Being up Front about Income Inequality

Brian Hill

Being up front about Income Inequality*

Brian Hill
CNRS & HEC Paris†

November 11, 2021

Abstract

This paper proposes universal provision of information about the income inequality involved in the creation of a good as a means of moderating income inequality. Recent evidence suggests that a significant section of the population would be willing to pay more for goods whose production involves less excessive income inequality. We show, on a simple model, that supplying inequality information to such a population under competitive markets will in general lead to a reduction in global income inequality. The effect tends to be stronger the more inequality averse the population. Moreover, the outcome will be socially efficient. Possibilities for (de-centralized) implementation are also discussed.

Keywords: Income inequality, information provision, externalities, inequality aversion.

JEL codes: D63, D62, D3, J3.

*Acknowledgements and thanks to be added.

†GREGHEC, 1 rue de la Libération, 78351 Jouy-en-Josas, France. E-mail: hill@hec.fr. URL: www.hec.fr/hill.

1 Introduction

Inequality is an increasingly pertinent, debated and topical issue, if anything even more so in the wake of Covid-19. Many studies have documented the rise in income and wealth inequality over the past 40 years (e.g. Katz and Autor, 1999; Acemoglu and Autor, 2011; Atkinson et al., 2011; Piketty, 2013). According to some commentators, this rise is not without wider consequences, being related to a range of recent social and political phenomena.

Income inequality, moreover, is not an issue that leaves the man on the street indifferent. Surveys (e.g. Kohut, 2013; Antunez and Papuchon, 2019) document that many citizens consider it an important issue. Studies suggest that, whilst people agree that some income inequality may be justified by skill, merit, desert or effort, most do not think that current levels of inequality in, say CEO vs. worker pay, can be justified on such grounds (see Section 2 for a brief survey). This adds up to a situation in which a significant section of the public seems uncomfortable with current levels of inequality.

The mismatch between public attitudes to and current levels of income inequality suggests that it may be an externality. One often finds similar economic tools being used to 'combat' undesirable inequalities and 'correct' externalities, taxation being a notable example. Considering income inequality as an externality, this paper sets out the theoretical case for a complementary tool: information. It considers *universal inequality information provision*, whereby potential consumers of every good or service are informed, at the point of purchase, of the income inequality across all those involved in the conception, production, financing, marketing and logistics leading to the existence of that good or service on the market. The paper also discusses implementation, suggesting that such information need not be mandated by regulation, but can be provided via a mobile phone application.

To theoretically evaluate the impact of blanket product-related inequality reporting, we use a simple perfect competition model with two markets, for goods and labour respectively. Our first main result indicates the effectiveness of inequality information provision in combatting inequality. It implies that whenever there is a proportion of consumers that are *inequality averse*—they are willing to pay a premium to avoid extreme inequality in the creation of the purchased good—inequality information provision will lead to a reduction in global income inequality. As we shall discuss, inequality averse consumers in this sense may fully accept moderate inequalities whilst being averse to extreme ones; indeed, existing experimental evidence points to a significant presence of such inequality aversion in the population (Section 2). Our second central result shows that, whenever markets are competitive, the equilibrium under information provision will be socially efficient: the resulting allocation of goods and inequality levels is Pareto optimal with respect to consumers' preferences.

These properties complement and tie into other qualities of information provision that arise from it being an essentially market-based mechanism. From an economic point of view, this renders it a relatively 'light' measure, in several senses. It is non-invasive, unlike taxes, and fully libertarian, insofar as it allows consumers and firms full freedom in their choice of purchase, production and investment, and ultimately their contribution to inequality. Aside from the information intervention, it leaves markets unaffected, which is key to obtaining social efficiency. Moreover, it embodies a libertarian approach to the issue at

the heart of policy choice regarding inequality: how to trade off equality against productivity. Rather than deciding this issue by fiat or by recourse to expert or politician judgements, information provision directly and explicitly leverages consumer preferences regarding productivity-inequality trade-offs, incorporating them in equilibrium by standard market mechanisms. If productivity is lower under information provision (as will typically be the case), it is because sufficiently many consumers consider that the putative productivity gains—and in particular their impact on prices—are outweighed by the accompanying inequality costs.

Beyond its purely economic advantages, the sort of information-based intervention considered here may also have behavioral and political dimensions. The universal and systematic nature of blanket inequality reporting contrasts with various sorts of 'labels', for instance, Corporate Social Responsability (CSR), Fair Trade or ecological labels, where it is the companies who decide whether to subscribe to them or not. Unlike voluntary labelling, it thus identifies and highlights the 'bad' cases. This is behaviorally relevant: going by the significant psychological evidence that negative cases tend to incite stronger reactions than positive ones (e.g. Kahneman et al., 1991; Rozin and Royzman, 2001; Baumeister et al., 2001), one might expect blanket reporting to have a larger effect on consumption choices.[1] It is also politically relevant, insofar as it can contribute to raising awareness of the issue of inequality and to correcting misperceptions, which appear to be widespread (Section 2). Moreover, it can be thought of as a way of empowering consumers: if they don't like the inequality in the distribution of the money spent on a good among those involved in creating it, they can immediately and individually do something about it, by altering their consumption choices.

This paper first quickly reviews some existing literature on attitudes to income inequality, and in particular willingness to pay to avoid large inequalities associated with purchased goods (Section 2). It then examines the effects of universal information provision in a simple economic model, showing that, under the aforementioned preferences, it has a moderating impact on overall income inequality. Moreover, the resulting inequality distribution is optimal, by the lights of consumers (Section 3). It discusses possibilities for implementation, drawing some general principles and making some tentative suggestions (Section 4). The paper ends situating the proposal with respect to the existing economic literature, and considering the potential of similar interventions for issues other than income inequality (Section 5). Proofs and other material are to be found in the Appendix.

[1] Research into existing voluntary labels generally tends to find a willingness to pay among a section of consumers for Fair Trade labeled products, for instance (e.g. Andorfer and Liebe, 2012), as also attested by a market for such products worth around $8.5 billion in 2017 (`https://www.fairtrade.net/impact/global-sales-overview`). The suggestion here is that when information about the parameter of interest is universally available—saying how good or bad the inequality (or how fair or unfair the trade) is for all products—then the impact can only be larger. Note that the universal character of the information contrasts with the potential information asymmetry involved in voluntary labelling; see Section 5 for more on the relationship between the current proposal and existing approaches in economics.

2 Attitudes to and opinions about income inequality: current evidence

A range of surveys, across of variety of countries, tells a consistent and by now well-documented story concerning opinions about and attitudes to income inequality (e.g. Osberg and Smeeding, 2006; Kiatpongsan and Norton, 2014; Clark and d'Ambrosio, 2015). Figure 1, drawn from Kiatpongsan and Norton (2014) presents, for a range of countries, typical findings concerning peoples' estimated and ideal pay ratio between CEOs of large national corporations and unskilled workers—which is a widely used measure in this literature, and serves as a reasonable proxy for income inequality for products produced by large firms. The Figure illustrates three typical findings.

Firstly, people have definite views on the ideal pay ratio, and these are not reductive. In particular, most people think that some inequality is justified.[2] This is consistent with them considering certain inequality levels to be acceptable on the grounds of fairness, merit or desert, as well as other factors (Fehr and Schmidt, 2003; Almås et al., 2010). Nevertheless, many state moderate levels of ideal inequality. For instance,the median ideal ratio in the global sample graphed in Figure 1 is $4.6:1$, (Kiatpongsan and Norton, 2014); in the US subsample, it is $6.7:1$.

Secondly, they generally think, correctly, that the actual pay ratio is higher than their ideal value: 73% of all subjects (81% of US subjects) estimated the pay ratio as higher than their ideal value.

Thirdly, people underestimate the extent of pay inequality, generally by a factor of ten or more. For instance, among US subjects, the median estimated pay ratio was 25:1, whereas the actual pay ratio was closer to 350:1 (Kiatpongsan and Norton, 2014). 88% of participants underestimated the pay ratio.

These facts suggest the potential effectiveness of an information-based intervention focussed on income inequality. The first two reveal that people have views on acceptable levels of inequality, which could inform their purchasing decisions. Moreover, they are aware that these views may not correspond to reality. The final fact—that people are typically misinformed about current levels of inequality—suggests that the information involved in the intervention is not already possessed by many people, hence leaving space for significant effects of information provision.

A long tradition of research in economics also attests to widespread sensitivity of preferences to inequality. Part of the behavioral literature focusses on 'comparative' or positional evaluations, where individuals' preferences are sensitive to how they compare to others, and hence where they lie in the income distribution (e.g. Fehr and Schmidt, 2003, 2006; Clark and d'Ambrosio, 2015). Others have studied 'disinterested' opinions or 'normative' evaluation of income inequality—whether a individual considers a certain inequality acceptable or justified, independently of how it affects them (Clark and d'Ambrosio, 2015)—revealing sophisticated preferences for inequality and fairness (e.g. Johansson-Stenman et al., 2002; Almås et al., 2020). Since the proposal here is to inform consumers of the inequality involved in the production of a good—not of how they situate with respect to it—it is perhaps closer to the latter literature.

[2]For instance, only 19% of all subjects (15% of US subjects) stated that a CEO should ideally earn the same or less than an unskilled worker.

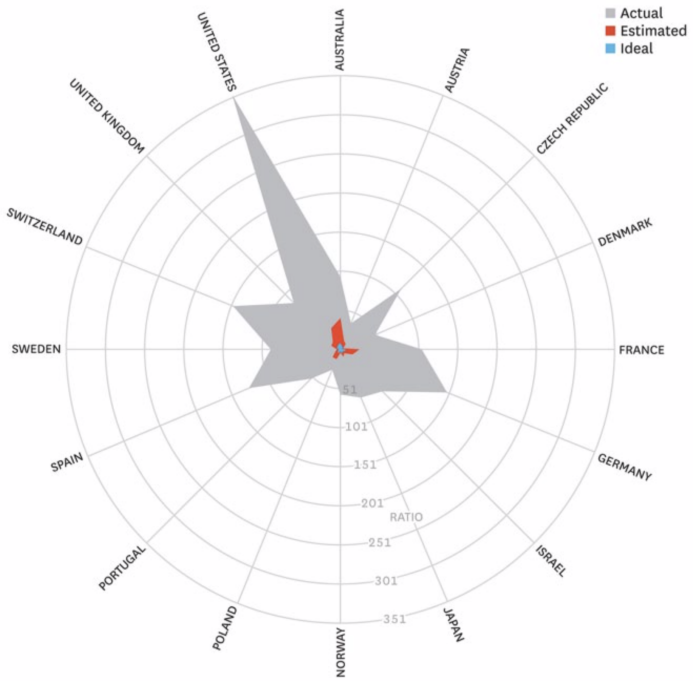

Figure 1: Actual, estimated and ideal pay ratios of CEOs to unskilled workers, across 16 countries (2009)
Data source: ISSP, 2009. Figure source: Kiatpongsan and Norton (2014)
This figure plots the median ideal and estimated CEO-unskilled worker pay ratios, as well as the actual ratios. The centre of the diagram represents a pay ratio of 1:1, the outmoster ring represents a pay ratio of 351:1. Readers wishing a zoom in on the estimated and ideal ratios (the red area and the barely visible blue area in the middle) are referred to Kiatpongsan and Norton (2014).
Survey participants are asked i. how much they think a chairman of large national corporation earns, ii. how much he / she should earn, iii. how much they think an unskilled factory worker earns, iv. how much he / she should earn. The estimated pay ratio is the answer to i. divided by the answer to iii.; the ideal ratio is ii. divided by iv.

Kuziemko et al. (2015) provide survey evidence for potential impacts of information on attitudes to public policies and issues, such as concern about inequality and redistributional preferences.

To our knowledge, the behavioral economics literature has not specifically studied the question most relevant here—namely, consumer willingness to pay for reduced inequality in the production of purchased goods. The several papers in other disciplines that have looked into this topic have found evidence that people are more willing to buy goods which, everything else being equal, are produced by firms with a lower CEO-to-median-worker pay ratio (e.g. Mohan et al., 2015). However, bar one notable exception (Mohan et al., 2018),[3] these

[3]Mohan et al. (2018) provide evidence from an incentivised field experiment suggesting a significant proportion of subjects opt to purchase from retailers with low CEO-to-average employee salary ratio (around 5 : 1) over retailers with a significantly higher ratio (of the

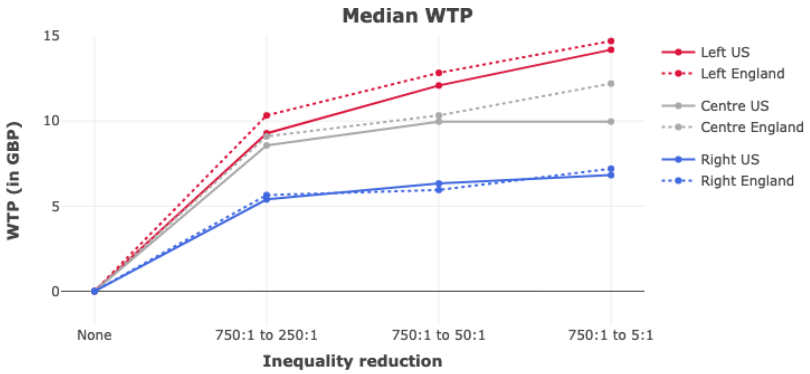

Figure 2: Median willingness to pay (WTP) for indicated inequality reductions, across self-reported political leanings and countries. Inequality is expressed as the CEO-to-median-worker pay ratio.
Data & Figure source: Hill and Lloyd (2021)
The Willingness to pay (WTP) is the value x such that the subject is indifferent between paying £30 for the high inequality good (i.e. 750:1) and paying £30+x for the lower inequality good (e.g. 250:1). Data was collected for representative populations in England and the US.

studies are typically not incentive compatible. Motivated by this, a sister paper (Hill and Lloyd, 2021) ran a choice experiment to elicit willingness to pay for various inequality reductions on representative samples of the English and US populations, using a standard Random Incentive Mechanism to ensure incentive compatibility. As suggested by Figure 2 drawn from that paper, this study finds a significant impact of the inequality involved in the production of a good on subjects' willingness to pay for the good in both populations. Over 89% of subjects were willing to pay a positive premium for some reduction in inequality in their purchased good, with the average premium rising to over £10 (on a £30 good), for the reduction in the CEO-to-median worker pay ratio from $750:1$ to $5:1$. Whilst it may vary with demographics, willingness to pay for reduced inequality was typically positive across the board, and, perhaps most surprisingly, on both sides of the political spectrum (Figure 2).

3 The impact of inequality information: a model

The evidence just reviewed suggests that a significant proportion of the population may be willing to pay moderate amounts for a reduction in extreme levels of income inequality among those involved in the creation of the products they purchase. We shall say that such consumers exhibit *inequality aversion*. Note that individuals that are inequality averse in this sense *fully accept* inequalities that could be justified by considerations such as merit, desert or fairness, and are not willing to pay to eliminate the corresponding inequality levels. They are *only* averse to those levels of inequality that could not, in their opinion, be justified on such grounds. For a population of consumers, some of which are inequality averse, what impact would providing information on the income

order of $700:1$).

inequality across all those involved in the creation of a good or service, for every product on the market, have on overall income inequality? In this section, we show, on a simple partial equilibrium model, that such an information intervention will generally reduce income inequality, will never cause it to rise, will tend to be stronger the more inequality averse the population, and will be socially optimal. We first set out the model (Section 3.1), then establish the main theoretical results (Section 3.2–3.4), before discussing a rough estimation of impact (Section 3.5).

3.1 Model

We limit attention to a single good, which can be produced with varying amounts of inequality, and consider two markets: the market for the good of interest (potentially with inequality information) and the market for the production factors (e.g. labour, capital) contributing to the creation of the good. Firms 'recruit' production factors on the latter market to produce the good, which they then sell to consumers on the former market.

Consumers The goods on the good market are identical except for their associated income inequality levels, which take values in a positive real interval I. Consumers are price-takers, where the price of the good may only depend on the inequality in production, of which they are fully and correctly informed. Each consumer can purchase zero or one unit of the good, so their choice is whether to purchase, and at which inequality level. A consumer k's preferences are represented by a utility function $U_k : (I \cup \{\}) \times \mathbb{R} \to \mathbb{R}$ of the following quasi-linear form:

$$U_k(i, n) = \begin{cases} n + (v_k - \psi_k(i)) & i \in I \\ n & i = \{\} \end{cases} \tag{1}$$

where n is the quantity of the numéraire, v_k—a positive real number—is the 'intrinsic' value of (one unit of) the good for the consumer, independent of inequality considerations, and ψ_k—an increasing real function of I—is the disutility of obtaining the good with inequality score i. $\{\}$ represents the outcome of not obtaining the good, which is normalised to have utility 0. So, for instance, a consumer with endowment $\hat{n_k}$ of the numéraire who pays $p(i)$ to purchase the good produced at inequality level i obtains utility

$$\hat{n_k} - p(i) + v_k - \psi_k(i) \tag{2}$$

The elements v_k and ψ_k are preference characteristics of consumer k. In particular, the latter reflects her willingness to pay to avoid (a given level of) inequality. Given the choice between a unit of the good offered at inequality i' and a unit offered at higher inequality i'', such a consumer would pay a premium of up to $\psi_k(i'') - \psi_k(i')$ for the lower inequality good.

