We also use the BFL database (Danish Statistics (2020b)) which reports similar information as IDAN but on a monthly level. We use the demographic registries to attain information on age (Danish Statistics (2020e)) and education (Danish Statistics (2020f)). The timing of worker deaths is obtained in the DOD registry (Danish Statistics (2020a)), which records the date of all deaths in Denmark.

For firm data, we use the FIRM registry (Danish Statistics (2020c)), which is an annual dataset that includes data on firm sales, value added, employment (measured in full time equivalents), gross salaries, and gross profits. The unit of observation is the firm which is identified with variable *cvrnr* rather than establishments which are identified with *arbnr*, as firm financial variables such as value added are reported only at the firm level. Financial data is reported from May of year t to April of year $t+1$.

In the section estimating the effect of worker deaths, we use data from 1,218 firms who experience a worker death, all of which have between 15 and 100 full-time equivalent employees. The summary statistics for these firms can be found in first three rows of Table 1. The firms in the sample on average employ 45.9 full-time equivalent workers and produce approximate $4.3 million in value added.

In our passthrough exercises, we use a much larger sample of workers and firms, including all full-time workers employed at a private sector firm that reports value added and industry, whose firm employed an average of at least 15 full-time equivalent employees. Summary statistics on the firms in this sample can be found in the bottom three rows of Table 1. There are 100,993 unique firms, with 8,081,081 worker-year observations. There are approximately 1 million worker-firm observations per year, out of a labor force of about 2.5 million.[19]

Grouping Occupations Danish occupations are categorized using the Danish International Standard Classification of Occupations, or DISCO codes, which has ten categories.

[17] The merged employer-employee data is described in detail in Bobbio and Bunzel (2018).

[18] Many workers' occupations change in 2009 but are the same in years 2008 and 2010. For workers whose employer is the same between 2008-2010 and the occupation is the same in 2008 and 2010, we impute the occupation in 2009.

[19] Public sector, part-time, and self-employment account for the remainder of the workforce.

Table 1: Firm Summary Statistics

	Deaths Sample - Firms			
N =1218	mean	sd		
Full Time Equivalent (FTE)	45.9	23.7		
Value Added ($, million)	4.327	3.894		
VA-salaries ($, million)	.851	2.936		
	Passthrough Sample - Firms			
Average taken across:	Firms		Workers	
N= 8,081,081	mean	sd	mean	sd
Full Time Equivalent (FTE)	75.9	346	1,794	3,966
Value Added ($, million)	9.14	63.1	221.6	661
VA/Worker ($, thous)	106.7	389	116.3	178

This table reports summary statistics for the firm samples used in the deaths and passthrough analyses. In the deaths analysis, we restrict our attention to firms with an average number of full time equivalent employees between 15 and 100 in the three years before the death. The average firm has 45.9 workers and produces approximately $4.3 million worth of value added each year (deflated to 2005 dollars). In the passthrough sample, we use larger firms, restricting attention to firms with an average of at least 15 full time equivalents over the sample. The average firm in the sample has 76 full time equivalent employees, but the mean of firm size across workers is 1,794 employees. Value added per worker is similar depending on if we weight by firms or workers, at about $100,000 per worker.

We exclude analysis on agricultural and armed forces occupations. Summary statistics on workers' wages, average tenure, differentiation score (which we will explain in Section B.5), average years of schooling, and share of employment by DISCO group in the 2008-2016 passthrough sample are shown in Table 2.

Throughout the paper, we divide workers into four broad occupation groups: managers (DISCO group 1), professionals/technicians (groups 2 and 3), crafts/assembly (groups 7 and 8) , and administrative/service/manual (groups 4, 5, and 9). At the top of the labor market, both managers and professionals exhibit high wages, high levels of education, and high levels of within-firm task differentiation. However, we treat these groups separately because we conjecture that managerial inputs may be more complementary in production than non-managerial inputs. We combine other occupation groups based on occupational characteristics and worker flows across occupation groups.

Institutional Setting Denmark is known for its "flexicurity" model of labor market policies, characterized by a generous welfare state and high union density, but with limited employment protections, high labor mobility, and industry-set minimum wages. The level of labor mobility, measured by the rate of hiring and separations, is high for OECD countries and is more similar to the US than other economies in continental Europe (Caldwell

Table 2: Mean Characteristics of 1-Digit Occupation Groups

	(1)	(2)	(3)	(4)	(5)
DISCO Group	Log Wage	Tenure	Diff. Score	Educ	Emp Share
Managers	5.77	7.6	1.89	14.8	5.9
	(.54)	(7.1)	(.45)	(2.2)	
Professionals	5.59	5.7	1.41	16.2	13.6
	(.36)	(5.6)	(.84)	(2.2)	
Technicians	5.38	6.3	0.62	14.5	15.7
	(.35)	(6.3)	(.77)	(1.9)	
Crafts	5.26	6.6	-0.01	13.5	13.2
	(.34)	(6.8)	(.84)	(1.9)	
Assembly	5.23	7.2	-0.87	12.4	11.5
	(.23)	(6.9)	(1.27)	(2.3)	
Administrative	5.15	6.4	-0.02	13.5	10.6
	(.30)	(6.6)	(.74)	(2.0)	
Service	4.91	3.6	-2.01	12.2	15.7
	(.37)	(4.3)	(.88)	(2.1)	
Manual	5.03	4.5	-1.51	11.9	13.8
	(.37)	(5.3)	(.89)	(2.5)	

This table reports summary statistics for the 1-digit occupation groups in Danish data, using the sample specified in Section 3. The data is from 2008-2016 including firms with on average at least 15 employees that also report value added. Values at the top of each cell are the mean, and values in parentheses report standard deviations.

and Harmon (2019)). There is no national minimum wage, but minimum wages are set at the industry level through collective bargaining agreements between labor unions and employer associations. In recent decades, bargaining at the firm level has become more important in determining wages (Dahl et al. (2013)), and firms have considerable ability to raise individual wages above collectively agreed minimums.

3.2 Burning Glass Technologies (BGT) Job Posting Data

We use the vacancy data from Burning Glass Technologies (Burning Glass Technologies (2018a), Burning Glass Technologies (2018b), henceforth "BGT"), which has nearly the entire universe of online job postings in the United States from 2010-2018. The job postings include information such as location, employer, and 6-digit SOC occupation codes. Further, the job postings include detailed information about the skill requirements posted with every vacancy, which BGT cleans and categorizes into broad skill categories.

4 Measuring Individual Hold-up Power

In this section, we measure individual hold-up power by occupation using two methods. First, we estimate the effect of worker deaths of firm profits, measured as value added less wage and salary payments, across occupations and across firms of different levels of productivity. This provides estimates of total turnover costs and hold-up power for four broad occupation groups. Second, we estimate a measure of within-firm, across-position task differentiation for each occupation from US job postings data and merge to Danish occupations, providing us a proxy of hold-up power at the detailed occupation level.

4.1 Identification of Parameters with Exogenous Worker Separations

In this section, we will show that we can identify the degree of hold-up power **h** if we observe exogenous separations of workers from firms by observing changes in firm-level value added less salaries. We are empirically estimating the surplus to the firm derived in Section 2.3. We put forward the following lemma:

Lemma 1 *If workers are exogenously separated and firms do not adjust the number of positions, then the per period total turnover costs* $\delta + \frac{h}{\alpha}$ *is identified by the change in reported value added following a worker separation:*

$$\Delta Y \approx \left(\Delta X \times N \times w_{\text{-}1}\right) \times \left(\delta + \frac{\boldsymbol{h}}{\alpha}\right)$$

If workers are exogenously separated and firms can adjust the number of positions N *holding the effective share of positions* X *fixed after separation has occurred, then total per period turnover costs* $\delta + \frac{\boldsymbol{h}}{\alpha}$ *are identified by the change in value added less wage and salary payments following an exogenous worker separation:*

$$\Delta(Y - wN) \approx \left(\Delta X \times N \times w_{\text{-}1}\right) \times \left(\delta + \frac{\boldsymbol{h}}{\alpha}\right). \tag{7}$$

$\boldsymbol{h}$ *and* δ *are separately identified with an interaction term on firm productivity:*

$$\Delta(Y - wN) \approx \underbrace{\left(\Delta X \times N \times w_{\text{-}1}\right)}_{\textit{Shock to firm}} \times \delta \; + \; \underbrace{\left(\Delta X \times N \times w_{\text{-}1} \times \frac{1}{\alpha}\right)}_{\textit{Interaction with firm productivity}} \times \boldsymbol{h}.$$

Proof: see Appendix A.8.

It is worth noting that value added less salaries, rather than value added, is the outcome of interest. The main reason is that firms may downsize following the death of a worker, and the incentive to downsize may differ based on the type of worker. Since a lower share of positions filled X decreases the productivity of other workers in the firm, deaths of

workers with higher hold-up power will cause larger declines in the productivity of the remaining workers. Therefore, we would expect that firms that experience a death of a particularly high hold-up worker will shrink the number of positions by more. With that said, firms are maximizing output net of costs, and if firms shrink their size, their wage costs fall as well.[20] Therefore, value added less salaries is the correct measure to relate to the firm's maximization problem and also addresses bias introduced if the outcome were value added due to differential incentive to downsize following a worker death.

The intuition for the third statement in Lemma 1 is that identifying **h** comes from the feature that for the same hold-up power **h**, output losses are larger at higher productivity firms. Using the exogenous separations of workers in the same occupation but across firms of different productivity, we can uncover the rate at which higher firm productivity generates higher output losses, thereby identifying **h**.

Sample Our sample uses deaths between 2008 and 2015 and firm outcomes from 2005 to 2017. We restrict our sample to deaths of non-owner workers[21] who were employed at the firm in the month prior to the death and who worked at least 100 hours in the prior month.[22] We select firms with 15 to 100 full time equivalent employees to focus on firms in which the effect of a death will be sufficiently large relative to the typical variance of output and profit. Due to our relatively short sample, the event window is six periods, denoted $s \in \{-3, 2\}$. The death occurs in period $s = 0$. Firms k are matched to five placebo firms on the exact year and 1-digit industry in period s = -1. The five selected control firms have the least sum of squares of differences in log firm size (measured in full time equivalent workers) and value added less salaries in periods s = -1 and s = -3. We average the outcomes of the control firms using equal weights. Treated firms are excluded from the pool of control firms.

Empirical Specification To estimate equation (7), we will run regressions of the form:

$$\tilde{Y}_{ks} = c^j + \psi_k^j + \sum_{s=-3}^{2} \beta_s^j \times \mathbb{1}\{s\} \times \tilde{w}_{k(i),\text{-}1}^j + \varepsilon_{ks}^j, \tag{8}$$

[20] Value added less salaries then provides a measure of accounting profit which does not account for payments to capital.

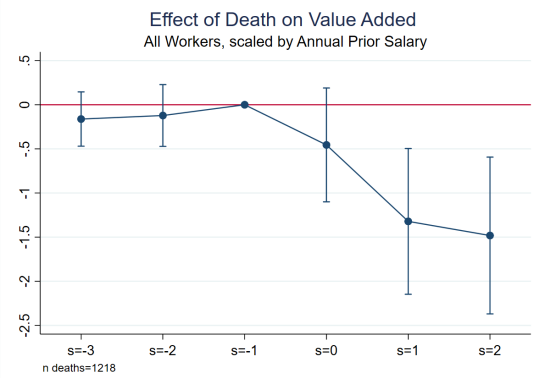
[21] Individuals in this sample are very unlikely to be employers: our deaths sample is dervied from the BFL database, where very few workers have a value of "A" for the variable *type*, indicating that the individual is the owner of a firm.

[22] Other studies such as Bertheau et al. (2021) and Jäger (2016) restrict the sample to exclude deaths where the worker had a long-term illness or had prior visits to the hospital. We show how the results in this sample differ in Appendix B.3.

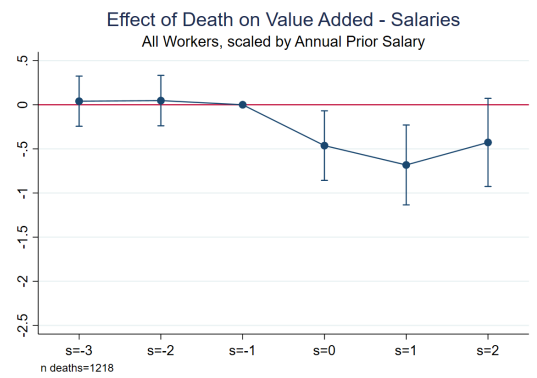
separately for each occupation j and pooling across occupations. The regression in equation (8) recovers a composite coefficient for total turnover costs: $(\beta_0^j + \beta_1^j + \beta_2^j)/3 = -(\delta + \frac{\mathbf{h}}{\alpha})$, where the units on the coefficients are the average annual flow loss in years of prior salary for three years. The outcome variables of interest Y is value added less salaries. We construct a within-match k, within-period s difference $Y_{ks}^{dif} = (Y_{ks}^{treated} - \bar{Y}_{ks}^{placebo})$, where $\bar{Y}_{ks}^{placebo}$ is the average value of the five placebo firms matched to treated firm k for period s. We winsorize these differences Y_{ks}^{dif} at the 10th and 90th percentiles separately for pre- and post-treatment observations, creating Y_{ks}^{win}.[23] We then divide both sides of the equation by the firm's size prior to the worker's death, generating $\tilde{Y}_{ks} = Y^{win}/N_{k,\text{-}1}$, where $N_{k,\text{-}1}$ is average number of workers in the three years prior to the shock. $w_{k(i),\text{-}1}^j$ is the prior annual wages of the the deceased worker i in occupation j at firm k, and $\tilde{w}_{k(i),\text{-}1}^j = w_{k(i),\text{-}1}^j/N_k$ scales the size of the shock to the firm by the prior size fo the firm. Monetary values are deflated into 2005 Danish kroner and converted to dollars using 6 DKK/USD. Standard errors are clustered at the match level, where each cluster has six observations of a treatment-placebo matched pair. Both the outcome variable and the wage variable are annual, allowing us to interpret coefficients as losses to value added less salaries in years of average prior wages.

Figure 2: Effect of a Worker Death on Firm Value Added and Value Added less Salaries

(a) Value Added



(b) Value Added - Salaries

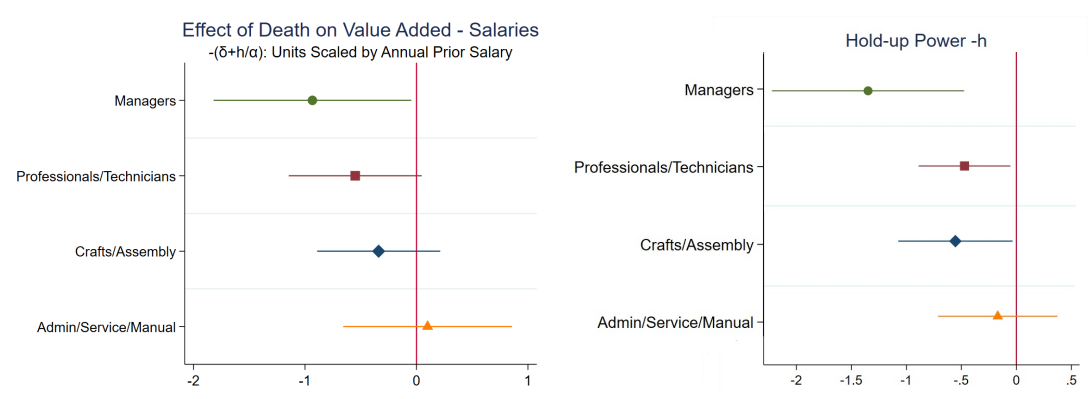


These figures report the event study for the specification pooling all occupations. Both wages and the outcome variables are annual, so the coefficients can be interpreted as flow losses per year in years of prior wages. Panel (a) reports the estimates when the outcome variable is value added. Panel (b) reports the estimates when the outcome variable is value added less salaries. Standard error bars report 95% confidence intervals.

