A Firm Link: Overall, Between- and Within-Firm Inequality Through the Lens of a Sorting Model*

Paweł Gola† and Yuejun Zhao‡

December 2023

Abstract

This paper provides a new theory of the observed co-movement between overall wage inequality and its between-firm component. We develop and solve analytically a frictionless sorting model with two-sided heterogeneity, in which firms consist of distributions of tasks, choose how many workers to employ and reward their workers both through wages and amenities. We show that, for empirically-relevant parameter ranges, overall and between-firm inequality are firmly linked: A change in any of the models' primitives increases overall wage inequality if and only if it also increases the ratio of between-firm to overall inequality. Subsequently, we calibrate the model to match the Norwegian economy and find that the increase in wage inequality from 1995 to 2014 had a different primary cause (raising span-of-control cost) than the accompanying rise in welfare inequality (increased skill variance), and that the apparent decrease in wage inequality after 2015 masked a continued increase in welfare inequality.

JEL Codes: C78, J21, J31, J32

Keywords: within-firm inequality, between-firm inequality, sorting, amenities, large firms, technological change

*We would like to thank Vasco Carvalho for inspiration and early discussions, as well as Jesper Bagger, Diego Battiston, Philipp Kircher, Espen R. Moen, Isaac Sorkin and the seminar audience in Edinburgh for helpful comments and suggestions. Gola gratefully acknowledges the support from Iceland, Liechtenstein, and Norway through the European Economic Area (EEA) Financial Mechanism 2014–2021 (Project "Micro-level responses to socio-economic challenges in face of global uncertainties", No S-BMT-21-8 (LT08-2-LMT-K-01-073)) under a grant agreement with the Research Council of Lithuania. Zhao acknowledges financial support from the Research Council of Norway, grant number 300917. Declarations of interest: none. All errors are ours.

†University of Edinburgh, Edinburgh, EH8 9JT, United Kingdom. E-mail: pawel.gola@ed.ac.uk

‡University of Edinburgh and the Ragnar Frisch Centre for Economic Research. E-mail: yuejun.zhao@ed.ac.uk

1 Introduction

Most developed countries experienced a well-documented increase in wage inequality from the 1960s until early 2010s. More recently, Barth, Bryson, Davis, and Freeman (2016); Barth, Davis, and Freeman (2018); Song, Price, Guvenen, Bloom, and von Wachter (2018); Sorkin and Wallskog (2023) have documented that in the United States, the increase in overall wage inequality has been driven mostly by a rise in between-firm wage inequality, with a much smaller contribution from changes in within-firm wage inequality; the same trend has been subsequently found to hold in a wide range of developed countries (Tomaskovic-Devey et al., 2020).

This empirical trend is quite intriguing. First, its drivers are unclear. The sustained increase in overall inequality has previously been convincingly linked to skill-biased technological change, that is, to changes in technology that increase the return to skill (see, e.g., Acemoglu, 2002). However, existing theories are silent on why skill-biased technological change would increase between-firm inequality by a larger proportion than within-firm inequality.[1] Second, paired with the recent findings by Sorkin (2018) that as much as 70% of firms' contribution to wage variance is caused by compensating differentials, the fact that most of the increase in overall wage inequality has been caused by its between-firm component creates the possibility that a significant part of the observed large increases in wage inequality may reflect changes in the compensating differentials paid by firms, and have no bearing on welfare inequality.

In this paper, we attempt to determine how welfare inequality has evolved in the past three decades, and what are the main drivers of the observed changes in overall, between-firm and within-firm wage inequality. We take a structural approach to this problem, by first developing a novel, extremely tractable model of worker's sorting between- and within-firms, and then calibrating it to Norwegian administrative data. Our model builds on Costrell and Loury (2004) and Eeckhout and Kircher (2018). As in Costrell and Loury (2004), firms consist of a hierarchy of tasks, that need to be performed in fixed proportions in order to produce output. As more complex tasks are complements with highly skilled workers, this implies that each firm hires a distribution of workers, and solves a within-firm problem of assigning workers to tasks. Further, we follow Eeckhout and Kircher (2018) in (a) allowing firms to be heterogeneous in their overall productivity and (b) choosing their size endogenously, subject to a span-of-control cost. Finally, workers care not just about wages, but also the effort they exert, and the level of amenities provided by the firm, and effort and amenities enter the firm's profit function differently than wages.

Of course, a general model of heterogeneous, hierarchical firms with endogenous

[1] Cortes, Lerche, Schönberg, and Tschopp (2023) point out that skill-biased technological change contributes to the observed increase in between-firm wage inequality in a Melitz (2003) model with search frictions, but even in this case it is unclear why this increase would be disproportionately large compared to the increase in within-firm inequality.

amenity provision would be entirely intractable. To ensure tractability, we impose three sets of simplifying assumptions. First, we assume that the type of the firm and the type of the task enter the production function through a one-dimensional index, which we call the *job* characteristics.[2] Second, we impose a set of standard functional form assumptions on the production and utility functions which ensure that, for a given distribution of jobs and the optimal solution of the effort exertion problem, the sorting problem reduces to the problem of sorting workers to jobs—that is, it effectively reduces to a standard Sattinger (1979) problem. However, the economy-wide distribution of jobs is a mixture of the within-firm job distributions weighted by the product of the number of firms of a given type and the firms' endogenous size, and thus remains an equilibrium object which can be found only by solving a complicated integral equation. Hence, and third, we assume that (a) workers' skill, (b) tasks within each firm, and (c) firms' productivity are all normally distributed with mean zero, which allows us to leverage the fact that a normal mixture of normal distributions is normal. Together with a standard Cobb-Douglas span-of-control cost function, the normality assumption ensures that the equilibrium distribution of jobs is always normal with mean zero; hence, one needs only to solve for the equilibrium variance of jobs, which is a problem that admits an analytical solution.

The equilibrium of the model reproduces a number of prominent empirical patterns. First, even though the model is perfectly competitive, larger firm pay higher wage premia in equilibrium (see, e.g., Bloom, Guvenen, Smith, Song, and von Wachter, 2018). Notably, while all workers receive the same utility in all firms, the composition of their compensation differs systematically with the firm type. In particular, higher amenity provision decreases the dollar-cost of effort exertion, which raises optimal effort exertion more than proportionately. Hence, firms that provide higher amenities must additionally compensate their workers for the effort by paying them more. Because effort and firm productivity act as complements, more productive firms benefit more from cheaper effort exertion, and hence find it optimal to provide more amenities. At the same time, more productive firms find it optimal to hire more workers, so that in equilibrium larger firms pay workers higher wage premia and provide more amenities, both as a compensation for higher effort exertion. Second, every type of firm hires all types of workers, but the workers hired by more productive firms are more skilled on average. Thus, the correlation between the skill of workers and the wage premia paid by firms is (a) equal to the correlation between workers skill and firm productivity, (b) strictly positive and (c) strictly smaller than 1. This means that the model is able to reproduce the findings from the literature on worker- and firm-fixed effect (started by Abowd, Kramarz, and Margolis, 1999), which—once the limited mobility bias is correctly accounted for—consistently finds positive, but small correlations between worker and firm fixed effects (see Bonhomme,

[2]Think of a job as a firm-task pair; our simplifying assumption is then to impose a total order on these firm-task pairs.

Holzheu, Lamadon, Manresa, Mogstad, and Setzler, 2023, for an in-depth discussion).

As the model admits an analytical solution, we are able to derive comprehensive comparative statics results. We consider four exogenous changes to the primitives: (i) changes in the distribution of workers' skill, (ii) changes in the distribution of firms' productivity, (iii) changes in the cost of amenity provision and (iv) changes in the span-of-control cost. The main insight emerging from these results is that—for empirically relevant parameter ranges—each of these exogenous changes affects overall, within- and between-firm wage and welfare inequalities *in the same direction*, but its effect on between-firm inequality is disproportionately large compared to its impact on within-firm inequality. In this sense, the empirical co-movement of wage inequality and its share explained by the between-firm component should not be surprising; indeed, in our model, increases in wage inequality are *not* driven by the between-firm component only if the economy is affected by two or more exogenous changes that offset each other.

To understand the intuition behind this strong result, suppose that amenity provision is prohibitively expensive, and thus firms pay no wage premia. In that case, a worker's wage is an increasing function of their skill only, and—because more productive firms hire better workers on average—the average wage paid by a firm is an increasing function of the firm's type. Naturally, then, one can always *re-normalise* the skill and productivity distribution in such a way that wages are linear in skill and average firm-level wages are linear in productivity.[3] The following relationship among between-firm wage inequality, overall wage inequality and equilibrium sorting follows then easily from basic properties of covariance and the definition of correlation:

$$\frac{\text{Between-firm Wage Inequality}}{\text{Overall Wage Inequality}} = \text{Corr}(\text{Skill}, \text{Productivity})^2. \tag{1}$$

Thus, the share of overall wage inequality that is explained by between-firm wage inequality is equal to the square of the equilibrium correlation between workers' skill and firms' productivity, both appropriately defined.[4]

In other words, the extent to which between-firm inequality accounts for overall inequality is very tightly linked to the strength of sorting between high skilled workers and high productivity firms. In our model, the strength of sorting is a function of the equilibrium supply of quality jobs only. The intuition for this is subtle, and thus we defer a detailed explanation until Section 5.2; broadly speaking, however, the reason is that higher productivity firms consist of a better distribution of jobs than lower productivity firms, and thus both the strength of sorting and the supply of quality jobs are determined by how many of the jobs available in the economy are offered by highly productive firms.

[3] In our case this is not necessary, as wages are linear in (squared) skill, and average firm-level wages are linear in (squared) productivity.

[4] That is, defined in such a way that, in equilibrium, wages are linear in skill, and average firm-level wage is linear in productivity.

Because high skill workers and quality jobs are complements, any improvement in the supply of quality jobs increases overall wage inequality. Therefore, if a change in one of the model's primitives increases the ratio of between-firm to overall wage inequality, then it must increase the strength of sorting, which is only possible if it improves the supply of quality jobs; thus, it must also increase overall wage inequality.

The presence of compensating differentials complicates the relationship between overall and between-firm wage inequality, because the wage earned by a worker depends now also on the firms' type. This implies, in particular, that the ratio of between-firm to overall wage inequality depends not only on the strength of sorting, but also on the sign and size of the compensating differential paid by more productive firms. Nevertheless, it remains true for empirically relevant parameter values that if a change in any of the four primitives considered by us increases wage inequality, it must also increase the ratio of between-firm to overall wage inequality.[5] Overall, therefore, our theoretical analysis reveals that the striking empirical co-movement of overall and between-firm wage inequality is the most natural outcome of the model, and can be caused by a wide range of exogenous changes. To shed some light on what its actual causes could have been, we need to bring the model to the data.

Specifically, we use our model to analyse the causes of changes in wage inequality between 1995 and 2019 in Norway. We start by documenting the empirical trends in the variance of wages, as well as within- and between-firm wage inequality in Norway in Figure 1. Between 1995 and 2014, which is the part of our panel that overlaps with the period analysed in Song et al. (2018), the trends were qualitatively similar to those in the United States.[6] Specifically, we find that overall wage variance increased by 3 log points,[7] within-firm wage inequality increased by 0.8 log points, between-firm wage inequality increased by 2.2 log points, implying that the share of overall inequality explained by between-firm wage inequality increased by 5 percentage points. The results for years 2015 to 2019, which we are first to document, show a reversal in some of these trends.[8] While within-firm wage inequality has continued with a modest increase of 0.3 log points, overall wage inequality experienced a decrease of 0.5 log points, and between-firm wage inequality has fallen by 0.8 log points, leading to a sharp decrease in the share of overall inequality explained by between-firm wage inequality by 3 percentage points.

We calibrate the model to matched employer-employee data in Norway. Our model can be identified using four key moments: within-firm wage inequality, between-firm

[5]Specifically, this result holds if (log) within-firm wage inequality is less than approximately 0.7, which is true for instance, in all countries studied by Tomaskovic-Devey et al. (2020).

[6]Of course, the levels are much lower in Norway than the US, but comparable to other Scandinavian countries such as Denmark (Friedrich, 2022) and Sweden (Bonhomme, Lamadon, and Manresa, 2019).

[7]We borrow the term 'log points' from Song et al. (2018). It means that the change in the variance of log earnings is multiplied by 100.

[8]Due to a change in the reporting scheme in 2015 (Statistics Norway, 2014), we divide our sample into two parts, 1995 to 2014, and 2015 to 2019. See Section 6.1 for more details.

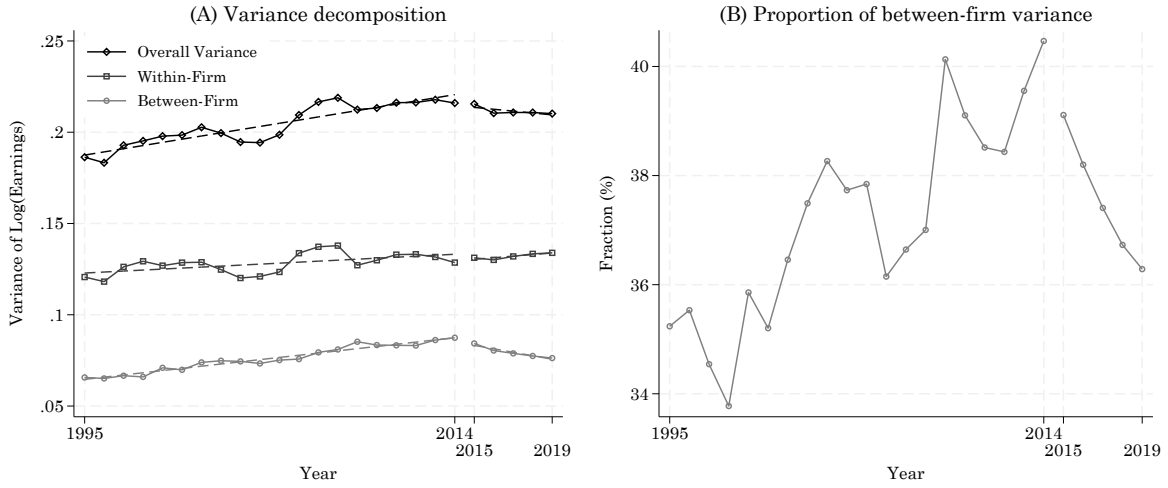


Figure 1: Decomposition of the variance of log annual earnings within and between firms in Norway from 1995 to 2019. *(A)* Overall variance, within-firm variance, and between-firm variance. *(B)* Proportion of between-firm variance.

wage inequality, the variance of within-firm wage variances and the average within-firm variance of wages *unweighted* by firm size.[9] We find that from 1995 to 2014, the variance of skill has sharply increased, whereas the span-of-control cost has sharply decreased, with little to no change in the distribution of firm productivity and the cost of amenity provision. Taken together, these changes have improved the equilibrium supply of quality jobs. After 2015, in contrast, the supply of quality jobs has decreased in the Norwegian economy, which was driven by a fall in the variance of productivity and an increase in the cost of amenity provision.

With a calibrated model in tow, we simulate changes in welfare inequality over the entire period and find that the trends in welfare inequality qualitatively match trends in wage inequality from 1995 to 2014. However, after 2015 overall welfare inequality continues with a modest increase. Hence, our calibration indicates that post-2014 the observed decreases in overall wage inequality reflect shifts in the composition of compensation between wages and compensating differentials, rather than actual decreases in underlying welfare inequality.

Finally, we perform counterfactual simulations to isolate the impact that each of the four channels had on wage and welfare inequality over the period of analysis. For 1995-2014 we find that while a fall in the span-of-control cost was the main driver of the increase in wage inequality (62%), changes in the distribution of workers skill were the main contributor to the increase in welfare inequality (52%). Given that changes in the distribution of skill are isomorphic to skill-biased technological change in our model, this finding reconciles the old and new literatures on inequality: Changes in the

[9]The difference between the first and fourth moment is that in the first moment is calculated by taking a weighted average of the within-firm variance of wages.

price/variability of skill were indeed the main driver of changes in *welfare* inequality, because a large proportion of the increase in between-firm wage inequality was welfare-neutral—specifically, out of the 62% of the increase in overall wage inequality that was caused by a fall in the span-of-control cost, more than three quarters reflects changes in the compensating differential for effort exertion.

Post-2014, the emerging deterioration in the quality of jobs in the Norwegian economy has dominated the continued increase in the variance of skill, leading to a decrease in overall wage inequality. However, for realistic parameter values, changes in the distribution of skill have a stronger effect on welfare than on wage inequality; thus, the increase in the variability of skill has caused a continued increase in welfare inequality post-2014.

The rest of the paper is structured as follows. Section 2 discusses the related literature. Section 3 develops the model and discusses our modelling choices. Section 4 solves the model. Section 5 discusses the behaviour that emerges in equilibrium and provides comparative statics results. Section 6 calibrates the model to the Norwegian data, simulates the evolution of welfare inequality, and provides a counterfactual analysis. Section 7 concludes. Appendix A contains omitted proofs and derivations. Online Appendix B discusses how to conduct an AKM variance decomposition in the context of our model. Online Appendix C details the resampling procedure used to calculate confidence intervals. Online Appendix D contains additional figures relating to the calibration exercise.

2 Related Literature

Theoretical Sorting Models By merging the models of Costrell and Loury (2004) and Eeckhout and Kircher (2018) our paper contributes to the literature on labour market sorting: It introduces firm heterogeneity and firm's size choice into the former and hierarchical firms (i.e., firms that solve a within-firm assignment problem) into the latter. Eeckhout and Kircher (2018) itself nests a large family of sorting models in which the size of a firm/number of workers performing a task is endogenous, as in Sattinger (1975); Teulings (1995, 2005); Costinot (2009); Costinot and Vogel (2010). Our paper provides a blueprint on how to introduce within-firm heterogeneity into these models, while maintaining tractability. Indeed, even with the addition of the within-firm sorting problem, our choice of normal skill and productivity distributions, coupled with a production function that makes equilibrium firm size exponential in the square of firm's type yields a more tractable model than most in this literature.[10]

This paper contributes to the nascent literature that uses sorting and/or search models to investigate the causes of the more-than-proportional increase in between-wage inequality. Boerma, Tsyvinski, and Zimin (2023) and Freund (2022) both develop models

[10]A notable exception is Teulings (1995), whose model also allows for a fully analytical solution, and from whose functional form choices we draw considerable inspiration.

of team formation in which the number of team members is fixed; hence, these models do not feature the connection before firms' size and equilibrium sorting, which is critical for the tight link between overall and between-firm wage inequality highlighted in this paper. Trottner (2022) and Cortes et al. (2023) study, respectively, the impacts of trade liberalisation and skill-biased technological change on overall, between- and within-firm inequality, in models that build on Melitz (2003). In both of these models, firms' production is determined endogenously and better firms tend to be both larger and hire better workers; both papers emphasise a similar mechanism, by noting that their shock of interest increases the employment in the most productive firms and through that raises between-firm inequality.[11] However, by developing a model with a simple analytical solution and considering a range of possible changes of primitives, we are able to make a more general point: The relationship between the size of productive firms and sorting means that most shocks (rather than just skill-biased technological change or trade liberalisation) which increase overall wage inequality do so by increasing between-firm inequality *by a larger proportion* than within-firm inequality. Our flexible framework allows us also to quantify the extent to which different types of changes have affected overall, within and between-firm inequality.