In the light of the evidence set out in Section 2, we assume that ψ_k has the following form:

$$\psi_k(i) = \begin{cases} 0 & i \leq \theta_k \\ \eta_k(i - \theta_k) & i > \theta_k \end{cases} \tag{3}$$

Consumers are insensitive to income inequality scores below a threshold θ_k: such scores are considered possibly justified by considerations such as fairness or merit. This *justifiable-inequality threshold* can capture the fact, found in the empirical data (Section 2, notably Figure 1), that most people accept inequalities below a certain 'ideal' level as potentially justified. They become sensitive above that point: this is reflected in the second clause of (3), where higher inequality leads to higher disutility. Higher η_k reflects more inequality aversion: individuals are willing to pay more to avoid (each given level of) inequality. Consumers with $\eta_k = 0$ are insensitive or neutral towards inequality: they are not willing to pay to reduce any inequalities.

Although we focus on the simple preference form (1) in the bulk of the paper, most of our results hold under more general additively-separable preferences (Appendix A). In particular, they do not require the disutility of inequality to be linear in the inequality level above the threshold (but can incorporate a wide range of shapes), nor that it be additive with respect to the utility for the numéraire (incorporating, for instance, inequality disutility that is relative to the price of the good). Interested readers are referred to Appendix A.1 for further details.

To focus on the effect of inequality aversion, we assume that all consumers have the same θ (justifiable-inequality threshold), but may differ in the inequality aversion η.[4] More specifically, we assume that there exist θ and $K > 1$ levels of η, $\eta_1 > \cdots > \eta_K = 0$ and say that a consumer is of type j if her disutility for inequality is as in (3) with (θ, η_j). We consider a continuum $P = [0, N] \subseteq \mathbb{R}_{\geq 0}$ of consumers, of measure N.[5] A sequence $\mu = (\mu_1, \ldots, \mu_K)$ of positive real numbers for which $\sum_{j=1}^{K} \mu_j = N$ is called an *inequality aversion distribution*. Each such μ represents the distribution of inequality attitude in the population: under μ, μ_j consumers have inequality aversion η_j, for each j. For a given inequality aversion distribution μ, P is partitioned into intervals $[0, \mu_1], \ldots, (\sum_{l=1}^{j-1} \mu_l, \sum_{l=1}^{j} \mu_l], \ldots, (\sum_{l=1}^{N-1} \mu_l, N]$, with jth interval containing all and only consumers of type j (and accordingly, this interval is of measure μ_j). This model can thus account for the varying inequality attitudes across the population noted in Section 2. We use μ^0 to denote the inequality aversion distribution where every consumer is inequality neutral, i.e. $\mu^0_K = N$; $\mu^0_j = 0$ for all $j \neq K$.

Each consumer k chooses (i, n) to maximise (1) under the budget constraint $\hat{n_k} \geq p(i) + n$, where $\hat{n_k}$ is the consumer's endowment. Note that purchasing a good with inequality level i at price $p(i)$ is only preferred to not purchasing if $v_k \geq \eta_k(i - \theta) + p(i)$; we call this the participation constraint. If the participation constraint is satisfied for every i, then the consumer basically maximises (2).

Production factors Production of the good requires two *types* of production factors, which we call L and H. Each production factor may admit a range of *levels*. Income inequality in the production process is driven by the different rates at which the types and levels involved are remunerated. This setup can interpreted in terms of production requiring two sorts of workers—for instance, more vs. less well-educated workers (e.g. managers vs. factory workers)—with

[4] As discussed in Appendix A.1, our results for a more general class of preferences can incorporate consumers with differing θ.

[5] Throughout, we adopt the Lebesgue measure on P.

each sort of worker admitting different talent or skill levels (e.g. differences in the talent levels of factory workers or or of managers). In this case, the relevant difference is that between their salaries. Since firms will not make profits in equilibrium because we consider perfect competition, L could alternatively represent labour and H capital, with the levels in the latter case representing the attractiveness of financing conditions, supply of financing, input of shareholders or owners, and so on. Here, the relevant difference is that between the rates of return on labour and capital. For the purposes of the exposition, we adopt the first interpretation, and speak of labour, wage differences, and so on.

For simplicity, we assume that there is only one skill level for the low type L, but a non-degenerate range of skill levels for the high type H.[6] For notational convenience, we label H-type skill levels by the productivity of workers of that type when matched with an equal number of L-type workers. More precisely, skill level $f \geq 1$ is that for which one unit of H-type labour at this skill level and one unit of L-type labour produce f units of the good. H-type skill levels (or alternatively productivity) thus take values in a real interval $[\underline{f}, \overline{f}] \subseteq [1, \infty)$.

Each worker type and level will be remunerated at a unit wage which depends only on the type and the level. We normalise wages (and prices) and set the wage for L-type labour to 1. In equilibrium, there will be a wage assigned to each hired H-type skill level: by the law of one wage it will be unique and continuous in skills. We model these assignments by a wage schedule: a continuous partial function $w : [\underline{f}, \overline{f}] \to \mathbb{R}_{\geq 0}$.[7] $w(f)$ denotes the (unit) wage for a H-type worker of skill or productivity level f.

To remain largely non-committal on the supply side of the labour market, we assume a sufficient supply of L-type labour at wage 1 to service the population of consumers, and consider a labour supply function $X : [\underline{f}, \overline{f}] \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, where $X(f, x)$ is the supply of H-type f-level labour when the wage offered for that level is x.[8] X is assumed to be differentiable, with $\frac{\partial X}{\partial x} > 0$ wherever X takes non-zero values: for a fixed skill level, labour supply increases with an increase in wages offered. To translate the fact that H-type workers are interpreted as high earners relative to L-type workers, we assume that $X(f, 1) = 0$ for all f—no H-type worker would work for the L-type wage.

Although we only make the aforementioned assumptions on X for most results, at one point we shall consider more structure on the labour supply function. For future reference, we state the following assumption.

Assumption 1. *For all* $f, f' \in [\underline{f}, \overline{f}]$, $x > 0$ *with* $X(f, x)$, $X(f', x) > 0$ *and* $f' > f$:[9]

[6] So the model developed here is naturally read as focussing on managerial pay in relation to a fixed (e.g. median) salary, or capital returns in relation to a fixed salary income. A similar analysis can be conducted for the opposite case of a single H level and a range of L levels—which is most naturally read as focussing on low wages or labour exploitation.

[7] A partial function $w : [\underline{f}, \overline{f}] \to \mathbb{R}_{\geq 0}$ is a function from a subset $X' \subseteq [\underline{f}, \overline{f}]$ to $\mathbb{R}_{\geq 0}$; it is continuous if the corresponding function on X' is continuous. A partial function w can be associated with the (ordinary) function $\dot{w} : [\underline{f}, \overline{f}] \to \mathbb{R}_{\geq 0}$ which coincides with w where defined and equals 0 elsewhere. Integrals of partial functions over the domain $[\underline{f}, \overline{f}]$ are identified with the integral of the associated ordinary function: e.g. $\int_{\underline{f}}^{\overline{f}} w df$ is defined to be $\int_{\underline{f}}^{\overline{f}} \dot{w} df$.

[8] The focus on the consumer side rather than the labour side sets our model apart from several existing studies of income inequality which concentrate on features of workers or suppliers in the labour market; see Section 5 for further discussion of the related literature.

[9] ∇X is the differential of X, and . denotes the scalar product.

$$-\frac{\frac{\partial X}{\partial f}(f,x)\,f}{X(f,x)} \geq 2 \tag{4}$$

and

$$\frac{\nabla X(f,x).(f,1+x)}{X(f,x)} \leq \frac{\nabla X(f',x).(f',1+x)}{X(f',x)} \tag{5}$$

Condition (4) implies that, for a fixed wage, labour supply is lower at higher skill levels. Moreover, it places a lower bound on the skill elasticity of labour—that is, the proportional drop in labour supply due to a unit proportional increase in the productivity of the high-skilled employee, when wages are kept fixed. When proposing the same wage to workers that are 10% more productive, the labour supply drops by at least 20%. For comparison, Card et al. (2018, p551) retains 4 as a typical value for the elasticity of labour supply with respect to the value added of workers; taking value added (in the two-skill level model used in the cited paper) as related to the worker productivity here suggests that this could serve as a first estimate for the skill elasticity of labour, which is consistent with (4).

Condition (5), on the other hand, involves a notion of elasticity involving changes in both the productivity of the hired worker and the wage offered. The left hand side is the elasticity of labour with respect to wage *and* a concomitant equi-proportional change in productivity, at (f,x). To see this, note that the total wage paid to the pair of a L-type and H-type worker is $1+x$ when the H-type worker is paid x. So, for a H-type worker of productivity f earning x, the proportional productivity change due to a increase of f units in the productivity of the hired worker is the same as the proportional total wage change due to a increase of $1+x$ in the H-type wage. $\nabla X(f,x).(f,1+x)$ is thus the change in labour supply brought on by concomitant equi-proportional changes in total wage and productivity. And $\frac{1}{X(f,x)}.\nabla X(f,x).(f,1+x)$ is the proportional change in labour supply per unit matching proportional changes in the total wage and productivity. We call this the *joint wage-productivity elasticity of labour*. The right-hand side term is the joint wage-productivity elasticity of labour at (f',x).

In the light of this, (5) just says that, as the skill level decreases (and hence the labour supply rises), the joint wage-productivity elasticity of labour becomes smaller. So the condition has the flavour of decreasing elasticity for higher levels of labour supply, but involves the joint wage-productivity elasticity, and applies to labour increases due to reductions in the skill level employed. The joint nature of the condition is central, and inhibits comparison with existing literature on the elasticity of labour to wages or skill levels taken separately. For instance, the condition is consistent with constant wage-elasticity of labour at each skill level, as well as constant skill-elasticity of labour at each wage level.[10] In the sequel, Assumption 1 will not be imposed unless explicitly stated.

Firms Firms recruit labour (production factors) on the labour market, and use them to produce goods, which are sold on the good market. We do not focus

[10]Indeed, it is satisfied by the family of separable labour supply functions—i.e. those of the form $X(f,x)=\psi(f).\phi(x)$—with constant skill elasticity of labour at each wage.

on firm market power in either market, and suppose that they are price-takers in both markets, which operate under perfect competition.

To produce the good, each firm recruits one unit of L-type labour and one unit of H-type labour at a single skill or productivity level: its only choice is thus the level of the unit of H-type labour recruited. Firms are fully and correctly informed of each worker's type and level.

The inequality involved in the production of the good by a firm choosing H-type level f is defined as the ratio between the wage paid for the H-type labour with skill level f—$w(f)$ under the wage schedule w—and the wage paid to the L-type worker—namely 1. So inequality associated with a firm hiring skill level f is fully characterised in this model by $w(f)$. The set of inequality levels is thus $I = [1, \infty)$.[11] Consumers in the good market are fully and correctly informed of the production inequality for the good offered by each firm, namely of $w(f)$ for the f chosen by the firm.

Firms choose the recruited skill level to maximise profits, so solve:

$$\arg \max_{f \in [\underline{f}, \overline{f}]} p(w(f)).f - (w(f) + 1) \tag{6}$$

Equilibrium Firms compete on both markets under perfect competition, with free entry. So an equilibrium for an inequality aversion distribution μ is defined as a set of prices $p^*_\mu : I \to \mathbb{R}_{\geq 0}$, specifying the price for each good with a given inequality score in the good market, a wage schedule $w^*_\mu : [\underline{f}, \overline{f}] \to \mathbb{R}_{\geq 0}$ specifying the wage for each skill level in the labour market, and $J^*_\mu : [\underline{f}, \overline{f}] \to \mathbb{R}_{\geq 0}$ specifying how many active firms are recruiting at each skill level, such that:

1. All firms maximise profits (6);
2. All consumers maximise utility (1) under the budget constraint $\hat{n_k} \geq p^*_\mu(i) + n$;
3. Closure of the labour market: $J^*_\mu(f) = X(f, w^*_\mu(f))$ for every $f \in [\underline{f}, \overline{f}]$;
4. Closure of the good market:

$$N' = \int_{\underline{f}}^{\overline{f}} fX\left(f, w^*_\mu(f)\right) df \tag{7}$$

   where $N' \leq N$ consumers purchase the good at some inequality level in equilibrium;

5. Free entry condition: for every $f \in [\underline{f}, \overline{f}]$, $p^*_\mu(w^*_\mu(f)).f - (w^*_\mu(f) + 1) = 0$.

We focus on cases where the labour market is comfortably deep enough to service all consumers whilst keeping within their budget and participation constraints. More precisely, we assume that there exists a wage schedule $\hat{w}$ such that

[11] Recall that by the basic properties of X, the H-type workers are always paid more than the L-type ones, whenever they are employed.

$$\int_{\underline{f}}^{\overline{f}} fX\left(f, \hat{w}(f)\right) df \geq N \tag{8}$$

and, for every consumer k and skill level f:

$$\hat{n}_k \geq \frac{\hat{w}(f)+1}{f} \tag{9}$$

$$v_k \geq \eta_k(\hat{w}(f)-\theta) + \frac{\hat{w}(f)+1}{f} \tag{10}$$

These conditions say that, under this wage schedule, enough goods are produced to service all the consumers (8), and each good satisfies the budget and participation constraints for each consumer, if firms sell at cost ((9) and (10) respectively).[12] In other words, there is a some way to service all consumers within their budget and participation constraints—though, of course, $\hat{w}$ may not satisfy the conditions for equilibrium above.

3.2 Solving the model: wage schedules in equilibrium

We show in Appendix A.2 that any equilibrium wage schedule for an inequality aversion distribution μ is characterised by a sequence of positive real numbers $C_k, \ldots, C_K$, for $0 \leq k \leq K$, and a sequence $\underline{f} \leq \underline{f_k} \leq \overline{f_k} \leq \theta \leq \underline{f_{k+1}} \leq \overline{f_{k+1}} \leq \cdots \leq \underline{f_K} \leq \overline{f_K} \leq \overline{f}$ with:

$$w^*_\mu(f) = \begin{cases} \frac{C_K f - 1}{\eta_K f + 1} & f \in [\underline{f_K}, \overline{f_K}] \\ \frac{C_{K-1} f - 1}{\eta_{K-1} f + 1} & f \in [\underline{f_{K-1}}, \overline{f_{K-1}}] \\ \ldots & \ldots \\ \frac{C_{k+1} f - 1}{\eta_{k+1} f + 1} & f \in [\underline{f_{k+1}}, \overline{f_{k+1}}] \\ C_k f - 1 & f \in [\underline{f_k}, \overline{f_k}] \end{cases} \tag{11}$$

where $\frac{C_{k+1}\underline{f_{k+1}}-1}{\eta_{k+1}\underline{f_{k+1}}+1} \geq \theta$ with $\frac{C_{k+1}\underline{f_{k+1}}-1}{\eta_{k+1}\underline{f_{k+1}}+1} = \theta$ whenever $\underline{f_k} \neq \overline{f_k}$, and $\underline{f_k} = \overline{f_k} = \underline{f_{k+1}}$ whenever $\frac{C_{k+1}\underline{f_{k+1}}-1}{\eta_{k+1}\underline{f_{k+1}}+1} > \theta$. Moreover, these sequences satisfy:

$$\mu_j = \int_{\underline{f_j}}^{\overline{f_j}} fX\left(f, \frac{C_j f - 1}{\eta_j f + 1}\right) df \tag{12}$$

for every $k+1 < j \leq K$ and

$$\sum_{j=1}^{k+1} \mu_j = \int_{\underline{f_k}}^{\overline{f_k}} fX\left(f, C_k f - 1\right) df + \int_{\underline{f_{k+1}}}^{\overline{f_{k+1}}} fX\left(f, \frac{C_{k+1} f - 1}{\eta_{k+1} f + 1}\right) df \tag{13}$$

The interested reader is referred to Appendix A.2 for a detailed discussion of this wage schedule; here, we simply report some properties that will be relevant

[12] Clearly, by (6), if a firm sells at cost, hiring skill level f at wage $\hat{w}(f)$, then it will sell at a price of $\frac{\hat{w}(f)+1}{f}$.

in the sequel. First, the wage w^*_μ can be shown to be strictly increasing in the skill level f, wherever it is defined. Second, the equilibrium wage schedule varies according to the inequality distribution in the population, μ. Finally, there is 'sorting' or 'self-selection' in the following sense: more inequality averse consumers will always buy from firms employing lower skilled workers. (Roughly, consumers of type j will buy from firms employing workers of skill level between $\underline{f_j}$ and $\overline{f_j}$ in equilibrium, and this yields the wage schedule of form (11).) However, the skill levels from which a consumer of given inequality aversion will purchase depend on the whole inequality aversion distribution μ, and not just on her individual inequality aversion.

Although not the focus here, note that a straightforward extension of the results in Appendix A.3 imply generic existence and uniqueness of equilibria (see Remark 1, Appendix A.3).

3.3 Inequality aversion and information provision

Based on the previous characterisation, we can look at how inequality across employed workers in the labour market varies with the inequality aversion of consumers. Each strictly increasing wage schedule w over H-type levels generates a (population-normalised) *wage distribution* g_w across all employed workers, defined by $g_w(1) = \frac{1}{2}$; $g_w(x) = \frac{1}{2}\frac{X(w^{-1}(x),x)}{w'(w^{-1}(x))\int_{\underline{f}}^{\overline{f}} X(f,w(f))df}$ for all $x > 1$; $g_w(x) = 0$ elsewhere.[13] For a wage distribution g and $x \in \mathbb{R}_{>0}$, $g(x)$ is the proportion of the employees with income x. We denote the set of such distributions by D and the cumulative distribution for $g \in D$ by G. In this section, we consider how the inequality in this wage distribution is impacted by information provision. To this end, we shall consider several inequality measures.