[23] We discuss data cleaning and the winsorization procedure in more detail in Appendix B.1.

Figure 3: Effect of Death on Firm Value Added net of Wages and Salaries, Years of Prior Wages per Year

(a) Turnover Costs $\delta + \mathbf{h}/\alpha$ (b) Hold-up Power $\mathbf{h}$



This figure shows the value added losses in years of prior salary per year. Panel (a) reports the the number year of value added salaries lost, in units of the average prior wage of workers, averaged across 3 years from period $s = 0$ to $s = 2$: $(\beta_0^j + \beta_1^j + \beta_2^j)/3$. A point estimate of -.55, for example for crafts/assembly, indicates that the estimated per-year losses in value added less salaries is equal to .55 years of average prior pay of workers in that occupation, implying a cumulative profit loss of over a year and a half worth of prior wages. These estimates come from 118 deaths of workers classified as managers; 315 deaths of workers classified as professionals/technicians, 452 deaths of workers classified as crafts/assembly, and 333 deaths of workers classified as administrative/service/manual. Panel (b) reports losses of value added less salaries, in years of prior wages, at high and low productivity firms for each occupation group, where firms are split into two evenly size groups. Standard errors report 95% confidence intervals.

Results Figure 2 reports the results of equation (8), pooling all occupation groups, and using both value added and value added less salaries as the outcome variable. On average over three years, firms that experience a worker death see a cumulative decline in value added and value added less salaries equivalent to 3.4 and 1.7 years of workers' wages, respectively. The results differ dramatically when the outcome variable is value added or value added less salaries, reflecting that some firms shrink considerably after a worker death.[24]

Figure 3a reports the estimates of value added less salaries following a death for each occupation group, measured in years of prior wages per year, averaging the coefficients from the the treatment years $\delta_j + \frac{\mathbf{h}_j}{\alpha} = -(\beta_0^j + \beta_1^j + \beta_2^j)/3$ using the event study specification

[24]The gap between the change in value added and the change in value added and less salaries is largest when deceased worker was a manager or professional, suggesting that the incentive for firms to downsize is indeed larger for occupations with higher complementarities or hold-up power.

in equation (8). The coefficients reflect losses in years of prior wages per year. For example, the point estimate on managers is -.93, indicating that the firm's losses of value added less salaries are equal to .93 years of the managers' prior wages in each year, implying a total cumulative loss nearly 3 years worth of prior wages.

Next we estimate the interaction regression specified in equation (9) to separately estimate regular turnover costs δ from hold-up power $\mathbf{h}$. To simplify the interpretation, we will modify the notation. On the right hand side, we pool the post-treatment periods into single indicators. We include an interaction term that multiplies the size of the treatment $\tilde{w}^j_{k(i),\text{-}1}$ by the inverse of the firm's labor share $\hat{\alpha}^{-1}_{k,\text{-}1}$ in the prior periods, where regression coefficient $\mathbf{h}^{est}_j$ directly corresponds to our estimate of the model parameter $\mathbf{h}_j$:

$$\tilde{Y}_{ks} = c^j + \psi^j_k + \sum_{s=-3}^{-1} \beta^j_s \times \mathbb{1}\{s\} \times \tilde{w}^j_{k(i),\text{-}1} - \delta^{est}_j \times \underbrace{\mathbb{1}\{s \geq 0\} \times \tilde{w}^j_{k(i),\text{-}1}}_{\text{baseline}} \tag{9}$$
$$- \mathbf{h}^{est}_j \times \underbrace{\mathbb{1}\{s \geq 0\} \times \tilde{w}^j_{k(i),\text{-}1} \times \hat{\alpha}^{-1}_{k,\text{-}1}}_{\text{interaction}} + \varepsilon_{ks},$$

The pooled term on the right hand side that is not interacted with the inverse of the firm's labor share has a regression coefficient that directly corresponds to δ_j. This can be interpreted as a kind of intercept: how large would profit losses be a "zero" productivity firm. In practice in the model, labor productivity measured in value added per wage bill is bounded below by 1: as firm the firm's production converges to being linear in labor, the wage bill share of value added converges to 1, and the wage is equal to average product.

To estimate this equation, we use a non-linear regression to restrict $\delta^{est}_j \geq 0$. We show in Appendix B.2 that when we run the standard linear regression with OLS, we estimate negative values of δ_j for each occupation. This is because the difference in estimated profit losses between high and low productivity firms is larger than the theory would predict given restrictions on the model parameters, particularly for managers and crafts/assembly.

Figure 3b shows the results of the non-linear estimation. We estimate the largest hold-up power for managers with $\mathbf{h} > 1$. Hold-up power in low-wage administrative/service/manual jobs is small and not significantly different from zero. The middle wage occupations show an intermediate level of hold-up power $\mathbf{h}$ around 0.5.

External Validity One concern is that the individual hold-up power of workers in a given occupation at a small firm may not be externally valid when applied to larger firms. When firms are larger, there may be more workers who can substitute in for any given co-worker, and larger firms may develop promotion structures that train and elevate particular workers when higher level positions become open. We partially address this

by using firms up to 100 workers in our analysis, which is larger than is standard for the literature (Bertheau et al. (2021), Jäger (2016)).[25]

A further concern is that deaths may be a more severe shock to the firm than the typical shock that determines firm wage setting, namely that a worker is poached by another firm. In this case, we would be overestimating the magnitude of costs from worker turnover.

4.2 Task Differentiation in US Job Postings

To construct an estimate of worker hold-up power for more detailed occupations, we construct a "differentiation score" for each occupation group using online job postings from Burning Glass Technologies (BGT). The measure is designed to capture how different a typical job posting's skills requirements are from the skill requirements of other jobs *within the firm*. Recall the two necessary conditions for individual hold-up power: (1) the output of positions are complementary, and (2) skills are position specific. Within-firm task dissimilarity is a good proxy for both of these conditions. First, the output of positions is less likely to be complementary if their tasks are homogenous, as task homogeneity tends to imply output substitutability. Second, both incumbent co-workers and outside hires are less likely to be able to replace tasks if the required combination of tasks of a position is uncommon. Details about the construction of the differentiation score and merging to Danish occupations is given in Appendix B.5.

Death Estimates by Task Differentiation Next we will combine our measures of hold-up power by estimating the extent to which high task differentiation occupations generate larger profit losses after separations and if task differentiation further predicts higher profit losses at productive firms. Equation (10) includes a pooled post-treatment term, as well as interactions on the inverse of firm k's labor share, the differentiation score of the occupation j, and the double interaction including these last two terms.

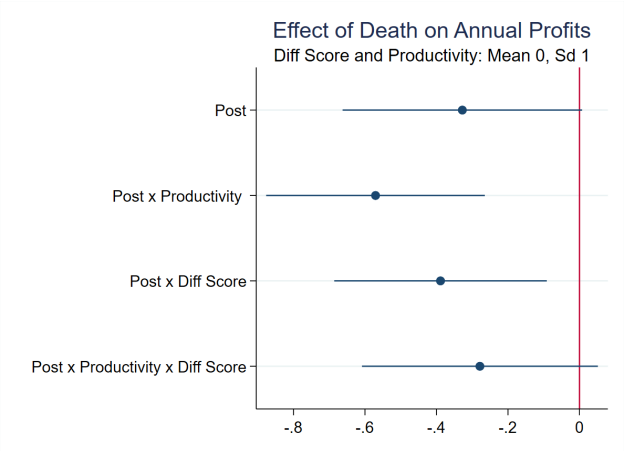
Figure 4 shows the results of a pooled regression containing all occupations. Both firm productivity and the task differentiation score are normalized to have a mean of 0 and standard deviation of 1. "Post" indicates that the year observed is in the treated periods. The "Post $\times$ Productivity" coefficient indicates that in general, profit losses are larger when a death occurs at a higher productivity firm. The "Post $\times$ Diff Score" coefficient indicates that deaths of workers in occupations with high task differentiation scores generate larger profit losses relative to prior wages, and the "Post $\times$ Productivity $\times$ Diff Score" coefficient indicates that the slope of profit losses with respect to firm

[25] A further theoretical extension that could justify similar levels of hold-up even at large firms could be that even in large firms, workers are organized into teams, where individual workers' output is likely complementary with that of other team members.

productivity is increasing in task differentiation. This final coefficient is the strictest test that our task differentiation measure is proxying for hold-up power, and therefore the task differentiation measure should be predictive of passthrough. The point estimate above .25 implies that the slope of profit losses with respect to firm productivity is much higher for high task differentiation occupations: for occupations one standard deviation above and below average task differentiation, an additional standard deviation's worth of firm productivity means that profit losses after a death is half a years worth of salary higher.

$$\begin{aligned}\tilde{Y}_{ks} = \alpha + \gamma_k + \sum_{s=-3}^{-1} \beta_s \mathbb{1}(period_s) \times \tilde{w}_{k(i),\text{-}1} \\ + \beta_{post} \underbrace{\mathbb{1}\{post\} \times \tilde{w}_{k(i),\text{-}1}}_{\text{baseline}} + \beta_{prod} \underbrace{\mathbb{1}\{post\} \times \tilde{w}_{k(i),\text{-}1} \times labsh_k^{-1}}_{\text{productivity}} \\ + \beta_{diff} \underbrace{\mathbb{1}\{post\} \times \tilde{w}_{k(i),\text{-}1} \times \text{Diff}_j}_{\text{Diff Score}} + \beta_{int} \underbrace{\mathbb{1}\{post\} \times \tilde{w}_{k(i),\text{-}1} \times \text{Diff}_j \times labsh_k^{-1}}_{\text{interaction}} + \epsilon_{ks}.\end{aligned} \tag{10}$$

Figure 4: Effect of Death On Profit with Task Differentiation Interaction



This figure shows the effect of a one standard deviation increase in productivity, one standard deviation effect of task differentiation, and their interaction on firm profits following a worker death. Task differentiation increases the cost of a death relative to prior wages significantly, such that an occupation with a differentiation score one standard deviation will generate approximately zero effect on firm profits, while an occupation with task differentiation one standard deviation above average will have yearly profit losses near -0.7 years of prior wages. Standard errors report 95% confidence intervals. While the interaction term on productivity and task differentiation is not significant at the 5% level, this figure indicates that the higher profit losses at productive firms generated by deaths of workers in high differentiation score occupations is economically significant.

5 Productivity-Wage Elasticity Estimation

In this section, we estimate the elasticity of wages to firm average product. We already derived a theoretical formula for this elasticity as a function of model parameters:

$$\varepsilon^{j}_{w,apl} = \frac{\mathbf{h}_j/\delta_j}{\gamma\alpha + (1+\gamma)\mathbf{h}_j/\delta_j}. \tag{11}$$

If this model is the underlying data generating process, we should be able to infer estimates of hold-up power $\mathbf{h}$ from the elasticity of wages to firm average product, as high $\mathbf{h}$ occupations will exhibit higher passthrough elasticities. We estimate passthrough for four broad occupation groups as well as estimate the effect of the task differentiation score on passthrough.

Measuring Firm Productivity for Passthrough Estimation In our model, we are interested in labor productivity: output per hour worked. However, a simple measure of value added per worker would include variation simply based on the occupational composition of the firm: firms with high wage workers should produce more output per worker, or else the firm would not be profitable. We account for the occupational composition of firms by constructing a residualized value added per worker measure. Details are given in Appendix B.4.

5.1 Passthrough Estimates

Sample We include firms that report value during the years 2008-2016 and have an average of at least 15 full-time equivalent employees. Summary statistics on the firms can be found in Section 3.1.

Specification and Results We estimate the passthrough of firm productivity changes to wages for stayers. In our main specification, we use a similar strategy as Card et al. (2018), Maibom and Vejlin (2021), and Juhn et al. (2018), regressing three year changes in wages on three year changes in our firm productivity measure $\hat{Y}_k$, using seven year changes in our valued added worker residual $\hat{Y}_k$ as instruments. This technique is used to extract permanent changes to productivity if productivity evolves according to a combination of permanent and transitory shocks. While we use a similar specification as Card et al. (2018), the interpretation is different. Through the lens of the model, long-run changes in TFP A (often interpreted as demand shocks) result in changes in firm size but not wages. The assumption is that firms are able to fully adjust to a new level of employment within two years of the shock: that is, if demand shocks increase the output per worker, the firm

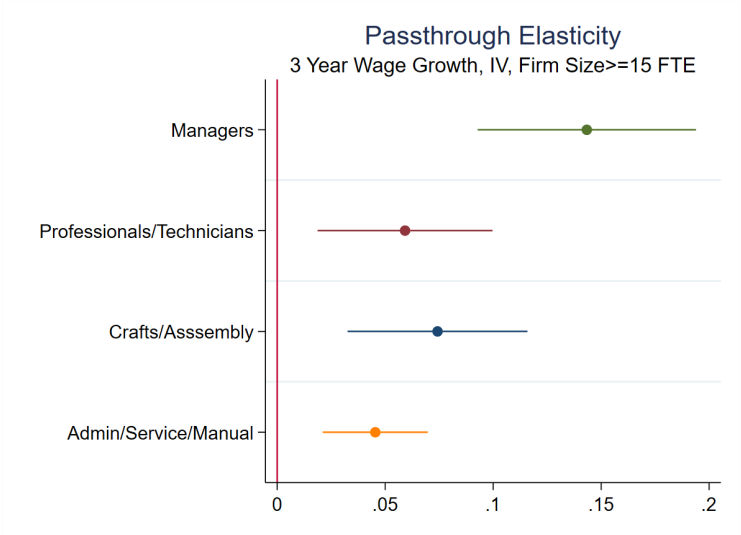
is able to hire a sufficient number of workers within two years to reach the same extensive marginal product that existed prior to the shock. Changes in output per worker that last longer than two years are due to changes in α, which governs the long-run ratio of average to extensive marginal product.

Our main specification includes an interaction term with dummies for each broad occupation, with a corresponding instrument for each interaction (suppressing the time subscript t):

$$\Delta \log(\mathrm{w}_{ijk,\text{-}1,s}) = \sum_j \tau_j \Delta \hat{Y}_{ik,\text{-}1,s} \times \mathbb{1}\{j\} + X_i\gamma + Z_k\tau + F_j + e_{ijk}, \tag{12}$$

where X_i is a vector of person controls, Z_k is a vector of firm controls, F_j is the set of 4-digit level occupation fixed effects, and $s = 2$. In our main specification, we divide occupations into the same four categories as in the deaths analysis: managers, professionals/technicians, crafts/assembly, and administrative/service/manual. Standard errors are clustered at the firm level. Figure 5 shows the main results.

Figure 5: Passthrough Elasticity of Productivity to Wages, Four Occupation Groups



This figure reports the results of the IV passthrough estimation using equation (12) separately for each occupation. Managers have the highest passthrough at .143, followed by professionals/technicians and crafts/assembly at .059 and .074, respectively, and administrative/service/manual occupations have the lowest passthrough at .045. Standard error bars report 95% confidence intervals.

As shown in Figure 5, managers have the highest estimated passthrough, with a point estimate of .143. Professionals/technicians and crafts/assembly have the next highest passthrough, estimated at .059 and .074, respectively. Administrative/service/manual

occupations have the lowest passthrough, with a point estimate of .045. We report various other specifications in Table 6 in Appendix B.6, including results after restricting the sample to firms with at least 50 full time equivalent workers to address concerns of reverse causality where changes in worker productivity affects both the worker's own wages and firm productivity. The passthrough estimates are similar but slightly attenuated.