Empirical Literature on Between-Firm Inequality Our paper contributes to the large literature documenting that the increase in overall wage inequality from the 1980s to early 2010s was disproportionately driven by the increase in its between-firm component. The rising influence of between-firm wage inequality is documented for 12 (out of 14) developed countries (including Norway and the US) in Tomaskovic-Devey et al. (2020); for the US in Barth et al. (2016, 2018); Song et al. (2018); Sorkin and Wallskog (2023), and (for more recent years) Haltiwanger, Hyatt, and Spletzer (2022, 2023); for Sweden in Håkanson, Lindqvist, and Vlachos (2021); for Germany in Card, Heining, and Kline (2013); Baumgarten, Felbermayr, and Lehwald (2020); Freund (2022); Lochner and Schulz (2022); and for Brazil (from 1986–1995) in Helpman, Itskhoki, Muendler, and Redding (2016).

Our paper contributes also to the growing body of work documenting that overall and between-firm wage inequality have both decreased in the second half of the 2010s in countries that have previously experienced increases in inequality. Specifically, overall wage inequality fell slightly between 2012 and 2020 in the US when measured using weekly earnings (Aeppli and Wilmers, 2022),[12] between 2014 and 2016 in Sweden (Engbom, Moser, and Sauermann, 2023), and between 2010 and 2017 in Germany (Lochner, Seth,

[11] Eeckhout and Pinheiro (2014) also develop a model in which all firms hire workers of many types, and more productive firms are larger and hire better workers. However, Eeckhout and Pinheiro (2014) do not apply their model to the study of wage inequality.

[12] With respect to hourly and annual earnings, Aeppli and Wilmers (2022) find that inequality has stabilised since 2012 in the US.

and Wolter, 2020; Freund, 2022). Accordingly, between-firm wage inequality experienced a slight decline between 2014 and 2018 in the US (Landefeld, Mortenson, and Willingham, 2023), and in the late 2010s in Germany (Freund, 2022).

Amenities and Inequality In our model, the only reason for firm wage premia are the compensating differentials for elevated effort; this modelling choice is motivated by the finding in Sorkin (2018) that compensating differentials account for majority of the variance in wage premia. Further, Ouimet and Tate (2023) document that the vast majority of variation in non-wage benefits—health, retirement, and leave benefits—is explained by the between-firm component; this finding has motivated our choice of modelling amenities to be the same for all workers within a firm. There is a number of recent papers focusing on how much inequality there exists in amenity provision, and how this amenity inequality affects measures of overall inequality (Kristal, Cohen, and Navot, 2020; Sockin, 2022; Bana, Bedard, Rossin-Slater, and Stearns, 2023; Maestas, Mullen, Powell, Von Wachter, and Wenger, 2023).[13] However, with the exception of Sockin (2022), these studies look at a limited subset of the amenities that matter in reality and that our model allows for. Indeed, the most comparable numbers are those provided by Sorkin (2018) and Taber and Vejlin (2020), in which a revealed preference approach is adopted. Sorkin (2018) and Taber and Vejlin (2020) find that 'real' inequality (i.e., inequality after accounting for the impact of compensating differentials) is about 23% lower and 125% higher, respectively, than observed wage inequality. In our calibrated model, welfare inequality is about 30% lower than observed wage inequality.[14]

Skill-Biased Changes There is a large literature linking changes in inequality to either skill-biased technological change, or increases in the heterogeneity of workers' skill.[15] An increase in the demand for cognitive skills caused by technological change has been identified as the main culprit behind the increase in wage inequality from the 1960s to the 1980s (Bound and Johnson, 1992; Katz and Murphy, 1992; Juhn, Murphy, and Pierce, 1993). More recently, Lindner, Muraközy, Reizer, and Schreiner (2022) have shown that skill-biased technological change increase the college premium by 6.1% in Norway from 2002 to 2013. Acemoglu and Restrepo (2022) document that more than half of the change in wage structure in the US in the past four decades was due to automation replacing

[13] In addition, Bagger, Moen, and Vejlin (2021) emphasise a less direct connection between amenities and inequality. They show, using a calibrated directed search model, that the presence of amenities amplifies sharply the deadweight loss of labour income taxation. Thus, in the presence of amenities the social planner should adopt lower tax rates, leading to more wage inequality.

[14] The fact that this number is lower in our calibrated model is expected, given that we interpret any variation in firm wage premia to be caused by compensating differentials. Indeed, if all of firm variation is excluded, then the decrease in inequality calculated by Sorkin (2018) is 28.5%, and thus very similar to the number produced by our calibration.

[15] In our model, skill-biased technological change is isomorphic to changes in the distribution of skill.

low-skilled jobs, whereas about 10% of the change resulted from skill-biased technological change. Finally, much of the AKM literature (e.g., Card et al., 2013; Song et al., 2018) finds sizeable increases in the variance of worker fixed effects, which can be most readily explained by increases in the variance of skill or the return to skill.[16]

Firm-Side Changes Autor, Dorn, Katz, Patterson, and Van Reenen (2020) document the rise of superstar firms, and show that (a) it explains declining labour shares in the United States and (b) that technological change is an important driver of the increasing concentration in certain industries. Mueller, Ouimet, and Simintzi (2017) provides direct evidence that the increases in inequality in the UK were caused by the increase in employment by the largest firms in the economy, which is consistent with the theoretical mechanism we highlight. Barth, Davis, Freeman, and McElheran (2023) show that capitalised software investments by firms cause increases in both within- and between-firm wage inequality.

3 Model

There are two populations: workers and firms. Workers differ in skill. Following Costrell and Loury (2004) we assume that each firm consists of a hierarchy of tasks. Following Eeckhout and Kircher (2018) we additionally assume that firms differ in productivity and choose endogenously how many workers to employ. Finally, on top of all that, we introduce amenities provision by the firms.

Workers There exists a unit measure of workers. Each worker is endowed with skill x, which is distributed normally with mean 0 and variance σ_x. A worker's utility is a product of the wage w they receive *per unit* of the effort e they provide and the number/quality of amenities a provided by the firm they are employed by, with

$$u(w, a) = \frac{w}{e} a.$$

Workers are paid competitive wages that depend on their skill, exerted effort and the amenity level provided by the firm. Their reservation utility is equal to 0.

Firms There exists a unit measure of profit-maximising firms. Firms differ in productivity θ, which is distributed normally with mean 0 and variance σ_θ. All firms consist of (the same) collection of tasks t. Each firm decides on the number and types of workers they hire, the assignment of workers to specific tasks, and the number/quality of amenities the firm provides. A worker of skill x hired by a firm of productivity θ to exert effort

[16] It is worth noting, however, that in our model changes in any other primitive could also increase the variance of worker fixed effects, albeit many of them only indirectly.

e in performing task t contributes

$$q(x,t,\theta,e) = A\sqrt{\exp(x(t+\theta))e},$$

to total output, where $A > 0$ is the total factor productivity (TFP).

Following Costrell and Loury (2004) we assume that for the final good to be produced, tasks need to be filled in an exogenously fixed proportion. Specifically, we assume that $\Phi(h)$ of workers employed by firm θ perform a task with $t \leq h$, where Φ denotes the cdf of a standard normal distribution. Firms face a span-of-control cost

$$C_l(L) \equiv L^{1+c_l},$$

and an amenity provision cost

$$C_a(a,L) = \frac{L}{c_a}\left(\frac{c_a a}{1+c_a}\right)^{1+c_a},$$

where c_l, c_a capture the convexity of their respective cost functions. The amenity provision cost, in particular, has been chosen in such a way that if $c_a \to \infty$ then, in equilibrium, all firms choose $a = 1$.

Denoting by $m : \mathbb{R} \to \mathbb{R}$ the assignment function of tasks to workers (i.e., $m(t)$ denotes the skill of the worker assigned to perform task t) and by $e : \mathbb{R} \to \mathbb{R}_{\geq 0}$ the effort function (i.e., $e(x)$ denotes the effort exerted by worker of skill x), the total output of firm θ is equal to the sum of the output's of all tasks and the two costs:

$$Q(L,m,\theta,e) = L\int_{-\infty}^{\infty} q(m(t),t,\theta,e(m(t)))\mathrm{d}\Phi(t) - C_L(L) - C_A(a,L).$$

The firms' reservation profit is 0.

Profit Maximisation and the Demand for Skill The demand for skill is determined by firms' hiring, assignment, effort exertion and amenity provision decisions. In order to hire a worker of skill x each firm has to provide them with the competitive utility $u(x)$, where $u : \mathbb{R} \to \mathbb{R}$. The need to provide worker of skill x with utility $u(x)$, implies that they have to be offered wage $w(x,a,e) = u(x)e/a$. Overall, firm θ earns profit $r(\theta)$, provides amenities $a(\theta)$, fills task t with a worker of skill $m(t;\theta)$, requires workers of skill x to exert effort $e(x,\theta)$ and hires $L(\theta)$ workers, where

$$r(\theta) = \max_{L,m,a,e} Q(L,m,\theta,e(m)) - L\int_{-\infty}^{\infty}\frac{u(m(t))e}{a}\,\mathrm{d}\Phi(t) \tag{2}$$

$$(L(\theta),m(\theta),a(\theta),e(\theta)) \in \arg\max_{L,m,a,e} Q(L,m,\theta,e(m)) - L\int_{-\infty}^{\infty}\frac{u(m(t))e}{a}\,\mathrm{d}\Phi(t). \tag{3}$$

We will define the demand for skill analogously to Gola (2021). The *demand for skill* x, $D(x)$, is equal to the measure of tasks filled by workers with skill of *at least* x:

$$D(x) = \int_{-\infty}^{\infty} L^*(\theta)\Pr(m^*(H;\theta) \geq x)\mathrm{d}\Phi(\theta). \tag{4}$$

Of course, the supply of skill x, $S(x)$, is defined as the measure of workers with skill greater or equal to x, with $S(x) = 1 - \Phi(x/\sigma_x)$.

The Competitive Equilibrium with Full Employment In the competitive equilibrium, firms maximise their profits given equilibrium utilities, and markets clear. For notational simplicity, and without loss of generality, we will restrict attention to equilibria with full employment.[17]

Definition 1 (Equilibrium with Full Employment)**.** A competitive equilibrium with full employment is characterised by:
(a) a *task assignment function* $m^* : \mathbb{R}^2 \to \mathbb{R}$, a *firm-size function* $L^* : \mathbb{R} \to \mathbb{R}$, an *effort function* $e^* : \mathbb{R}^2 \to \mathbb{R}$, and an *amenities provision function* $a^* : \mathbb{R} \to \mathbb{R}$, each consistent with firms' profit maximisation (Equations (2) and (3));
(b) a *utility function* $u : \mathbb{R} \to \mathbb{R}$, which clears the market, that is, equates the demand and supply for skill:

$$D(x) \equiv \int_{-\infty}^{\infty} L^*(\theta)\Pr(m^*(H;\theta) \geq x)\mathrm{d}\Phi(\theta) = 1 - \Phi(x/\sigma_x) \equiv S(x).$$

The wage paid in equilibrium to worker x who works for firm θ is pinned down by the utility function, the effort function, and the amenities provision function with

$$w(x,\theta) = \frac{u(x)e^*(\theta)}{a^*(\theta)}. \tag{5}$$

3.1 Discussion

Let us briefly discuss a few of our modelling and analysis choices.

3.1.1 The Role of Effort

Effort exertion plays two roles in the model. First, and most importantly, the introduction of effort—together with the exponential production function—ensures that the logarithm,

[17] Recall that an unemployed worker receives zero utility. Hence, they would be happy to work for any finite wage, receive any finite amenity and exert any finite effort. Similarly, any firm—regardless of the amenities it provides—would be happy to employ a worker of even the lowest skill, if the wage is low enough and the effort exerted is sufficiently high, as the revenue produced by such a worker would exceed the additional wage, amenity and span-of-control costs incurred. Thus, all workers must be employed in any equilibrium.

rather than the level, of wages and utility becomes the natural unit of analysis. This makes it much easier to calibrate the model to the variance of log wages, which is the most commonly used empirical measure of inequality. Second, the interplay between effort and amenity provision ensures that larger firms (a) pay higher wages and (b) provide better firm-level amenities in equilibrium, which aligns the model with recent empirical findings (Lamadon, Mogstad, and Setzler, 2022; Sockin, 2022).

It is wort noting that (the inverse of) 'effort' could be more generally interpreted as any amenity chosen at the task-firm-worker level, as opposed to 'amenities', which are chosen at the firm level. We chose the label 'effort' to succinctly differentiate between these two types of amenities.

3.1.2 Alternative Parameterisations

The model can be parameterised in a number of equivalent ways. It is well-known that in assignment/sorting models one can normalise either the distributions of characteristics or the production function. This remains true in our model. While we have decided to allow for changes in the variance of the distributions of skills and productivity, we could have equivalently assumed that skills and productivity are standard normally distributed, and—instead—the production function is as follows:

$$\bar{q}(x,t,\theta,e) = \sqrt{\exp(\sigma_x x(t+\sigma_\theta\theta))e}.$$

This isomorphism is the reason why we occasionally interpret changes in σ_x as skill-biased technological change—as it is equivalent to an improvement in the productivity of skills.

Further, once the alternative production function $\bar{q}$ has been written down, it becomes obvious that there exist many equivalent formulations, for example

$$\bar{q}(x,t,\theta,e) = \sqrt{\exp(\sigma_x x(t+\sigma_\theta\theta))e} = \sqrt{\exp(\gamma x(\beta t+(1-\beta)\theta))e} = \sqrt{\exp(x(\sigma_x t+\zeta\theta))e},$$

where $\sigma_x \equiv \gamma\beta$, $\sigma_\theta\sigma_x \equiv \gamma(1-\beta) \equiv \zeta$.

The rationale for choosing our baseline parametrisation was two-fold. First, we decided to keep the production function constant and change the distributions, because that makes σ_x easily interpretable as a change in the variability of skills. Second, we decided to normalise the distribution of tasks within firms to be standard normal, and allow for changes in the distributions of skill and productivity, because this was leading to by far the easiest algebraic formulations and cleanest intuition. Nevertheless, it remains true that some potential real-world changes—such as changes to the proportion of tasks needed to produce a good—would be best represented in the model through a combination of changes to σ_x and σ_θ.

3.1.3 The Role of σ_x, σ_θ

Finally, it is worth noting that while σ_x and σ_θ technically capture the variability of skill and productivity, respectively, they *de facto* also capture the average level of skill and productivity in the economy. The reason is that sorting between tasks and skills will be positive and assortative in equilibrium, and hence workers with negative skill will work for firms (and in tasks) with negative productivity. As the production function is multiplicative, this implies that workers with highly negative and highly positive skills are equally productive. In other words, it is x^2 and θ^2 that truly capture skill and productivity in this model; and, of course, the more variable x is, the higher is the mean of x^2.

The fact that the variance and level of the distributions of skill and productivity are captured by one parameter each is not ideal, but it is critical for the remarkable tractability of this model.

4 Solving the Model

In this Section, we will solve for the equilibrium of the model. Let us start by tackling the assignment of workers to *jobs* first, where a job is the sum the task and firm types:

$$h = t + \theta.$$

Note that the equilibrium distribution of jobs in the economy is given by

$$F(h) \equiv \int_{-\infty}^{\infty} L^*(\theta)\Phi(h-\theta))\mathrm{d}\Phi(\frac{\theta}{\sigma_\theta}). \tag{6}$$

We will *guess* that $F(h) = \Phi(\frac{h}{\sigma})$, where $\sigma > \max\{\sigma_x, 1\}$, that is, we will guess that there exists an equilibrium in which jobs are normally distributed with mean 0 and a high enough variance. In the following step, we will derive the wage functions holding in equilibrium given this guess. Finally, we will show that these wage functions lead to firm's size choices that indeed result in the jobs being normally distributed in the economy. In the process, we will also solve for the equilibrium variance of jobs σ^2.[18] Finally, note that for the same reasons as those discussed in Section 3.1.3, σ describes not only the variability of the job distribution but also, effectively, the average quality of supplied jobs. For that reason, we will typically refer to σ as the *supply of quality jobs*.

Effort exertion, Assignment and Utility Define the function $\mu : \mathbb{R}^2 \to \mathbb{R}$, which determines which worker performs job h at firm θ. Then the first order condition (the

[18] This equilibrium is unique in the class of equilibria in which jobs are normally distributed.

Euler-Lagrange condition) of the firm's problem with respect to the effort schedule e yields:

$$e^*(\mu(h),\theta) = \frac{A^2 \exp(\mu(h)h)a^2}{4(u(\mu(h,\theta))^2}.$$

Note that this implies that firms offering higher amenities require their workers to provide disproportionately higher effort. Hence, perhaps surprisingly, firms offering higher amenities will need to offer higher wages to workers. In other words, firms compensate workers for exerting higher effort both by paying them higher wages and by providing better amenities.