Base result Perhaps the simplest proxy for inequality in the labour market is the ratio of the maximum to minimum wage among all workers employed, which we call the *max-min wage ratio.*[14] In the model adopted here, it can be calculated directly from the wage schedule, with virtually no knowledge about the labour supply function X, and hence about the details of the labour market. Our first result concerns this inequality measure, and in particular the effect of increased inequality aversion in the population. To state it, let us say that an inequality aversion distribution μ *Inequality Aversion Dominates* another distribution μ' if, for every $1 \leq j \leq K$, $\sum_{i\leq j}\mu_i \geq \sum_{i\leq j}\mu'_i$. Recalling that lower j correspond to higher inequality aversion, this means that the proportion of the population having inequality aversion higher than a certain level is larger under μ than μ'. This is the standard notion of First Order Stochastic Dominance, applied to inequality aversion. We have the following result (see Appendix A.3 for the proof, and below for some intuition).

[13] To see the derivation of this formula, note that the proportion of employees with income between $\underline{x} > 1$ and $\overline{x}$ is $\int_{\underline{x}}^{\overline{x}} g_w(x)dx = \frac{1}{2}\int_{\underline{x}}^{\overline{x}} \frac{X(w^{-1}(x),x)}{w'\left(w^{-1}(x)\right)\int_{\underline{f}}^{\overline{f}} X(f,w(f))df}dx = \frac{1}{2}\frac{\int_{w^{-1}(\underline{x})}^{w^{-1}(\overline{x})} X(f,w(f))df}{\int_{\underline{f}}^{\overline{f}} X(f,w(f))df}$.

[14] Formally, this is $\frac{\sup(supp(g_w))}{\inf(supp(g_w))}$.

Theorem 1. *If μ Inequality Aversion Dominates μ', then the max-min wage ratio in equilibrium is lower under μ. Moreover, it is strictly lower if and only if the number of consumers purchasing the good at an inequality level higher than θ in equilibrium under μ' is strictly greater than $\sum_{i \geq \bar{j}} \mu_i$ where $\bar{j}$ is such that $\mu_{\bar{j}} \neq \mu'_{\bar{j}}$ and $\mu_i = \mu'_i$ for all $i > \bar{j}$.*

The $\bar{j}$ in the second clause of the theorem is the least inequality averse type at which the inequality aversion distributions μ and μ' diverge. Since less inequality averse types purchase higher-inequality goods (Section 3.2), the condition for strict inequality says that some consumers of this type are paying above the justifiable-inequality threshold.

The theorem attests to the impact of inequality aversion on income inequality in the labour market, under complete information about the inequality involved in the production of goods. Any inequality-aversion increasing shift in the population will reduce overall income inequality—as long as it involves some consumers which are sensitive to the inequality concerned, in the sense that they purchase goods with inequality above the justifiable-inequality threshold. So, for instance, even if the proportion of inequality neutral consumers in the population remains the same, but inequality averse consumers become more inequality averse, this will drive down inequality across the board—as measured by the gap between the highest and lowest incomes.

Information provision Theorem 1 also delivers insight into the potential impact of providing consumers with information about the inequality involved in the creation of the products purchased. A market without inequality information can be modelled as one where consumers' purchasing choices are independent of the level of inequality in goods; in other words, consumers act as if they were inequality neutral ($\eta = 0$). This market can thus be represented by the inequality aversion distribution μ^0 (Section 3.1), under which everyone is inequality neutral. Under information provision, by contrast, the equilibrium will be determined by the actual inequality aversion distribution of the population, μ. So the effect of information provision can be gleaned from the comparison of equilibrium wages in μ^0 and μ. On this comparison, Theorem 1 yields the following corollary.

Corollary 1. *Consider an inequality aversion distribution μ. The max-min wage ratio in equilibrium is lower under μ than under μ^0. Moreover, it is strictly lower if and only if there are strictly more consumers purchasing the good at a price above θ in equilibrium under μ^0 than inequality neutral consumers under μ.*

Providing consumers with product-level inequality information will thus never lead to an increase in inequality in the labour market. Moreover, it will lead to a strict decrease in inequality whenever the divergence from inequality neutrality involves some consumers that purchase goods at inequality levels to which they are sensitive, i.e. above the justifiable-inequality threshold θ. Current data—notably the finding that almost 90% of subjects are willing to pay a positive amount for some inequality reduction, and hence are not inequality neutral (Section 2), combined with the fact that all bar one S&P500 company have a CEO-median pay ratio above the ISSP median US ideal value (Section 2 and AFL-CIO, 2020)—suggest that this is typically the case, and hence that

information provision will have mitigating impact on inequality. Theorem 1 adds to this corollary the message that its impact is larger the more inequality averse the population.

Other inequality measures Whilst the max-min ratio is particularly convenient in the context of the current model, the general point does not rest on the use of it as a measure of inequality. For a more general treatment, we consider two families of inequality measures, where an inequality measure is a function $\iota : D \to \mathbb{R}_{\geq 0}$. For each $0 \leq a < 0.5$ and $0 < b \leq 0.5$, let $\iota^{quant}_{a,b}(g) = \frac{G^{-1}(1-a)}{G^{-1}(b)}$ and $\iota^{share}_{a,b}(g) = \frac{\int_a^1 G^{-1}(\tau)d\tau}{\int_0^b G^{-1}(\tau)d\tau}$.[15] The *quantile family* contains all inequality measures $\iota^{quant}_{a,b}$ (for $0 \leq a < 0.5$ and $0 < b \leq 0.5$): these measures reflect the ratio of the wage earned by the a% highest paid individual against that of the b% lowest paid one. Examples of measures in this family include the 20:20 ratio and the 90%-10% quantiles ratio reported by the OECD (2021), as well as the max-min ratio. The *share family* contains the measures $\iota^{share}_{a,b}$, which report the ratio of the shares of income received by the top a% vs. the bottom b% of the population. The Palma ratio—the ratio of the income share of the richest 10% to the poorest 40%—belongs to this family. Several researchers (e.g. Atkinson et al., 2011; Alvaredo, 2018) have advocated looking at quantiles and shares in the study of financial inequalities, suggesting the relevance of these families. The following result provides extensions of the previous findings to them.

Theorem 2. *Consider an inequality aversion distribution μ. There exists $0 \leq a' < 0.5$ such that for each inequality measure $\iota = \iota^{quant}_{a,b}$ or $\iota = \iota^{share}_{a,b}$ with $0 < b \leq 0.5$ and $a \leq a'$:*

($\star$) $\iota(g_{w^*_\mu}) \leq \iota(g_{w^*_{\mu^0}})$, with strict inequality if and only if there are strictly more consumers purchasing the good at a price above θ in equilibrium under μ^0 than inequality neutral consumers under μ.

Moreover, if X satisfies Assumption 1, then ($\star$) holds for every inequality measure in both the quantile and share families.

The central message of the previous results thus hold under a large range of currently-used inequality measures: providing product-level inequality information to a population with sufficiently many inequality averse consumers will have a moderating effect on inequality in equilibrium. In particular, the first clause confirms that, when focussing on the gap between top salaries and those in the bottom half of the distribution, overall inequality will typically be lower under information provision. As noted above, much current discussion of income and wealth inequality concentrates on this gap. Since the details of the labour supply function X determine the middle of the income distributions, a more general result involving measures looking at quantiles or shares further down the distribution cannot be had under the very minimal assumptions made up to this point. However, the second clause of the Theorem shows that the drop in inequality under information provision holds for all these inequality measures, whenever the productivity and joint wage-productivity elasticities of labour have the properties stated in Assumption 1. If, as suggested in Section 3.1, this can

[15] The inverse G^{-1} of a cumulative distribution G is defined by $G^{-1}(t) = \sup\{x : G(x) < t\}$.

reasonably be believed to be the case in a market, then information provision moderates inequality, no matter which of the measures in these families is used.

The results in this section are based on a simple insight. Consumers with little or no inequality aversion tend to prefer cheaper, higher inequality goods, so they support a demand for highly productive workers, which *ceteris paribus* can produce goods at lower unit cost. When there are many such consumers in the population, this translates to a significant demand for higher skilled or more productive workers, and drives up their wages. By contrast, inequality averse agents are willing to pay for lower inequality products, and so buy from firms employing lower skilled H-type workers. When a population has more of such consumers, this shifts the labour demand towards lower skilled H-type workers and away from higher skilled ones. This deflated demand leads to a drop in the highest wages (and an increase in mid-range wages), and hence less income inequality.

3.4 Social Efficiency

One might worry that introducing inequality information may lead to social inefficiencies.[16] Considering efficiency under the perfect competition, perfect inequality-information model set out above will allow evaluation of the extent to which such fears are driven by inequality information *per se,* or by other factors such as market imperfections. Indeed, since there is perfect competition, one might expect the First Welfare Theorem to hold, so that the equilibrium is Pareto optimal. Whilst it does not follow from the standard version of this theorem, because the wage in one market is a factor of differentiation in another, it will nevertheless turn out that there is social efficiency: the equilibrium is Pareto optimal.

To consider the issue of social efficiency, we first define allocations in our markets. An allocation (in the goods market) is a pair consisting of

- a (measurable) function $c : P \to (I \cup \{\}) \times \mathbb{R}_{\geq 0}$ specifying, for each consumer $k \in P$, the inequality level of the good received (or $\{\}$ if no good is received), $c_1(k)$, and the quantity of the numéraire obtained, $c_2(k)$.
- a (measurable) function $q : I \to \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}$ specifying, for each inequality level $i \in I$, the total quantity of the good produced with inequality i, $q_1(i)$, and the total cost of that production in numéraire terms, $q_2(i)$.

Moreover, the production allocation q must be generated from an assignment, to each skill level in the labour market, of how many firms hire at that skill level and the wage offered, i.e.:

- there exists a (measurable) function $r : [\underline{f}, \overline{f}] \to \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}$, specifying for each skill level $f \in [\underline{f}, \overline{f}]$, how many firms hire workers of skill level f, $r_1(f)$, and the wage offered to workers of skill level f, $r_2(f)$, such that:
  - $q_1(i) = \int_{\{f : r_2(f) = i\}} f r_1(f) df$ and
  - $q_2(i) = \int_{\{f : r_2(f) = i\}} (i+1) r_1(f) df$ for all $i \in I$.

[16] See Section 5 for a brief discussion of related literature.

For an allocation to be *feasible*, it must satisfy the market clearing conditions, namely: for each $i \in I$

$$\int_{c^{-1}(i\times\mathbb{R}_{\geq 0})} dk = q_1(i) \tag{14}$$

(i.e. the total amount of good consumed at inequality level i is equal to the total amount produced) and

$$\int_P c_2(k)dk = \int_P \hat{n}_k dk - \int_I q_2(i)di \tag{15}$$

(i.e. the total amount of numéraire consumed is equal to the total endowment minus the production costs) and, for all $f \in [\underline{f}, \overline{f}]$

$$X(f, r_2(f)) = r_1(f) \tag{16}$$

(ie. the total labour supply at every skill level equals the total labour demand). Any competitive equilibrium generates a feasible allocation in this sense.

We adopt the standard notion of Pareto optimality for a continuum of consumers (e.g. Hammond, 1979). A feasible allocation is *Pareto optimal* if there is no alternative feasible allocation under which no consumer has strictly lower utility and a set of consumers of strictly positive measure have strictly higher utility.

Theorem 3. *Any allocation generated by a competitive equilibrium is Pareto optimal.*

So providing inequality information in a perfectly competitive market leads to a socially efficient outcome in terms of the consumers' preferences, and in particular their preferences concerning the inequalities involved in the production of the goods they consume. By contrast, the equilibrium in a market lacking inequality information will typically not be Pareto optimal.

Proposition 1. *Consider any inequality aversion distribution μ such that there are strictly more consumers purchasing the good at a price above θ in equilibrium under μ^0 than inequality neutral consumers under μ. If* $\inf\left\{f \in [\underline{f}, \overline{f}] : w_{\mu^0}(f) > 0\right\} > \underline{f}$*, then any allocation consistent with $w^*_{\mu^0}$ is Pareto dominated.*

This proposition focusses on the cases where information provision leads to a drop in overall inequality (the first condition is drawn from the results in the previous section). Essentially, it amounts to the observation that, whenever there are some skill levels at the lower end of the scale that are not employed in equilibrium under μ^0, substituting firms employing such skill levels for some firms using higher skill levels will constitute a Pareto improvement (see Appendix A.5 for details). Although such a switch will typically involve productivity losses—because it assigns production away from the most productive workers—it may nevertheless be a Pareto improvement because the impacted consumers are willing to pay the increased costs brought about by lower productivity in order to obtain lower inequality.

Summing up, information provision guarantees Pareto efficiency, whereas allocations in the absence of information, though perhaps more productive, will typically not be efficient by the lights of consumers' inequality-sensitive preferences.

3.5 An impact estimation

The previous results show that information provision does have an impact on inequality. However, they give little indication of the size of this impact, which will depend on a range of factors, including the distribution of inequality aversion in the population and the supply of various skill levels in the labour market. We now perform some rough calculations to get an idea of an approximate potential size of the impact.

As concern consumer preferences, following the survey evidence on 'ideal' inequality levels (Section 2), we set $\theta = 10$. Given the experimental evidence that roughly 90% of the (US) population exhibits some degree of inequality aversion (Section 2), we consider economies where 10% of consumers are inequality neutral, whilst the rest have the same inequality aversion η. We compare how wage inequality, as measured by the max-min ratio, changes as the inequality aversion among this subpopulation increases from the $\eta = 0$ case modelling no inequality information.

For the labour supply function for H-type workers, we adopt the fairly standard form $X(f,x) = A_f \mathcal{P}(f)(x - b_f)^{\beta_f}$ from the literature, where $\mathcal{P}(f)$ is the proportion of workers of skill level f, b_f are baseline wage levels, β_f determine the wage elasticity of labour supply at skill level f, and A_f are (potentially skill-dependent) constants. (See e.g. Card et al., 2018, Sect V for a foundation in terms of worker preferences.) For all f, we adopt the typical value of 0.10 for β_f from Card et al. (2018) and set b_f to the justifiable-inequality threshold of 10. As concerns $\mathcal{P}(f)$, note firstly that in our model, worker productivity coincides with the employing firm's sales in equilibrium. To the extent that the latter is a fairly reasonable proxy for firm size—and indeed, the only performance-related one that can be directly mapped into our model (Gabaix and Landier, 2008, Sect III.A)—one can use the distribution of firm sizes as a proxy for the distribution of worker productivity levels. Given the evidence in favour of Zipf's law for firm sizes (Gabaix and Landier, 2008; Gabaix, 2016), this yields a distribution of worker productivity levels with density approximately $\mathcal{P}(f) \sim f^{-2}$.[17] Normalising and using the same A_f for all f, this gives $X(f,x) = A \max\left\{(x-10)^{0.10}, 0\right\} \frac{f^{-2}}{\int_{\underline{f}}^{\overline{f}} f^{-2} df}$.

Plugging this function into the equations from Section 3.2 and taking $[\underline{f}, \overline{f}] = [1, 1000]$, $N = 40000$ and $A = 5000$—i.e. the population of consumers to be 40 times larger than that of skill levels, and 8 times larger than that of workers (proxied by A)—we calculate estimates of the max-min wage ratio for various values of the inequality aversion of the inequality-sensitive section of the population, η. These are plotted in Figure 3. Note that, by the choice of the value of N and the other parameters, the max-min ratio at $\eta = 0$ is around 1600, which is the order of current highest CEO-to-median pay ratios among S&P500 firms (AFL-CIO, 2020).[18]

[17] Under Zipf's law, $P(F > f) \simeq Bf^{-1}$ for some constant B, where F is the worker productivity level, yielding the density in the text.

[18] The form displayed in this graph continues to hold under different assumptions about the size of the inequality-neutral subpopulation and the other parameters, though the max-min wage ratio values, and in particular the 'match' to S&P500 data, may depend on parameter choice. In that sense, the qualitative conclusion of this exercise—that information provision promises to have a significant effect on income inequality—is fairly independent of the specific assumptions made here.

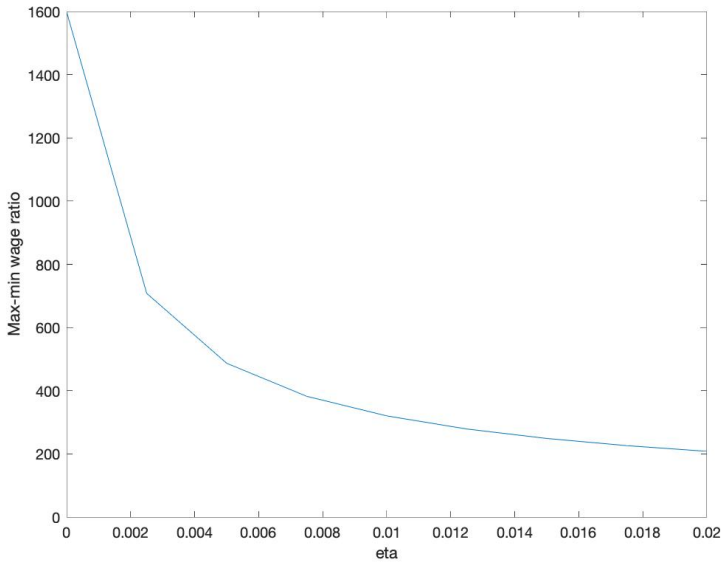

Figure 3: Max-min wage inequality vs. η, for markets where 90% of consumers have inequality aversion η and the rest have inequality aversion 0.
Calculated using (11) and (12), with $[\underline{f}, \overline{f}] = [0, 1000]$, $X(f, x) = 5000 \max\left\{(x-10)^{0.10}, 0\right\} \frac{f^{-2}}{\int_{\underline{f}}^{\overline{f}} f^{-2} df}$ and $N = 40000$.

Figure 3 suggests that even moderate amounts of inequality aversion may have significant effects on the degree of salary inequality. Even under an η of 0.005 among 90% of the population—which corresponds to them being willing to pay \$5 to fully eliminate an inequality of order 1000—full information will cut the overall inequality in equilibrium by more than half. The Hill and Lloyd (2021) data suggests a population-level average η of around 0.0175:[19] if 90% of the population were this inequality averse, information provision would divide the equilibrium inequality by six.