5.2 Levels and Passthrough: Task Differentiation vs. Education

In this section, we further test if task differentiation as a proxy of hold-up power is predictive of productivity passthrough to wages, as well as if hold-up power increases the level of wages in an occupation. To estimate the effect of task differentiaion on the level of wages, we regress log wages on the level of average firm productivity $\bar{Y}_k$, a measure of occupational average education, the task differentiation score, and covariates including individual education:

$$\log(w_{ijkt}) = \eta_1 \bar{Y}_{ik} + \eta_2 \overline{\text{Educ}}_j + \eta_3 \text{Diff}_j + X_i\gamma + Z_k\tau + u_{it} \tag{13}$$

To estimate the how the effect of firm productivity on wages depends on task differentiation, we estimate both a levels regression and a passthrough regression. In the passthrough estimation, the regressors are three year changes in the firm residualized output per worker $\Delta\hat{Y}_{k,\text{-}1,2}$, as well as with its interactions with the differentiation score and average occupational education attainment. The instruments are the seven year changes and corresponding interactions, and we include occupation fixed effects. We estimate:

$$\Delta \log(\text{w}_{ijk,\text{-}1,2}) = \chi_1 \Delta\hat{Y}_{ik,\text{-}1,2} + \chi_2 \Delta\hat{Y}_{ik,\text{-}1,s} \times \overline{\text{Educ}}_j + \chi_3 \Delta\hat{Y}_{ik,\text{-}1,2} \times \text{Diff}_j + X_i\gamma + Z_k\tau + F_j + e_{it} \tag{14}$$

The levels regression is identical, except that levels replace all first difference terms and is estimated with OLS.

Table 3 reports the results. The differentiation score is normalized to have a mean of 0 and standard deviation of 1, while years of education is only normalized to have a mean of 0. Columns 1, 2, and 3 of Table 3a report the cross-sectional levels regression, which together show that after controlling for individual demographic controls including education, both more highly educated and high task differentiation occupations pay a higher level of wages. Even after controlling for both individual and average occupational education levels, an occupation that is one standard deviation above average in task differentiation will pay 3.6 percent higher wages than an occupation with an average level of task differentiation.

Table 3: Task Differentiation vs. Mean Education: Wage Levels, Interaction, Passthrough

(a) Levels

	$\log(w)_t$ Levels - OLS		
	(1)	(2)	(3)
$\hat{Y}_t$	.077	.085	.075
	(.013)	(.014)	(.013)
$\overline{\text{Educ Years}}_j$	.076		.061
	(.002)		(.002)
Diff. Score_j		.088	.036
		(.003)	(.004)
Observations (Thous.)	4,808	4,808	4,808
R-squared	.508	.493	.511
Clustered SE's	Firm	Firm	Firm

(b) Interaction: Levels and Passthrough

	$\log(w)_t$ Levels - OLS			$d\log(w)_{t\text{-}1,t+2}$ Passthrough - IV		
	(1)	(2)	(3)	(4)	(5)	(6)
$\hat{Y}_t$	.077	.082	.076	.062	.056	.057
	(.006)	(.006)	(.006)	(.013)	(.013)	(.013)
$\hat{Y}_t \times \overline{\text{Educ Years}}_j$	.013		.001	-.002		-.011
	(.002)		(.003)	(.006)		(.006)
$\hat{Y}_t \times$ Diff. Score_j		.028	.026		.008	.021
		(.003)	(.004)		(.010)	(.010)
Occupation FE	N	N	N	Y	Y	Y
Observations (Thous.)	4,808	4,808	4,808	528	528	528
R-squared	.589	.590	.590	.080	.080	.079
Clustered SE's	Firm	Firm	Firm	Firm	Firm	Firm

Table 3a reports OLS results for equation (13). Sample sizes are smaller than in Table 6 because this table only includes occupations with successfully merged differentiation scores. Table 3b reports estimates for equation (14), first running the regression only in levels in columns 1-3 and then estimating passthrough in columns 4-6. Instruments include seven year changes in firm residual productivity, as well as interactions with the occupational variable(s) included in each specification. Controls in each regression include the firm age, workers' lagged tenure, individual years of education, years of education squared, potential experience, interactions of education and potential experience, as well as dummies for sex, year, and 1-digit industry code.

Table 3b reports the effect of firm productivity interacted with occupational differentiation scores and average educational attainment, reporting both levels and passthrough. Columns 1, 2, and 3 together show results of interactions level regression, showing that task differentiation is a much better predictor the slope of workers' wages to firm productivity than is the average educational attainment of an occupation. Column 4 reports

passthrough results only including the average years of education by occupation as an interaction term and shows that high education occupations do not have significantly higher passthrough than low education occupations. Column 6 includes passthrough results with both interaction terms. After conditioning the average educational attainment in an occupation, the coefficient on task differentiation is positive and statistically significant. This is likely driven by the fact that both managers and crafts/assembly workers have higher passthrough than professionals/technicians and admininstrative/service/manual occupations, respectively, while having similar education levels. These results suggest that it is the degree of task differentiation, rather than education, that is the most predictive of higher passthrough. We interpret these results as evidence that hold-up power is not a function of only traditional measure of skill: workers in highly educated occupations need not exhibit high hold-up power if the production process does not have complementarities or if skills are not position specific.

6 Implications

In the previous sections, we presented evidence that individual hold-up power differs across occupation groups. We can show that accounting for this heterogeneity has implications for the gender wage gap and the effect of superstar firms on wage inequality

Wage Level Premia and the Gender Wage Gap While the main tests of our model relate to the slope of wages to firm productivity, both our model and empirical evidence in Section 5.2 suggest that higher turnover costs and hold-up power increase the level of the wage in an occupation. We will now show that men's differential employment in higher hold-up occupations can account for approximately one fifth of the gender wage gap.[26]

One simple way to calculate the premia that men earn in higher hold-up occupations is to plug in the estimated turnover costs from Section 4 into our wage equation and weight the rents by employment shares. Recall that for a given outside options distribution characterized by γ and $\underline{w}$,[27] the wage level increases in the total turnover costs $\delta_j + \frac{\mathbf{h}_j}{\alpha}$. As a simple back-of-the-envelope calculation, we will compute the wage premium earned by managers and crafts/assembly workers relative to professional/technicians and

[26]The ratio of average women's to men's average hourly wage in our sample at private sector firms is approximately .81. The coefficient on the gender dummy in a standard mincer regression in our sample of private sector firms is .15 log points.

[27]Throughout the paper, this we focus on heterogeneity in *firm* outside options across workers in different types of positions, holding worker outside options constant wherever possible. Heterogeneity in worker outside options and its application to the gender wage gap has been explored in studies including Caldwell and Danieli (2021) and Caldwell and Harmon (2019).

admin/manual/service workers, respectively, as these two broad groups largely split workers into high- and low-education occupations. Table 4 reports the results from this exercise. Column 1 reports the point estimates of turnover costs from worker deaths. Column 2 reports the wage level premium using the equation $w_j^* = \left(\gamma\left(\delta_j + \frac{\mathbf{h}_j}{\alpha}\right)\right)^{\frac{1}{\gamma}} \underline{w}_j$.[28] Columns 3 and 4 report employment shares, and columns 5 and 6 add up total rents. In total, this exercise would predict that men earn $1.037/1.008 = 2.8\%$ higher wages due to employment in occupations with higher turnover costs.

Table 4: Hold-up Rents by Occupation and Gender

			Employment Shares		Rents	
	(1)	(2)	(3)	(4)	(5)	(6)
Occupation	$\widehat{\delta_j + \frac{\mathbf{h}_j}{\alpha}}$	Premium	Women	Men	Women	Men
Manager	.93	.14	2.6	6.7	.004	.009
Professional/Tech	.55	-	45.3	35.7	-	-
Crafts/Assembly	.34	.08	5.7	34.3	.004	.027
Admin/Serv/Manual	0	-	46.4	23.3	-	-
Total			100	100	.008	.037

This table calculates the level of rents that women and men earn in each occupation due to turnover costs and hold-up power by occupation. Column 1 reports the point estimates of turnover costs from worker deaths. Column 2 reports the wage level premium relative to admin/service/manual workers using the equation $w_j^* = \left(\gamma\left(\delta_j + \frac{\mathbf{h}_j}{\alpha}\right)\right)^{\frac{1}{\gamma}} \underline{w}_j$. Columns 3 and 4 report employment shares, and columns 5 and 6 add up total rents. The employment shares are calculated using all workers, not just workers at private sector firms.

The estimates from this exercise likely understates the role of different levels of hold-up power between men and women in contributing to the gender wage gap. In Appendix B.9, we show that even within 1-digit occupation groups, men are drastically overrepresented in high task-differentiation occupations relative to women. This suggest and even larger role for differences in average hold-up power in accounting for the gender wage gap.

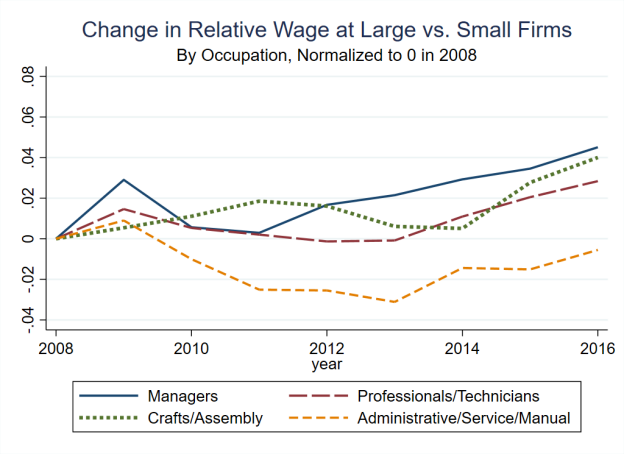
Superstar Firms and Wage Inequality Many advanced economies experienced significant divergence in productivity between larger leader firms and smaller firms. A range of studies have argued that these two trends can be reconciled by the rise of "superstar" firms, where sales or value added is increasingly concentrated in a small set of large, high-productivity, and low-labor share firms, and sales has shifted away from low-productivity,

[28] A technical condition in the firm's problem is that the firm must pay at least $\underline{w}$ in order to recruit new workers. Therefore, the optimal wage equation is better described as $w_j^* = \min\{\left(\gamma\left(\delta_j + \frac{\mathbf{h}_j}{\alpha}\right)\right)^{\frac{1}{\gamma}} \underline{w}_j, \underline{w}_j\}$

high-labor share firms.[29] Dividing workers evenly across larger and smaller firms, the ratio of value added per worker at large firms over the average productivity at small firms rose from 1.11 in 2008 to 1.31 to 2016 in Denmark.

In Figure 6, we show the wages of workers in these two groups within each occupation according to the size of their firm,[30] taking the ratio of average wages of workers at the larger firms divided by the average wage of workers in smaller firms. From 2008-2016, wages for the three higher-wage occupation groups (managers, professionals/technicians, crafts/assembly) increased in above-average sized firms relative to wages in smaller firms.[31] The trends in productivity and wages across firm size are therefore consistent with a story that the rise of superstar firms increased within-occupation wage inequality, but only for occupations that exhibit individual worker hold-up power.

Figure 6: Changes in Relative Wages at Large vs. Small Firms by Occupation



This figure shows the ratio wages of workers at "top half" firms relative to "bottom half" firms, where firms are split into two groups with equal total employment, and the top half of firms are the largest firms. Relative wages for workers at large firms rise for most occupation group by 3-4% from 2008 to 2016, except for workers in administrative/service/manual occupations.

[29]Kehrig and Vincent (2017) document this shift in manufacturing firms in the US. Gouin-Bonenfant et al. (2018) shows a similar change in all firms in Canada. Andrews et al. (2015) reports that these changes are present in services and are global. Autor et al. (2020) emphasize the reallocation from high- to low-labor share firms, driven by increasing competitive advantage of productive firms due to "winner take most" dynamics.

[30]The cutoff between large and small firms each year is approximately 150 full time equivalent employees.

[31]Our data provides information on occupation beginning in 2008.

7 Conclusion

In this paper, we show that the combination of individual production complementarities and position specific skills generates hold-up power for individuals workers, and firms are incentivized to share rents with workers in high hold-up positions. We measure hold-up power across occupations and estimate the passthrough of firm productivity innovations to wages. We find that managers exhibit both the highest hold-up power and passthrough, while low-wage service, manual, and administrative jobs have the smallest measure on both dimensions. We show that accounting for this heterogeneity in individual worker hold-up power has implications for the gender wage gap and the effect of superstar firms on the distribution of wages.

Establishing that workers have different degrees of individual hold-up power opens multiple avenues for future research. One direction to explore is if position specificity and hold-up power are changing over time. Among less educated workers, the composition of jobs has been shifting away from moderate hold-up jobs, like in crafts and assembly occupations, to low hold-up service occupations. A further question is if automation has not only decreased the share of work in blue collar occupations, like in manufacturing, but also decreased the specificity of workers in these occupations as well.

Individual worker hold-up power may also be a useful framework for exploring the borders of the firm, workplace fissuring, and internal pay equity constraints. The observation that industry pay premia in the US were once uniform across occupations (Dickens and Katz (1987)) but now firm premia in the US only benefit college educated workers (Bloom et al. (2018)), combined with the pay losses from outsourcing documented in Goldschmidt and Schmieder (2017), suggest that changes in internal pay equity constraints may be important for explaining changes in wage inequality. In particular, knowing which occupations have individual hold-up power may predict which workers benefit from erosion of internal pay constraints. Individual hold-up power may also be useful for understanding the boundaries of the firm, as employers may use the boundary of the firm to mitigate hold-up power while outsourcing the lower hold-up jobs.

References

Andrews, Dan, Chiara Criscuolo, and Peter N Gal. 2015. "Frontier Firms, Technology Diffusion and Public Policy: Micro Evidence from OECD Countries." OECD Future of Productivity Background Paper, 2015.

Autor, David, David Dorn, Lawrence F Katz, Christina Patterson, and John Van Reenen. 2020. "The Fall of the Labor Share and the Rise of Superstar Firms." *The Quarterly Journal of Economics* 135 (2): 645–709.

Baldwin, Carliss Y. 2018. "Design Rules, Volume 2: How Technology Shapes Organizations." *Harvard Business School Research Paper Series* 19–042.

Bennedsen, Morten, Francisco Pérez-González, and Daniel Wolfenzon. 2020. "Do CEOs Matter? Evidence from Hospitalization Events." *The Journal of Finance* 75 (4): 1877–1911.

Bertheau, Antoine, Pierre Cahuc, Simon Jager, and Rune Vejlin. 2021. "Turnover Costs: Evidence from Unexpected Worker Separations." Working Paper.

Blanchard, Olivier J, and Lawrence H Summers. 1986. "Hysteresis and the European Unemployment Problem." *NBER macroeconomics annual* 1 15–78.

Bloesch, Justin, and Birthe Larsen. 2022. "Which Firms Profit from Wage Sectting Power? New vs. Classical Monopsony." Working Paper.

Bloom, Nicholas, Fatih Guvenen, Benjamin S Smith, Jae Song, and Till von Wachter. 2018. "The Disappearing Large-Firm Wage Premium." In *AEA Papers and Proceedings*, Volume 108. 317–22.

Bobbio, Emmanuele, and Henning Bunzel. 2018. *The Danish Matched Employer-Employee Data*. Department of Economics and Business Economics, Aarhus BSS.

Burning Glass Technologies. 2018a. "Main (Job Postings), 2010-2018 [Database]." (Accessed December 11th, 2018).

Burning Glass Technologies. 2018b. "Skills (Skill Postings), 2010-2018 [Database]." (Accessed December 11th, 2018).

Cahuc, Pierre, Fabien Postel-Vinay, and Jean-Marc Robin. 2006. "Wage Bargaining with On-the-Job Search: Theory and Evidence." *Econometrica* 74 (2): 323–364.

Caldwell, Sydnee, and Oren Danieli. 2021. "Outside Options in the Labor Market." Working Paper.

Caldwell, Sydnee, and Nikolaj Harmon. 2019. "Outside Options, Bargaining, and Wages: Evidence from Coworker Networks." Job Market Paper, The University of California, Berkeley.