Substituting the expression for optimal effort exertion into the firm's problem yields:

$$r(\theta) = \max_{a,L} \max_{\mu} \; L \int_{-\infty}^{\infty} \frac{A^2 a \exp(\mu(h)h)}{4u(\mu(h,\theta)} \, \mathrm{d}\Phi(h-\theta)) - C_l(L) - C_a(a,L).$$

We can now solve the problem of assigning workers to jobs. Technically this is a two-dimensional problem, because the production function depends on jobs and—through amenity provision—on the firm's type. However, one can easily see that the the first order (Euler-Lagrange) condition with respect to the assignment μ implies that:

$$\frac{u'(\mu(h,\theta))}{u(\mu(h,\theta))} = h, \tag{7}$$

which does not depend on the firm's amenity provision. It must be the case, therefore, that the optimally chosen $\mu(\cdot;\theta)$ is independent of θ; henceforth, we will suppress θ when referring to μ. Clearly, by standard arguments (see, e.g., Sattinger, 1979), the second-order condition is satisfied only if $\mu(h)$ is increasing. Further, full employment implies that $\int_{-\infty}^{\infty} L^*(\theta)\mathrm{d}\Phi(\theta) = 1$ in equilibrium, so that

$$D(x) = 1 - \Phi(\frac{(\mu^*)^{-1}(x)}{\sigma}).$$

For the market to clear we require $D(x) = S(x) = 1 - \Phi(x/\sigma_x)$, implying that $\mu(h) = \frac{\sigma_x}{\sigma}h$ in equilibrium. Hence, by Equation (7), we can pin down the derivative of log utility, with $\frac{u'(\frac{h}{\sigma})}{u(\frac{h}{\sigma})} = h$. Integrating from 0 to x gives the difference between the log utilities enjoyed by workers of skill x:

$$\ln(u(x)) = \frac{\sigma}{\sigma_x}\frac{x^2}{2} + \ln(u(0)). \tag{8}$$

Amenities and Firms' Size Using the optimal assignment, the expression for equilibrium utility, and the fact that

$$E_h\left(\exp\left(\frac{\sigma_x h^2}{2\sigma}\right)\middle|\,\theta\right) = \frac{\exp\left(\frac{\theta^2}{2(\frac{\sigma}{\sigma_x}-1)}\right)}{\sqrt{1-\frac{\sigma_x}{\sigma}}},$$

firm's θ profit can be further rewritten as

$$r(\theta) = \max_{L,a} L \left(\frac{aA^2}{4u(0)} \frac{\exp\left(\frac{\theta^2}{2(\frac{\sigma}{\sigma_x}-1)}\right)}{\sqrt{1-\frac{\sigma_x}{\sigma}}} - \frac{1}{c_a}\left(\frac{c_a a}{1+c_a}\right)^{1+c_a} - L^{c_l} \right). \tag{9}$$

It is then straightforward to show that firm's θ optimal choice of amenities and size are

$$a^*(\theta) = (\frac{1}{c_a}+1)\left(\frac{A^2}{4u(0)} \frac{\exp\left(\frac{\theta^2}{2(\frac{\sigma}{\sigma_x}-1)}\right)}{\sqrt{1-\frac{\sigma_x}{\sigma}}} \right)^{\frac{1}{c_a}}, \tag{10}$$

$$L^*(\theta) = \frac{\exp\left[\alpha\left(\frac{\theta^2}{2(\frac{\sigma}{\sigma_x}-1)} + 2\ln A - 0.5\ln(1-\frac{\sigma_x}{\sigma}) - \ln(4u(0))\right)\right]}{(1+c_l)^{\frac{1}{c_l}}}, \tag{11}$$

where $\alpha \equiv \frac{1+c_a}{c_l c_a}$. The fact that the size chosen by a firm depends exponentially on the square of the firm's productivity is of critical importance, as it implies that (by Equation (6)), the distribution of jobs will resemble a normal mixture of normal distributions, and thus will itself be normal.

Job Distribution Revisited Specifically, by Equation (6) the density of the equilibrium job distribution is given by

$$f(h) = \int_{-\infty}^{\infty} \frac{L^*(\theta)}{\sigma_\theta} \phi(h-\theta)\phi(\frac{\theta}{\sigma_\theta})\,\mathrm{d}\theta, \tag{12}$$

where $\phi(\cdot)$ denotes the pdf of the standard normal distribution. After a few rearrangements, the density of the job distribution can be written as:

$$f(h) = \frac{L^*(0)}{\sqrt{1+\sigma_\theta^2\left(1-\frac{\alpha}{\sigma/\sigma_x-1}\right)}} \phi\left(h\sqrt{\frac{1}{1+\frac{\sigma_\theta^2}{1-\frac{\alpha\sigma_\theta^2}{\frac{\sigma}{\sigma_x}-1}}}} \right).^{19} \tag{13}$$

For our guess to be correct we need $f(h) = \frac{1}{\sigma}\phi(\frac{h}{\sigma})$ implying that:

$$1 = \frac{L^*(0)}{\sqrt{1-\frac{\alpha\sigma_\theta^2}{\frac{\sigma}{\sigma_x}-1}}}, \tag{14}$$

$$\sigma^2 = \frac{\sigma_\theta^2}{1-\frac{\alpha\sigma_\theta^2}{\frac{\sigma}{\sigma_x}-1}} + 1. \tag{15}$$

[19] The details of the derivation can be found in Appendix A.1.

Equation (15) can be written as a third-degree polynomial, which can be solved analytically, with a unique solution which meets the restrictions $\sigma > \max\{1, \sigma_x\}$; however, the solution is so complicated, as to be unhelpful. Instead, I will solve this equation for α, which yields:

$$\alpha = t^{-1}(\sigma; \sigma_x, \sigma_\theta) \equiv \left(\frac{\sigma}{\sigma_x} - 1\right)\left(\frac{1}{\sigma_\theta^2} - \frac{1}{\sigma^2 - 1}\right). \tag{16}$$

As $\sigma^2 > 1$ and $\alpha > 0$, it follows by inspection that $\frac{\partial}{\partial \sigma} t^{-1}(\sigma) > 0$; thus, an inverse of $t^{-1}(\cdot)$ is well-defined and also strictly increasing. I will denote this inverse as t; clearly, $t(\alpha; \sigma_x, \sigma_\theta) = \sigma$ and so gives us the standard deviation of the equilibrium distribution of jobs.

The last step is to find the level of utility that ensures that firms want to hire exactly a measure one of workers. Using Equations (11) and (14) it can be shown that

$$\ln(u(0)) = 2\ln(A) - 0.5\ln(1 - \frac{\sigma_x}{\sigma}) - \ln 4 - \frac{1}{\alpha}\left(\ln(\sigma_\theta) - 0.5\ln(\sigma^2 - 1)\right) - \frac{c_a}{1 + c_a}\ln(1 + c_l), \tag{17}$$

which closes the model. Overall, therefore, there exists a single equilibrium in which jobs are normally distributed.[20]

5 Jobs, Sorting and Inequality

In this Section, we first characterise the determinants of the supply of quality jobs σ (Section 5.1), the sorting between high skill workers and high productivity firms (Section 5.2), as well as welfare and wage inequality (Sections 5.3 to 5.5) in our economy.

5.1 Supply of Quality Jobs

It is natural to start our analysis by investigating how the primitives of the model determine σ, as this is the main endogenous variable in our model.

The key observation is that the impact of all parameters on the overall quality of jobs in the economy (σ) can be easily determined by differentiating the function $t^{-1}(\alpha; \sigma_x, \sigma_\theta)$.

[20] It remains an open questions whether there exist equilibria in which jobs are not normally distributed, although we conjecture that there are no such equilibria.

Specifically, we have that

$$\frac{\mathrm{d}}{\mathrm{d}c_a}\sigma = \frac{\partial \alpha}{\partial c_a}\frac{\partial}{\partial \alpha}t(\alpha;\sigma_x,\sigma_\theta) = -\frac{1}{c_l c_a^2}\frac{1}{\frac{\partial}{\partial \sigma}t^{-1}(\sigma;\sigma_x,\sigma_\theta)} < 0 \tag{18}$$

$$\frac{\mathrm{d}}{\mathrm{d}\sigma_x}\sigma = -\underbrace{\frac{\partial}{\partial \sigma_x}t^{-1}(\sigma;\sigma_x,\sigma_\theta)}_{<0}\underbrace{\frac{1}{\frac{\partial}{\partial \sigma}t^{-1}(\sigma;\sigma_x,\sigma_\theta)}}_{>0} > 0, \tag{19}$$

$$\frac{\mathrm{d}}{\mathrm{d}\sigma_\theta}\sigma = -\underbrace{\frac{\partial}{\partial \sigma_\theta}t^{-1}(\sigma;\sigma_x,\sigma_\theta)}_{<0}\underbrace{\frac{1}{\frac{\partial}{\partial \sigma}t^{-1}(\sigma;\sigma_x,\sigma_\theta)}}_{>0} > 0. \tag{20}$$

Hence, increases in the variances of skill (σ_x) and productivity (σ_θ), as well as amenity provision and span-of-control costs all lead to a better overall supply of quality jobs. The intuition for these results is straightforward, although it differs slightly across the different sources of change.

Let us start by focusing on changes in σ_x, c_a and c_l, as the mechanism through which they affect the supply of jobs is similar. Consider the firms' size choices—and the resulting job distribution—in a partial equilibrium that holds constant the utility a worker of a given skill receives. Each of these three types of changes would cause all firms to hire more workers in such a partial equilibrium. However, each of these changes would also cause the larger firms to expand by more. In the case of the span-of-control and amenity provision costs, the reason is that c_a and c_l govern not only the slope (the cost of expanding the workforce or amenity provision) but also the curvature of their respective cost functions; hence, relative to its initial level, a fall in c_l (c_a) decreases the cost of workforce (amenity provision) expansion by more for larger firms. In the case of an increase in the supply of high-skilled workers, the reason is the complementarity in production between high skill workers and high productivity firms.

Of course, given that the *number* of workers is fixed, it cannot be true that all firms expand in general equilibrium. Thus, the utility received by workers of all skills goes up, which increases firms' wage costs and causes a contraction of all firms. This contraction reverses the original expansion for low-productivity firms, but not for high-productivity firms. Overall, therefore, low-productivity firms contract and high-productivity firms expand; hence, the supply of quality jobs improves.

While an increase in the variance of productivity has the same final effect on the supply of quality jobs, the mechanism through which it operates is slightly different. An increase in σ_θ directly increases the number of quality jobs, by increasing the number of high productivity firms in existence (relative to the number of low productivity firms). Clearly, holding utility constant, this increases the demand for high skilled workers—who are mostly demanded by high-productivity firms—and decreases the demand for low-skilled workers, which translates into lower (higher) payoffs for low (high) skilled

workers. As a result, in general equilibrium low productivity firms expand and high productivity firms contract. However, this second-order intensive margin effect on job quality is dominated by the first-order extensive margin increase in the number of high-productivity firms, and thus the overall supply of quality jobs improves.

5.2 Sorting

How do workers sort with firms in our model? The key to answering this question is the observation that the sorting between workers and *jobs* is perfectly positive and assortative, and hence $\text{Corr}(X, H) = 1$. Hence, the correlation between workers and *firms* must be the same as the correlation between firms and jobs. We have already observed that the distribution of jobs within a firm is normal with mean θ and variance 1. It follows immediately that $\text{Corr}(H, \bar{\theta}) = \sqrt{\frac{\text{Var}(\bar{\theta})}{\text{Var}(H)}}$, where $\bar{\theta}$ denotes the distribution of firm productivity weighted by the size of each firm.[21] It is easy to see that

$$\text{Var}(H) = \sigma^2 = 1 + \text{Var}(\bar{\theta}),$$

so that $\text{Corr}(X, \bar{\theta}) \equiv \rho = \sqrt{1 - 1/\sigma^2}$. Indeed, one can also show that

$$\text{Corr}(X^2, \bar{\theta}^2) = \rho^2 = 1 - \frac{1}{\sigma^2}.^{22} \tag{21}$$

Thus, the strength of the positive sorting between workers and firms depends only on the overall supply of quality jobs in the economy; the better the jobs are, the stronger the sorting. In general, quality jobs are on offer in both high- and low-productivity firms: Even the least productive firm has some highly value-added tasks (such as management of the entire company) that need to be performed. If high productivity firms are relatively small (or there are relatively few of them), then there are few quality jobs, but a relatively large share of those few quality jobs is supplied by low-productivity firms. If, however, highly productive firms become very large (or there are very many of them), then most jobs are of high quality, and virtually all quality jobs are supplied by high-productivity firms; however, as low-productivity firms consist primarily of low quality jobs, a sizeable proportion of bad jobs will be still supplied by low-productivity firms, even though those firms are small overall (or there is few of them). As high skill workers fill quality jobs, this implies that the higher the overall quality of jobs, the stronger the sorting between high skilled workers and high-productivity firms.

[21] To see why, note that a simple regression of job on firm type would result in the following relation ship between h and θ: $h = \theta + \epsilon$, with ϵ standard normally distributed.

[22] This can be derived using the fact that $\text{Corr}(X^2, \bar{\theta}^2) = \text{Corr}(H^2, \bar{\theta}^2) = \text{Cov}((\bar{\theta} + T)^2, \bar{\theta}^2)/\sqrt{\text{Var}(H^2)\text{Var}(\bar{\theta})}$, where $T, \bar{\theta}$ are independent and jointly normal.

5.3 Welfare Inequality

The positive relationship between average job quality and sorting is of critical importance for the relationship between overall and between-firm welfare inequality.[23] To see why, first note that the law of total variance implies that

$$\mathrm{Var}_U\left(\ln U\right) = \underbrace{E_{\bar{\theta}}(\mathrm{Var}_U(\ln U|\bar{\theta}))}_{WFUI=\text{within-firm welfare inequality}} + \underbrace{\mathrm{Var}_{\bar{\theta}}E_U(\ln U|\bar{\theta}))}_{BFUI=\text{between-firm welfare inequality}}. \tag{22}$$

Bowsher and Swain (2012) show that for any random variables T and Z we have that $\frac{\mathrm{Var}_T(\mathrm{E}_Z(Z|T))}{\mathrm{Var}_Z(Z)} = \mathrm{Corr}(Z, \mathrm{E}_Z(Z|T))^2$. This has a simple, but profound implication for the relationship between sorting, between-firm and overall welfare inequality:

$$\frac{BFUI}{\mathrm{Var}(\ln U)} = \mathrm{Corr}\left(\ln U, \mathrm{E}_U(\ln U|\theta)\right)^2 = \mathrm{Corr}\left(\ln U(X), \mathrm{E}_X(\ln U(X)|\bar{\theta})\right)^2. \tag{23}$$

Equation (23) implies that the ratio of between-firm and overall welfare inequality is equal to the squared correlation between a function of skill ($\ln U(X)$) and a function of firms' type ($\mathrm{E}_X(\ln U(X)|\theta)$). Hence, as long as these two functions are monotone, this correlation is a measure of the sorting between workers and firms. This indicates that in a wide-range of competitive models, including any generalisation of the present model that would allow for general distributions of skills, tasks or productivities, the share of overall utility inequality explained by its between-firm component will be a function of an appropriately defined measure of the strength of sorting.

In our specific case, log utility is a quadratic function of skill (by Equation (8)) and the average of squared skill is a linear function of squared productivity, with

$$E(X^2|\theta) = \left(\frac{\sigma_x}{\sigma}\right)^2 (1+\theta^2).^{24} \tag{24}$$

Therefore, Equation (23) simplifies to

$$\frac{BFUI}{\mathrm{Var}(\ln U)} = \mathrm{Corr}\left(X^2, \bar{\theta}^2\right)^2 = \rho^4 = \left(1 - \frac{1}{\sigma^2}\right)^2. \tag{25}$$

Thus, the share of overall welfare inequality explained by its between-firm component depends positively on the overall supply of quality jobs in the economy and on nothing else! Furthermore, between-firm, within-firm and overall inequality all depend just on

[23] Of course, wage and welfare inequality coincide in the limit when $c_a \to \infty$.

[24] As discussed above, the distribution of H conditional on θ is normal with mean θ and variance 1. Thus, H^2 conditional on θ has a non-centralised Chi-squared distribution with 1 degree of freedom and centrality parameter θ^2. Thus, $\mathrm{E}_H(H^2|\theta) = 1+\theta^2$, and the relationship between skill and productivity follows from the assignment function $\mu(h)$.

the supply of quality jobs and the distribution of workers skills, with

$$\text{Var}(\ln U) = 0.5\sigma_x^2\sigma^2 \tag{26}$$

$$BFUI = \text{Var}(\ln U)\rho^4 = 0.5\sigma_x^2\sigma^2\left(1 - \frac{1}{\sigma^2}\right)^2 \tag{27}$$

$$WFUI = \text{Var}(\ln U)(1-\rho^4) = 0.5\sigma_x^2(2 - \frac{1}{\sigma^2}).^{25} \tag{28}$$

Hence, overall welfare inequality and its components all increase in the supply of quality jobs; overall inequality changes by the same proportion as σ^2, between-firm inequality changes by a larger proportion than σ^2, and within-firm welfare inequality changes by a lower proportion than σ^2.

The fact that all components of welfare inequality increase in job quality in the supply of high skill workers implies, together with the discussion in Section 5.1, that a change in any of the four primitives increases overall welfare inequality if and only if it also increases the share of welfare inequality that is explained by between-firm welfare inequality. This indicates that the trends in inequality observed in Norway and many other developed countries are exactly what one should expect; increases in overall welfare inequality will typically be driven by changes in between-firm rather than within-firm inequality. And while we will see that this co-movement of overall inequality and its share explained by the between-firm component is not inevitable, once we look at wages rather than welfare or allow for changes in more than one primitive at a time, it nevertheless strikes us as remarkable that the most basic mechanism present in our economy creates such a co-movement.

The following Proposition summarises the results in Sections 5.1 to 5.3.

Proposition 1. The supply of quality jobs σ^2, sorting ρ^2, overall welfare inequality $\text{Var}(\ln U)$, between-firm welfare inequality BFUI, within-firm welfare inequality WFUI and the share of overall welfare inequality explained by its between-firm component $\frac{BFUI}{\text{Var}(\ln U)}$ increase in σ_x, σ_θ and decrease in c_a, c_l.

5.4 Wage Inequality

Focusing on wages makes the link between sorting, and between-firm and overall inequality less straightforward. Let us start by decomposing overall wage inequality using the law of total variance:

$$\text{Var}_W(\ln W) = \underbrace{E_{\bar\theta}(\text{Var}_W(\ln W|\bar\theta))}_{WFWI=\text{within-firm wage inequality}} + \underbrace{\text{Var}_{\bar\theta}E_W(\ln W|\bar\theta))}_{BFWI=\text{between-firm wage inequality}}. \tag{29}$$

[25] $\text{Var}(\ln U)$ is computed from Equation (8), between-firm welfare inequality follows from Equation (25) and within-welfare inequality can be then computed from (22).