4 Implementation

The previous analysis suggests that blanket inequality reporting in the (goods) market could reduce inequality whilst enhacing social efficiency. It is worth stressing some central conditions for the success of such an intervention. A first, which has already been discussed and finds support from existing studies (Section 2), is a degree of inequality aversion among at least a section of the population. A second condition is that consumers understand the inequality report provided: if the report 'doesn't speak to' a consumer, she will not be able to incorporate it into her purchasing decisions. Thirdly, of course, consumers

[19] Regressions on their data of the WTP against inequality reduction in terms of CEO-to-median worker pay ratio reveals a coefficient of around 0.0175: this coefficient can serve as a proxy to the η in the model used here.

need to trust the source of inequality information: with no possibility of 'on the spot' verification, they need to adopt the reports as their beliefs about the goods on offer.

These conditions may have consequences for the options for the implementation of of product-level inequality information provision—an issue to which we now briefly turn.

4.1 Two phases of implementation

Information works—if and when it does—due to market mechanisms, so no information about individual preferences or assumptions about social ones are required. All that is needed is information about actual inequalities associated to the goods on the market. Given this, the challenge of implementation can naturally be separated into two parts: information collation and information provision.

Whilst part of the information required to calculate the inequality involved in the creation of a good or service is available to governments (e.g. in tax returns), at least in many developed countries, there is no centralised place where the inequality associated to every good or service on the market is collated and made publicly available in an objective, transparent way. Effectiveness of the intervention—and in particular, the previously noted need for trust in the information provided—depends on oversight of information collation and verification being assigned to a publicly trusted body. This could be a government body, but need not be.[20] It could, for instance, be a specifically created organisation, drawing on the competences of academic, governmental, non-governmental and business actors.

The other phase of implementation involves the provision or 'administering' of inequality information. For the sake of effectiveness—and in particular given the importance of consumers being able to understand and use the information—the information needs to be delivered in an appropriate, easy-to-use, easy-to-understand, accessible form at the point of purchase. Moreover, as emphasised above, it should be provided for all products, which means that voluntary labelling (such as CSR or eco-labels) cannot properly implement the proposal. Again, one possibility would involve government involvement: regulation could force all firms to inform customers of the inequality involved in a good or service, much as nutritional information for foodstuffs and origin labels for a range of products are mandatory in many markets. That said, current technology affords the possibility of a regulatorily lighter option, notably in the form of a mobile phone application which presents, on scanning the barcode of a product, its inequality report. Such an app could be run by any organisation; what counts is that the inequality figures reported are drawn from the database established and certified as described above.[21] Similar point-of-purchase applications giving, say, nutritional information about foodstuffs already exist.[22]

[20] Indeed, whilst government collaboration would be useful, insofar as they often can verify relevant earnings information, a single government will often not have all of the required information, especially for international companies and goods with cross-border production lines.

[21] Indeed, one could even imagine several apps presenting the same information, just as there are several weather apps drawing their information from the same source.

[22] One example is `https://yuka.io/en/`.

4.2 Inequality measures

These brief remarks still leave open the question of which measure(s) of inequality to report. A range of inequality measures have been developed and studied (e.g. Lambert, 2001; Chakravarty, 2009; Cowell, 2011), though relatively few with a mind to capturing the inequality involved in the creation of a product, as opposed to say the inequality in a country, or with the aim of public information, by contrast with guiding policy or furthering economic knowledge. It thus seems worth setting out some principles for the development of an appropriate measure of inequality for presentation to consumers at the point of purchase.

First of all, given the importance of public trust in the inequality reports (Section 4), it would seem desirable to use measures transparently based as far as possible on verifiable, *objective* information. This consideration pleads for inequality measures which rely as little as possible on difficult-to-access information, unverifiable self-reports or subjective judgements, or opaque treatments or weightings.[23]

Secondly, for information to impact income inequality 'across the board', the inequality reported to consumers should be as *exhaustive* as possible, ideally encompassing everyone contributing to the existence of the good or service on the market. This includes notably the stages of financing, conception, management, production, transport, marketing and sale. Hence product-level inequality should typically be used rather than firm-level inequality, when the two do not coincide.[24]

Finally, to incorporate inequality information into their decision making—and connect it to the basic intuitions and opinions on social justice that presumably underlie inequality attitudes—people need to understand it. Inequality thus should ideally be presented in a way that can be grasped easily and quickly, without specialist or theoretical knowledge. It must also be clear and unambiguous in meaning, to avoid undermining trust. These considerations plead in favour of *conceptual simplicity* in the inequality measure.

Simple inequality measures satisfying these constraints fairly well include many of those discussed in Section 3.3, when applied to the hourly revenue among all those involved in the financing, conception, management, production, transport, marketing and sales of the good. For instance, the max-min ratio is clearly objectively calculable on the basis of tax returns and salary slips, and, under the aforementioned application, exhaustive. It is conceptually easy to understand and grasp: everybody knows what it means for top management or owners to earn 1000 times more than the least well-paid worker. Indeed, experimental evidence suggests that the simplicity of measures such as these may be asset when it comes to inequality information provision. Hill and Lloyd (2021) find that subjects' willingness to pay for inequality reduction is typically larger when the inequality is reported as the CEO-to-median worker pay ratio as compared to the Gini Index, and that most subjects prefer inequality to be

[23] CSR or ESG ratings provide a cautionary tale on this score, with increasing recognition of the need for transparent, objective reporting standards to foster consumer and investor confidence.

[24] The focus on product-level inequality implies looking across all those firms involved in creating a good. In terms of the research on firm-level inequality, and the extent to which income inequality is driven by inequality within firms or firms (e.g. Mueller et al., 2017; Song et al., 2019), the proposal is thus closest to recent work on production networks (e.g. Huneeus et al., 2021) which suggest their importance as a determinant of income inequality.

reported using the former measure, considering it easier to understand and more informative. This suggests that trade-offs may have to be made between attractive axiomatic properties of inequality measures and the conceptual simplicity needed for consumers to effectively connect with the information. Further experimental research may provide insight into the contours of such trade-offs.

5 Discussion

5.1 Inequality information and some related economic literature

There is a significant literature in economics on the scale and sources of inequality, and the importance and role of information. We now make some brief remarks on connections with, and differences from, parts of this literature.

For consumers sensitive to the inequality involved in the creation of a good, the level of inequality is a 'credence quality' in the sense of Darby and Karni (1973)—it is expensive to judge, even after purchase. Such qualities naturally give rise to information asymmetries (firms typically know more about the inequality involved in production than consumers). Accordingly, much focus has been on the market effects of such asymmetries, and signalling techniques that firms could use to differentiate themselves. For instance, voluntary labelling policies (e.g. eco-labels, CSR) have been analysed in such a perspective (e.g. Baksi and Bose, 2007; Crifo and Forget, 2015; Manili, 2021). By contrast, universal inequality reporting implies complete removal of the information asymmetry and the associated market effects.

Under universal information, the inequality involved in production is a factor of differentiation in the goods market (e.g. Tirole, 1988). As noted in Section 3.2, consumers self-select according to the inequality levels of the goods, and hence the skill levels of the workers employed to produce them. Note also that in equilibrium goods are sold at cost and consumers with different degrees of inequality aversion prefer different inequality levels, so this is not a case of vertical differentiation in the sense of Shaked and Sutton (1983). Under free entry, as we have shown, markets remain competitive.

Policy proposals concerning income inequality—including evaluation of the need for policy—are sometimes suggested by mechanisms that are purportedly responsable for it. The literature on potential mechanisms driving the rise in income inequality is too large to survey here. For illustration, one part focusses on the rise of CEO salaries in the past decades, with suggestions that it could be due to incentivisation considerations in the face of moral hazard, managerial entrenchment or the structure of the firm-CEO matching at the upper tail of the talent distribution (e.g. Edmans and Gabaix, 2016 and references within). Rather than tapping into a theory about why there is upward pressure on income inequality, the current proposal focuses on a potential reason why the counterweight downward pressure is so weak: namely that inequality is not incorporated into the market and the consumption decisions of those who care about it. An analogy with pollution may be enlightening here. To the question 'Why has air pollution increased so much over the last two centuries?' one can cite upward pressures, such as technological change or population growth, as well as the lack of potential downward pressures, such as the fact that pollution

is and has been an externality in many markets. Whilst much of the aforementioned literature on inequality examines the (analogue of the) former sorts of reasons, the current proposal is inspired by reasons of the latter sort.

Although perhaps neglected recently,[25] the conception of inequality as an externality has a long history in economics (Thurow, 1971). However, to the best of our knowledge, the specific information intervention proposed here, which follows easily from such a perspective, has not been extensively studied to date. Indeed, the current proposal—to 'internalise' the inequality externality by universal information provision—differs from more classic interventions targetting externalities, such as Pigouvian taxes. It is most closely related with property-rights or tradeable-permit approaches. Indeed, one could reframe the proposal in terms of a particular allocation of special 'inequality-in-production' property rights. To produce a good with a given inequality level, a firm must acquire a permit to employ that inequality level in the production of that good. Since they specify the inequality level allowed, lets call these *specified* permits. Such permits are *non-amalgamable*: two 'medium-inequality' permits for a good only allow the firm to produce two units at that inequality level; they do not authorise it to produce one unit of the good at a higher inequality level. Whilst only specified permits can be traded, each consumer is allocated, for each unit of good purchased, a *specifiable* inequality-in-production permit for that good: a 'blank deed' that she must 'fill in' with the inequality level to which it gives rights before selling it on the market. So the nature of the permit—the inequality level to which it gives a right—is determined by the consumer prior to sale. It is clear that this market for goods and inequality-in-production permits is basically equivalent to the market set out and studied above: the inequality level at which a consumer purchases the good maps into the inequality level she puts on the permit she sells; the price at which the consumer purchases the good at a certain inequality level is the result of paying the market price for the good and receiving the proceeds of the sale of her specified inequality permit. So the equilibrium is the same, and the results carry over.

This reframing brings out several points on which the proposal differs from typical property-rights or tradeable-permit approaches to externalities. First, there is a simple allocation mechanism: according to good purchase. In particular, unlike standard marketable-permit (or 'cap and trade') approaches in, say, environmental policy, there is no need for a social planner or regulatory authority to decide on the optimal aggregate amount of inequality. Second, the permits here are non-amalgamable, in contrast to carbon markets for instance, where a firm can buy lots of permits from different actors to pollute more in the production of the same quantity of a good. Third, the allocation of specifiable permits that must be specified before sale has not, to our knowledge, been explored previously. These differences all contribute to clarifying each consumer's *responsibility* for inequality: she alone specifies the inequality level on the permit she sells, and is ensured that it will result in the production of at most one unit of the good at that level. As such, they tie into two previously mentioned contrasts with typical approaches, which are worth recalling.

One is the reliance on consumer preferences *concerning the inequalities in the production of the goods they purchase* (or, under the property-rights re-

[25] An notable exception is Støstad and Cowell (2021), who study the consequences of such a conception for optimal taxation.

framing, the inequality on their specified permit), rather than their preferences concerning the *overall level of inequality in society*. Since the level of inequality (or equality) in society could be considered a public good, and hence a 'nondepletable externality', standard analysis of property-rights or tradeable-permit allocation could be applied drawing on the latter preferences, with familiar related issues (e.g. Baumol and Oates, 1988).[26] By contrast, consumer preferences over the overall inequality in society play no role in our analysis, and inequalities associated to purchased goods are closer to private goods, a fact which is central to the social efficiency of the information provision intervention. Of course, as already emphasised, the intervention only has an impact if consumers' preferences are sensitive to these latter inequalities: existing empirical evidence, documented in Section 2, suggests that a significant proportion of them are.

The other specificity of our approach is the reliance on consumer inequality preferences *and nothing else*: in particular, there is no role for a social planner (beyond ensuring the proper functioning of the market). By contrast, a tradeable-permits approach to inequality would require the social planner to determine the optimal aggregate amount of inequality (that is, in the model above, the aggregate quantity of goods produced at each inequality level). Such a quota reflects the sorts of inequality versus productivity tradeoffs mentioned in the Introduction, and naturally poses questions of how the social planner can and should set these values. As noted previously, the current proposal avoids such issues.

5.2 Information beyond income inequality

In the light of the contrast with existing approaches to externalities, one might wonder whether the approach set out here could be applied to externalities other than income inequality. Certainly much of the theoretical analysis (Section 3) seems extendable to other externalities, and would seem to hold if income inequality were replaced by capital-to-labour share of proceeds, the lowest wages paid by the firm (or some other indicator of the degree of offshoring, dumping or unfair wages to low-paid workers), or income inequalities across gender or race in the firm, to name but a few examples. Similar theoretical points—impact if there is consumer sensitivity to these issues, whilst maintaining social efficiency—would thus hold for information provision on these issues. That said, for any issue, the potential for the proposal to be effectively applied to a given externality will ultimately depend on the extent to which the conditions noted at the beginning in Section 4 hold. Whilst there may be reason to suspect that these conditions hold for the previous examples, other cases are less straightforward.

A case where they could matter concerns universal information provision on the global-warming-related impacts stemming from the production of a good. Such information possibilities already exist, for instance in carbon footprint reporting,[27] and though they are typically voluntary in many sectors and re-

[26] In particular, a pure property-rights approach (with no aggregate inequality quota) will fail to be optimal (e.g. Mas-Colell et al., 1995); an approach involving tradeable permits and an aggregate quota may fail to be optimal in the presence of uncertainty (e.g. Weitzman, 1974).

[27] A carbon footprint is usually defined as the total emissions caused by an individual, event, organization or product, expressed as carbon dioxide equivalent.

gions, recent regulation is starting to impose blanket reporting, for instance in France.[28] Information interventions in this domain may face a significant challenge concerning consumer understanding (Section 4). Whilst in the case of income inequality, the intervention can tap into existing intuitions and opinions about social justice, it is less clear whether people have sufficiently developed views about atmospheric processes to incorporate, for instance, CO_2 emissions data into purchasing choices. Of course, recognising a challenge does not mean considering it insurmountable. It may well be possible to develop 'global warming impact measures' that are easily understandable, whilst also satisfying the other conditions set out above, such as objectivity. Moreover, understandability is a relative concept, depending on common knowledge in the community; one might thus expect that improvements in climate awareness and education may enhance the effectiveness of previously incomprehensible information.

Finally, whilst this paper has concentrated on information directed at consumers, similar questions have been studied for other 'targets'. For instance, the incorporation of Environmental, Social, and Governance (ESG) information into investment decisions involves an analogous information-based strategy aimed at investors, and recent papers have connected ESG-sensitive investment to investor preferences (e.g. Pedersen et al., 2021). To take another example, Card et al. (2012) focus on the impact of inequality information on the satisfaction and job-search intentions of employees.

6 Conclusion

This paper proposes universal information provision as a means of moderating income inequality, whilst retaining social efficiency. Current empirical evidence suggests a significant mismatch between current levels of income inequality and peoples' perceived and ideal levels; moreover, it suggests that many would be willing to pay more for a reduction in the income inequality involved in the creation of the goods they purchase. Tapping into these facts, we show on a simple model that informing all consumers about the inequality involved for each good on the market will lead to a drop in income inequality, even for only moderate levels of aversion to inequality among a limited section of the population. Higher willingness to pay to avoid large inequality levels in the population tends to lead to less income inequality in equilibrium. Moreover, information provision re-establishes social efficiency, incorporating in particular the inequality dimensions of consumer preferences.

Turning to implementation, we set out some principles for the choice of inequality measure to report and make some suggestions for information collation and provision.

This paper only focuses on the economic dimension of the proposed intervention—and in particular its impact on inequality levels in equilibrium—but this may not be the only one. In particular, as with any information intervention, there is a potential political dimension. Inequality information can correct misperceptions, which, as noted in Section 2, are widespread. It can improve awareness of the issue. Moreover, to the extent that it relates inequality levels to consumer choice, it involves an empowerment of citizens on this issue.

[28] Under the "Climat et Résilience" law passed in Spring 2021, disclosure of environmental information will become obligatory for a range of goods.

A Extensions and proofs

A.1 General inequality-attitude preferences

We prove the main results for a model that is more general than that presented in the bulk of the paper. The firm and worker structures are the same; the only difference is the use of more general forms of the utility function for consumers. More specifically, each consumer k has a utility function of the following additively separable form:

$$U_k(i,n) = \begin{cases} \xi_k(n) + (v_k - u_k(\mathcal{I}(i))) & i \in I \\ \xi_k(n) & i = \{\} \end{cases} \tag{17}$$

where v_k, as in the model in Section 3.1, is the 'intrinsic' value of the good, ξ_k—the utility function over the numéraire—is strictly increasing and twice differentiable, u_k—the disutility of inequality—is an increasing, twice differentiable function, and $\mathcal{I}(i)$ is the 'justifiable-threshold-corrected' inequality, given by:

$$\mathcal{I}(i) = \begin{cases} 0 & i \leq \theta \\ i - \theta & i > \theta \end{cases} \tag{18}$$

We assume without loss of generality that $u_k(0) = 0$ (the disutility of no inequality is zero). Note that, although the same θ is involved for all consumers, consumers with higher justifiable-inequality thresholds can be modelled by u_k functions which take the value zero up to a certain (higher) level (this is accommodated since these utility functions are not assumed to be strictly increasing).