Card, David, Ana Rute Cardoso, Joerg Heining, and Patrick Kline. 2018. "Firms and Labor Market Inequality: Evidence and Some Theory." *Journal of Labor Economics* 36 (S1): S13–S70.

Card, David, Ana Rute Cardoso, and Patrick Kline. 2016. "Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women." *The Quarterly Journal of Economics* 131 (2): 633–686.

Carnevale, Anthony P, Tamara Jayasundera, and Dmitri Repnikov. 2014. "Understanding Online Job Ads Data." *Georgetown University, Center on Education and the Workforce, Technical Report (April)*.

Chan, Mons, Sergio Salgado, and Ming Xu. 2020. "Heterogeneous Passthrough from TFP to Wages." Working Paper.

Cobb, J Adam, and Ken-Hou Lin. 2017. "Growing Apart: The Changing Firm-Size Wage Premium and its Inequality Consequences." *Organization Science* 28 (3): 429–446.

Dahl, Christian M, Daniel Le Maire, and Jakob R Munch. 2013. "Wage Dispersion and Decentralization of Wage Bargaining." *Journal of Labor Economics* 31 (3): 501–533.

Danish Statistics. 2020a. "Døde i Danmark (DOD, Deaths), 2001-2017 [Database]." Danmarks Statistiks Forskningsservice (Accessed January 15th, 2020).

Danish Statistics. 2020b. "Detaljeret lønmodtagerdata fra e-Indkomst (BFL, Detailed e-Indkomst Wage Data), 2008-2017 [Database]." Danmarks Statistiks Forskningsservice (Accessed January 15th, 2020).

Danish Statistics. 2020c. "Generel firmastatistik (FIRM, Firm Statistics), 2001-2017 [Database]." Danmarks Statistiks Forskningsservice (Accessed January 15th, 2020).

Danish Statistics. 2020d. "IDA ansættelser (IDAN, Employment), 2001-2017 [Database]." Danmarks Statistiks Forskningsservice (Accessed January 15th, 2020).

Danish Statistics. 2020e. "IDA persondata (IDAP, Person data), 2001-2017 [Database]." Danmarks Statistiks Forskningsservice (Accessed January 15th, 2020).

Danish Statistics. 2020f. "Uddannelse (UDDA, Education), 2001-2017 [Database]." Danmarks Statistiks Forskningsservice (Accessed January 15th, 2020).

Dickens, William, and Lawrence F Katz. 1987. "Inter-industry Wage Differences and Theories of Wage Determination."

Dube, Arindrajit, Eric Freeman, and Michael Reich. 2010. "Employee Replacement Costs." UC Berkeley Working Paper Series.

Edmond, Chris, and Simon Mongey. 2019. "Unbundling Labor." Working Paper.

Eisfeldt, Andrea L, and Dimitris Papanikolaou. 2013. "Organization Capital and the Cross-Section of Expected Returns." *The Journal of Finance* 68 (4): 1365–1406.

Friedrich, Benjamin, Lisa Laun, Costas Meghir, and Luigi Pistaferri. 2019. "Earnings Dynamics and Firm-Level Shocks." National Bureau of Economic Research Working Paper no. 25786.

Gabaix, Xavier, and Augustin Landier. 2008. "Why has CEO Pay Increased so Much?." *The Quarterly Journal of Economics* 123 (1): 49–100.

Garin, Andrew, and Filipe Silverio. 2018. "How Does Firm Performance Affect Wages? Evidence from Idiosyncratic Export Shocks." University of Illinois at Urbana-Champaign Working Paper.

Gibbons, Robert, and Lawrence Katz. 1992. "Does Unmeasured Ability Explain Inter-Industry Wage Differentials?." *The Review of Economic Studies* 59 (3): 515–535.

Gibbons, Robert, Lawrence F Katz, Thomas Lemieux, and Daniel Parent. 2005. "Comparative Advantage, Learning, and Sectoral Wage Determination." *Journal of Labor Economics* 23 (4): 681–724.

Goldin, Claudia, and Lawrence F Katz. 2016. "A Most Egalitarian Profession: Pharmacy and the Evolution of a Family-Friendly Occupation." *Journal of Labor Economics* 34 (3): 705–746.

Goldschmidt, Deborah, and Johannes F Schmieder. 2017. "The Rise of Domestic Outsourcing and the Evolution of the German Wage Structure." *The Quarterly Journal of Economics* 132 (3): 1165–1217.

Gouin-Bonenfant, Emilien et al. 2018. "Productivity Dispersion, Between-Firm Competition, and the Labor Share." In *Society for Economic Dynamics. Meeting Papers*, Volume 1171.

Gregory, Victoria et al. 2020. "Firms as Learning Environments: Implications for Earnings Dynamics and Job Search." Job Market Paper, New York University.

Guha, Sudipto, Rajeev Rastogi, and Kyuseok Shim. 2000. "ROCK: A Robust Clustering Algorithm for Categorical Attributes." *Information Systems* 25 (5): 345–366.

Haanwinckel, Daniel. 2018. "Supply, Demand, Institutions, and Firms: A Theory of Labor Market Sorting and the Wage Distribution." Job Market Paper, University of California, Berkeley.

Hall, Robert E, and Alan B Krueger. 2012. "Evidence on the Incidence of Wage Posting, Wage Bargaining, and On-the-Job Search." *American Economic Journal: Macroeconomics* 4 (4): 56–67.

Jäger, Simon. 2016. "How Substitutable are Workers? Evidence from Worker Deaths." Working Paper.

Jovanovic, Boyan. 1979a. "Firm-Specific Capital and Turnover." *Journal of Political Economy* 87 (6): 1246–1260.

Jovanovic, Boyan. 1979b. "Job Matching and the Theory of Turnover." *Journal of Political Economy* 87 (5, Part 1): 972–990.

Juhn, Chinhui, Kristin McCue, Holly Monti, and Brooks Pierce. 2018. "Firm Performance and the Volatility of Worker Earnings." *Journal of Labor Economics* 36 (S1): S99–S131.

Katz, Lawrence F. 1986. "Efficiency Wage Theories: A Partial Evaluation." *NBER Macroeconomics Annual* 1 235–276.

Kehrig, Matthias, and Nicolas Vincent. 2017. "Growing Productivity Without Growing Wages: The Micro-Level Anatomy of the Aggregate Labor Share Decline." *Economic Research Initiatives at Duke (ERID) Working Paper* (244): .

Kline, Patrick, Neviana Petcova, Heidi Williams, and Owen Zidar. 2019. "Who Profits from Patents? Rent-Sharing at Innovative Firms." *The Quarterly Journal of Economics* 1343 1404.

Kremer, Michael. 1993. "The O-ring Theory of Economic Development." *The Quarterly Journal of Economics* 108 (3): 551–575.

Krueger, Alan B, and Lawrence H Summers. 1988. "Efficiency Wages and the Inter-Industry Wage Structure." *Econometrica: Journal of the Econometric Society* 259–293.

Kuhn, Peter. 2004. "Is Monopsony the Right Way to Model Labor Markets? A Review of Alan Manning's Monopsony in Motion." *International Journal of the Economics of Business* 11 (3): 369–378.

Lachowska, Marta, Alexandre Mas, Raffaele Saggio, and Stephen A Woodbury. 2021. "Do Workers Bargain over Wages? A Test Using Dual Jobholders." National Bureau of Economic Research Working Paper no. 28409.

Lamadon, Thibaut, Magne Mogstad, and Bradley Setzler. 2019. "Imperfect Competition, Compensating Differentials and Rent Sharing in the US Labor Market." *American Economic Review* 112 (1): 169–212.

Lazear, Edward P. 2009. "Firm-Specific Human Capital: A Skill-Weights Approach." *Journal of Political Economy* 117 (5): 914–940.

Leping, Kristjan-Olari. 2009. "Measuring the Specificity of Human Capital: a Skill-based Approach." *Baltic Journal of Economics* 9 (1): 39–54.

Lindbeck, Assar, Dennis J Snower et al. 1989. "The Insider-Outsider Theory of Employment and Unemployment." *MIT Press Books* 1.

Lise, Jeremy, and Fabien Postel-Vinay. 2020. "Multidimensional Skills, Sorting, and Human Capital Accumulation." *American Economic Review* 110 (8): 2328–76.

Maibom, Jonas, and Rune Vejlin. 2021. "Passthrough of Firm Performance to Income and Employment Stability." IZA Discussion Paper no. 14131.

Manning, Alan. 1994. "How Robust is the Microeconomic Theory of the Trade Union?." *Journal of Labor Economics* 12 (3): 430–459.

Manning, Alan. 2011. "Imperfect Competition in the Labor Market." In *Handbook of Labor Economics*, Volume 4. 973–1041, Elsevier.

Montgomery, James D. 1991. "Equilibrium Wage Dispersion and Interindustry Wage Differentials." *The Quarterly Journal of Economics* 106 (1): 163–179.

Muehlemann, Samuel, and Mirjam Strupler Leiser. 2018. "Hiring Costs and Labor Market Tightness." *Labour Economics* 52 122–131.

Mueller, Holger M, Paige P Ouimet, and Elena Simintzi. 2017. "Wage Inequality and Firm Growth." *American Economic Review* 107 (5): 379–83.

Nagypál, Éva. 2001. "Match-Specific Learning and Unemployment." Working Paper.

Oi, Walter Y. 1962. "Labor as a Quasi-Fixed Factor." *Journal of Political Economy* 70 (6): 538–555.

Song, Jae, David J Price, Fatih Guvenen, Nicholas Bloom, and Till Von Wachter. 2019. "Firming up Inequality." *The Quarterly Journal of Economics* 134 (1): 1–50.

Stiglitz, Joseph E. 1974. "Alternative Theories of Wage Determination and Unemployment in LDC's: The Labor Turnover Model." *The Quarterly Journal of Economics* 88 (2): 194–227.

Stole, Lars A, and Jeffrey Zwiebel. 1996. "Intra-firm Bargaining under Non-binding Contracts." *The Review of Economic Studies* 63 (3): 375–410.

Tervio, Marko. 2008. "The Difference that CEOs make: An Assignment Model Approach." *American Economic Review* 98 (3): 642–68.

Yamaguchi, Shintaro. 2012. "Tasks and Heterogeneous Human Capital." *Journal of Labor Economics* 30 (1): 1–53.

A For Online Publication: Theory Appendix

A.1 Deriving Large Firms with Complementarities from Generalized O-ring

Generalized O-Ring We will briefly introduce a production function that generalizes the O-ring production function of Kremer (1993), and we show that its large firm approximation follows the form of $Y = F(N)g(m, X)$ as above.

The key intuition of the O-ring form of production is that labor inputs are divided into discrete positions, and that output losses are multiplicative when the tasks of a position are not performed. Let the firm produce using N tasks, and the maximum total output a firm can produce given N tasks is $F(N) = AN^{\alpha}$, where A is a productivity shifter and α governs the curvature of potential output with respect to the number of positions. Firm total output is equal to:

$$Y = AN^{\alpha} \prod_{i=1}^{N} \Big(1 - \frac{\alpha + m}{N}(1 - q_i)\Big), \tag{15}$$

where the second term is a cumulative product of N individual terms, $q_j \in \{0, 1\}$ is an indicator of whether tasks of position i were successfully performed, and m is the degree of complementarity. There are two main modifications to the original O-ring production. First, the size of the firm N appears in the denominator multiplying the $(1 - q_i)$ term. This means that the fraction of output lost from a non-performing position scales inversely with the size of the firm. Second, in the numerator of the multiplication are the terms $\alpha + m$ where α is the same as before and m is the degree of complementarity for position i. This production function reduces to the traditional O-ring when $m = N - \alpha$.

Consider the O-Ring production function described in Kremer (1993), in which production is composed of N tasks. Let $q_j \in \{0, 1\}$ be an indicator of whether task j was successfully performed. Then total firm output takes the form of:

$$Y = N^{\alpha} \prod_{j \in 1}^{N} q_j,$$

with $\alpha \in (0, 1]$. The defining feature of this production function is extreme complementarities that each position has with the rest of firm output. This generates the extreme result that total output is lost if a single worker fails to do their job:

$$\frac{\Delta Y}{\Delta q_j}\Big|_{q_i = 1 \forall i \neq j} = N^{\alpha}.$$

This is an extreme assumption that if a single task is not performed, the firm loses its entire output. Additionally, larger firms that produce with more tasks N have more to lose

from a single non-performing worker. To soften this extreme form of complementarities, we introduce a generalized O-ring production function.

Generalized O-Ring Consider a generalization of the O-Ring production function such that if a single worker fails to perform a given task j, the firm loses a fraction of total output in that scales inversely with the size of the firm, rather than total product. Consider a function of the form:

$$Y = AN^{\alpha} \prod_{j \in 1}^{N} \Big(1 - \frac{\alpha + m_j}{N}(1 - q_j)\Big).$$

There are two main modifications to the original O-ring production. First, the size of the firm N appears in the denominator multiplying the $(q_j - 1)$ term. This means that the fraction of output lost from a non-performing worker scales inversely with the size of the firm, which delivers convenient properties that we will explore shortly. Second, in the numerator of the multiplication are the terms $\alpha + m_j$ where α is the same as before and m_j is the degree of complementarity in production for position j.

Define the extensive marginal product MPL^E to be:

$$MPL^E\big|_{q_j=1 \forall j} = \frac{\Delta Y}{\Delta N}\bigg|_{q_j=1 \forall j} \approx A\alpha N^{\alpha-1},$$

evaluated at a sufficiently large N. The intensive marginal product is the change in output from a position going unperformed or unfilled, holding N constant: $\frac{\partial Y}{\partial q_j}$. Consider the intensive marginal product of labor MPL^I evaluated at $q_j = 1$ for all except a single j:

$$MPL^I\big|_{q_i=1 \ \forall i \neq j} = \frac{\Delta Y}{\Delta q_j}\bigg|_{q_i=1 \ \forall i \neq j} = A(\alpha + m)N^{\alpha-1} = \underbrace{A\alpha N^{\alpha-1}}_{MPL^E} + \underbrace{AmN^{\alpha-1}}_{\substack{\text{Multiplicative} \\ \text{losses}}}.$$

We can see that there are two components of intensive marginal product: the first is just equal to the extensive marginal product: the level of output loss that is similar to as when a firm gets smaller. But crucially, there is a second term governed by m_j, which is fixed for position type j and does not vary with α. This says that the absence of a worker in occupation j creates an output loss of m_j/N share of output, regardless of the productivity or the size of the firm. This is the channel through which output losses of specific workers interacts with inframarginal product.