Of course, it remains true that

$$\frac{BFWI}{\text{Var}(\ln W)} = \text{Corr}\left(\ln W, \text{E}_W(\ln W|\theta)\right)^2. \tag{30}$$

In contrast to utility, however, the wage a worker receives is a function of not just their own type, but also the amenities provided by the firm they work for, and the amount of effort the worker exerts. Specifically, Equations (5), (7) and (8) imply jointly that the log wage received by a worker of skill x working for firm θ is equal to:

$$\ln(w(x; a^*(\theta))) = \ln(u(x)) + \ln(a^*(\theta)) = \frac{\sigma}{2\sigma_x}x^2 + \frac{\theta^2}{2c_a\left(\frac{\sigma}{\sigma_x} - 1\right)} + B \tag{31}$$

where B is a non-stochastic term.[26] From this and Equation (24) follows immediately that the average wage paid by a firm is linear in the square of the firm's productivity θ:

$$\text{E}(\ln W|\theta) = 0.5\frac{\sigma_x}{\sigma}\theta^2\left(1 + \frac{1}{c_a\left(1 - \frac{\sigma_x}{\sigma}\right)}\right) + B + 0.5\frac{\sigma_x}{\sigma}.$$

Hence, while it remains true that the average wage paid by a firm is strictly increasing in the firm's type, it is not anymore true that more skilled workers *always* receive a higher wage. The reason is that more productive firms pay higher wages to all of their workers. Indeed, a highly skilled worker who works for a low-productivity firm may well receive a lower wage and fewer amenities than a low skilled worker who works for a high productivity firm, because the low-productivity firm will ask its workers to exert very little effort. In terms of between-firm and overall wage inequality, this implies that the correlation between a worker's wage and the average wage paid at their firm is *higher* than the correlation between their utility and the average utility of workers at their firm, with

$$\frac{BFWI}{\text{Var}(\ln W)} = \frac{1}{1 + \frac{\frac{1}{\rho^4} - 1}{\left(1 + \frac{1}{c_a\left(1 - \frac{\sigma_x}{\sigma}\right)}\right)^2}}. \tag{32}$$

Of course, if providing amenities is prohibitively expensive ($c_a \to \infty$) then wage and welfare inequality coincide, and the expression above tends to ρ^4. If amenities are being provided, however, then the relationship between the supply of quality jobs and the share of overall wage inequality explained by its between-firm component becomes more complex. To understand why, suppose that jobs are becoming better because more productive firms grow their workforce. Of course, the resulting increase in sorting still contributes to a more than proportional increase in between-firm wage inequality. However, there is also an additional channel present, which can dominate in some regions of the parameter

[26] Specifically, $B \equiv \ln(u(0)) + \ln(\frac{1}{c_a} + 1) + \frac{c_l}{1+c_a}\left(\ln(\sigma_\theta) - 0.5\ln(\sigma^2 - 1) + \frac{\ln(1+c_l)}{c_l}\right)$.

space. When more productive firms expand, they must start hiring worse workers on average—both in absolute terms, and relative to the quality of workers hired by low-productivity firms. Because wages and amenities are complements in the workers' utility function, and better workers are paid higher wages (on average and within each firm), the smaller the difference in the average skill of workers hired by high- and low-productivity firms, the smaller the difference in amenities provided by those firms. However, as a fall in amenity provision causes a more than proportional decrease in the effort exertion required by the firm, smaller differences in amenities translate then into smaller compensating differentials paid by more productive firms, which decreases the correlation between individual workers' wages and the average wage paid by the firm they work for.

Overall wage inequality can be computed from Equation (31), between-firm wage inequality can be calculated using Equation (32) and within-firm inequality obtains then from (29):

$$\operatorname{Var}(\ln W) = \operatorname{Var}(\ln U)\left(1 + \frac{\rho^4}{c_a\left(1-\frac{\sigma_x}{\sigma}\right)}\left(2 + \frac{1}{c_a\left(1-\frac{\sigma_x}{\sigma}\right)}\right)\right), \tag{33}$$

$$BFWI = \operatorname{Var}(\ln U)\rho^4\left(1 + \frac{1}{c_a\left(1-\frac{\sigma_x}{\sigma}\right)}\right)^2, \tag{34}$$

$$WFWI = WFUI. \tag{35}$$

Within-firm wage inequality coincides with within-firm welfare inequality because amenity provision is firm specific rather than worker-firm specific. Thus, it remains true that an increase in the supply of quality jobs increases within-firm wage inequality. However, its effect on overall inequality is ambiguous in general.

Proposition 2. (i) Overall wage inequality $\operatorname{Var}(\ln W)$, between-firm wage inequality BFWI, and the share of overall wage inequality explained by its between-firm component $\frac{BFWI}{\operatorname{Var}(\ln W)}$ increase in σ_x and decrease in c_a. (ii) Further, if $\sigma_x < 1$ ($\sigma_x < 0.99$) then BFWI and $\operatorname{Var}(\ln W)$ ($\frac{BFWI}{\operatorname{Var}(\ln W}$) increase in σ_θ and decrease in c_l. If, however, $\sigma_x > 1$ then BFWI, $\operatorname{Var}(\ln W)$ and $\frac{BFWI}{\operatorname{Var}(\ln W)}$ are all non-monotone in σ_θ and c_l.

Hence, changes in the distribution of workers' skill and the cost of amenity provision affect between-firm, within-firm and overall welfare and wage inequality in the same direction. Changes in firms' type distribution and span-of-control cost, can—in theory—affect wage and welfare inequality differently in certain regions of the parameter space. Similarly, in theory it is possible that, say, a decrease in the span-of-control cost may increase wage inequality but decrease the share of wage inequality that is explained by its between-firm component.

Critically, however, the condition $\sigma_x < 0.99$ is extremely likely to be satisfied in practice: Virtually all calibrations of the model to real-world data would estimate σ_x to

be lower than 0.99. To see why, recall that $\sigma > 1$ and note that Equations (28) and (35) imply that

$$\sigma_x < \frac{\sqrt{2}\sqrt{WFWI}}{\sqrt{2 - \frac{1}{\sigma^2}}} < \sqrt{2}\sqrt{WFWI}.$$

For developed countries, the empirical estimates of within-firm wage inequality are well below 0.7, implying that σ_x would indeed take values lower than 0.99 if calibrated to match this moment.[27] This indicates that, in practice, any change in a single primitive of our model changes overall wage inequality and its share explained by the between-wage component in the same direction.

5.5 Differential Impact of Changes in Primitives

The preceding discussion implies that, in practice, overall wage inequality and its share explained by the between-firm component may *not* co-move only if two or more changes in primitives are counteracting each other.

As an example, suppose that the variance of skill σ_x and the cost of amenity provision have both decreased at the same time, so that these two changes affect all outcomes in opposite directions. Suppose further, that their impacts on overall welfare inequality exactly offset each other. In that case, the ratio of welfare inequality explained by its between-firm component would increase, because for a decrease in σ_x and c_a to have an effect on welfare inequality of the same magnitude, the fall in c_a must have had a stronger effect on the equilibrium supply of quality jobs and thus also on sorting. One can similarly, if much more laboriously, show that overall wage inequality would also increase in such a scenario, as would its share explained by the between-firm component.

Proposition 3. For a given set of parameters $\sigma_x, \sigma_\theta, c_a, c_l$ define two scalars y_{σ_x}, y_{c_l} such that

$$\frac{\mathrm{d}}{\mathrm{d}\sigma_x}\mathrm{Var}(\ln U) = y_{c_a}\frac{\mathrm{d}}{\mathrm{d}c_a}\mathrm{Var}(\ln U) = y_{c_l}\frac{\mathrm{d}}{\mathrm{d}c_l}(\ln U).^{28}$$

Then the relative (local) impact of changes in σ_x, c_a, c_l on $\frac{BFUI}{\mathrm{Var}(\ln U)}$, $\frac{BFUI}{\mathrm{Var}(\ln U)}$ and $\mathrm{Var}(\ln W)$ is as follows:

$$\text{(i)} \quad y_{c_a}\frac{\mathrm{d}}{\mathrm{d}c_a}\frac{BFUI}{\mathrm{Var}(\ln U)} = y_{c_l}\frac{\mathrm{d}}{\mathrm{d}c_l}\frac{BFUI}{\mathrm{Var}(\ln U)} > \frac{\mathrm{d}}{\mathrm{d}\sigma_x}\frac{BFUI}{\mathrm{Var}(\ln U)}$$

$$\text{(ii)} \quad y_{c_a}\frac{\mathrm{d}}{\mathrm{d}c_a}\frac{BFWI}{\mathrm{Var}(\ln W)} > \max\{\frac{\mathrm{d}}{\mathrm{d}\sigma_x}\frac{BFWI}{\mathrm{Var}(\ln W)}, y_{c_l}\frac{\mathrm{d}}{\mathrm{d}c_l}\frac{BFWI}{\mathrm{Var}(\ln W)}\}$$

$$\text{(iii)} \quad y_{c_a}\frac{\mathrm{d}}{\mathrm{d}c_a}\mathrm{Var}(\ln W) > \max\{\frac{\mathrm{d}}{\mathrm{d}\sigma_x}\mathrm{Var}(\ln W), y_{c_l}\frac{\mathrm{d}}{\mathrm{d}c_l}\mathrm{Var}(\ln W)\}.$$

[27] Tomaskovic-Devey et al. (2020) report overall and between-firm wage inequality for 14 developed countries (Table S3.3), with the largest value of overall wage variance being 0.585 (Canada in 2007). In our data, within-firm wage inequality in Norway peaks in 2008 at 0.144.

[28] We omit σ_θ as its impact on the outcomes considered in this Proposition must trivially be the same as the impact of c_l.

Furthermore, if $\text{Var}(\ln U) < 0.5$ then

$$\text{(iv)} \qquad y_{c_l} \frac{\mathrm{d}}{\mathrm{d}c_l} \frac{BFWI}{\text{Var}(\ln W)} > \frac{\mathrm{d}}{\mathrm{d}\sigma_x} \frac{BFWI}{\text{Var}(\ln W)},$$

$$\text{(v)} \qquad y_{c_l} \frac{\mathrm{d}}{\mathrm{d}c_l} \text{Var}(\ln W) > \frac{\mathrm{d}}{\mathrm{d}\sigma_x} \text{Var}(\ln W).$$

Proposition 3 normalises the size of changes in the primitives by requiring that each affects overall welfare inequality equally, and then compares their impact on selected other outcomes. The impact on the ratio of between-firm to overall welfare inequality is the most straightforward one. Because c_l and c_a affect overall welfare inequality only through job quality/sorting, and σ_x affects it directly as well, if c_l/c_a have the same impact on overall welfare inequality as σ_x then their impact on sorting must be greater than that of σ_x.

The relative impact that different changes in primitives have on overall wage inequality and on its share that is explained by the between-firm component is less obvious. It should be fairly clear why changes in the cost of amenity provision have the strongest impact, as c_a has both a strong effect on sorting and impacts $\text{Var}(\ln W)$ ($\frac{BFWI}{\text{Var}(\ln W)}$) directly. In general, it is ambiguous whether changes in the skill distribution or changes in the span-of-control cost have greater impact on overall wage inequality and on its share that is explained by the between-firm component; on the one hand, the increase in σ_x has a direct positive effect on $\text{Var}(\ln W)$ ($\frac{BFWI}{\text{Var}(\ln W)}$) through $1/(1-\sigma_x/\sigma)$, on the other hand, the increase in σ_x has a much weaker impact on sorting/supply of quality jobs (compared to a fall in c_l that has equal impact on overall welfare inequality). However, one can show quite easily that if $\text{Var}(\ln U) < 0.5$ then the impact of a fall in the span-of-control cost must be stronger; to see why, let us rewrite Equation (33):

$$\text{Var}(\ln W) = \text{Var}(\ln U) \left(1 + \frac{1 - \frac{1-\sqrt{2\text{Var}(\ln U)}}{\sigma^2 - \sqrt{2\text{Var}(\ln U)}}}{c_a} \right) \left(2(1 - \frac{1}{\sigma^2}) + \frac{1 - \frac{1-\sqrt{2\text{Var}(\ln U)}}{\sigma^2 - \sqrt{2\text{Var}(\ln U)}}}{c_a} \right).$$

Partialling out the impact of changes in $\text{Var}(\ln U)$, $\text{Var}(\ln W)$ increases in σ if $\text{Var}(\ln U) < 0.5$; and because a fall in c_l increases σ by more, its impact on overall wage variance must be greater as well. The ratio of between-firm to overall wage inequality can be rewritten analogously.

Critically, again, overall wage inequality is smaller than 0.5 in the vast majority of the 14 developed countries studied in Tomaskovic-Devey et al. (2020), with the only exceptions being the US, Israel and Canada. Most importantly, our estimate of overall wage inequality in Norway peaks at 0.217 in 2007. Therefore, we expect that in our calibration changes in the span-of control cost will have a stronger impact on overall wage inequality (relative to their impact on overall welfare inequality) than changes in

the variance of skill.

6 Calibration Exercise

We first describe the data used in the calibration exercise (Section 6.1), then identify the model (Section 6.2), calibrate it (Section 6.3), and finally simulate counterfactual scenarios with an eye on establishing the main drivers of the observed changes the composition of inequality (Section 6.4).

6.1 Data

The main data used for calibration are matched employer-employee data in Norway, drawn from administrative registers maintained by Statistics Norway (Statistics Norway, 2020a,c,d). The data contain the universe of workers and firms recorded in the registers. Each worker is assigned a unique person identifier, which allows us to track workers over time. Similarly, each firm is identified by a unique firm identifier that does not change over time. When a worker matches to a firm in a given year, an entry for the worker-firm pair is created for that year. For each worker-firm-year combination, we have details regarding the worker's date of birth, the firm's industry sector, and crucially, the worker's wage at the firm, along with other attributes. Since we observe all workers within a firm, as well as all firms in the economy, we can reliably compute firm sizes and determine wage distributions for each year.

Earnings records for workers pertain to remuneration, which include fixed salary, holiday pay, sickness and maternity benefits, and other cash benefits paid to the worker deemed as remuneration by the Norwegian Tax Administration (Norwegian Tax Administration, 2018). All earnings are adjusted to 2013 values using the Consumer Price Index published by Statistics Norway (Statistics Norway, 2023).[29]

Employment and wage data prior to 2015 are available on an annual basis. Statistics Norway links the employee register with individual tax files to provide wages and labour information. Since 2015, employment statistics are reported on a monthly basis. Employers submit salary and employment information directly to the employee register, the tax office, and Statistics Norway (Statistics Norway, 2020e). To ensure wages are measured at the same frequency, we aggregate monthly earnings within each year, starting from 2015, to generate annual earnings. Due to this change in the reporting scheme in 2015 (Statistics Norway, 2014), we divide our sample into two parts, 1995 to 2014, and 2015 to 2019.[30]

[29] We chose 2013 as the reference year following Song et al. (2018).

[30] In our baseline sample, we do not observe structural breaks in key statistics such as the average number of workers, number of firms, and average wage, from 2014 to 2015. Nonetheless, as a precautionary measure, we divide the sample period into pre-2014 and post-2015 segments.

What constitutes a firm in our sample? We consider businesses listed in the Central Register of Establishments and Enterprises. A business is an entity that primarily operates within a specific industrial classification (Statistics Norway, 2018). It is the level at which employers submit work and pay information for tax purposes (see, for instance, a walk-through at Norwegian Tax Administration (2023a)). A business is a subdivision of the legal unit, an enterprise. As businesses and enterprises can be consistently identified from 1995 in our data (e.g., Statistics Norway, 2020b), we start our sample period from 1995. We conclude our sample period in 2019 since it is the last year unaffected by the COVID-19 pandemic.

Our analytic sample consists of individuals aged between 20 and 60, following Song et al. (2018). To identify which workers are weakly attached to the labour force or to a certain job, we introduce the concept of Basic Amount ("Grunnbeløp") (e.g., Norwegian Tax Administration, 2023b). The Basic Amount is updated each year by the Parliament. It is the legal amount at which the national insurance scheme becomes applicable.[31] We remove individuals who earn less than two times the Basic Amount in that year, following common practice (e.g., Markussen and Røed, 2019; Hoen et al., 2022).[32] Regarding firms, we remove firms with less than 5 employees because it is not meaningful to calculate within-firm variance for extremely small firms. We exclude firms and workers in the following sectors: public administration, defence, and social security schemes subject to public administration; education; paid work in private households; and international organisations and bodies. Wages in these sectors are often flat according to preset wage schedules. These restrictions result in a sample of 1.3 million workers and 54,000 firms per year on average. Throughout the sample period, the average annual growth rate in the number of workers is 1.5%, and that for the number of firms is 2.5%.

6.2 Identification

Let us denote the vector of exogenous parameters by $\Sigma \equiv (\sigma_x, \sigma_\theta, c_a, c_l, A)$. We will identify Σ using five moments of the wage distribution:

1. within-firm wage inequality $(\mathrm{E}_{\bar{\theta}}\left(\mathrm{Var}W(\ln W|\bar{\theta})\right))$,
2. the variance of the within-firm wage variances of wages $(\mathrm{Var}_{\bar{\theta}}\left(\mathrm{Var}W(\ln W|\bar{\theta})\right))$,
3. the unweighted within-firm wage inequality $(\mathrm{E}_{\theta}\left(\mathrm{Var}W(\ln W|\theta)\right))$,
4. variance of log wages $(\mathrm{Var}_W(\ln(W)))$,

[31] The Basic Amount is similar to CPI in concept. It is is updated according to wage growth, while CPI is calculated on the price of consumer goods and services. The two do not always align. From 1995 to 2019, the Basic Amount has grown by 155%, whereas CPI has only grown by 65%.

[32] On average, Norwegians earn approximately six times the Basic Amount.

5. and the expected (log) wage in the economy ($\mathrm{E}_W(\ln(W))$).