Clearly, the utility function presented in Section 3.1, Eqns (1) and (3) is the special case where ξ_k is the identity and $u_k(x) = \eta_k x$. However, the functional form (17) is considerably more flexible, accommodating a range of 'shapes' of the disutility of inequality, including an higher sensitivity to inequality increases at higher inequality levels, as well as the opposite more acute sensitivity to changes at low inequality levels. Note also that using log utilities, this form encompasses inequality disutility that is relative to wealth or the price of the good (i.e. multiplicative), rather than absolute (additive).[29]

Define the net utility of a payment for the good in numéraire units, $\bar{\xi}_k$, by $\bar{\xi}_k(x) = \xi_k(\hat{n}_k - x)$. We assume that consumers share the same differences in (not necessarily linear) utility for numéraire payments, i.e. there exists a strictly decreasing and twice differentiable $\bar{\xi}$ with $\bar{\xi}(x) - \bar{\xi}(x') = \bar{\xi}_k(x) - \bar{\xi}_k(x')$ for all $x, x' \in \mathbb{R}_{\geq 0}$ and every consumer k. (It follows that $\bar{\xi}' = \xi_k'$ for all k.) Clearly this holds for the special case presented in Section 3.1. We say that consumer k is *more inequality averse* than k' if, for every pair of inequality levels $i_1 > i_2$, the difference in k's disutility between them is larger than for k': i.e. $u_k(i_1) - u_k(i_2) \geq u_{k'}(i_1) - u_{k'}(i_2)$. When this holds with strict inequality for some i_1, i_2 we write $u_k >_{I.A.} u_{k'}$. More inequality averse consumers obtain a sharper jump in disutility from any increase in inequality. Clearly, for the

[29] For instance, taking ξ_k and u_k to be the appropriate multiple of logarithms, it is clear that the utility function

$$U_k(i,n) = \begin{cases} An^{\alpha}.\mathcal{I}(i)^{\beta} & i \in I \\ n^{\alpha} & i = \{\} \end{cases}$$

characterises preferences belonging to the family represented by (17) and to which the results proved below apply.

special case considered in the paper, higher η_k implies more inequality aversion in this sense.

To formulate the generalised version of Theorem 1, we assume that all consumers are ordered according to inequality aversion: i.e. there exists $u_1 >_{I.A.} \cdots >_{I.A.} u_K$ where u_K is the constant function taking the value zero, such that each consumer's utility over inequality is given by one of these functions.[30] A consumer of type j has (dis)utility for inequality u_j. Inequality aversion distributions and other related notions are defined as in Sections 3.1 and 3.3 with this notion of consumer type. Given that, the formulations of Theorems 1 and 3 in the context of this general model are identical to those in Sections 3.3 and 3.4.

We first present an analysis of the wage schedule and then prove our results, mostly under this general model. The versions for special case presented in the paper follow immediately.

A.2 Derivation of equilibrium wage schedule

General case

We first derive the equilibrium wage schedule under the general utility form (17). As a point of terminology, we say that a skill level $f \in [\underline{f}, \overline{f}]$ *services* a consumer of type j in equilibrium if there is strictly positive demand for goods with inequality $w^*(f)$ among consumers of type j.

The average cost of production of a good of inequality $w(f)$ is $\frac{w(f)+1}{f}$. By standard reasoning, in equilibrium if there is any demand for goods at inequality level $w^*(f)$, then $p^*(w^*(f)) = \frac{w^*(f)+1}{f}$.[31] So, in equilibrium, wages and prices are connected.

For a consumer of type j faced with prices p and wages w, the FOC for an interior solution above the threshold θ are given by:

$$-\bar{\xi}'(p(w(f))\frac{dp}{d(w(f))} = -\bar{\xi}'(p(w(s))\left(\frac{1}{f} - \frac{p(w(f))}{fw'(f)}\right) = -u_j'(w(f)-\theta) \quad (19)$$

Plugging in the form of equilibrium p^*, this can be rewritten as:

$$(w^*)'(f) = \frac{w^*(f)+1}{f\left(1 - \frac{u_j'(w^*(f)-\theta)}{\bar{\xi}'(\frac{w^*(f)+1}{f})}f\right)} \quad (20)$$

Because ξ, u_j are twice differentiable, $\bar{\xi}'(x) < 0$ and $u_j'(x) \geq 0$, the functional on the right hand side (considered as a functional of f and $w^*(f)$) is uniformly Lipschitz continuous in f and $w^*(f)$. Hence, by the Picard-Lindelöf Theorem, for any initial value for w^* (i.e. specification of $w^*(t)$ for some $t \in [\underline{f}, \overline{f}]$), there

[30] In the special case considered in the paper, this assumption is automatically satisfied, since the η_j are ordered.

[31] Suppose that $p^*(w^*(f)) > \frac{w^*(f)+1}{f}$ in equilibrium: then a firm entering the market and recruiting at skill level f would make a strictly positive profit, violating the free entry condition. On the other hand, if $p^*(w^*(f)) < \frac{w^*(f)+1}{f}$, firms recruiting at skill level f would make strictly negative profits, and hence drop out of the market (i.e. this would be a violation of the free entry condition).

exists a unique, continuously differentiable solution $w^*(f)$ for the initial value problem given by (20) and the initial value. We write such solutions as functions $\Psi(C_j, u_j)$ where C_j is a constant (real number) encoding the initial value. So, in equilibrium, the wage schedule has the form

$$w^*(f) = \Psi(C_j, u_j)(f) \tag{21}$$

for all f servicing consumers of type j above the threshold θ. Note that, since ξ and u_j are increasing (strictly in the former case), it follows from (20) that $(w^*)'(f) > 0$ for all f: $\Psi(C_j, u_j)(f)$ is strictly increasing in f.

Whenever a consumer with disutility for inequality u_j purchases the good with inequality below the threshold θ, she is minimising price under the condition that the inequality is below θ, so across f servicing such customers, $\bar{\xi}\left(\frac{w^*(f)+1}{f}\right)$ is constant. Therefore, in equilibrium:

$$w^*(f) = Cf - 1 \tag{22}$$

for some constant C.

Furthermore, higher skilled workers service less inequality averse consumers in equilibrium, as demonstrated in the following two claims.

Claim 1. For all consumers of types $j < k$, if they are serviced by $s, t \in [\underline{f}, \overline{f}]$ respectively in equilibrium, with inequalities strictly greater than the threshold θ, then $s \leq t$.

Proof. Consider $j < k$ with inequality utility functions $u_j >_{I.A.} u_k$ and suppose for reductio that s services j but not k and t services k but not j, with $s > t$ and s and t producing goods at inequality above the threshold θ. Since j prefers the good produced by firms employing s to that produced by firms employing skill level t, we have:

$$\bar{\xi}(p^*(w^*(s)) - \bar{\xi}(p^*(w^*(t)) > u_j(\mathcal{I}(w^*(s))) - u_j(\mathcal{I}(w^*(t)))$$

whereas since k prefers the good produced by firms employing t to that produced by firms employing skill level s:

$$\bar{\xi}(p^*(w^*(s)) - \bar{\xi}(p^*(w^*(t)) < u_k(\mathcal{I}(w^*(s))) - u_k(\mathcal{I}(w^*(t)))$$

If $w^*(s) > w^*(t)$, then it follows from the two inequalities that $u_k(\mathcal{I}(w^*(s)) - u_k(\mathcal{I}(w^*(t)) > u_j(\mathcal{I}(w^*(s)) - u_j(\mathcal{I}(w^*(t))$, contradicting the fact that $u_j >_{I.A.} u_k$. If $w^*(s) \leq w^*(t)$, then $p^*(w^*(s)) = \frac{w^*(s)+1}{F(s)} < \frac{w^*(t)+1}{F(t)} = p^*(w^*(t))$, since F is strictly increasing. But then firms employing s produce goods which are cheaper and have less inequality than those employing skill level t, and hence are preferred by all consumers; this contradicts consumer k's preferences. So, for all consumers of types $j < k$ purchasing goods at inequality above the threshold θ, if they are serviced by s and t respectively in equilibrium, then $s \leq t$: the higher skilled workers service the less inequality averse consumers.

□

Claim 2. If there exist consumers of types j, k, with $j < k$, purchasing the good at inequality levels strictly greater than the threshold θ in equilibrium, then no consumer of type k purchases the good at inequality level less than θ.

Proof. Consider $j < k$ as specified with inequality utility functions $u_j >_{I.A.} u_k$, and suppose for reductio that there exist consumers of type k serviced by $s, t \in [\underline{f}, \overline{f}]$ such that $s > t$, $w^*(s) > \theta$ and $w^*(t) \leq \theta$. Since consumers of type k are indifferent between the goods produced by these two skill levels, we have that:

$$\bar{\xi}(p^*(w^*(s)) - u_k(\mathcal{I}(w^*(s))) = \bar{\xi}(p^*(w^*(t))$$

By the assumption, there exist consumers of type j serviced by $t' \in [\underline{f}, \overline{f}]$ with $s > t' > t$ and $w^*(t') > \theta$. By the argument in the proof of the previous claim, and the fact that $u_j >_{I.A.} u_k$:

$$\begin{aligned}\bar{\xi}(p^*(w^*(s)) - u_k(\mathcal{I}(w^*(s))) &> \bar{\xi}(p^*(w^*(t')) - u_k(\mathcal{I}(w^*(t'))) \\ &> \bar{\xi}(p^*(w^*(t')) - u_j(\mathcal{I}(w^*(t')))\end{aligned}$$

Combining this with the previous equality, it follows that consumers of type j strictly prefer purchasing goods at inequality level $w^*(t) \leq \theta$ under w^*, contradicting the assumption that some consumers of type j purchase the good with inequality above θ. So no consumer of type k purchases the good at inequality levels less than θ, as required. □

Given this, the equilibrium wage schedule is characterised by a sequence of real numbers $C_k, \ldots, C_K$, for $0 \leq k \leq K$, and a sequence of skill levels $\underline{f} \leq \underline{f_k} \leq \overline{f_k} \leq \underline{f_{k+1}} \leq \overline{f_{k+1}} \leq \cdots \leq \underline{f_K} \leq \overline{f_K} \leq \overline{f}$ with:

$$w^*(f) = \begin{cases} \Psi(C_K, u_K)(f) & f \in [\underline{f_K}, \overline{f_K}] \\ \Psi(C_{K-1}, u_{K-1})(f) & f \in [\underline{f_{K-1}}, \overline{f_{K-1}}] \\ \ldots & \ldots \\ \Psi(C_{k+1}, u_{k+1})(f) & f \in [\underline{f_{k+1}}, \overline{f_{k+1}}] \\ C_k f - 1 & f \in [\underline{f_k}, \overline{f_k}] \end{cases} \tag{23}$$

where $\Psi(C_{k+1}, u_{k+1})(\underline{f_{k+1}}) \geq \theta$ with $\Psi(C_{k+1}, u_{k+1})(\underline{f_{k+1}}) = \theta$ whenever $\underline{f_k} \neq \overline{f_k}$, and $\underline{f_k} = \overline{f_k} = \underline{f_{k+1}}$ whenever $\Psi(C_{k+1}, u_{k+1})(\underline{f_{k+1}}) > \theta$. By Claim 1, for each $k+1 \leq j \leq K$, only consumers of type j are serviced by workers with skill levels in $\left(\underline{f_j}, \overline{f_j}\right)$. By Claim 2, only consumers of type $j' \leq k+1$ are serviced by workers with skill levels in $\left(\underline{f_k}, \overline{f_k}\right)$. It follows from these observations and the closure of the good market that, if all consumers purchase the good in equilibrium (i.e. their budget and participation constraints are satisfied), the aforementioned sequences satisfy:

$$\mu_j = \int_{\underline{f_j}}^{\overline{f_j}} fX\left(f, \Psi(C_j, u_j)(f)\right) df \tag{24}$$

for every $k+1 < j \leq K$, and

$$\sum_{j=1}^{k+1} \mu_j = \int_{\underline{f_k}}^{\overline{f_k}} fX\left(f, C_k f - 1\right) dx + \int_{\underline{f_{k+1}}}^{\overline{f_{k+1}}} fX\left(f, \Psi(C_{k+1}, u_{k+1})(f)\right) df \tag{25}$$

Note that, for any j such that $\mu_j = 0$, $\underline{f_j} = \overline{f_j}$. We now show that the budget and participation constraints are always satisfied in equilibrium, so that these equalities hold.

Claim 3. In equilibrium, all consumers' budget and participation constraints are satisfied.

Proof. For each consumer k and $f \in [\underline{f}, \overline{f}]$, $b_k(f) = \sup\left\{x \in [1,\infty) : \hat{n}_k \geq \frac{x+1}{f}\right\}$ and $p_k(f) = \sup\left\{x \in [1,\infty) : v_k \geq \bar{\xi}(0) - \bar{\xi}(\frac{x+1}{f}) + u_k(\mathcal{I}(x))\right\}$ are the supremum wages that can be offered to labour of skill level f whilst satisfying the budget and participation constraints for consumer k. Let $\rho_k(f) = \min\{b_k(f), p_k(f)\}$.

For reductio, suppose that in equilibrium w^*, there is a consumer type j such that the corresponding condition (24) or (25) is not satisfied. Consider the case where $j > k+1$; the other case is treated similarly. It follows from Claim 1 that

$$\mu_j > \int_{\underline{f_j}}^{\overline{f_j}} fX\left(f, \Psi(C_j, u_j)(f)\right) df$$

If, for there exists no $[\underline{f^*}, \overline{f^*}] \subseteq [\underline{f}, \overline{f}] \setminus [\underline{f_j}, \overline{f_j}]$ with $w^*(f) < \rho_j(f)$ for all $f \in [\underline{f^*}, \overline{f^*}]$ then:

$$\begin{aligned} N - \mu_j &\geq \int_{[\underline{f},\overline{f}]\setminus[\underline{f_j},\overline{f_j}]} fX(f, w^*(f)) df \\ &\geq \int_{[\underline{f},\overline{f}]\setminus[\underline{f_j},\overline{f_j}]} fX(f, \hat{w}(f)) df \\ &\geq N - \int_{[\underline{f_j},\overline{f_j}]} fX(f, \hat{w}(f)) df \end{aligned}$$

where the first inequality follows from the previous observations concerning the servicing of consumers; the second follows from the assumption, the conditions (9) and (10) defining $\hat{w}$ and the fact that X is strictly increasing wherever it it non-zero; and the last inequality holds by (8). So $\int_{[\underline{f_j},\overline{f_j}]} fX(f, \hat{w}(f)) df \geq \mu_j$. Since the budget and participation constraints for j are satisfied by $\hat{w}(f)$ for all f, it follows that firms could enter the market, pay up to $\hat{w}(f)$ for $f \in [\underline{f_j}, \overline{f_j}]$, sell to the remaining consumers of type j, in which case w^* would not be an equilibrium. So there exists $[\underline{f^*}, \overline{f^*}] \subseteq [\underline{f}, \overline{f}] \setminus [\underline{f_j}, \overline{f_j}]$ with $w^*(f) < \rho_j(f)$ for all $f \in [\underline{f^*}, \overline{f^*}]$. In any such $[\underline{f^*}, \overline{f^*}]$, consumers of type j would buy goods produced by firms employing these skill levels, contradicting Claim 1; moreover, firms could enter the market, employ workers of such skill levels to accommodate non-serviced consumers of type j, contradicting the fact that w^* is an equilibrium. So there is no such type j, as required.

□

Moreover, for any $k < j < j' < K$, if $\overline{f_j} = \underline{f_{j'}}$, then by continuity of w^* (the law of one price), C_j and $C_{j'}$ are related by

$$w^*(\overline{f_j}) = \Psi(C_j, u_j)(\overline{f_j}) = \Psi(C_{j'}, u_{j'})(\overline{f_j}) \tag{26}$$

(I.e. $\Psi(C_j, u_j)$ and $\Psi(C_{j'}, u_{j'})$ solve their respective initial value problems with the same initial value.) By contrast, for $k < j < K$ with $\mu_j > 0$, if $\overline{f_j} < \underline{f_{j+1}}$, then $\overline{f_j} = \sup\{f : X(f, \Psi(C_j, u_j)(f)) > 0\}$. This is because, for any f such that $X(f, \Psi(C_j, u_j)(f)) > 0$, if $\overline{f_j} < f$ a firm would be able to enter the market, hire workers with skill f at wage $\Psi(C_j, u_j)(f)$, and sell to consumers with inequality utility function u_j. So, in equilibrium, $\overline{f_j}$ must be greater than or equal to the supremal such f; but since above the supremum there is no labour supply, $\overline{f_j} = \sup\{f : X(f, \Psi(C_j, u_j)(f)) > 0\}$. Hence $\overline{f_j}$ is determined by C_j, u_j and the functional form Ψ as the highest skill level for which there is positive labour supply under this wage pattern. A similar argument establishes that $\underline{f_{j+1}} = \inf\{f : X(f, \Psi(C_{j+1}, u_{j+1})(f)) > 0\}$ whenever $\overline{f_j} < \underline{f_{j+1}}$ and $\mu_{j+1} > 0$. Similar arguments establish that $\overline{f_K} = \sup\{f : X(f, \Psi(C_K, u_K)(f)) > 0\}$ and $\underline{f_k} = \inf\{f : X(f, C_k f - 1) > 0\}$ whenever there is positive demand for the good at the corresponding inequality levels. So, in equilibrium, the wage schedule is entirely characterised, modulo the functional form Ψ, by the inequality aversion distribution μ and the sequence $C_k, \ldots, C_K$.

Note that by (26) and the discussion immediately following it, as well as the fact that each $\Psi(C_j, u_j)(f)$ is strictly increasing in f, w^* is strictly increasing in the skill level f.