To find a larger firm approximation of our firm with a discrete number of positions, we will use the limit definition of e:

$$\lim_{N\to\infty}\left(\frac{N-1}{N}\right)^N \to e^{-1}.$$

With this in hand, consider a firm in which 20% of positions are unfilled. How much output is lost relatively to the benchmark of the firm having 100% of its positions filled and productive? For each incremental position that is unfilled, output is multiplied by $\frac{N-1}{N}$, but only $.2N$ times. Therefore, the total multiplicative factor multiplying the potential output of the firm is $\left(\frac{N-1}{N}\right)^{.2N} = \left(\frac{N-1}{N}\right)^{(1-.8)N} = \left(\left(\frac{N-1}{N}\right)^N\right)^{1-.8} \xrightarrow[N\to\infty]{} e^{-1(1-.8)} = e^{(.8-1)}$. Therefore, as we take the large firm limit and N represents a mass, we have a new important variable X, the share of positions filled. We show in the appendix that for a general m, the production function becomes

$$Y = AN^{\alpha}e^{-(m+\alpha)(1-X)}.$$

A.2 Deriving the Elasticity of Wages to Average Product $\varepsilon_{w,apl}$

We begin with the continuous production function:

$$Y = N^{\alpha}e^{-(\alpha+m)(1-X)}$$

$$\frac{Y}{N} = APL = N^{\alpha-1}e^{-(\alpha+m)(1-X)}$$

$$\frac{\partial Y}{\partial N} = MPL^E = \alpha N^{\alpha-1}e^{-(\alpha+m)(1-X)}$$

$$\frac{\partial Y}{\partial X}\times\frac{1}{N} = MPL^I = (\alpha+m)N^{\alpha-1}e^{-(\alpha+m)(1-X)}.$$

Thus we still have:

$$MPL^I = \big(1+\frac{m}{\alpha}\big)MPL^E$$

$$APL = \frac{MPL^E}{\alpha}.$$

From the firms problem, we have

$$\max_{N,w} N^{\alpha}e^{-(\alpha+dm)(1-X)} - wN \text{ s.t.}$$

$$X = F(w).$$

Using $F(w) = 1 - w^{-\gamma}$, the problem is written in one line as:

$$\max_{N,w} N^{\alpha}e^{-(\alpha+dm)w^{-\gamma}} - wN.$$

Taking first order conditions:

$$FOC_N : \underbrace{\alpha N^{\alpha-1} e^{(\alpha+dm)(w)^{-\gamma}}}_{MPL^E} - w = 0$$
$$FOC_w : \gamma w^{-\gamma-1}(\alpha+dm)N^{\alpha} e^{-(\alpha+dm)(w)^{-\gamma}} - N = 0.$$

Notice that we have

$$MPL^E = w^*.$$

Rerranging the two first order conditions, we get the optimal wage expression:

$$w^* = \left(\gamma\Big(1+\frac{dm}{\alpha}\Big)\right)^{\frac{1}{\gamma}}.$$

In logs:

$$\log(w^*) = \frac{1}{\gamma}\bigg(\log(\gamma) + \log\Big(1+\frac{dm}{\alpha}\Big)\bigg).$$

Taking

$$\frac{\partial \log(w^*)}{\partial \alpha^{-1}} = \frac{1}{\gamma}\frac{dm}{1+\frac{dm}{\alpha}}$$
$$\frac{1}{\alpha}\frac{\partial \log(w^*)}{\partial \alpha^{-1}} = \frac{1}{\gamma}\frac{dm}{1+\frac{dm}{\alpha}}\frac{1}{\alpha}.$$

Simplifying, we get:

$$\varepsilon_{w,\alpha^{-1}} = \frac{1}{\gamma}\frac{dm}{\alpha+dm}.$$

Returning to the results that:

$$APL = \frac{MPL^E}{\alpha} = \frac{w^*(\alpha)^{-1}}{\alpha} = w^*(\alpha^{-1})\alpha^{-1}$$

In logs, then

$$\log(APL) = \log(w^*(\alpha^{-1})) + \log(\alpha^{-1})$$

$$\frac{\partial \log(APL)}{\partial \alpha^{-1}} = \frac{\frac{\partial w^*}{\partial \alpha^{-1}}}{w^*} + \frac{1}{\alpha^{-1}}$$
$$\alpha^{-1}\frac{\partial \log(APL)}{\partial \alpha^{-1}} = \Big(\frac{\frac{\partial w^*}{\partial \alpha^{-1}}}{w^*} + \frac{1}{\alpha^{-1}}\Big)\alpha^{-1}.$$

Thus yielding:

$$\varepsilon_{apl,\alpha^{-1}} = \varepsilon_{w^*,\alpha^{-1}} + 1.$$

What's going on economically, at this point? Average product APL and α^{-1} are clearly related. If extensive marginal products were equalized across firms, then APL is

exactly α^{-1}. But, of firms with different α^{-1} , wages are not exactly equalized, as higher α^{-1} will lead to higher wages. Hence the elasticity of APL to α^{-1} will be $\varepsilon_{w^*,\alpha^{-1}} + 1$.

For the last step, we have $\varepsilon_{w^*,\alpha^{-1}}$ and $\varepsilon_{apl,\alpha^{-1}}$, and we want $\varepsilon_{w,apl}$. Thus, we will have

$$\begin{aligned}
\varepsilon_{w,apl} =& \varepsilon_{w^*,\alpha^{-1}} \varepsilon_{\alpha^{-1},apl} \\
\varepsilon_{w,apl} =& \varepsilon_{w^*,\alpha^{-1}} \frac{1}{\varepsilon_{apl,\alpha^{-1}}} \\
\varepsilon_{w,apl} =& \frac{\varepsilon_{w^*,\alpha^{-1}}}{\varepsilon_{w^*,\alpha^{-1}} + 1}.
\end{aligned}$$

Plugging in

$$\varepsilon_{w,\alpha^{-1}} = \frac{1}{\gamma} \frac{dm}{\alpha + dm},$$

our expression of interest becomes

$$\varepsilon_{w,apl} = \frac{\frac{1}{\gamma}\frac{dm}{\alpha+dm}}{\frac{1}{\gamma}\frac{dm}{\alpha+dm} + 1} = \frac{dm}{dm + \gamma(\alpha + m)} = \frac{dm}{\gamma\alpha + (\gamma + 1)dm}.$$

A.3 Firm's Dynamic Problem with Slow Learning On the Job

The cost of worker turnover costs may also depend on the length of time it takes for a replacement worker to be fully productive. In the main body of the text, the assumption is that a replacement worker is less productive for only one period. This section generalizes the model so that each period, untrained workers have some probability p that they become productive. To do so, we consider a firm in a dynamic setting.

Consider a firm that is maximizing the flow of future profits:

$$\max_{\{N_t\},\{w_t\}} \sum_{t=0}^{\infty} \beta^t \Big(N_t^{\alpha} e^{-(\alpha+md)(1-X_t)} - w_t N_t \Big), \text{ s.t.}$$

$$X_t = r(w)\Big(X_{t-1} + p(1 - X_{t-1}) \Big),$$

where p is the probability that a worker who is in a specific position but is not yet skilled uprades to a fully productive worker. As $\beta \to 1$, the choice of $\{N_t\}, \{w_t\}$ is the level of N and w that maximizes the per-period profit.

If firms are maintaining a constant choice of N and w, then the state variable X, the share of positions filled with trained workers, will be in steady state. To find the steady-state value, we can rearrange the the law of motion for X assuming that $X_t = X_{t-1} = X$:

$$X(r(w)) = \frac{pr(w)}{1 - r(w)(1 - p)}.$$

Examining the case of $\beta = 1$, we can just solve the firm's static maximization problem:

$$\max_{N,w} N^{\alpha} e^{-(\alpha+dm)(1-X(w))} - Nw \text{ s.t.}$$

$$X = \frac{pr(w)}{1 - r(w)(1-p)}.$$

$$\begin{aligned} FOC_N: &\qquad \alpha N^{\alpha-1} e^{-(\alpha+dm)(1-X(w)} - w = 0 \\ FOC_w: \; X'(r(w))\gamma w^{-\gamma-1}(\alpha+dm) &\qquad N^{\alpha} e^{-(\alpha+dm)(1-X(w))} - N = 0. \end{aligned}$$

Since

$$X'(w) = \frac{-pr'(w)}{\big(1 - r(w)(1-p)\big)^2},$$

rearranging the first order conditions yields:

$$\frac{-pr'(w)}{\big(1 - r(w)(1-p)\big)^2}\gamma(\alpha+dm)w = \alpha.$$

It is worth noting that in a monthly model, $r(w)$ will be close to 1. If we take the approximation of this equation by setting $r(w) = 1$, we get

$$\begin{aligned} \frac{-pr'(w)}{\big(1 - (1-p)\big)^2}(\alpha+dm)w \approx& \alpha \\ \frac{-r'(w)}{p}(\alpha+dm)w \approx& \alpha. \end{aligned}$$

Rearranging and substituting back in that $r'(w) = \gamma w^{-\gamma-1}$ we get our optimal wage expression:

$$w^* \approx \left(\frac{\gamma}{p}\Big(1 + \frac{dm}{\alpha}\Big)\right)^{\frac{1}{\gamma}}.$$

Now the probability p that an untrained worker becomes a fully productive position-specific worker enters into the wage setting in a very similar way as the productivity of replacement workers. As p falls, and the expected training duration $1/p$ increases, the cost of turnover increases, incentivizing the firm to pay higher wages.

Notice, however, that our function for the elasticity of wages to average product will be the same: $\varepsilon_{w,apl} = \mathbf{h}/(\gamma\alpha + (\gamma+1)\mathbf{h})$. That is, training time $1/p$ affects only the level of wages but not the long-run passthrough of productivity to wages.

A.4 Relaxing Equal Wages between Trained and Untrained Workers

In Section 2.2 of the main text, we assume that new hires who are untrained receive the same wage as fully trained workers with position specific skills. We make this assumption to derive a closed form passthrough elasticity of average product to wages. In this section, we relax that assumption and show that main results of the passthrough elasticities are unaffected. Consider a firm facing the same problem as in Section 2.2, where the firm chooses the number of positions N and the wage of trained workers w, but can rehire untrained workers for some exogenous wage w_u. The firm's problem is then:

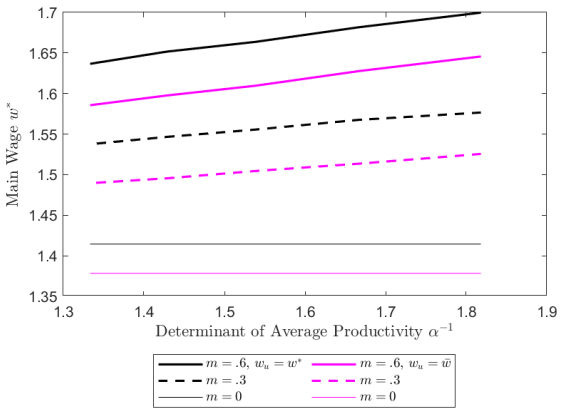
$$\max_{N,w} N^{\alpha} e^{-(\delta\alpha+h)(1-X)} - wNX - w_u N(1-X)$$

s.t.

$$X = F(w),$$

where w is the wage paid to trained workers and w_u is the wage paid to untrained workers. Using the functional form $F(w) = 1 - (w/\bar{w})^{-\gamma}$ and setting $\bar{w} = 1$. The firm pays the trained wage w to NX workers and pays the untrained wage w_u to $N(1-X)$ workers.

Figure A.7: Optimal Wage under $w_u = w^*$ and $w_u = \bar{w}$



This figure plots the optimal wage for trained workers w^* under different values of the degree of complementarity m, the underlying curvature that determines average product α, and wage policies for untrained workers. The black lines indicate the assumption in the main text, that untrained workers receive the same wage as trained workers. The magenta (lighter shade) lines indicate that untrained workers are paid the outside minimum wage $\bar{w} = 1$. For any value of m, firms pay a slightly lower wage level to trained workers if the untrained workers get the low wage. However, the slope of wages with respect to α^{-1} are functionally identical across the assumptions on wages for untrained workers.

Figure A.7 shows the optimal wage for trained workers w^* under different values of m,

α^{-1}, and the wage setting assumptions of untrained workers. The black (darker shade) lines show the optimal wage when untrained workers earn the same wage as trained workers. The magenta (lighter shade) lines show the optimal wage for trained workers w^* when untrained workers are paid $w_u = \bar{w} = 1$. As the figure shows, the wage policies for trained workers are quite similar across the two scenarios, but the level is shifted down when untrained workers are paid the minimum wage $\bar{w}$. Since untrained workers are cheaper when they are paid $\bar{w}$, turnover becomes less expensive for the firm, and the firm can pay trained workers less. However, the important result from this exercise is that the slope of wages with respect to changes in the underlying concavity α^{-1} is nearly identical across different values of the degree of complementarity m.

A.5 Endogenous Hold-up Power

Suppose that firms can hire a flow of consulting services who are able to help the firm despecify workers, thereby making $\mathbf{h}$ a choice variable. The consultants charge the firm a fraction of the wage bill based on the resulting level of h, in total charging $wN(\mathbf{h}^{\frac{-1}{\kappa}} - 1)$.

$$\begin{aligned}\max_{N,w,m} & N^{\alpha} e^{-(\alpha+\mathbf{h})(1-X)} - wN(1 + \mathbf{h}^{\frac{-1}{\kappa}} - 1) \\ = & N^{\alpha} e^{-(\alpha+\mathbf{h})(1-X)} - wN\mathbf{h}^{\frac{-1}{\kappa}}\end{aligned}$$

subject to $X = 1 - w^{-\gamma}$.

$$\begin{aligned}FOC_N :& \alpha N^{\alpha-1} e^{-(\alpha+\mathbf{h})w^{-\gamma}} - w\mathbf{h}^{\frac{-1}{\kappa}} = 0 \\ FOC_w :& \gamma(\alpha + h) w^{-\gamma-1} N^{\alpha} e^{-(\alpha+\mathbf{h})w^{-\gamma}} - N\mathbf{h}^{\frac{-1}{\kappa}} = 0 \\ FOC_m :& - w^{-\gamma} N^{\alpha} e^{-(\alpha+\mathbf{h})w^{-\gamma}} - \kappa w N \mathbf{h}^{\frac{-1}{\kappa}-1} = 0.\end{aligned}$$

Combining the first order conditions on the wage w and hold-up power $\mathbf{h}$ yields:

$$\mathbf{h}^* = \frac{\kappa\gamma\alpha}{1 - \kappa\gamma},$$

with a technical restriction that $0 < \kappa < \gamma^{-1}$. Note that firms with higher average product will choose lower values of $\mathbf{h}$, all else equal: $\varepsilon_{apl,\alpha} < 0$ and $\partial \mathbf{h}^* / \partial \alpha > 0$. This arises naturally from the fact that firms with steeper concavity (low α) and hence higher average product have more to lose from turnover and will pay higher costs to despecify workers.

Given that we typically calibrate γ to be 4, it's worth considering the interpretation of κ when $\kappa < \gamma^{-1} = .25$. When κ is very low, it is very cheap for the firm to choose an $\mathbf{h}$ close to 0. Therefore, $\mathbf{h}^*$ falls as κ falls. The optimal wage equation is:

$$w^* = \left(\frac{\gamma}{1-\kappa\gamma}\right)^{\frac{1}{\gamma}},$$

which also indicates that wages are lower when κ is small and $\partial w^*/\partial\kappa > 0$. Therefore, in total, firms that face higher costs of despecifying workers κ settle on higher values of hold-up $\mathbf{h}^*$ and pay higher wages. Higher productivity (low α) firms end up choosing lower hold-up $\mathbf{h}^*$ to counteract the higher hold-up power. Interestingly, low α firms do not pay higher wages, though their per-worker costs are higher because choosing a lower $\mathbf{h}$ is costly, though this result is likely due to the particular choice of functional form.

A.6 Firms with Multiple Occupations

In this section, we demonstrate that the simplification that each firm utilizes positions of only one type of occupation is not restrictive, and that our mapping of hold-up power $\mathbf{h}_j = d_j m_j$ to wages is robust to firms with multiple occupations.

Consider a firm that produces using positions of two different occupations a and b, where the number of positions are N_a and N_b, respectively. The outer production structure is Cobb-Douglas, where σ_a and σ_b are the elasticities of output (within the parentheses) with respect to the quantity of positions a and b, and α is the same overall concavity. The firm maximizes revenue minus wages:

$$\max_{N_a,N_b,w_a,w_b} = \left(N_a^{\sigma_a} N_b^{\sigma_b}\right)^{\alpha} e^{-\sum_j \sigma_j(\delta_j\alpha+\mathbf{h}_j)(1-X_j)}, \quad j \in \{a,b\}$$

subject to:

$$X_j = F(w_j).$$

If the outside offer distribution for each occupation j is similar to before with the form $F(w_j') = 1 - \left(\frac{w_j'}{\underline{w}_j}\right)^{-\gamma}$, for $w_j \geq \underline{w}_j$, then it is simple to show that we end up with the same wage expression:

$$w_j^* = \left(\gamma\Big(\delta_j + \frac{\mathbf{h}_j}{\alpha}\Big)\right)^{\frac{1}{\gamma}} \underline{w}_j, \quad j \in \{a,b\}.$$

The elasticity of optimal wages to α^{-1} is:

$$\varepsilon_{w_j,\alpha^{-1}} = \frac{1}{\gamma}\frac{h_j}{\delta\alpha + h_j},$$

yielding an total passthrough elasticity:

$$\varepsilon_{w_a,apl} = \frac{h_a}{\gamma(\delta\alpha + h_a) + \sigma_a h_a + \sigma_b h_b \frac{\alpha+h_a}{\alpha+h_b}},$$

where σ_j is the wage bill share of occupation j. Therefore, this expression does show that the elasticity of wages to measured average product will depend on the wage share of each occupation. This is because the firm's average product responds endogenously to the optimal wages: if workers have a lot of hold-up power and increases in α^{-1} lead to an increase in wages, the firm will pull back on employment, thereby raising average product. Therefore, the shares of wages going to occupations of different levels of hold-up power matters for how much total average product responds to underlying changes in α.