Of course, with just 5 parameters the model is very much over-identified. We chose these 5 particular moments because amenity and effort exertion (both of which are not fully observed in the data) do not affect the variance of wages within any firm; and thus, the first three moments depend on the cost of amenity provision parameter c_a only through its impact on $\alpha \equiv (1+1/c_a)/c_l$ which itself affects only the equilibrium supply of quality jobs σ^2.[33] The first three moments can be thus used to identify σ_x, σ_θ and α; the cost of amenity provision is then identified from the overall wage variance, which it affects directly—together with α, this identifies c_l; finally, the TFP parameter A is identified from the average (log) wage in the economy.

The variance of the within-firm wage variances of wages and the unweighted within-firm wage inequality are equal to

$$\mathrm{Var}_{\bar{\theta}}\left(\mathrm{Var}W(\ln W|\bar{\theta})\right) = \left(\sqrt{2}\sigma_x^2\left(1-\frac{1}{\sigma^2}\right)\right)^2, \tag{36}$$

$$\mathrm{E}_{\theta}\left(\mathrm{Var}W(\ln W|\theta)\right) = 0.5\left(\frac{\sigma_x}{\sigma}\right)^2(1+2\sigma_\theta^2).^{34} \tag{37}$$

Equations (36) and (35) jointly identify the equilibrium supply of quality jobs, with

$$\sigma^2 = \frac{\sqrt{2}WFWI - 0.5\sqrt{\mathrm{Var}_{\bar{\theta}}\left(\mathrm{Var}_X(\ln(W)|\bar{\theta})\right)}}{\sqrt{2}WFWI - \sqrt{\mathrm{Var}_{\bar{\theta}}\left(\mathrm{Var}_X(\ln(W)|\bar{\theta})\right)}}. \tag{38}$$

Observe that the right-hand side of this equation increases in the variance of within firm wage dispersion; hence, to rationalise a large variance of within-firm wage dispersion in our economy, the equilibrium supply of quality jobs must be large as well. This is because σ is small only if most jobs are offered by firms with low θ^2, which are all similar to each other. Therefore, if σ is small and we weight the population by firm size, then the variance of within firm wage variance does not differ much across firms.

Once we know σ^2, the variance of skill (high productivity firms) can be identified from

[33] A sensible alternative to using one of these 3 moments would be to estimate an AKM-like fixed effects regression, and identify σ from the correlation between the worker and firm fixed effects. We have decided against it, as there are many modelling choices that go into an AKM regression and such moment would be significantly more 'processed'. However, in Section 6.3 we will compare the AKM covariance estimates found in the literature to those resulting from our calibration as an non-targeted moment, to validate the quality of calibration.

[34] To see how these expressions are derived, recall that the jobs offered by firm θ have a normal distribution with mean θ and variance 1 (see footnote following Equation (24)). It follows from the properties of non-centralised Chi-squared distribution that the variance of H^2 offered by firm θ is equal to $\mathrm{Var}_H(H^2|\theta) = 2+4\theta^2$. Hence, from the facts that $X^2 = \left(\frac{\sigma_x}{\sigma}H\right)^2$ and $\ln W = \frac{\sigma}{2\sigma_x}X^2 + \frac{\theta^2}{2c_a\left(\frac{\sigma}{\sigma_x}-1\right)} + B$ in equilibrium it follows that $\mathrm{Var}_W(\ln(W)|\theta) = 0.5(\frac{\sigma_x}{\sigma})^2(1+2\theta^2)$. Finally, $\mathrm{E}_\theta(\theta^2) = \sigma_\theta^2$ and $\mathrm{Var}_{\bar{\theta}}(\theta^2) = 2(\sigma^2-1)^2$ from which Equations (36) and (37) follow.

Equations (35) (and (37)):

$$\sigma_x^2 = \frac{2WFWI}{2 - \frac{1}{\sigma^2}}, \tag{39}$$

$$\sigma_\theta^2 = 0.5 \left(\frac{\mathrm{E}_\theta \left(\mathrm{Var} W(\ln W|\theta) \right)}{WFWI} (2\sigma^2 - 1) - 1 \right). \tag{40}$$

Within-firm inequality depends on σ and σ_x only, and thus, for a given σ, the variance of skill must be high for the model to rationalise a larger degree of within-firm inequality. The ratio of within firm wage inequality *unweighted* by firm size to within firm wage inequality *weighted* by firm size depends only on σ and σ_θ. Weighted within-firm inequality can be large compared to unweighted inequality, only if there are few high productivity firms, but those firms are large.

With σ, σ_x and σ_θ identified, α follows immediately from the equilibrium condition (Equation (15)). The cost of amenity provision c_a can then be easily identified from overall wage inequality, which it affects directly, with

$$c_a = \frac{\left(\frac{\sqrt{2\mathrm{Var}_W(\ln(W)) - WFWI}}{\sigma_x \sigma (1 - \frac{1}{\sigma^2})} - 1 \right)^{-1}}{1 - \frac{\sigma_x}{\sigma}}. \tag{41}$$

Essentially, the value of c_a in our calibration is determined by the gap between the actual value of overall wage inequality and the value that the model produces without the presence of amenities. The larger this gap is, the lower the cost of providing amenities must be.

The span-of-control cost is then given by α and c_a:

$$c_l = \frac{1 + \frac{1}{c_a}}{\left(\frac{\sigma}{\sigma_x} - 1 \right) \left(\frac{1}{\sigma_\theta^2} - \frac{1}{\sigma^2 - 1} \right)}. \tag{42}$$

The span of control cost is, effectively, calibrated as a remainder: It is the value of c_l that ensures that the calibrated σ equalibrates the model given the estimates of all the other parmaters.

Finally, note that—by Equation (31)—the average (log) wage in the economy equals

$$E_W(\ln(W)) = 0.5\sigma_x \left(\sigma \left(1 + \frac{1}{c_a(1 - \frac{\sigma_x}{\sigma})} \right) - \frac{1}{c_a(\sigma - \sigma_x)} \right) + B,$$

which can be used to identify the TFP parameter A using Equation (17) and the definition of B.

6.3 Calibration and Simulation

The targeted moments are calculated using the Norwegian administrative data for the years 1995 to 2019, as discussed in Section 6.1. Instead of using the raw moments, we produce 1500 samples for each year using a bootstrap-like resampling procedure, and then calculate the moments and calibrate the model separately for each of these samples; the reported central estimates of our moments and parameters simply represent the average value over the 1500 samples, and the confidence intervals correspond to the 2.5th and 97.5th percentiles of the 1500 estimates. Furthermore, in the case of the variance of within firm variances we de-noise the raw estimate. The raw estimate overestimates the *true* variation in the variances of within-firm wage distributions because it is also affected by the variance of the within-firm variance estimator; as some firms in our sample are small, this bias can be substantial.[35] We estimate the size of this bias and subtract this estimate from the raw data moment of the variance of within-firm wage variances. The details of both our resampling procedure and the adjustment applied to the variance of within-firm variances are described in Online Appendix C.

The targeted moments are reported in Figure 2. Three observations are in order. First, from 1995 to 2014 overall wage inequality and both of its components have increased, but the increases in BFWI were larger in both absolute and relative terms. Second, since 2015 both overall wage variance and between-firm wage inequality have been decreasing, the latter quite sharply; however, within-firm wage inequality has continued to rise. Third, changes in the variance of within-firm wage variance have mimicked closely the changes in WFWI between 1995 and 2014; since 2014, however, the variance of within-firm wage variance has remained constant.

The calibrated parameters are graphed in Figure 3.[36] All targeted moments are matched exactly within each of the 1500 year-specific samples, and thus also on average. The calibration reveals that the variance of productivity and the cost of providing amenities remained relatively stable between 1995 and 2014; the change from 1995 to 2014 is insignificant in the case of σ_θ and only marginally significant in the case of c_a. After 2015, the variance of productivity seems to have fallen slightly, and the cost of amenity provision has slightly increased.

The changes in the remaining parameters are large and statistically significant. The variance of skill has reached its peak in 2007, but has remained significantly higher than its 1995 level in 2014; this increase in σ_x has continued after 2015. The span-of-control

[35] To understand the source of this bias better, suppose that all firms were of the same type, and thus with an infinite population would hire the same distribution of workers. As our sample is finite, the firms would nevertheless differ in the type of workers they end up hiring in the data, and thus the observed variance of wages would differ across firms, even though the 'true' variance of variances is 0.

[36] The change in the estimates from year 1995 (for estimates from 1996 to 2014) and 2015 (for estimates from 2016 to 2019), together with the corresponding confidence intervals, is depicted in Figure OA.2 in the Online Appendix.

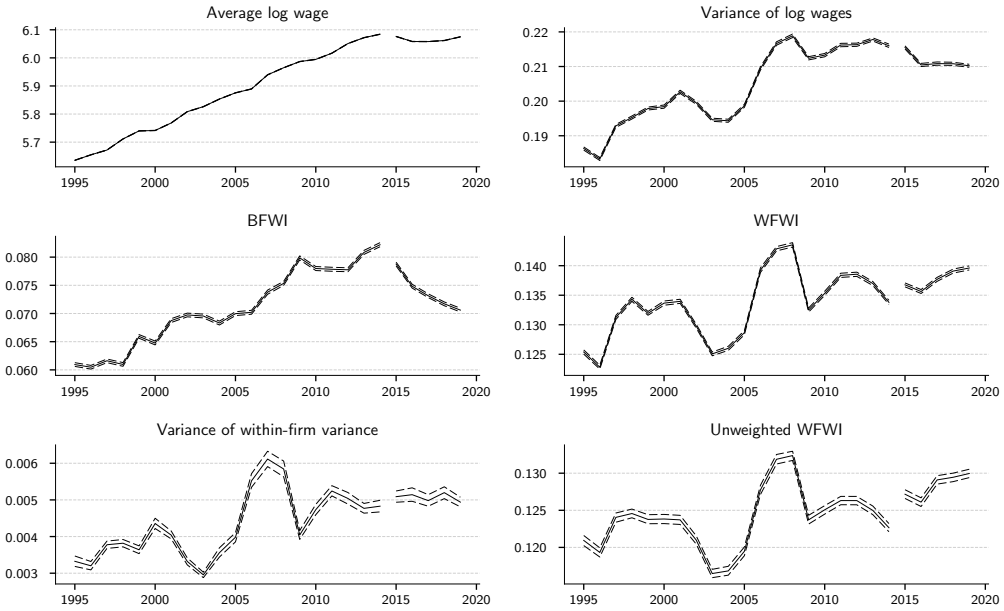


Figure 2: Moments used in calibration

Notes: Plot of the data moments and confidence used for the calibration. Top panel: (1) Mean log wage; (2) Variance of log wages. Middle panel: (1) The across firms, firm-size weighted variance of (within firm) mean log wages; (2) The across firms, firm-size weighted mean of the variance of log wages. Bottom panel: (1) The across firm, firm-size weighted variance of the (within-firm) variance of log wages; (2) The across firms, size-unweighted mean of the variance of log wages. Both the moments and the confidence intervals were obtained through the resampling procedure discussed in detail in Online Appendix C.

cost plummeted between 1996 and 1998, but has remained relatively stable since. Finally, the TFP of the Norwegian economy has been steadily climbing between 1995 and 2014, but has levelled off since 2015.

The changes in $\sigma_\theta, \sigma_x, c_a$ and c_l have caused an increase in the equilibrium supply of quality jobs σ between 1995 and 2014 (with a peak in 2007).[37] Since 2015, however, the supply of quality jobs has actually decreased in the Norwegian economy.

We use our calibrated parameter values to simulate a number of moments that are not directly observed in the data. Most importantly, we use the calibration to decompose between-firm wage inequality into the variance of the compensating differential for effort, the variance of average within-firm worker utility, and the covariance between compensating differentials and utility (Figure 4). This exercise is analogous to decomposing BFWI into the variance of firm fixed effects (FFE), firm-level average worker fixed effects

[37]The fact that a number of our calibrated parameters peaks in 2007 is an artefact of the fact that our calibration interprets any observed changes in inequality as caused by a change in the primitives; thus, the spike in within-firm variance just before the financial crisis, which was likely caused by rents in reality, our model interprets as being caused by changes in σ_θ, σ_x and c_a. This is partly why, when discussing the results, we focus on the 1995-2014 comparison, rather than the evolution of the parameters throughout the entire period.

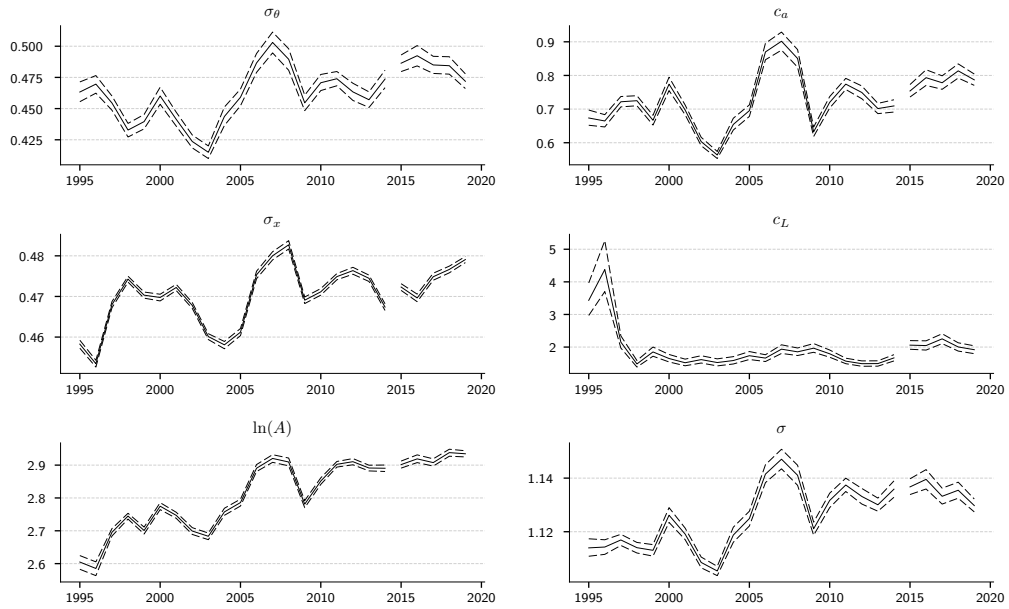


Figure 3: Calibrated parameters

Notes: The calibrated parameters from 1995 to 2019. Top panel: (1) the variance of productivity σ_θ, (2) cost of amenity provision c_a. Middle panel: (1) variance of skill σ_x, (2) span-of-control cost c_l. Bottom panel: (1) log of total factor productivity $\ln A$, (2) equilibrium supply of quality jobs σ. The dashed lines depict the 95% confidence intervals.

(WFE), and twice the covariance between these two (see Equation (5) and Table IV in Song et al., 2018). The reason is that, as a direct consequence of Equation (31), in our economy a worker's utility is exactly equal to that worker's fixed effect in an AKM regression; and the compensating differential paid by a firm is exactly equal to both the amount of amenities provided by the firm and the firm's fixed effect in an AKM regression. See Online Appendix B for a discussion on how to perform the AKM variance decomposition in our model.

Our simulation indicates that between 1995 and 2014 all three components of between-firm wage inequality have increased. In relative terms, the changes in all components are very similar in magnitude. In absolute terms, the increases in the variance of firm fixed effects and twice the covariance between worker and firm fixed effects are of similar magnitude, and dwarf the changes in the variance of average firm-level worker fixed effects. The (non-targeted) estimates of the variance of firm fixed effects and the covariance between worker and firm fixed effects can be used to validate our calibration, by comparing them to empirical estimates from the literature. Specifically, Bonhomme et al. (2023) report the results of an AKM fixed effects regression in Norway for the years 2009–2014 (Figure F.10). After correcting for the limited mobility bias, their estimates (the FE-HE, FE-HO and CRE estimators) of the ratio of twice the covariance between firm- and

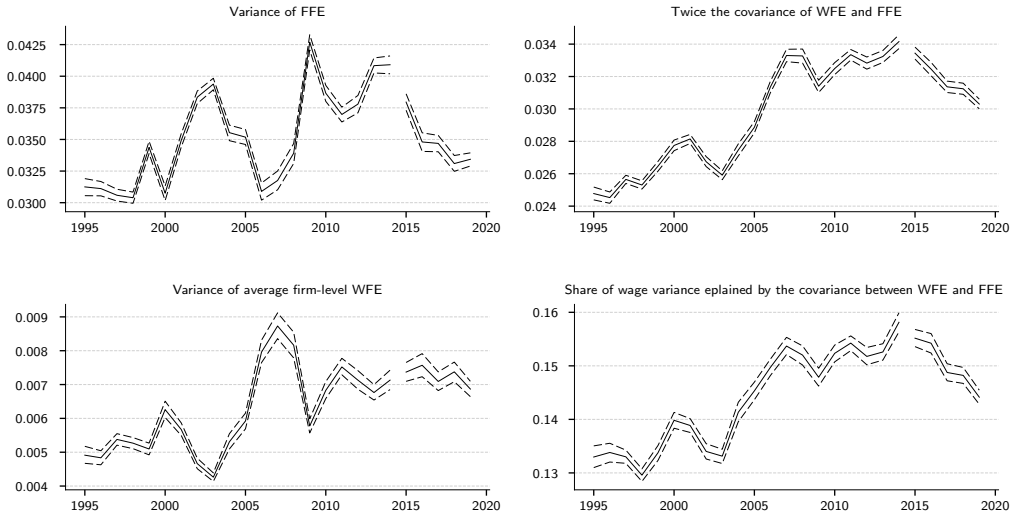


Figure 4: Decomposition of between-firm wage inequality

Notes: Components of between-firm wage inequality from 1995 to 2019. Top panel: (1) variance of firm fixed effects, (2) twice the covariance of worker fixed effects and firm fixed effects. Bottom panel: (1) variance of average firm-level worker fixed effects, (2) the ratio of twice the covariance between worker fixed effects and firm fixed effects to overall wage variance. The dashed lines depict the 95% confidence intervals.

worker-fixed-effects and overall annual wage variance ranges between 13%–17%. In our calibration, the central simulation of the same moment for the years 2009–2014 ranges between 14.8% and 15.8%. Since this moment was not targeted by us, we find it reassuring that it is in a similar range to that reported in the literature. More worryingly, however, Bonhomme et al. (2023) report that the variance of firm fixed effects explained between 9% and 10% of overall annual wage variance in Norway between 2009 and 2014, whereas in our calibrated model it explains between 17% and 20%, which is closer to the uncorrected estimates from Bonhomme et al. (2023).[38]

Of the moments depicted in Figure 4, the between-firm differences in utility/WFE are particularly interesting, as—within our model—they correspond exactly to between-firm utility inequality. This allows us then to compute overall utility inequality—by adding between-firm utility inequality to within-firm wage inequality—and decompose it into its within- and between-firm components (Figure 5). Interestingly, between-firm utility inequality was responsible for less than 9% of overall utility inequality throughout our sample.[39] The changes in utility inequality from 1994 to 2014 were broadly in the same direction as the changes in wage inequality, with between-firm welfare inequality

[38] Admittedly, Bonhomme et al. (2023) use a different sample of firms and individuals. For example, they use full-time equivalent wages to measure labour market attachment.