Finally, as a point of notation, although a wage schedule w^* defined according to (23) is a partial function, defined on $\bigcup_{i=k}^{K}[\underline{f_i}, \overline{f_i}]$, it can be extended to a function on $[\underline{f}, \overline{f}]$ by setting it equal to 0 outside $\bigcup_{i=k}^{K}[\underline{f_i}, \overline{f_i}]$. Recall that, without loss of generality, we use w^* to denote this function where convenient, notably for writing integrals involving w^*. Nevertheless, the minimum and maximum are defined with respect to $\bigcup_{i=k}^{K}[\underline{f_i}, \overline{f_i}]$: $\min w^* = w^*(\underline{f_k})$ and $\max w^* = w^*(\overline{f_K})$.

Special case: utility of form (1)

In the special case presented in the body of the paper, the FOC (19) for consumers of type j buying above the threshold θ simplifies to

$$\frac{dp}{d(w(s))} = \frac{1}{f} - \frac{p(w(s))}{fw'(f)} = -\eta_j \tag{27}$$

For equilibrium p^* and w^*, this can be solved analytically as:

$$w^*(f) = \frac{C_j f - 1}{\eta_j f + 1} \tag{28}$$

for all f servicing such consumers. The utility obtained by the consumer is $\hat{n} + v + \theta\eta_j - C_j$.[32] As noted (and as can be verified directly from $(w^*)'(f) = \frac{(C_j+\eta_j)}{(\eta_j f+1)^2}$), the wage is strictly increasing in f.

Plugging this into the general solution form derived above, we obtain that the equilibrium wage schedule is characterised by a sequence of positive real numbers $C_k, \ldots, C_K$, for $0 \leq k \leq K$, and a sequence $\underline{f} \leq \underline{f_k} \leq \overline{f_k} \leq \theta \leq \underline{f_{k+1}} \leq \overline{f_{k+1}} \leq \cdots \leq \underline{f_K} \leq \overline{f_K} \leq \overline{f}$ with:

[32] Plugging in the form of p^*, (27) implies that $\frac{(w^*)'(f)}{w^*(f)+1} = \frac{1}{f(1+\eta_j f)}$; solving this differential equation yields (28).

$$w^*(f) = \begin{cases} \frac{C_K f - 1}{\eta_K f + 1} & f \in [\underline{f_K}, \overline{f_K}] \\ \frac{C_{K-1} f - 1}{\eta_{K-1} f + 1} & f \in [\underline{f_{K-1}}, \overline{f_{K-1}}] \\ \dots & \dots \\ \frac{C_{k+1} f - 1}{\eta_{k+1} f + 1} & f \in [\underline{f_{k+1}}, \overline{f_{k+1}}] \\ C_k f - 1 & f \in [\underline{f_k}, \overline{f_k}] \end{cases} \tag{29}$$

where $\frac{C_{k+1}F(\underline{f_{k+1}})-1}{\eta_{k+1}F(\underline{f_{k+1}})+1} \geq \theta$ with $\frac{C_{k+1}F(\underline{f_{k+1}})-1}{\eta_{k+1}F(\underline{f_{k+1}})+1} = \theta$ whenever $\underline{f_k} \neq \overline{f_k}$, and $\underline{f_k} = \overline{f_k} = \underline{f_{k+1}}$ whenever $\frac{C_{k+1}F(\underline{f_{k+1}})-1}{\eta_{k+1}F(\underline{f_{k+1}})+1} > \theta$. Moreover, as for the general solution above, these sequences satisfy:

$$\mu_j = \int_{\underline{f_j}}^{\overline{f_j}} fX\left(f, \frac{C_j f - 1}{\eta_j f + 1}\right) df \tag{30}$$

for every $k+1 < j \leq K$ and

$$\sum_{j=1}^{k+1} \mu_j = \int_{\underline{f_k}}^{\overline{f_k}} fX\left(f, C_k f - 1\right) df + \int_{\underline{f_{k+1}}}^{\overline{f_{k+1}}} fX\left(f, \frac{C_{k+1} f - 1}{\eta_{k+1} f + 1}\right) df \tag{31}$$

Note that, for any $j > k$ such that $\mu_j = 0$, $\underline{f_j} = \overline{f_j}$, and similarly for $j \leq k$. Of course, the other properties of the general solution (e.g. characterisation by μ and the sequence $C_k, \dots, C_K$) are inherited in this special case.

A.3 Proof of Theorem 1

We now prove Theorem 1 under the general utility form for consumers, (17). The statement is precisely as in Section 3.3.

Let $w^*_{\mu'}$, the equilibrium wage schedule under μ', be as in (23), satisfying (24) and (25), for the sequences $C_k, \dots, C_K$ and $\underline{f} \leq \underline{s_k} \leq \overline{s_k} \leq \underline{s_{k+1}} \leq \overline{s_{k+1}} \leq \dots \leq \underline{s_K} \leq \overline{s_K} \leq \overline{f}$. If $k = K$ and all wages are below θ, then this same wage schedule satisfies the equilibrium conditions under μ, and has the same inequality. We henceforth suppose that not all wages are below θ, so $k < K$. We consider the case where $\underline{s_j} = \overline{s_{j-1}}$ for all $k < j \leq K$, where the $\overline{s_j}, \underline{s_j}$ are as in (23): an argument similar to that below holds for μ_j for the highest j for which $\underline{s_j} \neq \overline{s_{j-1}}$, hence establishing the other case. Moreover, we assume that $\mu'_K > 0$: again, if this is not the case, the same argument can be run starting from the highest j such that $\mu'_j > 0$.

Let $\mathcal{S} = [\underline{f}, \overline{f}] \times \left\{X(\overline{f}, x) : x \in [1, \infty)\right\}$; we denote a typical element by $(\tilde{f}, q)$. Now consider the following construction, which is based on the insight that, given the inequality aversion distribution, a candidate equilibrium solution is determined by the highest skill level employed if this level is not $\overline{f}$, or the quantity of the labour hired by firms employing workers of skill $\overline{f}$ otherwise.

Definition 1. For an inequality aversion distribution μ and $(\tilde{f}, q) \in \mathcal{S}$, define the *sequences* $D_l, \dots, D_K$ *and* $\underline{f} \leq \underline{t_l} \leq \overline{t_l} = \underline{t_{l+1}} \leq \overline{t_{l+1}} = \dots \leq \underline{t_K} \leq \overline{t_K} \leq \overline{f}$ *generated by* $(\tilde{f}, q)$ *with respect to* μ inductively by:

- if $\tilde{f} \neq \overline{f}$, then $\overline{t_K} = \tilde{f}$, D_K is the unique C satisfying[33]

$$\sup\{t : X(t, \Psi(C, u_K)(t)) > 0\} = \tilde{f} \tag{32}$$

and $\underline{t_K}$ is the unique t satisfying

$$\mu_K = \int_t^{\overline{t_K}} fX(f, \Psi(D_K, u_K)(f))\, df \tag{33}$$

- if $\tilde{f} = \overline{f}$, then $\overline{t_K} = \overline{f}$, D_K is the unique C satisfying

$$X\left(\overline{f}, \Psi(C, u_K)(\overline{f})\right) = q \tag{34}$$

and $\underline{t_K}$ is the unique t satisfying

$$\mu_K = \int_t^{\overline{t_K}} fX(f, \Psi(D_K, u_K)(f))\, df \tag{35}$$

- for $j > 1$, if D_{j+1} and $\underline{t_{j+1}}$ are such that $\Psi(D_{j+1}, u_{j+1})(\underline{t_{j+1}}) > \theta$, then let $\overline{t_j} = \underline{t_{j+1}}$, define D_j as the constant in the solution of (20) for u_j with initial value:

$$\Psi(D_j, u_j)(\overline{t_j}) = \Psi(D_{j+1}, u_{j+1})(\overline{t_j})$$

and define $\underline{t_j} = \max\{t_{j1}, t_{j2}, t_{j3}\}$ where:

  - t_{j1} is the maximal t satisfying

$$\mu_j = \int_t^{\overline{t_j}} fX(f, \Psi(D_j, u_j)(f))\, df \tag{36}$$

    if such a t exists, and $\underline{f}$ if not;
  - t_{j2} is the unique t satisfying $\Psi(D_j, u_j)(t) = \theta$ if such a t exists,[34] and $\underline{f}$ if not;
  - $t_{j3} = \inf\{t : X(t, \Psi(D_j, u_j)(t)) > 0\}$.

- for $j = 1$, if D_{j+1} and $\underline{t_{j+1}}$ are such that $\Psi(D_{j+1}, u_{j+1})(\underline{t_{j+1}}) > \theta$, then let $\overline{t_j} = \underline{t_{j+1}}$, define D_j as the constant in the solution of (20) for u_j with initial value:

$$\Psi(D_j, u_j)(\overline{t_j}) = \Psi(D_{j+1}, u_{j+1})(\overline{t_j})$$

and define $\underline{t_j} = \max\{t_{j2}, t_{j3}\}$ where:

  - t_{j2} is the unique t satisfying $\Psi(D_j, u_j)(t) = \theta$ if such a t exists, and $\underline{f}$ if not;
  - $t_{j3} = \inf\{t : X(t, \Psi(D_j, u_j)(t)) > 0\}$.

[33] Such a C is unique because of the uniqueness of the solutions of initial value problems given by (20) and the fact that X is continuous and strictly increasing in x for each f.

[34] Such a t is unique because $\Psi(D_j, u_j)(s)$ is continuous and strictly increasing in s.

- for $j \geq 0$, if D_{j+1} and $\underline{t_{j+1}}$ are such that $\Psi(D_{j+1}, u_{j+1})(\underline{t_{j+1}}) = \theta$, then let $\overline{t_j} = \underline{t_{j+1}}$, define D_j by

$$D_j \overline{t_j} - 1 = \Psi(D_{j+1}, u_{j+1})(\overline{t_j}) = \theta$$

  and $\underline{t_j} = \inf\{t : X(t, D_j t - 1) > 0\}$. In this case, set $l = j$ and the induction (construction of the sequences) is complete.

- if D_1 and $\underline{t_1}$ are such that $\Psi(D_1, u_1)(\underline{t_1}) > \theta$, then let $\underline{t_0} = \overline{t_0} = \underline{t_1}$ and define D_0 by $D_0 \overline{t_0} - 1 = \Psi(D_1, u_1)(\overline{t_0})$. Set $l = 0$ and the induction (construction of sequences) is complete.

Let $w_{\mu,(\tilde{f},q)}$ be the wage schedule defined according to (23) with $D_l, \ldots, D_K$ *and* $\underline{f} \leq \underline{t_l} \leq \overline{t_l} = \underline{t_{l+1}} \leq \overline{t_{l+1}} = \cdots \leq \underline{t_K} \leq \overline{t_K} \leq \overline{f}$.

Consider the sequences generated by $\left(\overline{s_K}, X(\overline{f}, \Psi(C_K, u_K)(\overline{f}))\right)$ with respect to μ, which we denote by $D_l, \ldots, D_K$ and $\underline{f} \leq \underline{t_l} \leq \overline{t_l} = \underline{t_{l+1}} \leq \overline{t_{l+1}} = \cdots \leq \underline{t_K} \leq \overline{t_K} \leq \overline{f}$. For conciseness, we denote the wage schedule they generate according to (23) by $\widehat{w_\mu}$; i.e. $\widehat{w_\mu} = w_{\mu,(\overline{s_K}, X(\overline{f}, \Psi(C_K, u_K)(\overline{f})))}$. Moreover, by construction, this wage schedule is such that $\Psi(D_K, u_K)(\overline{t_K}) = \Psi(C_K, u_K)(\overline{s_K})$—the maximum wage is the same as under the equilibrium for μ'—and it satisfies the condition (24) for μ for all inequality aversion levels greater than $l^* + 1$, where l^* is the lowest j such that $\underline{t_j} \neq \overline{t_j}$.

Proposition 2. $\sum_{j=1}^{K} \mu_j \leq \int_{\underline{f}}^{\overline{f}} fX\left(f, \widehat{w_\mu}(f)\right) ds = \sum_{j=l+1}^{K} \int_{\underline{f_j}}^{\overline{f_j}} fX\left(f, \Psi(D_j, u_j)(f)\right) df + \int_{\underline{f_l}}^{\overline{f_l}} fX\left(f, D_l f - 1\right) df$. *Moreover, the inequality is strict if and only if*

$$\sum_{i \geq \bar{j}} \mu_i < \int_{\left\{t : w^*_{\mu'}(t) > \theta\right\}} fX\left(f, w^*_{\mu'}(f)\right) df \tag{37}$$

where $\bar{j} = \max\left\{j : \sum_{i=j}^{K} \mu_i \neq \sum_{i=j}^{K} \mu'_i\right\}$ *(i.e.* $\bar{j}$ *is such that* $\mu_{\bar{j}} \neq \mu'_{\bar{j}}$ *and* $\mu_i = \mu'_i$ *for all* $i > \bar{j}$*).*

Proof. First of all, since $\overline{t_K} = \overline{s_K}$ and $\Psi(D_K, u_K)(\overline{t_K}) = \Psi(C_K, u_K)(\overline{s_K})$ if $\overline{t_K} = \overline{s_K} = \overline{s_H}$, it follows from the equilibrium solution under μ' (Section A.2), the fact that $\frac{\partial X}{\partial x} > 0$ where X is non-zero, and Definition 1 that $D_K = C_K$. Since μ Inequality Aversion Dominates μ', $\mu'_K \geq \mu_K$, which implies that $\underline{s_K} \leq \underline{t_K}$. Moreover, the latter inequality is strict whenever the former is and $\underline{s_K} \neq \overline{s_k}$.[35] Since $\underline{s_K} \geq \inf\{f : X(f, \Psi(C_K, u_K)(f)) > 0\}$, it follows that the same holds for $\underline{t_K}$. We first show that $\Psi(D_{K-1}, u_{K-1})(f) > \Psi(C_K, u_K)(f)$ for all $\underline{f} \leq f < \underline{t_K}$. Since $u_{K-1} >_{I.A.} u_K$, it clearly follows that the derivatives are ordered according to $u'_{K-1}(x) \geq u'_K(x)$, for all x. Hence, for every f such that $\Psi(D_{K-1}, u_{K-1})(f) = \Psi(C_K, u_K)(f)$, since these functions solve (20) with u_{K-1} and u_K respectively and this initial value, it follows from this differential equation and the aforementioned ordering of u'_{K-1} and u'_K that

[35] Note that, by definition, if $\mu'_K = N$ and $\underline{s_K} = \inf\{f : X(f, \Psi(C_K, u_K)(f)) > 0\}$, then $\overline{s_j} = \underline{s_j} = \underline{s_K}$ for all $j < K$.

$\Psi(D_{K-1},u_{K-1})'(f) < \Psi(C_K,u_K)'(f)$. By definition, $\Psi(D_{K-1},u_{K-1})(\underline{t_K}) = \Psi(C_K,u_K)(\underline{t_K})$, so $\Psi(D_{K-1},u_{K-1})'(\underline{t_K}) < \Psi(C_K,u_K)'(\underline{t_K})$. It follows from a standard argument that there exists no $\underline{f} \leq f < \underline{t_K}$ with $\Psi(D_{K-1},u_{K-1})(f) = \Psi(C_K,u_K)(f)$,[36] so $\Psi(D_{K-1},u_{K-1})(f) > \Psi(C_K,u_K)(f)$ for all such f, as required. Since $\frac{\partial X}{\partial x} > 0$ where X is non-zero, it follows that, if $\underline{s_K} < \underline{t_K}$, then

$$\int_{\underline{s_K}}^{\underline{t_K}} fX\left(f,\Psi(D_{K-1},u_{K-1})(f)\right)df > \int_{\underline{s_K}}^{\underline{t_K}} fX\left(f,\Psi(C_K,u_K)(f)\right)df$$

Now note that by the uniqueness of the solutions defining $\Psi(D_{K-1},u_{K-1})$, if $\underline{s_K} = \underline{t_K}$, then $\Psi(D_{K-1},u_{K-1}) = \Psi(C_{K-1},u_{K-1})$. If $\underline{s_K} < \underline{t_K}$, then by the previous observation, $\Psi(D_{K-1},u_{K-1})(\underline{s_K}) > \Psi(C_K,u_K)(\underline{s_K}) = \Psi(C_{K-1},u_{K-1})(\underline{s_K})$. However, if there exists $f < \underline{s_K}$ with $\Psi(D_{K-1},u_{K-1})(f) = \Psi(C_{K-1},u_{K-1})(f)$, then since these functions solve the same differential equation with the same initial value (at f), by the uniqueness of the solution they must be identical, contradicting the strict inequality at $\underline{s_K}$. Hence there is no such f, and $\Psi(D_{K-1},u_{K-1})(f) > \Psi(C_{K-1},u_{K-1})(f)$ for all $f \leq \underline{s_K}$. Repeating the previous arguments if necessary and drawing on the definitions of $\widehat{w_\mu}$ and $w^*_{\mu'}$, we have that $\widehat{w_\mu}(f) \geq w^*_{\mu'}(f)$ for all $\underline{s_{K-1}} \leq f \leq \overline{s_K}$, with strict inequality for $f' < \underline{t_K}$ whenever $\mu_K < \mu'_K$. In the light of this and the previous inequalities, we have:

$$\int_{f'}^{\overline{f}} fX\left(f,\widehat{w_\mu}(f)\right)df \geq \int_{f'}^{\overline{f}} fX\left(f,w^*_{\mu'}(f)\right)df$$

for all $f' \geq \underline{s_{K-1}}$, with strict inequality for $f' < \underline{t_K}$ whenever $\mu_K < \mu'_K$. Since $\mu_{K-1}+\mu_K \leq \mu'_{K-1}+\mu'_K = \int_{\underline{s_{K-1}}}^{\overline{f}} fX\left(f,w^*_{\mu'}(f)\right)df$ it follows that $\underline{t_{K-1}} \geq \underline{s_{K-1}}$, with strict inequality whenever either $\mu_K < \mu'_K$ or $\mu_{K-1}+\mu_K < \mu'_{K-1}+\mu'_K$.