Proof When firms have multiple occupations workers, we need to redefine average productivity. We define average product to be output divided by wage-weighted employment, where the weights on the employment of each occupation are the average wages of each occupation in the economy as a whole, $\bar{w}_a$ and $\bar{w}_b$. Average product is then:

$$\widetilde{APL} = \frac{Y}{N_a\bar{w}_a + N_b\bar{w}_b}.$$

From the firm's first order conditions, we also have that the total wage bill is only fraction α of total output:

$$Y = \frac{N_a w_a + N_b w_b}{\alpha}.$$

Combining the last two expressions and rewriting all of the firm's endogenous variables as functions of α^{-1} ($\bar{w}_a$ and $\bar{w}_b$ are exogenous to the firm), we have:

$$\widetilde{APL}(\alpha^{-1}) = \frac{\big(N_a(\alpha^{-1})w_a(\alpha^{-1}) + N_b(\alpha^{-1})w_b(\alpha^{-1})\big)\alpha^{-1}}{N_a(\alpha^{-1})\bar{w}_a + N_b(\alpha^{-1})\bar{w}_b}.$$

In logs, this is:

$$\log(\widetilde{APL}(\alpha^{-1})) = \log\big(N_a(\alpha^{-1})w_a(\alpha^{-1}) + N_b(\alpha)w_b(\alpha^{-1})\big) + \log(\alpha^{-1}) - \log\big(N_a(\alpha^{-1})\bar{w}_a + N_b(\alpha^{-1})\bar{w}_b\big).$$

Deriving with respect to α^{-1} and suppressing the (α^{-1}) of the endogenous variables, we have:

$$\frac{\partial \log(\widetilde{APL}(\alpha^{-1}))}{\partial \alpha^{-1}} = \frac{N_a\frac{\partial w_a}{\partial \alpha^{-1}} + w_a\frac{\partial N_a}{\partial \alpha^{-1}} + N_b\frac{\partial w_b}{\partial \alpha^{-1}} + w_b\frac{\partial N_b}{\partial \alpha^{-1}}}{N_a w_a + N_b w_b} + \alpha - \frac{\bar{w}_a\frac{\partial N_a}{\partial \alpha^{-1}} + \bar{w}_b\frac{\partial N_b}{\partial \alpha^{-1}}}{N_a w_a + N_b w_b}.$$

Thus, we have an expression for the elasticity average product to α^{-1} for firm paying average wages. Evaluating this term at $w_a = \bar{w}_a$ and $w_b = \bar{w}_b$, terms will cancel, yielding:

$$\frac{\partial \log(\widetilde{APL}(\alpha^{-1}))}{\partial \alpha} = \frac{N_a\frac{\partial w_a}{\partial \alpha^{-1}} + N_b\frac{\partial w_b}{\partial \alpha^{-1}}}{N_a w_a + N_b w_b} + \alpha.$$

$$\varepsilon_{\widetilde{APL},\alpha^{-1}} = \frac{N_a w_a \varepsilon_{w_a,\alpha^{-1}} + N_b w_b \varepsilon_{w_b,\alpha^{-1}}}{N_a w_a + N_b w_b} + 1.$$

Taking the inverse of this expression, we have:

$$\varepsilon_{\alpha^{-1},\widetilde{APL}} = \frac{N_a w_a + N_b w_b}{N_a w_a(\varepsilon_{w_a,\alpha^{-1}} + 1) + N_b w_b(\varepsilon_{w_b,\alpha^{-1}} + 1)}.$$

We can simplify by dividing by the wage bill and using the fact that the wage bill share parameters σ_h and σ_l are constant: $\sigma_j = \frac{N_j w_j}{N_a w_a + N_b w_b}$:

$$\varepsilon_{\alpha^{-1},\widetilde{APL}} = \frac{1}{\sigma_a(\varepsilon_{w_a,\alpha^{-1}} + 1) + \sigma_b(\varepsilon_{w_b,\alpha^{-1}} + 1)}.$$

Using again that $\varepsilon_{w_j,\alpha^{-1}} = \frac{1}{\gamma}\frac{h_j}{\delta\alpha + h_j}$:

$$\varepsilon_{w_a,\widetilde{APL}} = \varepsilon_{w_a,\alpha^{-1}}\varepsilon_{\alpha^{-1},\widetilde{APL}} = \frac{h_a}{\gamma(\delta\alpha + h_a) + \sigma_a h_a + \sigma_b h_b \frac{\delta_a\alpha + h_a}{\delta_b\alpha + h_b}}.$$

This does imply that the inference of h_j from a given elasticity of wages to measured average product will depend on the occupational composition of the firm. However, in the following table, we will show that this effect is small. Consider a firm with two occupations, a and b, where the degrees of complementarity for the two occupations are $h_a = .5$ and $h_b = .1$, respectively. Then varying the wage bill share θ_a of occupation a (which is itself a function of h_a, h_b, and σ_a), we can see how the passthrough elasticity is affected.

Table 5: Passthrough Elasticity as Occupation Composition Changes, $h_a = .5$, $h_b = .1$

θ_a	0	.5	1
$\varepsilon_{w_a,apl}$	.114	.106	.099
$\varepsilon_{w_b,apl}$	.031	.033	.036

A.7 Comparison with Alternative Models

Our model of individual worker hold-up power differs from other common models of imperfect labor market competition and firm wage premia.

Upward Sloping Labor Supply/Convex Hiring Costs - A common model applied to questions of firm premia is the upward sloping labor supply as in Card et al. (2018) and Lamadon et al. (2019) that use upward sloping labor supplies to generate firm wage premia. In these models, firms must pay higher wages in order to be large, dipping deeper into a pool of workers that are less willing or more difficult to be employed. The main distinguishing features of individual hold-up power is that the incentive for a firm to pay high wages need not be a function of firm size. A related model to the upward sloping

labor supply is convex adjustment costs, where in response to shocks, firms increase wages more for types of labor that have more steeply convex adjustment costs, as in Kline et al. (2019). However, as argued by Kuhn (2004), if the convex costs are in percent terms relative to firm size, rather than in absolute terms, then wages should equilibrate after firms have grown to their desired size. Therefore, individual worker hold-up power can match the patterns of productivity wage premia that (i) do not depend on firm size, and (ii) exist both in response to shocks and in steady state, which neither the upwards sloping labor supply nor convex adjustment cost models can match.

Stole and Zwiebel (1996) - Another common model originates from Stole and Zwiebel (1996), in which multi-worker firms bargain bilaterally with each worker. In this setting, firm average productivity affects the bargained wage: the firm considers that if negotiations break down, the firm will have to bargain with a subsequent worker who has greater marginal product. Following this logic, the inframarginal productivity of hypothetical workers at deeper levels of bargaining affect the bargaining *position* of the firm and thus affects what a worker at the first layer of bargaining can extract, given a worker's bargaining power β. However, to the extent that there is heterogeneity in bargaining power among types of workers, and therefore the ability to extract higher wages from more productive firms, is an assumption in a Stole and Zwiebel (1996)-type setting. In our model of complementarities and position specific skills, the firm's outside option differs across types of workers due to heterogeneous ability to disrupt production across positions. With wage posting and imperfect contracts, heterogeneous wage premia at productive firms is a result. In a similar way that wage posting and incomplete contracts generates isomorphic outcomes as bargaining (i.e. Manning 2011), we generate results that look like heterogeneous bargaining power but with firm wage setting and incomplete contracts.

Search and Matching - A third common model is the standard Diamond-Mortensen-Pissarides search model, in which firms pay a vacancy posting cost, and after meeting, firms and workers bargain over the surplus of the match. Combined with a free entry condition, the surplus of a match is pinned down by the cost of a vacancy and the per-period job filling rate for the firm. Therefore in this benchmark model, the most that a worker can hold up is the expected flow of vacancy of costs. With complementarities and position specificity, workers can hold up a portion of firm output in addition to the expected cost of a vacancy, as a worker's sudden departure disrupts the output of other factors in the firm.

Firm Specific Skills - When skills are only firm specific, workers are internally substitutable. As we argued in Section 2.2, if workers are substitutable, firms are never at

risk of losing inframarginal product from individual worker turnover. As such, wage setting decisions take into account only the productivity of labor at the margin, and therefore inframarginal productivity is irrelevant for wage setting, resulting in no wage premium at productive firms. However, as we showed in Section 2.3, non-multiplicative turnover costs do raise the level of the wage. In total, firm specific skills raise the level of an occupation's wage but do not generate a slope of wages with respect to productivity.

Rent Sharing Models - Many studies such as Card et al. (2016) put forward a simple rent sharing model of the form $w_{ikt} = w_0 + \gamma_j S_{kt}$, where S_{kt} is the surplus per worker at firm k and w_0 is an outside competitive wage. The rent sharing parameter γ_j may differ by occupation j (or by gender). Our study argues instead that the variation across occupations will come from variation in the surplus per worker S due to heterogeneous complementarities and position specificity, instead of differences in the rate at which surplus is shared γ. For example, low-wage service workers receive low passthrough not because their bargaining power is low, but because there is little surplus in the match to begin with.

Efficiency Wage Models - Our model of individual worker hold-up power is most closely related to efficiency wage models. In efficiency wage models, firms get more "effective" labor from the same number of workers by paying a higher wage. We achieve a similar result where firms have a higher share of trained workers by paying higher wages $X'(w) > 0$, where the mechanism is turnover and incomplete contracts. A model in which effort were increasing in the wage and interacts multiplicatively with average product would be isomorphic. The contribution of our paper to the efficiency wage literature is to isolate a particular set of conditions (i.e., individual complementarities and position specificity) where efficiency wage motivations (i) differ across occupations and (ii) interact with firm productivity heterogeneously across occupations.

Assignment Models and Sorting - In assignment models such as Tervio (2008) and Gabaix and Landier (2008), two main results emerge: (1) better quality workers are assigned to positions where the stakes are largest, and (2) small differences in worker quality can generate large differences in wages. While we abstract from ex-ante worker heterogeneity in this paper, our generalization of O-ring production provides a framework for extending assignment models to settings of multi-worker firms.

A.8 Proof of Identification

In this section, we prove that if firms do not choose to change the number of positions in response to worker deaths, the correct outcome measure is the change in value added. If firms systematically decrease employment after a separation shock based on the hold-

up power of a worker, then using value added to estimate underlying parameters will be biased, and the correct outcome measure to use is value added less wage and salary costs.

Firms Hold Constant Number of Positions N - Suppose that in response to an exogenous separation of a worker, the firm's share of fill positions X falls by ΔX:

$$\begin{aligned}\Delta Y =& N^{\alpha}e^{-(\delta\alpha+\mathbf{h})(1-X)} - N^{\alpha}e^{-(\delta\alpha+\mathbf{h})(1-(X-\Delta X))}\\ =& N^{\alpha}e^{-(\delta\alpha+\mathbf{h})(1-X)} - N^{\alpha}e^{-(\delta\alpha+\mathbf{h})(1-X)}e^{-(\delta\alpha+\mathbf{h})\Delta X}\\ =& \big(N^{\alpha}e^{-(\delta\alpha+\mathbf{h})(1-X)}\big)\big(1-e^{-(\delta\alpha+\mathbf{h})\Delta X}\big).\end{aligned}$$

Using the first order approximation that $e^{-z} \approx 1 - z$ for small z, we get that

$$\begin{aligned}\Delta Y =& \big(N^{\alpha}e^{-(\delta\alpha+\mathbf{h})(1-X)}\big)\big(1-e^{-(\delta\alpha+\mathbf{h})\Delta X}\big)\\ \approx& \frac{N}{\alpha}\big(\alpha N^{\alpha-1}e^{-(\delta\alpha+\mathbf{h})(1-X)}\big)\big((\delta\alpha+\mathbf{h})\Delta X\\ =& \frac{N}{\alpha}w^{*}\big((\delta\alpha+\mathbf{h})\Delta X,\end{aligned}$$

where the last line uses the first order condition on the number of positions prior to the shock: $w^* = \alpha N^{\alpha-1}e^{-(\delta\alpha+\mathbf{h})(1-X)}$. Rearranging, we end up with:

$$\frac{\big(\frac{\Delta Y}{\Delta X}\big)\times\frac{1}{N}}{w^{*}} = \delta + \frac{\mathbf{h}}{\alpha}.$$

Therefore, if in response to the shock, the firm does not change the number of positions N and hires a replacement worker that offsets δ of regular losses and d of multiplicative losses, then the outcome variable used to identify these parameters is change in value addd.

Firms Reoptimizing the Number of Positions N - If a worker suddenly leaves, and the firm's share of positions filled with productive workers X falls, then productivity of other workers is negatively affected and the firm may find it optimal to decrease its size. Suppose that after a separation shock occurs, the firm is able to change N, but maintains a new lower level of the share of filled positions $X_2 = X_1 - \Delta X$:

We will show that if firms (a) hire a replacement worker who partially offsets the losses of the worker who left, and (b) the firm can freely adjust N down post-shock while holding post-shock X fixed, then the correct dependent variable of a separations regression in the data is value added less wage and salary payments.

$$\Delta(Y - wN) = N_1^{\alpha} e^{-(\delta\alpha+\mathbf{h})(1-X_1)} - N_2^{\alpha} e^{-(\delta\alpha+\mathbf{h})(1-(X_1-\Delta X))} - (w_1 N_1 - w_2 N_2). \quad (16)$$

Before going further, we can show that

$$\begin{aligned} FOC_N : \alpha N^{\alpha-1} e^{-(m+\alpha)(1-d)(1-X+\Delta X)} - w = 0 \\ FOC_w : \gamma w^{-\gamma-1}(m+\alpha)(1-d) N^{\alpha} e^{-(m+\alpha)(1-d)(1-X+\Delta X)} - N = 0. \end{aligned}$$

Like in the standard problem, the inclusion of ΔX in the exponent does not matter and cancels out, so we have that the pre- and post-separation optimal wage for remaining workers is identical

$$w_1^* = w_2^* = \Big(\gamma\big(\delta + \frac{\mathbf{h}}{\alpha}\big)\Big)^{\frac{1}{\gamma}}.$$

Next, we can show how the optimal number of positions N changes now that the firm faces a lower share of postions filled X. From the first order condition on N above, we have:

$$N^* = \left(\frac{\alpha e^{-(\delta\alpha+\mathbf{h})(1-X+\Delta X)}}{w^*}\right)^{\frac{1}{1-\alpha}}.$$

Taking the ratio from before and after the shock and putting this expression in logs, we have:

$$\begin{aligned} \frac{N_1}{N_2} &= \left(\frac{e^{-(\delta\alpha+\mathbf{h})(1-X)}}{e^{-(\delta\alpha+\mathbf{h})(1-X+\Delta X)}}\right)^{\frac{1}{1-\alpha}} \\ \log\left(\frac{N_1}{N_2}\right) &= \frac{1}{1-\alpha}\Big((\delta\alpha+\mathbf{h})\Delta X\Big) \\ \Rightarrow \frac{\Delta \log(N)}{\Delta X} &= \frac{\delta\alpha+\mathbf{h}}{1-\alpha}. \end{aligned}$$

Thus, when a worker is lost, the productivity of the remaining workers declines. The firm still finds it optimal to pay the same wage to minimize wage and turnover costs, and so the firm wants to lower employment to raise the extensive marginal product MPL^E in order to make the marginal position worth the cost. This force becomes stronger if workers' absence is more disruptive due to a high δ or $\mathbf{h}$. However, since employment and therefore additional value added declines will change depending on the parameter values, we no longer can use changes in value added as the outcome variable.