[39] This is, perhaps, not so surprising, given our assumption that all differences in wages paid to workers of identical skill are compensating differentials. However, given the findings in Sorkin (2018), this assumption may not be drastically off.

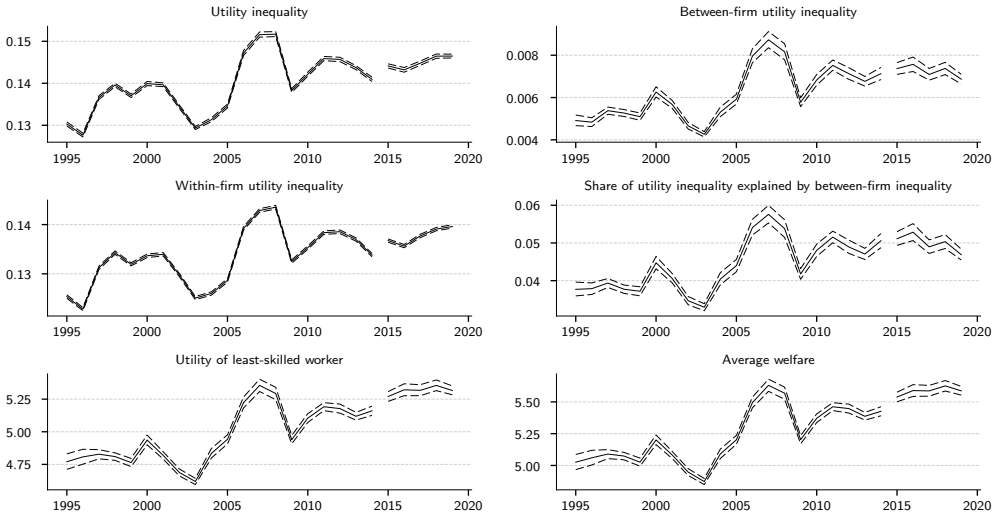


Figure 5: Decomposition of utility inequality

Notes: Components of utility inequality from 1995 to 2019. Top panel: (1) utility inequality, (2) between-firm utility inequality. Middle panel: (1) within-firm utility inequality, (2) share of utility inequality explained by between-firm inequality. Bottom panel: (1) utility of least-skilled worker, (2) average welfare. The dashed lines depict the 95% confidence intervals.

responsible for 65% of the overall increase in welfare inequality over this time-frame. The magnitudes, however, are quite different. As we report in Table 1 below, overall wage inequality increased by 0.03, whereas overall utility inequality increased by only 0.011; thus, only a little more than a third of the overall increase in wage inequality reflects an actual increase in welfare inequality. As we have learnt from Figure 4 and the corresponding discussion, the remaining two thirds of the increase in wage inequality reflect (in equal measure) an increase in the variance in the compensating differential for effort exertion and twice the covariance between workers' utility and the compensating differentials.

Strikingly, between 2015 and 2019 the trends in overall wage and welfare inequality have diverged: The decrease in overall wage inequality masks a continued increase in overall welfare inequality. In the broadest of sense, the reason for this divergence of trends is the fact that between-firm welfare inequality is much smaller than between-firm wage inequality in our calibrated model, whereas within-firm wage and welfare inequalities are always the same. Thus, the continued increase in within-firm welfare inequality dominates the fall in between-firm welfare inequality, but the fall in between-firm wage inequality dominates the increase in within-firm wage inequality. Of course, this raises a deeper question: Why are within- and between-firm inequality changing in opposite direction since 2015; we will attempt to answer this question in Section 6.4 below.

The last two panels Figure 5 depict the changes in (a) the utility of the worst-off

worker and (b) the average welfare in the economy, both of which were much higher in 2014 than in 1995, and have remained stable after 2015. Thus, the technological changes responsible for the observed increases in inequality have, in fact, made everyone better off.

6.4 Counterfactual Analysis

In this section, we will decompose the changes in wage and welfare inequality into the share explained by the change in each of the four parameters. However, because the impact that a change in any parameter has on the outcomes depends on the other parameters, the order of change matters for the result: The impact of a change in, say, σ_x may be drastically different if all the other parameters are at their 1995 levels than if the other parameters are at their 2014 levels. We deal with this issue by simulating the effect of changes in parameters for each of the 24 permutations of the order in which the vector $(\sigma_\theta, \sigma_x, c_a, c_l)$ may have changed between any two points in time, and then averaging over these permutations.

Table 1 reports how the changes from 1995 to 2014 and from 2015 to 2019 in each of the four main primitives of our models affected 7 outcomes of interest: overall wage inequality, within-firm wage inequality, between-firm wage inequality, the ratio of between-firm to overall wage inequality, overall welfare inequality, between-firm welfare inequality, and the ratio of between-firm to overall welfare inequality. Starting with the 1995 to 2014 adjustments, the bulk of the overall increase in overall wage and welfare inequality can be attributed to changes in σ_x and the span-of-control cost c_l; this is unsurprising given that the changes in the other two parameters were, at best, marginally significant. Having said that, note that changes in the distribution of high productivity firms (σ_θ) nevertheless explain a non-trivial share of the change in between-firm wage and welfare inequality (29% in the case of BFWI and 24% in the case of BFUI). Changes to the cost of amenity provision had a significantly *negative* effect on between-firm wage inequality, which almost exactly offset the impact of changes in σ_θ; however, changes in c_a have little impact on any outcomes that do not depend on c_a directly, that is on within-firm inequality and any type of welfare inequality.

The fall in the span-of-control cost c_l increased welfare inequality by 0.004, whereas the increase in the variance of skill/skill-biased technological change parameter σ_x raised welfare inequality by 0.06; overall, the change in σ_x explained the majority of the change in welfare inequality. However, as we have learnt in Section 5.5, conditioning on their impact on welfare inequality, changes in the span-of-control cost have a greater impact on wage inequality than changes in σ_x, provided that overall welfare inequality is as low as in Norway. Thus, it is the change in the span-of-control cost that has caused the bulk of the increase in overall wage inequality (0.018 or 62%), with the increase in σ_x contributing

		σ_θ	σ_x	c_a	c_l	Overall
VarW	1995-2014	0.007	0.010	-0.006	0.018	0.030
	conf. interval	[-0.001,0.015]	[0.009,0.012]	[-0.011,-0.001]	[0.015,0.022]	[0.029,0.030]
	share explained	24.1	34.6	-20.8	62.2	100.0
	conf. interval	[-1.7,49.0]	[29.9,39.4]	[-37.3,-3.7]	[50.9,72.9]	[100.0,100.0]
	2015-2019	-0.010	0.008	-0.005	0.002	-0.005
	conf. interval	[-0.017,-0.004]	[0.007,0.009]	[-0.009,-0.001]	[-0.001,0.005]	[-0.006,-0.005]
	share explained	194.3	-150.1	99.5	-43.7	100.0
	conf. interval	[73.9,312.3]	[-185.4,-117.8]	[23.3,181.4]	[-99.8,12.3]	[100.0,100.0]
WFWI	1995-2014	0.001	0.005	-0.000	0.002	0.008
	conf. interval	[-0.000,0.002]	[0.005,0.006]	[-0.000,-0.000]	[0.002,0.003]	[0.008,0.009]
	share explained	10.9	62.2	-1.3	28.3	100.0
	conf. interval	[-0.8,22.1]	[54.0,70.4]	[-2.3,-0.2]	[22.8,33.7]	[100.0,100.0]
	2015-2019	-0.001	0.004	-0.000	0.000	0.003
	conf. interval	[-0.002,-0.001]	[0.003,0.005]	[-0.000,-0.000]	[-0.000,0.001]	[0.002,0.003]
	share explained	-49.8	142.4	-3.8	11.1	100.0
	conf. interval	[-85.4,-17.0]	[120.1,167.7]	[-6.5,-1.0]	[-3.1,25.7]	[100.0,100.0]
BFWI	1995-2014	0.006	0.005	-0.006	0.016	0.021
	conf. interval	[-0.000,0.013]	[0.004,0.006]	[-0.011,-0.001]	[0.013,0.019]	[0.021,0.022]
	share explained	29.3	23.6	-28.5	75.6	100.0
	conf. interval	[-2.1,59.8]	[20.3,27.2]	[-51.8,-5.0]	[61.6,89.2]	[100.0,100.0]
	2015-2019	-0.009	0.004	-0.005	0.002	-0.008
	conf. interval	[-0.014,-0.003]	[0.003,0.005]	[-0.009,-0.001]	[-0.001,0.005]	[-0.009,-0.008]
	share explained	109.5	-47.7	62.8	-24.6	100.0
	conf. interval	[40.4,177.5]	[-55.6,-39.7]	[15.5,110.1]	[-56.0,6.8]	[100.0,100.0]
$\frac{\text{BFWI}}{\text{Var}W}$	1995-2014	0.018	0.007	-0.019	0.047	0.054
	conf. interval	[-0.001,0.037]	[0.006,0.008]	[-0.034,-0.003]	[0.039,0.055]	[0.052,0.055]
	share explained	34.2	12.7	-35.5	88.6	100.0
	conf. interval	[-2.4,69.7]	[10.7,14.8]	[-64.8,-6.3]	[72.0,104.4]	[100.0,100.0]
	2015-2019	-0.025	0.005	-0.015	0.006	-0.030
	conf. interval	[-0.040,-0.009]	[0.004,0.006]	[-0.027,-0.004]	[-0.002,0.013]	[-0.031,-0.028]
	share explained	84.3	-17.5	52.0	-18.8	100.0
	conf. interval	[30.7,137.6]	[-20.2,-14.8]	[13.0,91.2]	[-43.1,5.3]	[100.0,100.0]
VarU	1995-2014	0.001	0.006	-0.000	0.004	0.011
	conf. interval	[-0.000,0.003]	[0.005,0.006]	[-0.000,-0.000]	[0.003,0.004]	[0.010,0.011]
	share explained	13.7	52.2	-1.7	35.8	100.0
	conf. interval	[-1.1,27.0]	[43.7,61.1]	[-2.9,-0.3]	[28.6,43.0]	[100.0,100.0]
	2015-2019	-0.002	0.004	-0.000	0.001	0.002
	conf. interval	[-0.004,-0.001]	[0.004,0.005]	[-0.000,-0.000]	[-0.000,0.001]	[0.002,0.003]
	share explained	-103.9	188.4	-7.4	22.9	100.0
	conf. interval	[-204.3,-29.9]	[134.8,266.9]	[-11.5,-2.4]	[-6.0,57.2]	[100.0,100.0]
BFUI	1995-2014	0.001	0.000	-0.000	0.001	0.002
	conf. interval	[-0.000,0.001]	[0.000,0.000]	[-0.000,-0.000]	[0.001,0.002]	[0.002,0.003]
	share explained	24.0	14.1	-3.0	64.9	100.0
	conf. interval	[-2.2,44.2]	[10.4,19.0]	[-4.8,-0.7]	[48.6,84.7]	[100.0,100.0]
	2015-2019	-0.001	0.000	-0.000	0.000	-0.001
	conf. interval	[-0.001,-0.000]	[0.000,0.000]	[-0.000,-0.000]	[-0.000,0.000]	[-0.001,-0.000]
	share explained	188.8	-65.0	20.5	-44.3	100.0
	conf. interval	[127.7,304.1]	[-164.1,-34.9]	[1.8,81.5]	[-137.6,15.9]	[100.0,100.0]
$\frac{\text{BFUI}}{\text{Var}U}$	1995-2013	0.004	0.000	-0.000	0.009	0.013
	conf. interval	[-0.000,0.007]	[0.000,0.001]	[-0.001,-0.000]	[0.008,0.011]	[0.010,0.016]
	share explained	26.6	3.8	-3.3	72.9	100.0
	conf. interval	[-2.6,48.3]	[2.6,5.3]	[-5.2,-0.8]	[53.5,98.4]	[100.0,100.0]
	2015-2019	-0.005	0.000	-0.000	0.001	-0.004
	conf. interval	[-0.009,-0.002]	[0.000,0.001]	[-0.001,-0.000]	[-0.000,0.003]	[-0.007,-0.002]
	share explained	128.7	-10.9	12.4	-30.2	100.0
	conf. interval	[84.8,170.2]	[-19.4,-7.2]	[1.6,38.1]	[-77.5,9.8]	[100.0,100.0]

Table 1: The average impact of changes in $\sigma_\theta, \sigma_x, c_a$ and c_l on 7 outcomes of choice.

Notes: All effects are calculated by calibrating the model to each draw of our resampling procedure; the headline effects are then just the average across the 1500 draws for each year, whereas the confidence intervals are calculated by ranking the outcomes from lowest to highest and reporting the 0.025 and 0.975 percentiles.

just 35% (0.01 in absolute terms). In other words, over three quarters of the increase in wage inequality caused by the fall in the span-of-control cost reflected changes in the compensating differential for effort; in the case of the increase in the supply of high-skilled workers, the majority (60%) of the increase in wage inequality reflects actually changes in welfare inequality, with only 40% being caused by changes to compensating differentials.

The final aspect of the adjustments between 1995 and 2014 that is worth noting, is the fact that the change in σ_x is responsible for just 3.8% of the overall increase in the share of overall inequality that is explained by its between-firm component. Recall from Section 5.3 that this ratio is equal to ρ^4, which itself is a function of only the equilibrium supply of quality jobs; hence, the change in σ_x has a minimal impact on σ, and thus the vast majority of its large impact on other outcomes is direct.

From 2015 to 2019, we can observe large impact of changes in all primitives except perhaps for the span-of-control cost c_l. As discussed in Section 6, overall wage inequality has decreased after 2015, but overall welfare inequality continued its climb. Interestingly, the impact of changes in σ_x especially (but also in c_l to a smaller degree) remained broadly the same as in the pre-2014 period. What has changed, is that post-2015 the inequality increases caused by σ_x and c_l have been counteracted by a decrease in all inequality components spurned by an increase in c_a and a fall in σ_θ. However, as we have discussed at length in Section 5.5, relative to their impact on welfare inequality changes in σ_x have a weak effect on wage inequality and on sorting. Thus, while the continued increase in the variance of skill was enough to ensure that welfare inequality continued to climb, the fall in the variance of productivity and the increase in the cost of amenity provision have caused a fall in wage inequality and sorting.

7 Concluding Remarks

This paper develops a novel, extremely tractable model of workers' sorting within and across firms, and uses the model to shed some light on the drivers of the observed changes in overall, within- and between-firm wage inequality in Norway from 1995 to 2019.

The tractability of the model allows us to derive a comprehensive set of comparative statics results with respect to changes in four primitives: the variance of skill, the variance of productivity, the cost of amenity provision and the span-of-control cost. The main theoretical insight is the existence of a firm link between changes in overall wage inequality, and its share explained by the between-firm component. Specifically, we find that—for realistic parameter values—if a change in one of our four primitives increases overall wage inequality, it also must increase the ratio of between-firm to overall wage inequality. While it is possible for overall and between-firm inequality to not co-move, if changes in two or more primitives counteract each other, it appears that their co-movement should be treated as the default behaviour, rather than a surprise.

We then use our model to analyse the causes of changes in wage inequality between 1995 and 2019 in Norway. We document that overall and between-firm inequality have indeed co-moved in Norway throughout this period; prior to 2014 both were increasing, and after 2015 both were decreasing. After calibrating the model to the Norwegian data, we find that welfare inequality has been evolving qualitatively similarly to wage inequality before 2014, but post 2015 welfare inequality continued to increase. Given that changes in any of our four primitives can match the observed changes in wage inequality qualitatively, we perform a counterfactual exercise to find which of these four potential changes was quantitatively responsible for the observed changes. For 1995 to 2014, we find that a decrease in the span-of-control cost drove most of the change in observed wage inequality, while the increase in the variance of skill was responsible for the majority of the increase in welfare inequality.

A Omitted Proofs and Derivations

A.1 Derivation of Equation 13

First, define $p = 1 + \frac{1}{\sigma_\theta} - \frac{\alpha}{\frac{\sigma}{\sigma_x}-1}$. Using this and Equation (11), Equation (12) can be rewritten as

$$\begin{aligned}
f(h) &= \frac{L^*(0)}{2\pi\sigma_\theta} \cdot \int_{-\infty}^{\infty} \exp\left(\frac{1+\frac{1}{\sigma_\theta}-p}{2}\theta^2 - \frac{1}{2}(h-\theta)^2 - \frac{1}{2}\left(\frac{\theta}{\sigma_\theta}\right)^2\right) d\theta \\
&= \frac{L^*(0)}{2\pi\sigma_\theta} \cdot \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\left[\theta^2 p - 2h\theta + h^2\right]\right) d\theta \\
&= \frac{L^*(0)}{2\pi\sigma_\theta} \cdot \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\left(\theta\sqrt{p} - \sqrt{\frac{1}{p}}h\right)^2 - \frac{1}{2}h^2\frac{p-1}{p}\right) d\theta \\
&= \frac{L^*(0)}{\sigma_\theta\sqrt{p}} \cdot \phi\left(h\sqrt{1-\frac{1}{p}}\right) \underbrace{\int_{-\infty}^{\infty} \sqrt{p}\phi\left(-\sqrt{p}\left(\theta - \frac{1}{p}h\right)\right) d\theta}_{=1},
\end{aligned}$$

which is trivially equal to Equation 13.