The previous argument implies, for every $k+1 < j < K$ with $j \geq 1$, that if $\Psi(D_{j+1},u_{j+1})(\underline{t_{j+1}}) \geq \Psi(C_{j+1},u_{j+1})(\underline{t_{j+1}})$ and $\underline{t_{j+1}} \geq \underline{s_{j+1}}$ with strict inequality (in both inequalities) whenever there exists $\hat{j} \geq j+1$ with $\sum_{j'=\hat{j}}^K \mu_{j'} < \sum_{j'=\hat{j}}^K \mu'_{j'}$, then $\widehat{w_\mu}(f) \geq w^*_{\mu'}(f)$ for all $\underline{s_j} \leq f \leq \underline{t_{j+1}}$, $\Psi(D_j,u_j)(\underline{t_j}) \geq \Psi(C_j,u_j)(\underline{t_j})$, and $\underline{t_j} \geq \underline{s_j}$ with strict inequality (for some f in the first inequality) whenever there exists $\hat{j} \geq j$ with $\sum_{j'=\hat{j}}^K \mu_{j'} < \sum_{j'=\hat{j}}^K \mu'_{j'}$. Hence, by induction, for all $\underline{t_{k+2}} \leq f \leq \overline{s_K}$, $\widehat{w_\mu}(t) \geq w^*_{\mu'}(t)$, and $\widehat{w_\mu}(\underline{t_{k+2}}) = \Psi(D_{k+2},u_{k+2})(\underline{t_{k+2}}) \geq \Psi(C_{k+2},u_{k+2})(\underline{t_{k+2}}) = w^*_{\mu'}(\underline{t_{k+2}})$ and $\underline{t_{k+2}} \geq \underline{s_{k+2}}$, with strict inequality (for some f in the first inequality) whenever there exists $j \geq k+2$ with $\mu_j \neq \mu'_j$.

---

[36] For reductio, suppose there exists $\underline{f} \leq f < \underline{t_K}$ with $\Psi(D_{K-1},u_{K-1})(f) = \Psi(C_K,u_K)(f)$, and let f be the largest such one. Then, by the previous fact, $\Psi(D_{K-1},u_{K-1})'(f) < \Psi(C_K,u_K)'(f)$ and so, for $t > f$ sufficiently close to f, $\Psi(D_{K-1},u_{K-1})'(t) < \Psi(C_K,u_K)'(t)$ and $\Psi(D_{K-1},u_{K-1})(t) < \Psi(C_K,u_K)(t)$. Hence $\left\{f < t < \underline{t_K} : \Psi(D_{K-1},u_{K-1})(t) < \Psi(C_K,u_K)(t),\ \Psi(D_{K-1},u_{K-1})'(t) < \Psi(C_K,u_K)'(t)\right\}$ is non-empty. Since $\Psi(D_{K-1},u_{K-1})'(\underline{t_K}) < \Psi(C_K,u_K)'(\underline{t_K})$ and $\Psi(D_{K-1},u_{K-1})(\underline{t_K}) = \Psi(C_K,u_K)(\underline{t_K})$, for $t < \underline{t_K}$ sufficiently close to $\underline{t_K}$, $\Psi(D_{K-1},u_{K-1})'(t) < \Psi(C_K,u_K)'(t)$ and $\Psi(D_{K-1},u_{K-1})(t) > \Psi(C_K,u_K)(t)$, so $\left\{f < t < \underline{t_K} : \Psi(D_{K-1},u_{K-1})(t) > \Psi(C_K,u_K)(t),\ \Psi(D_{K-1},u_{K-1})'(t) < \Psi(C_K,u_K)'(t)\right\}$ is non-empty. By the intermediate value theorem, there exists $\sup\left\{f < t < \underline{t_K} : \Psi(D_{K-1},u_{K-1})(t) \leq \Psi(C_K,u_K)(t), \Psi(D_{K-1},u_{K-1})'(t) < \Psi(C_K,u_K)'(t)\right\} < f' < \inf\left\{f < t < \underline{t_K} : \Psi(D_{K-1},u_{K-1})(t) > \Psi(C_K,u_K)(t), \Psi(D_{K-1},u_{K-1})'(t) < \Psi(C_K,u_K)'(t)\right\}$ with $\Psi(D_{K-1},u_{K-1})(f') = \Psi(C_K,u_K)(f')$, contradicting the maximality of f.

To complete the proof, first consider the case where condition (37) is not satisfied, so $\sum_{i\geq\bar{j}}\mu_i \geq \int_{\left\{t:w^*_{\mu'}(t)>\theta\right\}} fX\left(f,w^*_{\mu'}(f)\right)df$ where $\bar{j}$ is as defined in the statement of the Proposition. It follows that $\mu_j = \mu'_j$ for all $j \geq k+2$, so the sequences $D_{k+2},\ldots,D_K$ and $C_{k+2},\ldots,C_K$ are identical, as are $\underline{t_{k+2}} \leq \overline{t_{k+2}} = \cdots \leq \underline{t_K} \leq \overline{t_K}$ and $\underline{s_{k+2}} \leq \overline{s_{k+2}} = \cdots \leq \underline{s_K} \leq \overline{s_K}$, and hence $\widehat{w_\mu}(f) = w^*_{\mu'}(f)$ for $\underline{s_{k+2}} \leq f \leq \overline{s_K}$. If $\Psi(C_{k+1},u_{k+1})(\underline{s_{k+1}}) > \theta$, then $\overline{s_k} = \underline{s_k} = \underline{s_{k+1}}$ by (23). Since $\Psi(C_{k+1},u_{k+1})(\underline{s_{k+1}}) > \theta$, in tandem with $\sum_{i\geq\bar{j}}\mu_i \geq \int_{\left\{t:w^*_{\mu'}(t)>\theta\right\}} fX\left(f,w^*_{\mu'}(f)\right)df$ and the definition of $\bar{j}$, also implies that $\mu_{k+1} = \mu'_{k+1}$, it follows from Definition 1, (23), and the observations following (26) that the sequences $D_{k+1},\ldots,D_K$ and $C_{k+1},\ldots,C_K$ are identical, as are $\underline{t_k} = \overline{t_k} = \underline{t_{k+1}} \leq \overline{t_{k+1}} = \cdots \leq \underline{t_K} \leq \overline{t_K}$ and $\underline{s_k} = \overline{s_k} = \underline{s_{k+1}} \leq \overline{s_{k+1}} = \cdots \leq \underline{s_K} \leq \overline{s_K}$. So $\widehat{w_\mu}(f) = w^*_{\mu'}(f)$ for $\underline{s_k} \leq f \leq \overline{s_K}$. So $\widehat{w_\mu}$ satisfies (25), since $w^*_{\mu'}$ does, and the inequality in the Proposition holds with equality. If $\overline{s_k} \neq \underline{s_k}$, then by (23), $\Psi(C_{k+1},u_{k+1})(\underline{s_{k+1}}) = \theta$. It follows from the arguments above that $D_{k+1} = C_{k+1}$. Moreover, $\sum_{i\geq\bar{j}}\mu_i \geq \int_{\left\{t:w^*_{\mu'}(t)>\theta\right\}} fX\left(f,w^*_{\mu'}(f)\right)df$ implies that $\mu_{k+1} \geq \mu'_{k+1}$. Hence $\widehat{w_\mu}(\underline{t_{k+1}}) = \Psi(D_{k+1},u_{k+1})(\underline{t_{k+1}}) = \theta = \Psi(C_{k+1},u_{k+1})(\underline{s_{k+1}})$, $\underline{t_{k+1}} = \underline{s_{k+1}}$. It follows that $D_k\underline{t_{k+1}} - 1 = \theta$, so $D_k = C_k$, $\underline{t_k} = \underline{s_k}$ and $\sum_{j'=1}^k \mu_{j'} = \sum_{j'=1}^k \mu'_{j'}$ satisfy (25). So the inequality in the Proposition holds with equality.

Now suppose that condition (37) is satisfied. We distinguish three cases. If $\Psi(C_{k+1},u_{k+1})(\underline{s_{k+1}}) > \theta$, then $\overline{s_k} = \underline{s_k} = \underline{s_{k+1}}$ and $\mu_j \neq \mu'_j$ for some $j > k+1$ (for if not, that would imply that $\mu_{k+1} = \mu'_{k+1}$ and hence $\mu_j = \mu'_j$ for all j, contradicting (37)). So $\Psi(D_{k+2},u_{k+2})(\underline{t_{k+2}}) > \Psi(C_{k+2},u_{k+2})(\underline{t_{k+2}})$ and, $\overline{t_{k+1}} = \underline{t_{k+2}} > \underline{s_{k+2}} = \overline{s_{k+1}}$. Reapplying the argument above, we have that $\widehat{w_\mu}(f) > w^*_{\mu'}(f)$ for all $\underline{s_{k+1}} \leq f < \underline{t_{k+2}}$. In particular, since $w^*_{\mu'}(\underline{s_{k+1}}) = \Psi(C_{k+1},u_{k+1})(\underline{s_{k+1}}) > \theta$, $\underline{s_{k+1}} = \inf\{t : X(t,\Psi(C_{k+1},u_{k+1})(\overline{t})) > 0\}$, and there exists $l < k^* \leq k+1$ with $\underline{t_{k^*}} \leq \underline{s_{k+1}} \leq \overline{t_{k^*}}$, so $\widehat{w_\mu}(\underline{s_{k+1}}) = \Psi(D_{k^*},u_{k^*})(\underline{s_{k+1}}) > \Psi(C_{k+1},u_{k+1})(\underline{s_{k+1}}) > \theta$. By the previous argument applied again, $\widehat{w_\mu}(f) \geq \Psi(D_{k^*},u_{k^*})(f)$ for all $\underline{t_{l+1}} \leq f \leq \underline{s_{k+1}}$. It thus follows that $\underline{t_{l+1}} \leq t^* < \underline{s_{k+1}}$, where $t^* = \max\{\inf\{f : X(f,\Psi(D_{k^*},u_{k^*})(f)) > 0\}, \inf\{f : \Psi(D_{k^*},u_{k^*})(f) > \theta\}\}$. Hence

$$\begin{aligned}
\int_{\underline{f}}^{\overline{f}} fX\left(f,\widehat{w_\mu}(f)\right)df &\geq \int_{t^*}^{\overline{t_K}} fX\left(f,\widehat{w_\mu}(f)\right)df \\
&> \int_{\underline{s_k}}^{\overline{s_K}} fX\left(f,\widehat{w_\mu}(f)\right)df \\
&> \int_{\underline{s_k}}^{\overline{s_K}} fX\left(f,w^*_{\mu'}(f)\right)df \\
&= \sum_{j=1}^K \mu'_j
\end{aligned}$$

where the first two inequalities follow from the fact that $\underline{t_{l+1}} \leq t^* < \underline{s_{k+1}}$, and the final one follows from the fact that w_μ dominates $w^*_{\mu'}$ for $f \geq \underline{s_k} = \underline{s_{k+1}}$, strictly for $f \in [\underline{s_k},\underline{t_{k+2}})$. Since $\sum_{j=1}^K \mu'_j = \sum_{j=1}^K \mu_j$, this establishes the desired strict inequality.

Now consider the case where condition (37) is satisfied, $\overline{s_k} \neq \underline{s_k}$ (so, by (23), $\Psi(C_{k+1}, u_{k+1})(\underline{s_{k+1}}) = \theta$), and there exists $j \geq k+2$ with $\mu_j \neq \mu'_j$. It follows that $\Psi(D_{k+2}, u_{k+2})(\underline{t_{k+2}}) > \Psi(C_{k+2}, u_{k+2})(\underline{t_{k+2}})$ and, $\overline{t_{k+1}} = \underline{t_{k+2}} > \underline{s_{k+2}} = \overline{s_{k+1}}$. Applying the previous argument, we have that $\widehat{w_\mu}(f) > w^*_{\mu'}(f)$ for all $\underline{s_{k+1}} \leq f < \underline{t_{k+2}}$, so $\widehat{w_\mu}(\underline{s_{k+1}}) > w^*_{\mu'}(\underline{s_{k+1}}) = \Psi(C_{k+1}, u_{k+1})(\underline{s_{k+1}}) = \theta$. Hence there exists $l < k^* \leq k+1$ with $\underline{t_{k^*}} \leq \underline{s_{k+1}} \leq \overline{t_{k^*}}$; for such k^*, $\widehat{w_\mu}(\underline{s_{k+1}}) = \Psi(D_{k^*}, u_{k^*})(\underline{s_{k+1}}) > \theta$. By definition, for all D_j with $k^* \geq j > l$, $\Psi(D_j, u_j)(f) > \theta \geq C_k f - 1$ for all $\underline{t_j} < f \leq \overline{t_j}$, so $\widehat{w_\mu}(f) > C_k f - 1 = w^*_{\mu'}(f)$ for all $\underline{t_{l+1}} \leq f < \underline{s_{k+1}}$. Moreover, since $C_k \overline{t_k} - 1 = \theta$, $C_k \overline{t_l} - 1 < \theta = D_l \overline{t_l} - 1$, so $D_l > C_k$, and hence $\widehat{w_\mu}(f) = D_l f - 1 > C_k f - 1 = w^*_{\mu'}(f)$ for all $\underline{s_k} \leq f \leq \overline{t_l}$. So $\widehat{w_\mu}(f) \geq w^*_{\mu'}(f)$ for all $\underline{s_k} \leq f \leq \overline{s_K}$, with strict inequality on a non-degenerate interval. Since, for every $f \leq \overline{t_l}$, if $X(f, C_k f - 1) > 0$, then $X(f, D_l f - 1) > 0$, $\underline{t_l} = \inf\{f : X(f, D_l f - 1) > 0\} < \inf\{t : X(f, C_k f - 1) > 0\} = \underline{s_k}$. Hence

$$\begin{aligned}
\int_{\underline{f}}^{\overline{f}} fX(f, \widehat{w_\mu}(f))\, df &\geq \int_{\underline{t_l}}^{\overline{t_K}} fX(f, \widehat{w_\mu}(f))\, df \\
&> \int_{\underline{s_k}}^{\overline{s_K}} fX(f, \widehat{w_\mu}(f))\, df \\
&> \int_{\underline{s_k}}^{\overline{s_K}} fX\left(f, w^*_{\mu'}(f)\right) df \\
&= \sum_{j=1}^{K} \mu'_j
\end{aligned}$$

Since $\sum_{j=1}^K \mu'_j = \sum_{j=1}^K \mu_j$, this establishes the desired strict inequality.

Finally, consider the case where condition (37) is satisfied, $\overline{s_k} \neq \underline{s_k}$ (so, by (23), $\Psi(C_{k+1}, u_{k+1})(\underline{s_{k+1}}) = \theta$), and $\mu_j = \mu'_j$ for all $j \geq k+2$. It follows that the sequences $D_{k+2}, \ldots, D_K$ and $C_{k+2}, \ldots, C_K$ are identical, as are $\underline{t_{k+2}} \leq \overline{t_{k+2}} = \cdots \leq \underline{t_K} \leq \overline{t_K}$ and $\underline{s_{k+2}} \leq \overline{s_{k+2}} = \cdots \leq \underline{s_K} \leq \overline{s_K}$, and hence $\widehat{w_\mu}(f) = w^*_{\mu'}(f)$ for $\underline{s_{k+2}} \leq f \leq \overline{s_K}$. So $\sum_{i \geq k+2} \mu_i = \int_{\underline{t_{k+2}}}^{\overline{t_K}} fX(f, \widehat{w_\mu}(f))\, df = \int_{\underline{s_{k+2}}}^{\overline{s_K}} fX\left(f, w^*_{\mu'}(f)\right) df$, whence, by condition (37), $\mu_{k+1} < \int_{\underline{s_{k+1}}}^{\overline{s_{k+1}}} fX\left(f, w^*_{\mu'}(f)\right) df \leq \mu'_{k+1}$. Hence $\underline{t_{k+1}} > \underline{s_{k+1}}$, and by the previous argument, $\widehat{w_\mu}(f) = \Psi(D_k, u_k)(f) > \Psi(C_{k+1}, u_{k+1})(f) = w^*_{\mu'}(f)$ for $\underline{s_{k+1}} \leq f < \underline{t_{k+1}}$; so $\widehat{w_\mu}(\underline{s_{k+1}}) > w^*_{\mu'}(\underline{s_{k+1}}) = \theta$. By the argument in the previous case, it follows that $\widehat{w_\mu}(f) \geq w^*_{\mu'}(f)$ for all $\underline{s_k} \leq f \leq \overline{s_K}$, with strict inequality on a non-degenerate interval, and $\underline{t_l} < \underline{s_k}$. Hence $\int_{\underline{f}}^{\overline{f}} fX(f, \widehat{w_\mu}(f))\, df > \sum_{j=1}^K \mu'_j$ by the argument in the previous case, thus establishing the desired strict inequality.

□

For a fixed inequality aversion distribution μ, define $\pi_\mu : \mathcal{S} \to \mathbb{R}$ by: for every $(\tilde{f}, q) \in \mathcal{S}$, $\pi(\tilde{f}, q) = \sum_{j=l+1}^K \int_{\underline{t_j}}^{\overline{t_j}} fX(f, \Psi(D_j, u_j)(f))\, df + \int_{\underline{t_l}}^{\overline{t_l}} fX(f, D_l f - 1)\, ds$ where $D_l, \ldots, D_K$, $\underline{s_H} \leq \underline{t_l} \leq \overline{t_l} = \underline{t_{l+1}} \leq \overline{t_{l+1}} = \cdots \leq$

$\underline{t_K} \leq \overline{t_K} \leq \overline{s_H}$ are the sequences generated by $(\tilde{f}, q)$ with respect to μ (Definition 1).