Before going on with the proof, we know that incremental changes in X yield incremental changes in N. We can approximate the previous line with

$$\frac{\Delta \log(N)}{\Delta X} = \frac{\Delta \log(N)\frac{1}{N}}{\Delta X \times \frac{1}{N}} \approx \frac{\Delta N}{\Delta X \times \frac{1}{N}}.$$

Thus, both the numerator and denominator are in terms of individual workers, since X is scaled by $1/N$. Returning now to equation (16), we can plug in the common wage w:

$$\begin{aligned}
\Delta(Y - wN) =& N_1^{\alpha} e^{-(\delta\alpha+\mathbf{h})(1-X_1)} - N_2^{\alpha} e^{-(\delta\alpha+\mathbf{h})(1-(X_1-\Delta X))} - w(\Delta N) \\
=& (N_1^{\alpha} - N_2^{\alpha}) e^{-(\delta\alpha+\mathbf{h})(1-X_1)} + N_2^{\alpha} e^{-(\delta\alpha+\mathbf{h})(1-X_1)} - N_2^{\alpha} e^{-(\delta\alpha+\mathbf{h})(1-X_1)} e^{-(\delta\alpha+\mathbf{h})\Delta X} - w(\Delta N) \\
=& (N_1^{\alpha} - N_2^{\alpha}) e^{-(\delta\alpha+\mathbf{h})(1-X_1)} + N_2^{\alpha} e^{-(\delta\alpha+\mathbf{h})(1-X_1)} (1 - e^{-(\delta\alpha+\mathbf{h})\Delta X}) - w(\Delta N) \\
\approx& \Delta N \alpha N_1^{(\alpha-1)} e^{-(\delta\alpha+\mathbf{h})(1-X_1)} + N_2^{\alpha} e^{-(\delta\alpha+\mathbf{h})(1-X_1)} ((\delta\alpha + \mathbf{h})\Delta X) - w(\Delta N).
\end{aligned}$$

Again using the feature that $w^* = \alpha N^{\alpha-1} e^{-(\delta\alpha+\mathbf{h})(1-X)}$, we can replace the first term:

$$\begin{aligned}
\Delta(Y - wN) =& \Delta N w + w \frac{N_2}{\alpha} (-(\delta\alpha + \mathbf{h})\Delta X) - w(\Delta N) \\
\Delta(Y - wN) =& w N_2 (-(\delta + \frac{\mathbf{h}}{\alpha})\Delta X) \\
\Rightarrow \frac{\frac{\Delta(Y-wN)}{\Delta X \times N}}{w} =& \delta + \frac{\mathbf{h}}{\alpha}.
\end{aligned}$$

Thus, losses of value added less salaries, relative to the prior wage, identify $\delta + \mathbf{h}/\alpha$.

A.9 Relation to Jäger (2016)

Jäger (2016) finds that in response to worker deaths, the earnings of workers in similar occupations increases, suggesting that workers in similar occupations are substitutes. This evidence may seem at odds with our model, where every worker's ouptut is complementary with the output of the entire rest of the firm. However, these need not be in conflict: all that matters is that incumbent workers of the same occupation of the deceased are the *best* substitute for the deceased workers in the short run. Consider a firm with concavity α and occupation with complementarity degree m, giving $MPL^I = \frac{m+\alpha}{\alpha} MPL^E$ and $MPL^E = w^*$. Suppose the firm has two options: do not respond to a worker separation (beyond hiring an untrained replacement worker, who is not yet productive), or offering incumbent workers with related skills to replace part of the separated worker's output. Suppose to do so, however, the firm has to pay an overtime rate $w^{ot} = w^*(1 + \psi^{ot})$, with $\psi^{ot} > 0$. Suppose a coworker's replacement productivity is $d^{coworker} \in (0, 1)$. The firm is strictly better off paying the co-workers to replace lost output if:

$$d^{coworker} MPL^I > w^*(1+\psi^{ot})$$
$$d^{coworker} MPL^I / w^* > (1+\psi^{ot})$$
$$d^{coworker}(1+\frac{m}{\alpha}) > (1+\psi^{ot}).$$

Thus, as long as output is sufficiently complementary (high m), co-worker productivity $d^{coworker}$ is sufficiently high, and the overtime rate ψ^{ot} is not too high, the firm will find it optimal to pay overtime to the co-workers of the deceased worker to partially replace lost output. However, paying workers overtime is not a profitable long term solution, as fully trained replacement workers are more cost effective than paying overtime to imperfectly substitutable coworkers from other positions:

$$\frac{w^*}{MPL^I} < \frac{w^*(1+\phi^{ot})}{d^{coworker} MPL^I},$$

by $\phi^{ot} > 1$ and $d^{coworker} < 1$. Thus higher MPL^I (and therefore higher complementarity m) is consistent with larger output losses and increased wages for co-worker in response to a worker death.

B For Online Publication: Empirical Appendix

B.1 Data Cleaning for Deaths Sample

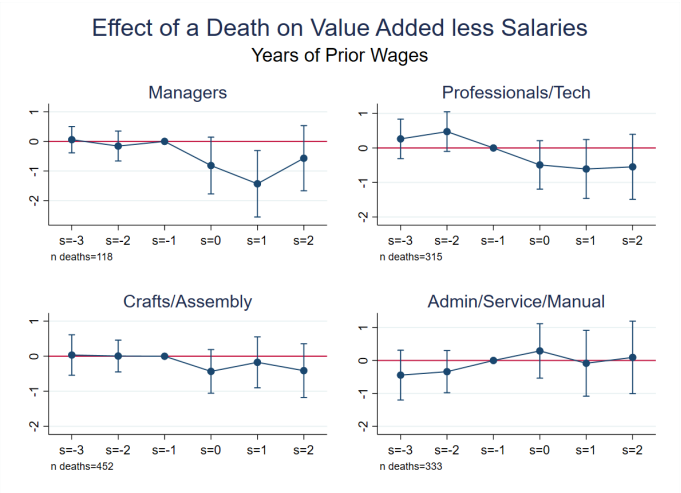
To estimate the effect of worker deaths on firm profits, we clean the data in two major ways: imputing missing values and winsorizing first differences between treated firms and the average of matched placebo firms.

Imputations - Some of the observations in value added and salaries variables are missing. To address this, we impute values from within same firm, within the same pre- or post-treatment subsample. For example, for missing observations in period s=-3, we impute the value using period s=-2. If the missing period is s=-1, we also impute using s=-2, rather than impute with post-treatment observations..

Winsorizing - As described in Section 4.1, we winsorize the difference between the treated firm and the average of control firms $\tilde{Y}$. As mentioned in the main text, we consider observations for the periods before treatment (s=-3 to s=-1) and periods including and after treatment (s=0 to s=3) separately. However, the variance of Y^{dif} differs substantially across firms of different size. Therefore, we group firms into categories with roughly equal number of firms based on the average full time equivalents prior to treatment: <30, 30-50, and 50-100. We then winsorize observations separately for each (pre/post)-by-(firm size group) subsample at the 10th and 90th percentile.

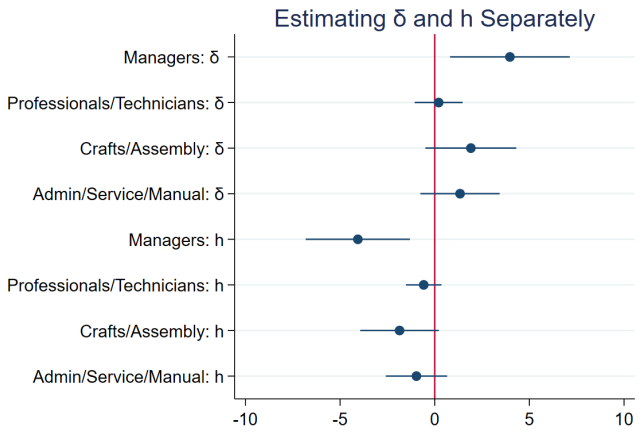
B.2 Additional Figures for Death Estimates

Figure B.1: Effect of a Worker Death on Value Added - Salaries by Occupation



This figure reports the estimates of the dynamic difference-in-difference from equation (8) but for each occupation separately. Pooling the post-treatment coefficients produces the Figure 3a.

Figure B.2: Worker Deaths Unrestricted Interaction Regression



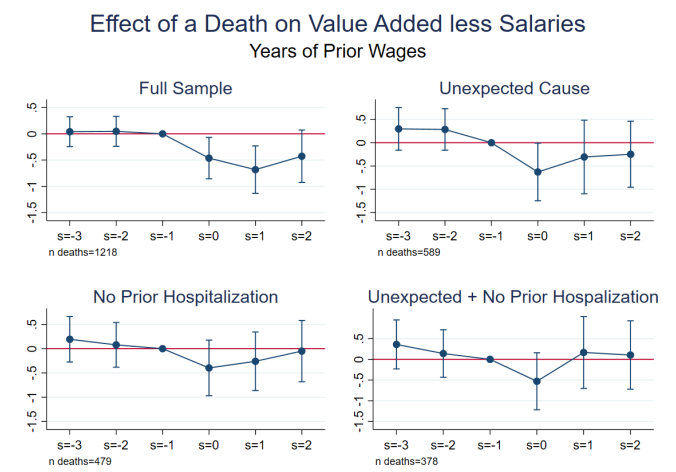
This figure reports the unrestricted estimates of equation (9), with the sign of the coefficients reversed. A negative value of $\mathbf{h}^{est}$ and positive value of δ^{est} in this figure indicate that the slope of profit losses with respect to firm productivity (measured as the inverse of the labor share) is very steep. The theory predicts that δ cannot be negative, which would predict both non-positive values for both regression coefficients in this figure.

Restricting to unexpected deaths may produce different estimates from including all deaths for various reasons. First, if the death is expected, the firm may have more time to adjust, thereby making unexpected deaths have larger effects. However, workers who die unexpected tend to be younger and may be in less critical roles, thereby lowering the effect of unexpected deaths relative to all deaths.

B.3 Unexpected Deaths

In our data, we have access to hospital records as well as medical cause of death. It is common to restrict to workers who (i) had an unexpected cause such as a heart attack or an accident, and (ii) did not have a hospital admission in a prior time window. If we make these restrictions our sample shrinks dramatically and in general decreases the size of our estimates, as shown in Figure B.3.

Figure B.3: Effect of "Unexpected" vs. All Deaths on Value Added - Salaries



B.4 Residual Productivity Measure

In Appendix A.6 we show that residualizing for occupational composition is the correct empirical measure if changes in observed average product are due to changes in underlying curvature α or shifts in occupational composition. As a first step, we compute the predicted wage bill based on the occupational composition of hours worked in the firm:

$$\widehat{\text{wage bill}}_{kt} = \sum_j \text{hours}_{jkt} \bar{w}_{jt},$$

where hours_{jkt} is the total number of hours reported worked by workers in occupation j at firm k in year t, and $\bar{w}_{jt}$ is the economy-wide average of wages of workers in occupation

j in year t. We then run an OLS regression to obtain a predicted level of log value added, given the predicted wage bill in the firm and controls:

$$\widehat{\log(VA)}_{kt} = \xi_0 + \xi_1 \log(\widehat{\text{wage bill}})_{kt} + \xi_2 Z_{kt},$$

where Z_{kt} are controls including industry and year fixed effects, as well as the log number of hours worked by workers without an assigned occupation. then:

$$\hat{Y}_{kt} = \log(VA)_{kt} - \widehat{\log(VA)}_{kt},$$

where $\log(VA)_{kt}$ is the observed log value added reported by firm k in year t.

B.5 Construction of the Differentiation Score

To construct the differentiation score, we apply a clustering algorithm that allocates job postings for a given establishment to clusters based on the set of skills listed in the job posting. The algorithm is called ROCK (Guha et al. (2000)), or "a RObust Clustering using linKs", and is used for clustering categorical data.[32] In the job posting data, BGT cleans and categorizes the skills found in job postings into skill groups and skill group families.[33] The intermediate level, skill groups, is the level of categorization that we use in the ROCK clustering algorithm. Examples of how skills are categorized are provided in Appendix Figure B.4; we use the middle column of skills used, of which there are over 700 in total. An example allocation of job postings to clusters can be seen in Appendix Figure B.5.

The ROCK clustering algorithm occurs in two steps: (1) forming "links" between inputs, and (2) creating clusters based on links. Two observations are linked if the Jaccard coefficient between two observations is above a chosen threshold. We choose the threshold to be equal to 0.8 to form links. Then, using this measure of links across points, the algorithm categorizes observations into clusters based on the number of shared links that job postings have in common.

To compute the differentiation score by detailed (6-digit soc) occupation, we first compute two statistics: (i) what share of job postings are successfully put into a cluster, and (ii) conditional on being in a cluster larger than one, the log of the average size of that cluster. Taking the first principal component of these two metrics gives us our differentiation score. Both components of this score are intended to measure how often and to what degree a given job has a very different set of tasks from other jobs in the firm.

[32] A common application is to cluster transaction data in which sets of goods are often purchased together.

[33] BGT refers to these groups as "skill clusters" and "skill cluster family", but we will use the term "skill group" to avoid using the word cluster with two different meanings.

The two measure are moderately correlated (ρ=0.42), and the first principle component removes noise in each of the two measures.

Summary statistics for the differentiation score by occupation, aggregated to the 1-digit DISCO occupation level, can be found in column 3 of Table 2. We compute these by merging US soc occupations to the DISCO codes in the Danish administrative data,[34] and then taking a weighted average of these scores within broad DISCO occupation groups. Managerial and professional jobs are the most differentiated, followed by technicians. Crafts and assembly occupations have moderately high levels of differentiation relative to their low levels of average education. Service and manual jobs have the lowest differentiation scores. Some examples of differentiation scores at the detailed occupation level include: "Computer and Information Systems Managers" 1.58, "Pharmacists" -0.49, "Construction Equipment Operators" .67, "Cooks, fast food" -1.86.

There are two concerns regarding the representativeness of job postings data to measure the within-firm task differentiation of employed workers. First, online job posting data tends to overrepresent white collar and STEM jobs that require a bachelor's degree (Carnevale et al. (2014)). Overrepresentation of these high-wage occupations in the job postings would tend to bias down the differentiation scores of these occupations, as there should be a greater number of similar postings to be put into these jobs' clusters. The second concern is that low-wage jobs will tend to have a higher turnover, meaning that the actual number of employed workers who are high-wage will be higher given the number of job postings. This will tend to offset the first bias, and the overall direction of the bias in ambiguous.

[34] A special thanks to Eskil Heinesen to providing the crosswalk of US SOC occupations to Danish ISCO codes.