A.2 Proof of Proposition 2

(i) Let us start by showing that

$$\frac{\mathrm{d}}{\mathrm{d}c_a} c_a\left(1 - \frac{\sigma_x}{\sigma}\right) > 0, \frac{\mathrm{d}}{\mathrm{d}\sigma_x} c_a\left(1 - \frac{\sigma_x}{\sigma}\right) < 0. \tag{43}$$

The former follows as soon as we rewrite Equation (16) as

$$c_a(1-\frac{\sigma_x}{\sigma}) = \frac{\frac{1+c_a}{c_l}\frac{\sigma_x}{\sigma}}{\frac{1}{\sigma_\theta}-\frac{1}{\sigma^2-1}}. \tag{44}$$

The right-hand side increases in c_a both directly, and indirectly, because it is decreasing in σ (which is, itself, decreasing in c_a). To see the latter, denote $\frac{\sigma}{\sigma_x}$ by $\bar{\sigma}$, and define the function

$$h(\bar{\sigma};\sigma_x,\sigma_\theta) \equiv (\bar{\sigma}-1)\left(\frac{1}{\sigma_\theta}-\frac{1}{\bar{\sigma}^2\sigma_x^2-1}\right).$$

As h is increasing in $\bar{\sigma}$, in equilibrium it must be the case that $h^{-1}(\bar{\sigma};A,\beta)=\bar{\sigma}$. Given this and the fact that $\frac{\partial}{\partial\sigma_x}h>0$, it follows that $\frac{\partial}{\partial\sigma_x}\bar{\sigma}<0$, implying that $\frac{\mathrm{d}}{\mathrm{d}\sigma_x}c_a(1-\frac{1}{\bar{\sigma}})<0$ as well.

Part (i) follows then immediately from Proposition 1, and Equations (33), (34), (35) and (43).

(ii)

We will start with the results regarding BFWI. Define $b\equiv\frac{\sqrt{BFUI}}{c_a(1-\frac{\sigma_x}{\sigma})}$ and note that

$$\frac{\mathrm{d}}{\mathrm{d}\sigma}b = \frac{\sqrt{2}\sigma_x\left(\sigma^2-2\sigma_x\sigma+1\right)}{2c_a(\sigma-\sigma_x)^2}.$$

It is easy to see that if $\sigma_x<1$ then $\frac{\mathrm{d}}{\mathrm{d}\sigma}b>0$, and thus $\frac{\mathrm{d}}{\mathrm{d}\sigma}BFWI>0$. Hence, by Proposition 1, if $\sigma_x<1$ then BFWI increases in σ_θ and decreases in c_a.

Suppose, instead, that $\sigma_x>1$. In that case, $\frac{\mathrm{d}}{\mathrm{d}\sigma}b>0$ if and only if $\sigma>\sigma_x+\sqrt{\sigma_x^2-1}$, which implies that $\frac{\mathrm{d}}{\mathrm{d}\sigma}BFWI>0$ if $\alpha>\sqrt{1-\frac{1}{\sigma_x^2}}\left(\sigma_\theta^{-1}-1/\left(2\left(\sigma_x^2-\sigma_x\sqrt{\sigma_x^2-1}-1\right)\right)\right)$. In other words, BFWI decreases in c_l for c_l smaller than $(1+\frac{1}{c_a})\left(\sqrt{1-\frac{1}{\sigma_x^2}}\left(\sigma_\theta^{-1}-1/\left(2\left(\sigma_x^2-\sigma_x\sqrt{\sigma_x^2-1}-1\right)\right)\right)\right)^{-1}$ and increases in σ_θ for $\sigma_\theta>\left(\alpha/\sqrt{1-\frac{1}{\sigma_x^2}}+1/\left(2\left(\sigma_x^2-\sigma_x\sqrt{\sigma_x^2-1}-1\right)\right)\right)^{-1}$. The fact that $\frac{\mathrm{d}}{\mathrm{d}\sigma}b>0$ if and only if $\sigma>\sigma_x+\sqrt{\sigma_x^2-1}$ implies also that $\lim_{\sigma\to\sigma_x}\frac{\mathrm{d}}{\mathrm{d}\sigma}b=-\infty$, as the denominator of $\frac{\mathrm{d}}{\mathrm{d}\sigma}b$ goes to zero as $\sigma\to\sigma_x$. One can easily infer from Equation (16) that $\sigma\to\sigma_x$ as $c_l\to\infty$ or $\sigma_\theta\to0$. As

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\sigma}BFWI &= 2(\frac{\mathrm{d}}{\mathrm{d}\sigma}\sqrt{BFUI}+\frac{\mathrm{d}}{\mathrm{d}\sigma}b)\sqrt{BFUI}\\ \frac{\mathrm{d}}{\mathrm{d}\sigma}\mathrm{Var}(\ln W) &= \frac{\mathrm{d}}{\mathrm{d}\sigma}WFWI+\frac{\mathrm{d}}{\mathrm{d}\sigma}BFWI,\end{aligned}$$

and both $\frac{\mathrm{d}}{\mathrm{d}\sigma}\sqrt{BFUI}$ and $\frac{\mathrm{d}}{\mathrm{d}\sigma}WFWI$ have finite limits as $\sigma\to\sigma_x$, it follows that BFWI and $\frac{\mathrm{d}}{\mathrm{d}\sigma}\mathrm{Var}(\ln W)$ decrease in σ_θ for very small σ_θ, and increase in c_l for very large values of c_l.[40]

[40] A careful reader may be surprised at this result. It is quite clear that for $\sigma_\theta=0$ the model reduces

To show the results regarding $\frac{BFWI}{\text{Var}W}$ let us define $e \equiv \frac{\rho^2}{\sqrt{1-\rho^4}} \frac{1}{1-\sqrt{1-\rho^2}\sigma_x}$, and note that

$$\frac{\mathrm{d}}{\mathrm{d}\rho^2} e = \frac{\left(\rho^6 + \rho^2 - 2\right)\sigma_x + 2\sqrt{1-\rho^2}}{2(\rho^2-1)^2(\rho^2+1)^{3/2}\left(\sqrt{1-\rho^2}\sigma_x - 1\right)^2}.$$

One can show that for $\sigma_x < 2/(2.24(2\sqrt{5}/5)) < 1$ the numerator of this expression is always positive. Thus, if $\sigma_x < 0.99$ then $\frac{\mathrm{d}}{\mathrm{d}\rho^2} e > 0$. As we can write

$$\frac{BFWI}{\text{Var}W} = \frac{1}{1 + \frac{1}{\left(\frac{1}{\sqrt{\frac{1}{\rho^4}-1}} + e\right)^2}} = \frac{1}{1 + \frac{WFWI}{BFWI}}$$

it follows by Proposition 1 that if $\sigma_x < 0.99$, then $\frac{BFWI}{\text{Var}W}$ increases in σ_θ and decreases in c_l. Next, note that because $\frac{\mathrm{d}}{\mathrm{d}\sigma} WFWI > 0$, it follows that if $\frac{\mathrm{d}}{\mathrm{d}\sigma} BFWI < 0$ then $\frac{\mathrm{d}}{\mathrm{d}\sigma} \frac{BFWI}{\text{Var}W} < 0$. Thus, it follows from the discussion above that if $\sigma_x > 1$, then $\frac{BFWI}{\text{Var}W}$ decreases in σ_θ for very small σ_θ, and increases in c_l for very large values of c_l. Finally, it is easy to see that $\frac{\mathrm{d}}{\mathrm{d}\rho^2} e > 0$ if an only if $\sigma_x < \frac{1}{\sqrt{1-\rho^2}(\rho^4+\rho^2+2)}$; for any value of σ_x, there must exist some values of ρ that are close enough to 1 such that this condition is satisfied. It is easy to see by inspection of Equations (16) and (21) that if $\sigma_\theta \to \infty$ or $c_l \to 0$, then $\rho \to 1$. Overall, therefore, if $\sigma_x > 1$, then $\frac{BFWI}{\text{Var}W}$ is non-monotone in both c_l and σ_θ, as required.

A.3 Proof of Proposition 3

Throughout this proof, we will denote $\text{Var}(\ln U)$ by V.

(i) Note that sorting can be rewritten as

$$\rho^2 = 1 - \frac{\sigma_x^2}{2V}, \tag{45}$$

which immediately yields that

$$y_{c_a} \frac{\mathrm{d}}{\mathrm{d}c_a}\rho^4 = y_{c_l} \frac{\mathrm{d}}{\mathrm{d}c_l}\rho^4 = 4y_{c_l} \frac{\mathrm{d}}{\mathrm{d}c_l} V \frac{\sigma_x^2}{2V^2}\rho^2 > 4\left(\frac{\mathrm{d}}{\mathrm{d}\sigma_x} V \frac{\sigma_x^2}{2V^2} - \frac{\sigma_x}{2V}\right)\rho^2 = \frac{\mathrm{d}}{\mathrm{d}\sigma_x}\rho^4,$$

and part (i) follows from Equation (25).

to a Costrell and Loury (2004) model, in which between-firm wage inequality is equal 0. How can it thus be that BFWI decreases for low σ_θ? The reason is that if $\sigma_x > 1$ then the equilibrium σ is discontinuous at $\sigma_\theta = 0$. For $\sigma_\theta = 0$, we have that $\sigma = 1$; however, for any $\sigma_\theta > 0$, σ is strictly greater than σ_x. We find this discontinuity fascinating, but not of first-order importance, as $\sigma_x > 1$ would imply a much higher level of within-firm wage inequality than those observed in the data.

(ii) Define $\delta \equiv \frac{1}{c_a(1+\sigma_x/\sigma)}$. Using Equations (44) and (45) we can rewrite it as

$$\delta = \left(\frac{\frac{1+c_a}{c_l} \frac{\frac{\sqrt{2}}{2}\sigma_x^2}{\sqrt{V}}}{\frac{1}{\sigma_\theta} - \frac{1}{\frac{2V}{\sigma_x^2}-1}} \right)^{-1}.$$

Hence, conditioning on V, δ decreases in c_a and σ_x, and increases in c_l. The ratio of between-firm to overall wage inequality can be readily rewritten as

$$\frac{BFWI}{\text{Var}(\ln W)} = \frac{1}{1 + (\frac{1}{(1-\frac{\sigma_x^2}{2V})^2} - 1)(1+\delta)^{-2}}.$$

Again, conditioning on V, this expression decreases in c_a and σ_x, and increases in c_l. Thus, it follows that a fall in c_a that has the same impact on V as an increase in σ_x (decrease in c_l) will increase $\frac{BFWI}{\text{Var}(\ln W)}$ by more.

(iii) The expression for $\text{Var}(\ln W)$ can be rewritten as

$$\text{Var}(\ln W) = V\left(1 + \delta(1 - \frac{\sigma_x^2}{2V})^2(2+\delta)\right).$$

Clearly, this expression also decreases in c_a and σ_x, and increases in c_l, and thus part (iii) follows by the same reasoning as above.

(iv) Let us now define $\bar{\delta} \equiv \frac{1-\frac{1}{\sigma^2}}{c_a(1-\sigma_x/\sigma)}$, and note that

$$\bar{\delta} = \frac{1-\frac{1}{\sigma^2}}{c_a(1-\frac{\sqrt{2V}}{\sigma^2})} = c_a^{-1}\left(1 + \frac{\sqrt{2V}-1}{\sigma^2 - \sqrt{2V}}\right) = c_a^{-1}\left(1 + \frac{\sqrt{2V}-1}{\frac{2V}{\sigma_x^2} - \sqrt{2V}}\right).$$

Hence, if $V < 0.5$ and conditioning on V, then $\bar{\delta}$ decreases in σ_x and does not depend on c_l. We can then clearly write

$$\frac{BFWI}{\text{Var}(\ln W)} = \frac{1}{1 + (1 - (1-\frac{\sigma_x^2}{2V})^2)\left(1 - \frac{\sigma_x^2}{2V} + \bar{\delta}\right)^{-2}},$$

so that if $V < 0.5$ (and conditioning on V) then $\frac{BFWI}{\text{Var}(\ln W)}$ decreases in σ_x and does not depend on c_l, and part (iv) follows.

(v) Finally, we can rewrite the expression for $\text{Var}(\ln W)$ as

$$\text{Var}(\ln W) = V(1 + \bar{\delta}(2(1 - \frac{\sigma_x^2}{2V}) + \bar{\delta}))$$

from which it follows immediately that (conditioning on V), if $V < 0.5$ then $\text{Var}(\ln W)$ decreases in σ_x but does not depend on c_l; hence, part (v) follows as well.

References

Abowd, J. M., F. Kramarz, and D. N. Margolis (1999). High wage workers and high wage firms. *Econometrica 67*(2), 251–333.

Acemoglu, D. (2002, March). Technical change, inequality, and the labor market. *Journal of Economic Literature 40*(1), 7–72.

Acemoglu, D. and P. Restrepo (2022). Tasks, automation, and the rise in u.s. wage inequality. *Econometrica 90*(5), 1973–2016.

Aeppli, C. and N. Wilmers (2022, October). Rapid wage growth at the bottom has offset rising US inequality. *Proceedings of the National Academy of Sciences 119*(42), e2204305119.

Autor, D., D. Dorn, L. F. Katz, C. Patterson, and J. Van Reenen (2020, May). The fall of the labor share and the rise of superstar firms. *The Quarterly Journal of Economics 135*(2), 645–709.

Bagger, J., E. R. Moen, and R. Vejlin (2021). Equilibrium worker-firm allocations and the deadweight losses of taxation. CEPR Discussion Papers 16735, C.E.P.R. Discussion Papers.

Bana, S., K. Bedard, M. Rossin-Slater, and J. Stearns (2023, April). Unequal use of social insurance benefits: The role of employers. *Journal of Econometrics 233*(2), 633–660.

Barth, E., A. Bryson, J. C. Davis, and R. Freeman (2016, April). It's where you work: Increases in the dispersion of earnings across establishments and individuals in the united states. *Journal of Labor Economics 34*(S2), S67–S97.

Barth, E., J. Davis, and R. B. Freeman (2018, January). Augmenting the human capital earnings equation with measures of where people work. *Journal of Labor Economics 36*(S1), S71–S97.

Barth, E., J. C. Davis, R. B. Freeman, and K. McElheran (2023, April). Twisting the demand curve: Digitalization and the older workforce. *Journal of Econometrics 233*(2), 443–467.

Baumgarten, D., G. Felbermayr, and S. Lehwald (2020, January). Dissecting between-plant and within-plant wage dispersion: Evidence from germany. *Industrial Relations 59*(1), 85–122.

Bloom, N., F. Guvenen, B. S. Smith, J. Song, and T. von Wachter (2018). The disappearing large-firm wage premium. In *AEA Papers and Proceedings*, Volume 108, pp. 317–322. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

Boerma, J., A. Tsyvinski, and A. P. Zimin (2023). Sorting with teams.

Bonhomme, S., K. Holzheu, T. Lamadon, E. Manresa, M. Mogstad, and B. Setzler (2023). How much should we trust estimates of firm effects and worker sorting? *Journal of Labor Economics 41*(2), 291–322.

Bonhomme, S., T. Lamadon, and E. Manresa (2019). A distributional framework for matched employer employee data. *Econometrica 87*(3), 699–739.

Bound, J. and G. Johnson (1992). Changes in the structure of wages in the 1980's: An evaluation of alternative explanations. *The American Economic Review 82*(3).

Bowsher, C. G. and P. S. Swain (2012). Identifying sources of variation and the flow of information in biochemical networks. *Proceedings of the National Academy of Sciences 109*(20), E1320–E1328.

Card, D., J. Heining, and P. Kline (2013). Workplace Heterogeneity and the Rise of West German Wage Inequality. *Quarterly Journal of Economics*.

Cortes, G. M., A. Lerche, U. Schönberg, and J. Tschopp (2023, April). Technological change, firm heterogeneity and wage inequality.

Costinot, A. (2009). An elementary theory of comparative advantage. *Econometrica 77*(4), 1165–1192.

Costinot, A. and J. Vogel (2010, 08). Matching and Inequality in the World Economy. *Journal of Political Economy 118*(4), 747–786.

Costrell, R. M. and G. C. Loury (2004, December). Distribution of Ability and Earnings in a Hierarchical Job Assignment Model. *Journal of Political Economy 112*(6), 1322–1363.

Eeckhout, J. and P. Kircher (2018). Assortative Matching with Large Firms. *Econometrica 86*(1), 85–132.

Eeckhout, J. and R. Pinheiro (2014). Diverse organizations and the competition for talent. *International Economic Review 55*(3), 625–664.

Engbom, N., C. Moser, and J. Sauermann (2023, April). Firm pay dynamics. *Journal of Econometrics 233*(2), 396–423.

Freund, L. (2022). Superstar teams: The micro origins and macro implications of coworker complementarities. *Available at SSRN 4312245*.

Friedrich, B. U. (2022, July). Trade shocks, firm hierarchies, and wage inequality. *The Review of Economics and Statistics 104*(4), 652–667.

Gola, P. (2021). Supply and Demand in a Two-Sector Matching Model. *Journal of Political Economy 129*(3), 940–978.

Håkanson, C., E. Lindqvist, and J. Vlachos (2021). Firms and skills the evolution of worker sorting. *Journal of Human Resources 56*(2), 512–538.

Haltiwanger, J., H. R. Hyatt, and J. R. Spletzer (2023, October). Increasing earnings inequality: Reconciling evidence from survey and administrative data. *Journal of Labor Economics 41*(S1), S61–S93.

Haltiwanger, J. C., H. R. Hyatt, and J. Spletzer (2022, April). Industries, mega firms, and increasing inequality.

Helpman, E., O. Itskhoki, M.-A. Muendler, and S. J. Redding (2016). Trade and Inequality: From Theory to Estimation. *Review of Economic Studies*.

Hoen, M. F., S. Markussen, and K. Røed (2022, October). Immigration and economic mobility. *Journal of Population Economics 35*(4), 1589–1630.

Juhn, C., K. M. Murphy, and B. Pierce (1993). Wage inequality and the rise in returns to skill. *Journal of Political Economy 101*(3), 410–442.

Katz, L. and K. M. Murphy (1992). Changes in relative wages, 1963—1987: Supply and demand factors. *The Quarterly Journal of Economics 107*(1), 35–78.

Kristal, T., Y. Cohen, and E. Navot (2020, April). Workplace compensation practices and the rise in benefit inequality. *American Sociological Review 85*(2), 271–297.

Lamadon, T., M. Mogstad, and B. Setzler (2022, January). Imperfect competition, compensating differentials, and rent sharing in the US labor market. *The American Economic Review 112*(1), 169–212.