Figure B.4: Sample Skills in Burning Glass Data

	skillclusterfamily	skillcluster	skill
1	Analysis	Natural Language Processing (NLP)	Speech Recognition
2	Analysis	Statistics	Multivariate Testing
3	Architecture and Construction	Carpentry	Basic Tools
4	Engineering	Engineering Practices	Prototype Design Development
5	Environment	Water Testing and Treatment	Water Quality Modeling
6	Finance	Financial Risk Management	Strategic Risk Management
7	Finance	Financial Trading	Exchange-Traded Products (ETP)
8	Finance	Mortgage Lending	Mortgage Loan Origination
9	Finance	Tax	CorpTax
10	Health Care	Alternative Therapy	Therapeutic Procedures
11	Health Care	Basic Patient Care	Telemedicine
12	Health Care	Cardiology	Electrocardiogram (EKG) Equipment
13	Health Care	Clinical Data Management	Clinical Programming
14	Health Care	Clinical Research	Clinical Trial Design
15	Health Care	Emergency and Intensive Care	Critical Care Nursing
16	Health Care	Medical Records	Medical Records Management
17	Health Care	Oncology	Bone Marrow Transplant
18	Health Care	Physical Therapy	Physical Disability
19	Health Care	Routine Examination Tests and Procedures	Treatment Recommendation
20	Human Resources	Occupational Health and Safety	Health and Safety Compliance
21	Industry Knowledge	Supply Chain and Logistics Industry Knowledge	Export Industry Knowledge
22	Industry Knowledge	Transportation Industry Knowledge	Shipping Industry Knowledge
23	Information Technology	Data Management	IBM InfoSphere Optim
24	Information Technology	Database Administration	Shareplex
25	Information Technology	Database Administration	Oracle Database Administration (DBA)
26	Information Technology	Extraction, Transformation, and Loading (ETL)	Data Transformation
27	Information Technology	Networking Hardware	DSLAM
28	Information Technology	Programming Principles	Functional Programming
29	Information Technology	SAP	SAP CRM
30	Information Technology	Scripting Languages	JRuby
31	Information Technology	Systems Administration	Backup administration
32	Information Technology	Test Automation	Test Development
33	Manufacturing and Production	Computer-Aided Manufacturing	Siemens Plc
34	Marketing and Public Relations	Online Advertising	Internet Advertising
35	Sales	General Sales	Cold Calling
36	Sales	Prospecting and Qualification	Product Promotion
37	Sales	Specialized Sales	Furniture Sales
38	Supply Chain and Logistics	Facility Management and Maintenance	Facility Repair
39	Supply Chain and Logistics	General Shipping and Receiving	Shipping Methods
40	Supply Chain and Logistics	Material Handling	Forklift Operation

This figure provides an example of the Burning Glass skill data. In the leftmost column is the most aggregated categories. In the middle column called "skillcluster", which in the main text refer to as a "skill group", is the level of skills that we use.

Figure B.5: Sample Skills in Burning Glass Data

Medical and Health Services Managers											
bgtjobid	3.80E+10	3.78E+10	3.81E+10	3.78E+10	3.82E+10	3.81E+10	3.82E+10	1.03E+08	3.78E+10	2.38E+08	3.82E+10
soc	119111	119111	119111	119111	119111	119111	119111	119111	119111	119111	119111
Cluster	12	12	12	12	141	141	141	1180	9999	9999	9999
scheduling	1	1	0	0	1	1	1	1	0	0	0
emergency and intensive care	1	1	1	1	0	0	0	1	1	0	0
general medicine	1	1	1	1	0	0	0	1	0	0	0
advanced patient care	1	1	1	1	0	0	0	0	1	0	0
basic patient care	1	1	1	1	0	0	0	0	1	0	0
medical support	1	1	1	1	0	0	0	0	0	0	0
oncology	1	1	0	0	0	0	0	0	0	0	0
patient education and support	1	1	1	1	0	0	0	0	0	0	0
surgery	1	1	1	0	0	0	0	0	0	0	0
health information management	0	0	0	0	1	1	1	1	0	1	0
general administrative and clerical	0	0	0	0	1	1	1	1	0	0	0
medical billing and coding	0	0	0	0	0	0	0	1	0	1	0
medical records	0	0	0	0	0	0	0	1	0	1	0
data analysis	0	0	0	0	0	0	0	1	0	0	0
data management	0	0	0	0	0	0	0	1	0	0	0
data techniques	0	0	0	0	0	0	0	1	0	0	0
management information systems	0	0	0	0	0	0	0	1	0	0	0
medical documentation and abstrac...	0	0	0	0	0	0	0	1	0	0	0
process improvement	0	0	0	0	0	0	0	1	0	0	0
budget management	0	0	0	0	0	0	0	0	1	1	1
business process and analysis	0	0	0	0	0	0	0	0	0	1	0
people management	0	0	0	0	0	0	0	0	1	1	1
quality assurance and control	0	0	0	0	0	0	0	0	0	1	0
auditing	0	0	0	0	0	0	0	0	0	0	1
businessstrategy	0	0	0	0	0	0	0	0	0	0	1
contract management	0	0	0	0	0	0	0	0	0	0	1
healthcare procedure and regulation	0	0	0	0	0	0	0	0	0	0	1
occupational health and safety	0	0	1	0	0	0	0	0	1	0	0
	Direct Caregiving				Admininstrative			Admin. Analysis	Care Analysis	Management	Regulation

This figure shows an example of how the skill groups listed for each job are used to cluster job posting within the firm. Every column in this figure is a different job posting in the same firm, and each job is categorized as the "Medical and Health Service Managers" occupation. An entry is a 1 if a job had a skill requirement that fell in one of the skill groups, which make up the rows of the figure. The row labeled "cluster" indicates the (arbitrary) cluster ID assinged to job postings that are clustered together. The cluster ID of 9999 indicates that the job posting was not successfully put into a large cluster, i.e., it is in a cluster of size 1. Job postings within a cluster tend to have very similar skills requirements.

B.6 Cross-Sectional and Passthrough Estimates by Occupation Group

Table 6 reports both cross-sectional and passthrough estiamtes using various specifications. Columns 1, 2, and 3 show OLS cross-sectional specifications that estimate the relationship between firm productivity, size, and wages. In column 1, the cross-sectional elasticity between productivity and wages for all occupations is 0.071. With a standard deviation of firm productivity of 0.63, this suggests that workers at firms that are one standard deviation above average earn 9% higher hourly wages than workers at a firm one standard deviation below average. When looking at the results for firm size in column 2, a worker at a large firm earns only 4% more than at a small firm, with an elasticity of 0.01 and standard deviation of 1.96. Column 3 includes both firm size and productivity, but firm productivity is interacted with a dummy for each occupation group. Managers have the highest cross-sectional elasticity of productivity to wages at .224, and administrative/service/manual jobs have the lowest value at .061. Notably, the coefficient on firm size is now negative at -.009, suggesting that in the cross-section, firm size is not predictive of wages after conditioning on productivity.

Column 4 of Table 6 reports the estimates of an OLS specification that regresses three year changes in wages on three year changes in productivity, and columns 5 and 6 show the results reported in Figure 5. Compared to the IV estimates, the coefficients in the OLS specification are much smaller, with magnitudes of one fourth to one third of the size. This reflects that the IV strategy addresses measurement error that will still be present in the OLS specification, as well as the OLS regression includes some components of transitory shocks that are likely to have lower passthrough.

Comparing columns 5 and 6 with column 3, we can infer what share of the cross-sectional elasticity of wages to productivity is due to rents, with the remainder due to sorting of higher wage workers to productive firms within occupation groups. For both administrative/service/manual and crafts/assembly occupations, the passthrough elasticity is approximately three-quarters of the cross-sectional elasticity. For professionals/technicians and managers, rents account for about two-thirds of the cross-sectional elasticity, leaving a larger role for sorting.

Table 6: Cross-Sectional and Passthrough Estimates by Occupation Group

	$\log(w)_t$			$d\log(w)_{t\text{-}1,t+2}$		
		OLS		OLS	IV	
	Main	Main	Main	Main	Main	Size≥50
	(1)	(2)	(3)	(4)	(5)	(6)
Productivity $\hat{Y}$	.071	.	.	.	.	.
	(.005)					
$\log(FTE)$	.	.010	-.009	.	.	.
		(.003)	(.003)			
$\hat{Y} \times \mathbb{1}$ {Managers}	.	.	.224	.029	.143	.103
			(.018)	(.003)	(.026)	(.022)
$\hat{Y} \times \mathbb{1}$ {Professional/Technician}	.	.	.089	.014	.059	.049
			(.007)	(.003)	(.021)	(.023)
$\hat{Y} \times \mathbb{1}$ {Crafts/Assembly}	.	.	.091	.021	.074	.064
			(.008)	(.002)	(.021)	(.025)
$\hat{Y} \times \mathbb{1}$ {Admin/Service/Manual}	.	.	.061	.014	.045	.034
			(.007)	(.002)	(.012)	(.013)
Occ FE Digit	4	4	4	4	4	4
Observations (Thous.)	5,989	5,989	5,989	2,441	625	492
R-squared	.575	.569	.579	.119	.071	.073
Clustered SE's	Firm	Firm	Firm	Firm	Firm	Firm

Standard errors in parentheses
p<0.01, * p<0.05

This table reports estimates for equation (12) in both levels and changes. "Main" indicates the full worker sample. Columns 1, 2, and 3 show the OLS levels regression, where the firm size and productivity variables are averages over the entire sample period. The first column reports only firm productivity, and the second column reports only firm size. The standard deviations for average $\hat{Y}$ and the log of FTE are 0.63 and 1.96, respectively. The third column includes firm productivity interacted with dummies for each broad occupation group, as well as firm size. The levels regressions include fixed effects at the 4-digit occupation level. Column 4 shows the OLS first-differences over three year changes, and columns 5 and 6 show the IV estimates, where three year changes in log value added residuals $\Delta\hat{Y}_{k,t\text{-}1,t+2}$ are instrument with five year changes $\Delta\hat{Y}_{k,t\text{-}3,t+4}$. The sample in column 6 restricts to sample to firms that had at least 50 full time equivalent employees in year $t-1$. All regressions include individual worker controls include the firm age, workers' lagged tenure, years of education, years of education squared, potential experience, interactions of education and potential experience, as well as dummies for sex, year, and 1-digit industry code.

B.7 Productivity vs. Size

Using the wage changes of job switchers, we can also test if firm size is predictive of wages. This provides a test of hold-up power as a source of wage premia, where wages should be a function of productivity but not firm size, against models that predict wage premia at large firms, regardless of productivity. We regress changes in wages of switchers on changes in the average productivity and average size of the worker's two employers, using

the specification with s=3:

$$\Delta \log(\mathrm{w}_{ij,t,t+s}) = \chi_0^j + \chi_1^j \Delta \bar{Y}_{i,k(t),k(t+s)} + \chi_2^j \Delta \log(\overline{FTE})_{i,k(t),k(t+s)} + X_i^j \gamma + e_{it}, \quad (17)$$

We use this specification to make use of the common assumption that observing the same worker across firms will net out worker fixed effects. The coefficients in the regression will be unbiased under the assumption of exogenous mobility.

Table 7: Firm Productivity and Size Wage Effects, Levels and Switchers

Dependent:	$log(w)_t$ Cross-Sectional			$d\log(w)_{k(t),k(t+3)}$ Switchers			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Workers	All	All	All	All	All	All	Executives
$\bar{Y}$	.071		.090	.028		.034	.015
	(.005)		(.006)	(.002)		(.002)	(.035)
$\log(\overline{FTE})$		.010	-.010		.004	-.003	.041
		(.003)	(.003)		(.001)	(.001)	(.021)
Observations (Thous)	5,989	5,989	5,989	70.9	70.9	70.9	.326
R-squared	.575	.569	.579	.050	.043	.051	.063

Robust standard errors in parentheses. $corr(\bar{Y}, \log(\overline{FTE}))$=.69.
p<0.01, * p<0.05. $sd(\bar{Y}) = .59$, $sd(\log(\overline{FTE})) = 1.96$.

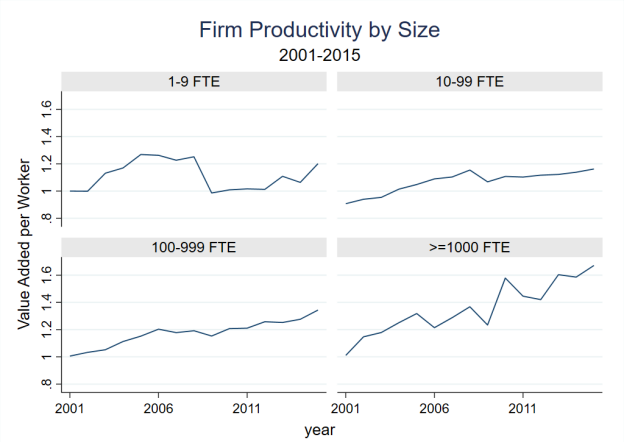
This table reports estimate of the effect of firm productivity and firm size on wages. Columns 1, 2, and 3 report OLS estimates in the cross-section. Columns 4, 5, and 6 report esimates from switchers. The switcher sample includes workers who switch firms between year t and year $t+1$. Therefore, the wage reading at the leaving firm is from the final November of the workers tenure. The wage reading at the arrival firm is in the third November, so the worker has at least two full years of tenure in the firm.

Table 7 shows the results of the above regression, both in the cross-section and for switchers. Focusing on column 5, we find that workers who switch into larger firms receive only very small increases in wages, with an elasticity of less than 1 percent. This presents a challenge for models of long-run upward sloping labor supply as a source of heterogeneity in firm premia, but it is consistent with a model where rents are shared by productive firms to workers with hold-up power, independently of firm size.

In the last column of Table 7, we report the effect of firm size and productivity on workers who are listed as a top executive in both firms before and after the switch. We estimate that when an executive moves to a firm that is 100 log points larger, their pay increases by 4.1 log points. This is consistent with the idea that production complementarities may be increasing in firm size for executives.

B.8 Productivity Gains by Danish Large Firms

Figure B.6: Changes in Average Labor Productivity by Firm Size, Denmark 2001-2015

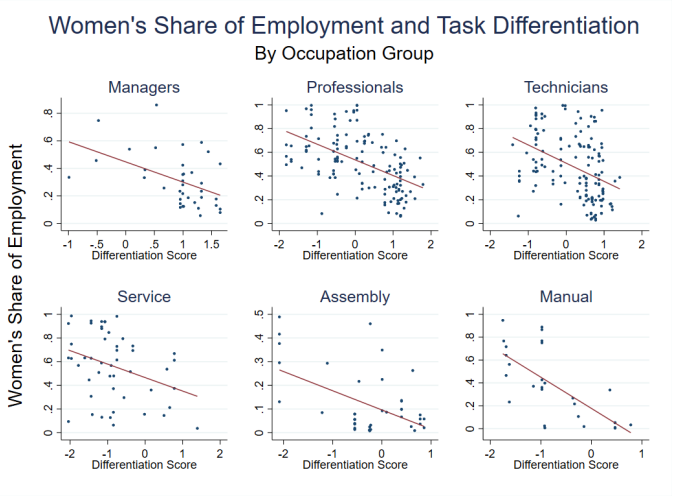


This figure shows changes in value added per worker (measured as value added per full time equivalent) from 2001 to 2015. Value added per worker is deflated by the Danish consumer price index and normalized relative to the level of the smallest firms (1-9 FTE) in 2001. Productivity growth over this period is monotonically related to firms size, and value added per worker at the largest firms grows approximately 65%.

Figure B.6 reports the inflation-adjusted average value added per worker in firms of different sizes from 2001 to 2015, relative to the productivity of the smallest firms (1-9 full time equivalents) in 2001. This figure shows that larger firms have seen faster productivity growth, with firms between 100-999 FTE increasing output per worker by 38% and the largest firms with over 1000 FTE increasing productivity by 65%. The shares of employment across these groups has been relatively constant across this time period, implying that an increasing share of value added is produced by these larger firms. The trend in labor shares (not shown) mirrors the changes in productivity: the labor share in the largest firms, those with at least 1000 FTE, fell from .71 to .55 between 2001 and 2015, while the aggregate labor share in all private sector firms with at least one employee fell from .73 to .64.

B.9 Task Differentiation and Within-Occupation Gender Wage Gap

Figure B.7: Task Differentiation and Women's Employment Share by Occupation



This figure reports the share of workers in an occupation who are women and the occupation's differentiation score, separately for detailed 4-digit DISCO codes within broad 1-digit DISCO occupation groups. Managers, professionals, assembly, and manual occupations show the steepest negative relationship between task differentiation and the share of workers in each occupation who are women.