Landefeld, P., J. Mortenson, and T. Willingham (2023). Firms and the us earnings distribution.

Lindner, A., B. Muraközy, B. Reizer, and R. Schreiner (2022, July). Firm-level Technological Change and Skill Demand. CEPR Discussion Papers 17421.

Lochner, B. and B. Schulz (2022, September). Firm productivity, wages, and sorting. *Journal of Labor Economics*.

Lochner, B., S. Seth, and S. Wolter (2020, September). Decomposing the large firm wage premium in germany. *Economics Letters 194*, 109368.

Maestas, N., K. J. Mullen, D. Powell, T. Von Wachter, and J. B. Wenger (2023). The value of working conditions in the united states and implications for the structure of wages. *American Economic Review 113*(7), 2007–2047.

Markussen, S. and K. Røed (2019, September). Economic mobility under pressure. *Journal of the European Economic Association 18*(4), 1844–1885.

Melitz, M. J. (2003, November). The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. *Econometrica 71*(6), 1695–1725.

Mueller, H. M., P. P. Ouimet, and E. Simintzi (2017, May). Wage inequality and firm growth. *American Economic Review 107*(5), 379–383.

Norwegian Tax Administration (2018). Kodespeil for lønnssystem A02 [Code for payroll system A02]. https://www.skatteetaten.no/bedrift-og-organisasjon/arbeidsgiver/a-meldingen/veiledning/kodespeil/.

Norwegian Tax Administration (2023a). The a-melding. https://www.skatteetaten.no/en/business-and-organisation/employer/the-a-melding/.

Norwegian Tax Administration (2023b). National insurance scheme basic amount. https://www.skatteetaten.no/en/rates/national-insurance-scheme-basic-amount/.

O'Neill, B. (2014). Some useful moment results in sampling problems. *The American Statistician 68*(4), 282–296.

Ouimet, P. and G. Tate (2023, July). Firms with benefits? nonwage compensation and implications for firms and labor markets.

Sattinger, M. (1975). Comparative Advantage and the Distributions of Earnings and Abilities. *Econometrica 43*(3), 455–468.

Sattinger, M. (1979, March). Differential Rents and the Distribution of Earnings. *Oxford Economic Papers 31*(1), 60–71.

Sockin, J. (2022). Show me the amenity: Are higher-paying firms better all around?

Song, J., D. J. Price, F. Guvenen, N. Bloom, and T. von Wachter (2018, 10). Firming Up Inequality*. *The Quarterly Journal of Economics 134*(1), 1–50.

Sorkin, I. (2018, 01). Ranking Firms Using Revealed Preference*. *The Quarterly Journal of Economics 133*(3), 1331–1393.

Sorkin, I. and M. Wallskog (2023, October). The slow diffusion of earnings inequality. *Journal of Labor Economics 41*(S1), S95–S127.

Statistics Norway (2014). Fra bedrift til virksomhet [From company to business]. https://www.ssb.no/virksomheter-foretak-og-regnskap/artikler-og-publikasjoner/fra-bedrift-til-virksomhet.

Statistics Norway (2018). Virksomheter [Businesses]. https://www.ssb.no/virksomheter-foretak-og-regnskap/virksomheter-og-foretak/statistikk/virksomheter.

Statistics Norway (2020a). The a-scheme: monthly employment and wage data from january 2015 to november 2022 (ameld_statdata_yyyy_mm). https://www.ssb.no/data-til-forskning/utlan-av-data-til-forskere/variabellister/a-ordningen.

Statistics Norway (2020b). Arbeidsmarked [Labour market]. https://www.ssb.no/data-til-forskning/utlan-av-data-til-forskere/variabellister/arbeidsmarked.

Statistics Norway (2020c). Constant person characteristics (tab_flan_rev). https://www.ssb.no/data-til-forskning/utlan-av-data-til-forskere/variabellister/befolkning.

Statistics Norway (2020d). Employer and employee register (aa register) and the wage and withdrawal register (lto): Annual employment and wage data from 1995 to 2014 (atmltoyyyy). https://www.ssb.no/data-til-forskning/utlan-av-data-til-forskere/variabellister/arbeidsmarked.

Statistics Norway (2020e). Om datagrunnlaget for registerbasert sysselsettingsstatistikk [About the data base for register-based employment statistics]. https://www.ssb.no/data-til-forskning/utlan-av-data-til-forskere/variabellister/a-ordningen/om-datagrunnlaget-for-registerbasert-sysselsettingsstatistikk.

Statistics Norway (2023). Consumer price index. https://www.ssb.no/en/priser-og-prisindekser/konsumpriser/statistikk/konsumprisindeksen.

Taber, C. and R. Vejlin (2020). Estimation of a roy/search/compensating differential model of the labor market. *Econometrica 88*(3), 1031–1069.

Teulings, C. (2005). Comparative Advantage, Relative Wages, and the Accumulation of Human Capital. *Journal of Political Economy 113*(2), 425–461.

Teulings, C. N. (1995, April). The Wage Distribution in a Model of the Assignment of Skills to Jobs. *Journal of Political Economy 103*(2), 280–315.

Tomaskovic-Devey, D., A. Rainey, D. Avent-Holt, N. Bandelj, I. Boza, D. Cort, O. Godechot, G. Hajdu, M. Hällsten, L. F. Henriksen, A. S. Hermansen, F. Hou, J. Jung, A. Kanjuo-Mrčela, J. King, N. Kodama, T. Kristal, A. Křížková, Z. Lippényi, S. M. Melzer, E. Mun, A. Penner, T. Petersen, A. Poje, M. Safi, M. Thaning, and Z. Tufail (2020, April). Rising between-workplace inequalities in high-income countries. *Proceedings of the National Academy of Sciences 117*(17), 9277–9283.

Trottner, F. (2022). Who gains from scale? trade and wage inequality within and between firms. https://trottner.me/files/WGS.pdf. Accessed: 2023-11-15.

ONLINE APPENDIX

B Firm and Worker Fixed Effects

Since Abowd et al. (1999) a large empirical literature has focused on estimating worker fixed effects (WFE) and firm fixed effects (FFE) using matched employer-employee data sets, and then using the estimated effects to perform various variance decomposition exercises. As a direct consequence of Equation (31), in our economy the worker fixed effect estimated through any AKM-like regressions would be exactly equal to the worker's utility, and the firm fixed effects would be exactly equal to the firm's amenity provision. The standard decomposition of overall wage variance into the variance of firm fixed effects, worker fixed effects and twice the covariance between these two types of fixed effects takes the following form in our model

$$\text{Var}(\ln W) = \underbrace{0.5(\sigma_x\sigma)^2}_{\text{Variance of WFE}} + \underbrace{0.5\left(\frac{\sigma_x\sigma\rho^2}{c_a(1-\frac{\sigma_x}{\sigma})}\right)^2}_{\text{Variance of FFE}} + \underbrace{\frac{(\sigma_x\sigma\rho^2)^2}{c_a(1-\frac{\sigma_x}{\sigma})}}_{\text{2*covariance between WFE and FFE}}. \tag{OA.1}$$

The proportion of overall wage variance explained by (twice) the covariance of worker and firm fixed effects ($2 * \text{Cov}(WFE, FFE)/\text{Var}(\ln W)$) is of particular interest in this literature, and is used to measure the contribution of the sorting between high wage workers and high wage firms to overall wage inequality. Indeed, this ratio of such importance, that it is often referred to as simply 'sorting' and changes in this ratio are referred to as 'changes in sorting' (see, for example, Song et al., 2018; Bonhomme et al., 2023). An advantage of our simple model is that it clarifies that conflating changes in $2 * \text{Cov}(WFE, FFE)/\text{Var}(\ln W)$ with changes in sorting is not always warranted.

To see why, note that the correlation between X^2 and $\bar{\theta}^2$ (ρ^2) is clearly a correct measure of the strength of sorting in our model, because it captures only the actual interdependence between true worker and firm types.[41] One can see immediately from Equation (OA.1) that $2 * \text{Cov}(WFE, FFE)/\text{Var}(\ln W)$ will not, in general, be equal to ρ^2. However, the problem runs deeper, because *the share of variance explained by the covariance between worker and firm fixed effects can change in the opposite direction than the correlation between worker and firm types does.* To keep the resulting expressions simple, let us write out the inverse of $2 * \text{Cov}(WFE, FFE)/\text{Var}(\ln W)$:

$$\frac{\text{Var}(\ln W)}{2 * \text{Cov}(WFE, FFE)} = 1 + \frac{0.5}{\rho^2}\left(\sqrt{\frac{\text{Var}(WFE)}{\text{Var}(FFE)}} + \sqrt{\frac{\text{Var}(FFE)}{\text{Var}(WFE)}}\right),$$

[41] Of course, other measures of interdependence between X^2 and $\bar{\theta}^2$, such as Kendell's tau or Spearman's rho, could be also used to correctly measure the strength of sorting.

which, in our model, reduces to

$$\frac{\text{Var}(\ln W)}{2 * \text{Cov}(WFE, FFE)} = 1 + 0.5\frac{c_a(1-\sqrt{1-\rho^2}\sigma_x)}{\rho^4} + 0.5\frac{1}{c_a(1-\sqrt{1-\rho^2}\sigma_x)}.$$

The third term decreases in ρ^2, and hence contributes to $2*\text{Cov}(WFE, FFE)/\text{Var}(\ln W)$ changing in the same direction as ρ^2. The second term, however, can increase in ρ^2 for ρ^2 sufficiently close to 1. This indicates that the *empirical measure* of sorting decreases in the *theoretically correct measure* of sorting if c_a is large but finite, and the true degree of sorting in the economy is large. Note that in these circumstances both the variance of firm fixed effects and the ratio $2 * \text{Cov}(WFE, FFE)/\text{Var}(\ln W)$ would be low, which is consistent with many empirical estimates. Fortunately, this issue can be easily remedied by using the *correlation* of firm and worker fixed effects to measure the strength of sorting. As worker fixed effects are linear in x^2 and firm fixed effects are linear in θ^2 in our economy, this correlation is equal to ρ^2. Of course, any value of this correlation that lies between 0 and 1 can be explained by our model.

Song et al. (2018) provide a further decomposition of between-firm wage inequality into the variance of the firm fixed effect, twice the covariance between the worker and firm fixed effects and the variance of the average worker fixed-effect within each firm. In our model, this last component is clearly equal to between-firm utility inequality, as workers within the same firm receive the same amenities. Hence, in our economy the Song et al. (2018) decomposition reduces to

$$BFWI = \underbrace{0.5\left(\frac{\sigma_x\sigma\rho^2}{c_a(1-\frac{\sigma_x}{\sigma})}\right)^2}_{\text{Variance of FFE}} + \underbrace{\frac{\left(\sigma_x\sigma\rho^2\right)^2}{c_a(1-\frac{\sigma_x}{\sigma})}}_{2*\text{covariance between WFE and FFE}} + \underbrace{0.5\rho^4\left(\sigma_x\sigma\rho^2\right)^2}_{\text{Variance of average WFE=BFUI.}}$$

C Random sampling

The confidence intervals reported in all tables and figures from Sections 6.3 and 6.4 were calculated through random sampling. For each year t in our data, denote the number of workers in the original sample by N_t. First, for each year, we use a bootstrapping procedure to sample with replacement N_t workers from the original data. Next, we remove duplicated workers to create a sample without replacement. These steps ensure that the percentage of workers that remain is randomly defined for each draw. Then, we recompute the size of each firm, calculating the moments necessary for our calibration for the new sample. The procedure was repeated 1,500 times for each year in the data. After that, we calibrate the model and perform counterfactual analysis for each sample. Finally, we rank the results of the calibration and counterfactual analysis from the smallest to the largest. The 2.5th and 97.5th percentile of the results simply represent the bounds

of the 95% confidence interval for the parameter or moment in question.

The reason we perform bootstrapping on workers rather than on firms is because the latter would lead to an underestimation of the spread of within-firm variance. We remove duplicated workers after bootstrapping because they would bias the within-firm variance estimator downward. Furthermore, we do not change worker-firm matching because randomly assigning workers to firms would undermine the sorting mechanism in the data.

For each sample, approximately two thirds of the workers and over 99% of firms remain after each random sampling procedure. Unweighted estimates are based on the firm sample, while weighted estimates are derived from the paired worker-firm sample. To ensure variance estimators are unbiased, we use Bessel's correction, scaling the sum of squared mean deviations by $N-1$ or $n-1$, where N denotes worker-firm pairs for overall variance, and n represents firm size for within-firm variance. We compute between-firm variance by subtracting within-firm variance from overall variance. To derive the standard deviation of within-firm wage variance, we adjust for the uncertainty in the estimation of the variance itself. Following, for example, O'Neill (2014), we use adjustment factor $\widehat{\mu}_4/n - \widehat{\mu}_2^2(n-3)/(n^2-n)$, where $\widehat{\mu}_k$ is the k's sample moment for a given firm and n is the firm size. We subtract this firm-specific adjustment factor from each firm's within-firm variance, and then calculate the variance of these adjusted within-firm variances.

After the adjustment, our resampling procedure produces central estimates of the variance of within-firm variances that are very close to, but not identical to the adjusted moment obtained from the original sample (see Figure OA.1 below, which plots the central estimates derived from the data with the central estimates obtained from averaging over the 1500 bootstrap samples). Note that the raw estimates are comfortably contained within the confidence intervals produced by the resampling procedure.

D Additional Figures

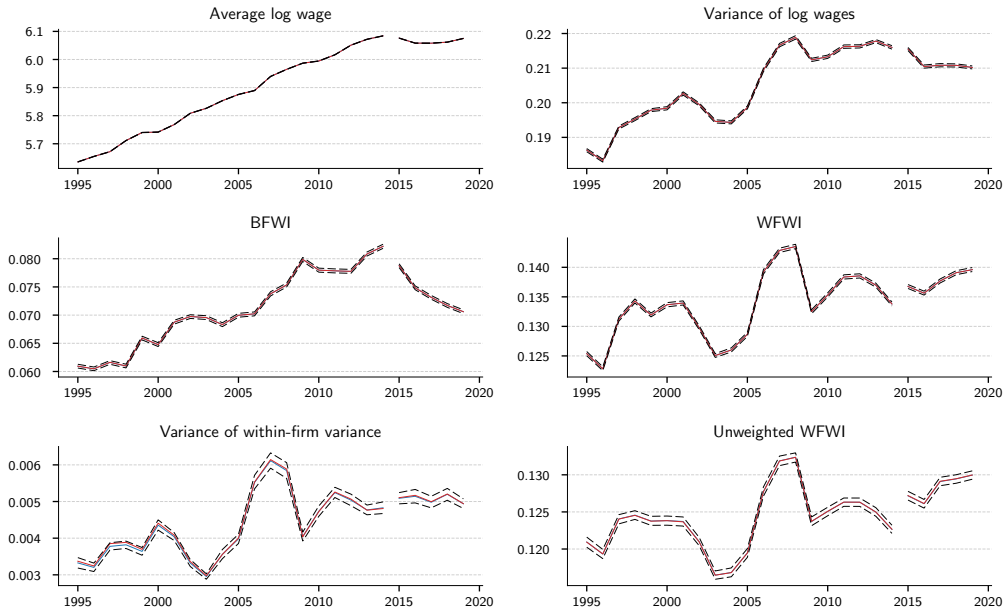


Figure OA.1: Moments used in calibration

Notes: This plots the data moments from the original sample (red), the average over the 1500 bootstrap samples (blue) and the confidence interval (dashed black). Top panel: (1) Mean log wage; (2) Variance of log wages. Middle panel: (1) The across firms, firm-size weighted variance of (within firm) mean log wages; (2) The across firms, firm-size weighted mean of the variance of log wages. Bottom panel: (1) The across firm, firm-size weighted variance of the (within-firm) variance of log wages; (2) The across firms, size-unweighted mean of the variance of log wages.

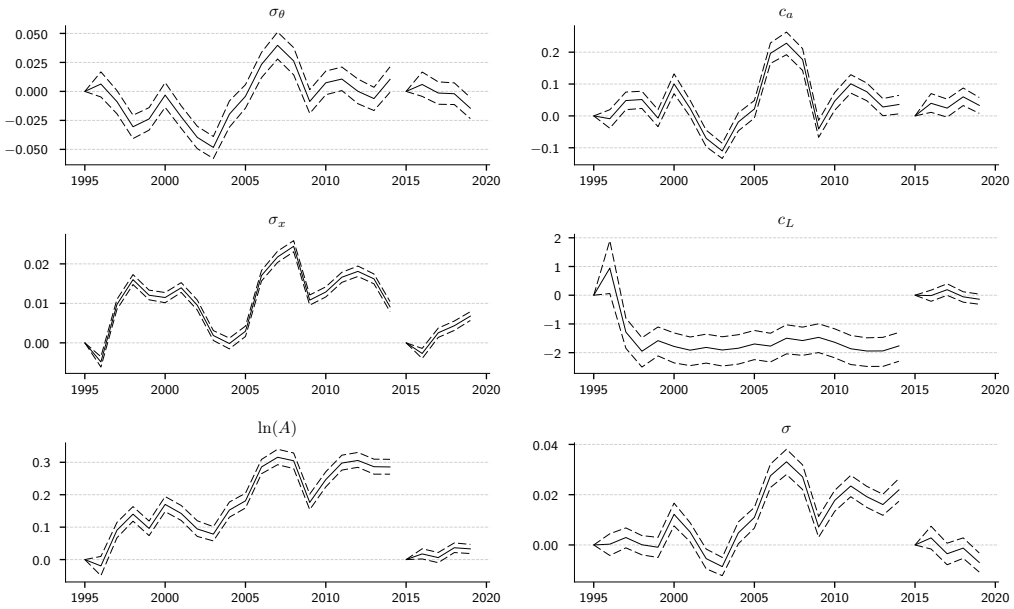


Figure OA.2: Change in calibrated parameters

Notes: The difference in the calibrated parameters between the year in question and 1995. Top panel: (1) the variance of productivity σ_θ, (2) cost of amenity provision c_a. Middle panel: (1) variance of skill σ_x, (2) span-of-control cost c_l. Bottom panel: (1) log of total factor productivity $\ln A$, (2) equilibrium supply of quality jobs σ. The dashed lines depict the 95% confidence intervals.