Proposition 3. *For every μ, π_μ is strictly increasing in $(\tilde{f}, q)$: i.e. $\pi(\hat{f}, \hat{q}) > \pi(\tilde{f}, q)$ whenever $\hat{f} > \tilde{f}$ or $\hat{f} = \tilde{f} = \overline{f}$ and $\hat{q} > q$.*

Proof. Consider $(\hat{f}, \hat{q}), (\tilde{f}, q)$ with $\hat{f} > \tilde{f}$ or $\hat{f} = \tilde{f} = \overline{f}$ and $\hat{q} > q$, and let $D_k, \ldots, D_K$, $\underline{s_H} \leq \underline{t_k} \leq \overline{t_k} = \underline{t_{k+1}} \leq \overline{t_{k+1}} = \cdots \leq \underline{t_K} \leq \overline{t_K} \leq \overline{s_H}$, $\hat{D}_l, \ldots, \hat{D}_K$, $\underline{s_H} \leq \underline{\hat{t}_l} \leq \overline{\hat{t}_l} = \underline{\hat{t}_{l+1}} \leq \overline{\hat{t}_{l+1}} = \cdots \leq \underline{\hat{t}_K} \leq \overline{\hat{t}_K} \leq \overline{s_H}$ be the sequences generated by $(\tilde{f}, q)$ and $(\hat{f}, \hat{q})$ respectively with respect to μ. Let w and $\hat{w}$ be the corresponding wage schedules generated according to (23). By the argument in the proof of Proposition 2, $l \leq k$, $\overline{\hat{t}_{k+1}} > \overline{t_{k+1}}$ and for every $\underline{t_{k+1}} \leq t \leq \overline{\hat{t}_{k+1}}$, $\hat{w}(t) > w(t)$. So $\hat{w}(\underline{t_{k+1}}) = \Psi(\hat{D}_{l^*}, u_{l^*})(\underline{t_{k+1}}) > \Psi(D_{k+1}, u_{k+1})(\underline{t_{k+1}}) = w(\underline{t_{k+1}}) \geq \theta$ for $k+1 \geq l^* > l$ with $\underline{\hat{t}_{l^*}} \leq \underline{t_{k+1}} \leq \overline{\hat{t}_{l^*}}$. We separate two cases.

If $\underline{t_k} = \overline{t_k} = \underline{t_{k+1}}$, then $\mu_j = 0$ for all $j \leq k$, and $l^* = k+1$. Moreover, by the previously noted fact $\underline{\hat{t}_{k+1}} \leq t^* = \max\left\{\inf\left\{f : X\left(f, \Psi(\hat{D}_{l^*}, u_{l^*})(f)\right) > 0\right\}, \inf\left\{f : \Psi(\hat{D}_{l^*}, u_{l^*})(f) > \theta\right\}\right\} < \inf\{f : X(f, \Psi(D_{k+1}, u_{k+1})(f)) > 0\} = \underline{t_{k+1}}$. This coupled with the previous observations implies that

$$\begin{aligned}
\int_{\underline{\hat{t}_{k+1}}}^{\overline{\hat{t}_{k+1}}} fX\left(f, \Psi(\hat{D}_{k+1}, u_{k+1})(f)\right) df &= \int_{\underline{\hat{t}_{k+1}}}^{\overline{\hat{t}_{k+1}}} fX\left(f, \hat{w}(f)\right) df \\
&> \int_{\underline{t_{k+1}}}^{\overline{\hat{t}_{k+1}}} fX\left(f, \hat{w}(f)\right) df \\
&> \int_{\underline{t_{k+1}}}^{\overline{t_{k+1}}} fX\left(f, w(f)\right) df \\
&= \int_{\underline{t_{k+1}}}^{\overline{t_{k+1}}} fX\left(f, \Psi(D_{k+1}, u_{k+1})(f)\right) df
\end{aligned}$$

If $\underline{t_k} \neq \overline{t_k}$, then $\Psi(D_{k+1}, u_{k+1})(\underline{t_{k+1}}) = w(\underline{t_{k+1}}) = \theta$. By the argument at the end of the proof of Proposition 2, $\hat{w}(f) \geq w(f)$ for every $\underline{t_k} \leq f \leq \underline{\hat{t}_{k+1}}$. Moreover, by that argument again, $\underline{\hat{t}_l} = \inf\left\{f : X\left(f, \hat{D}_l f - 1\right) > 0\right\} < \inf\{f : X(D_k f - 1) > 0\} = \underline{t_k}$. This coupled with the previous observations

implies that:

$$\begin{aligned}\textstyle\sum_{j=l+1}^{k+1}\int_{\underline{\hat{t}_j}}^{\overline{\hat{t}_j}} fX\left(f,\Psi(\hat{D}_j,u_j)(f)\right)df \\ +\textstyle\int_{\underline{\hat{t}_l}}^{\overline{\hat{t}_l}} fX\left(\hat{D}_lF(f)-1\right)df\end{aligned} &= \int_{\underline{\hat{t}_l}}^{\overline{\hat{t}_{k+1}}} fX\left(f,\hat{w}(f)\right)df \\ &> \int_{\underline{t_k}}^{\overline{t_{k+1}}} fX\left(f,\hat{w}(f)\right)df \\ &> \int_{\underline{t_k}}^{\overline{t_{k+1}}} fX\left(f,w(f)\right)df \\ &= \begin{array}{l}\int_{\underline{t_{k+1}}}^{\overline{t_{k+1}}} fX\left(f,\Psi(D_{k+1},u_{k+1})(f)\right)df \\ +\int_{\underline{t_k}}^{\overline{t_k}} fX\left(f,D_kf-1\right)df\end{array}$$

Since, by Definition 1, $\int_{\underline{t_j}}^{\overline{t_j}} fX\left(f,\Psi(\hat{D}_j,u_j)(f)\right)df = \int_{\underline{t_j}}^{\overline{t_j}} fX\left(f,\Psi(D_j,u_j)(f)\right)df = \mu_j$ for all $j > k+1$, the result follows. □

Since X and $\Psi(\bullet,\bullet)$ are continuous, the sequences constructed in Definition 1 are continuous (pointwise) in $(\tilde{f},q)$, as is π_μ. It follows from Proposition 2 that any equilibrium under μ' is an equilibrium under μ, whenever condition (37) is not satisfied. Whenever this condition is satisfied, it follows from Propositions 2 and 3 that any equilibrium w^*_μ under μ is such that either the highest hired skill level is lower than for the equilibrium $w^*_{\mu'}$ under μ'—i.e. $\sup\left\{f : w^*_\mu(f) > 0\right\} < \sup\left\{f : w^*_{\mu'}(f) > 0\right\}$—or they are both hire at the highest skill level $\overline{f}$ but the labour supply there is strictly lower under μ—i.e. $X(\overline{f}, w^*_\mu(\overline{f})) < X(\overline{f}, w^*_{\mu'}(\overline{f}))$. By Proposition 3, any such equilibrium is unique; moreover, by the previously noted continuity of π_μ, there exists such an equilibrium. It follows from the form of the solution (Section A.2) and the fact that $\frac{\partial X}{\partial x} > 0$ where X takes non-zero values that the maximum wage is higher under μ' than under μ, and strictly so precisely when (37) holds. This establishes Theorem 1.

Remark 1. Note that the deduction of the equilibrium wage schedule (Section A.2), Proposition 3 and the continuity of the functions involved establish the generic existence of an equilibrium; Proposition 3 implies that it is unique.

A.4 Proof of Theorem 2

We prove Theorem 1 under the utility specification (1) used in the text; see Remark 2 on extensions to the general specification (17).

By Proposition 2 and the arguments in the proof of Theorem 1, if there are (weakly) more inequality neutral consumers under μ than consumers purchasing the good at a price higher than θ in equilibrium under μ^0, then $w^*_\mu = w^*_{\mu^0}$, so $g_{w^*_\mu} = g_{w^*_{\mu^0}}$, and the inequality is the same in the two cases, for any measure. Henceforth we consider the case where there are strictly fewer inequality neutral consumers under μ than consumers purchasing the good at a price higher than θ in equilibrium under μ^0. It follows from Theorem 1 that $\max w^*_\mu < \max w^*_{\mu^0}$. The proof rests on two main propositions; we begin with several auxiliary lemmas.

Lemma 1. w^*_μ crosses $w^*_{\mu^0}$ once from above: i.e. there exists $\hat{f}$ such that $w^*_\mu(f) > w^*_{\mu^0}(f)$ for $f < \hat{f}$ and $w^*_\mu(f) < w^*_{\mu^0}(f)$ for $f > \hat{f}$.

Proof of Lemma 1. As noted above, by the specification of the case and Theorem 1, $\max w^*_\mu < \max w^*_{\mu^0}$. However, if $w^*_\mu(f) \leq w^*_{\mu^0}(f)$ for all $f \in [\underline{f}, \overline{f}]$, then, since $\frac{\partial X}{\partial x} > 0$ where X takes non-zero values, $\int_{\underline{f}}^{\overline{f}} fX(f, w^*_\mu(f))df < \int_{\underline{f}}^{\overline{f}} fX(f, w^*_{\mu^0}(f))df = N$, contradicting $\int_{\underline{f}}^{\overline{f}} fX(f, w^*_\mu(f))df = N$. So there exists $f \in [\underline{f}, \overline{f}]$ with $w^*_\mu(f) > w^*_{\mu^0}(f)$, i.e. the wage schedules cross. Now consider any crossing point, that is, $f \in [\underline{f}, \overline{f}]$ with $w^*_\mu(f) = w^*_{\mu^0}(f)$. By the uniqueness of the solution of (20) with a given initial value of w, it follows that the type of consumer serviced by workers of level f under w^*_μ must be different from the type serviced by workers of level f under $w^*_{\mu^0}$, which in this case implies that the former will not be K. However, by (20), and noting that under μ^0, the wage schedule has the same form above and below the threshold θ, it follows that, for any crossing point f with $w^*_\mu(f) > \theta$, $\frac{dw^*_{\mu^0}}{df}(f) < \frac{dw^*_\mu}{df}(f)$. Hence there is exactly one crossing point f such that $w^*_\mu(f) = w^*_{\mu^0}(f) > \theta$; by the form of w^*_μ below θ (i.e. (23)), it follows that there is exactly one crossing point, as required. □

Note that this result implies that $w^*_\mu(f) > w^*_{\mu^0}(f)$ for all sufficiently low f. It follows, by the constraints (24) and (25) on w^*_μ and the fact that $\frac{\partial X}{\partial x} > 0$ when X is non-zero, that $\min w^*_\mu < \min w^*_{\mu^0}$.

Lemma 2. $\int_{\underline{f}}^{\overline{f}} X(f, w^*_\mu(f))df > \int_{\underline{f}}^{\overline{f}} X(f, w^*_{\mu^0}(f))df$.

Proof of Lemma 2. By (24) and (25), $\int_{\underline{f}}^{\overline{f}} fX(f, w^*_\mu(f))df = \int_{\underline{f}}^{\overline{f}} fX(f, w^*_{\mu^0}(f))df$, so $\int_{\underline{f}}^{\hat{f}} f\left(X(f, w^*_\mu(f))df - X(f, w^*_{\mu^0}(f))\right)df = \int_{\hat{f}}^{\overline{f}} f\left(X(f, w^*_{\mu^0}(f))df - X(f, w^*_\mu(f))\right)df$, where $\hat{f}$ is the crossing point for the wage schedules (Lemma 1). Define $f^+ = \frac{\int_{\hat{f}}^{\overline{f}} f\left(X(f,w^*_{\mu^0}(f))df - X(f,w^*_\mu(f))\right)df}{\int_{\hat{f}}^{\overline{f}} X(f,w^*_{\mu^0}(f))df - X(f,w^*_\mu(f))df}$ and $f^- = \frac{\int_{\underline{f}}^{\hat{f}} f\left(X(f,w^*_\mu(f))df - X(f,w^*_{\mu^0}(f))\right)df}{\int_{\underline{f}}^{\hat{f}} X(f,w^*_\mu(f))df - X(f,w^*_{\mu^0}(f))df}$. $X(f, w^*_{\mu^0}(f)) - X(f, w^*_\mu(f)) > 0$ for $f > \hat{f}$ and $X(f, w^*_\mu(f)) - X(f, w^*_{\mu^0}(f)) > 0$ for $f < \hat{f}$, because $\hat{f}$ is the crossing point and $\frac{\partial X}{\partial x} \geq 0$. It follows that $f^+ > \hat{f} > f^-$. Hence

$$\int_{\underline{f}}^{\hat{f}} X(f, w^*_\mu(f))df - X(f, w^*_{\mu^0}(f))df = \frac{f^+}{f^-}\int_{\hat{f}}^{\overline{f}} X(f, w^*_{\mu^0}(f))df - X(f, w^*_\mu(f))df$$
$$> \int_{\hat{f}}^{\overline{f}} X(f, w^*_{\mu^0}(f))df - X(f, w^*_\mu(f))df$$

so $\int_{\underline{f}}^{\overline{f}} X(f, w^*_\mu(f))df > \int_{\underline{f}}^{\overline{f}} X(f, w^*_{\mu^0}(f))df$ as required. □

Lemma 3. *Under Assumption 1, for any interval $J \subseteq [\min w^*_{\mu^0}, \max w^*_{\mu}]$ such that $\left(w^*_{\mu}\right)^{-1}(x) < \left(w^*_{\mu^0}\right)^{-1}(x)$ for all $x \in J$, $\frac{X(\left(w^*_{\mu}\right)^{-1}(x),x)}{X(\left(w^*_{\mu^0}\right)^{-1}(x),x)} . \frac{w'_{\mu^0}\left(\left(w^*_{\mu^0}\right)^{-1}(x)\right)}{w'_{\mu}\left(\left(w^*_{\mu}\right)^{-1}(x)\right)}$ is monotonically decreasing across J (i.e. for $x, x' \in J$ if $x > x'$, then $\frac{X(\left(w^*_{\mu}\right)^{-1}(x),x)}{X(\left(w^*_{\mu^0}\right)^{-1}(x),x)} . \frac{w'_{\mu^0}\left(\left(w^*_{\mu^0}\right)^{-1}(x)\right)}{w'_{\mu}\left(\left(w^*_{\mu}\right)^{-1}(x)\right)} \leq \frac{X(\left(w^*_{\mu}\right)^{-1}(x'),x')}{X(\left(w^*_{\mu^0}\right)^{-1}(x'),x')} . \frac{w'_{\mu^0}\left(\left(w^*_{\mu^0}\right)^{-1}(x')\right)}{w'_{\mu}\left(\left(w^*_{\mu}\right)^{-1}(x')\right)}$).*

Proof of Lemma 3. Note that since $\min w^*_{\mu} < \min w^*_{\mu^0}$ and $\max w^*_{\mu} < \max w^*_{\mu^0}$, both $X(\left(w^*_{\mu}\right)^{-1}(x), x)$ and $X(\left(w^*_{\mu^0}\right)^{-1}(x), x)$ are differentiable (as functions of x) on the interval $[\min w^*_{\mu^0}, \max w^*_{\mu}]$. $\frac{d}{dx}\left(\frac{X(\left(w^*_{\mu}\right)^{-1}(x),x)}{X(\left(w^*_{\mu^0}\right)^{-1}(x),x)} . \frac{\left(w^*_{\mu^0}\right)'\left(\left(w^*_{\mu^0}\right)^{-1}(x)\right)}{\left(w^*_{\mu}\right)'\left(\left(w^*_{\mu}\right)^{-1}(x)\right)}\right) \leq 0$ if and only if $\frac{d}{dx}\left(log\left(\frac{\left(X(\left(w^*_{\mu}\right)^{-1}(x),x)\right)}{\left(w^*_{\mu}\right)'\left(\left(w^*_{\mu}\right)^{-1}(x)\right)}\right)\right) \leq \frac{d}{dx}\left(log\left(\frac{X(\left(w^*_{\mu^0}\right)^{-1}(x),x)}{\left(w^*_{\mu^0}\right)'\left(\left(w^*_{\mu^0}\right)^{-1}(x)\right)}\right)\right)$ which holds if and only if $\frac{d}{dx}\left(\frac{X(\left(w^*_{\mu}\right)^{-1}(x),x)}{\left(w^*_{\mu}\right)'\left(\left(w^*_{\mu}\right)^{-1}(x)\right)}\right)\frac{\left(w^*_{\mu}\right)'\left(\left(w^*_{\mu}\right)^{-1}(x)\right)}{X(\left(w^*_{\mu}\right)^{-1}(x),x)} \leq \frac{d}{dx}\left(\frac{X(\left(w^*_{\mu^0}\right)^{-1}(x),x)}{\left(w^*_{\mu^0}\right)'\left(\left(w^*_{\mu^0}\right)^{-1}(x)\right)}\right)\frac{\left(w^*_{\mu^0}\right)'\left(\left(w^*_{\mu^0}\right)^{-1}(x)\right)}{X(\left(w^*_{\mu^0}\right)^{-1}(x),x)}$. So it suffices to show that this inequality holds for all $x \in J$. Note that

$$\begin{aligned}
&\frac{d}{dx}\left(\frac{X(\left(w^*_{\mu}\right)^{-1}(x),x)}{\left(w^*_{\mu}\right)'\left(\left(w^*_{\mu}\right)^{-1}(x)\right)}\right)\frac{\left(w^*_{\mu}\right)'\left(\left(w^*_{\mu}\right)^{-1}(x)\right)}{X(\left(w^*_{\mu}\right)^{-1}(x),x)} \\
&= \left(\frac{\nabla X\left(\left(w^*_{\mu}\right)^{-1}(x),x\right).(\frac{1}{\left(w^*_{\mu}\right)'(\left(w^*_{\mu}\right)^{-1}(x))},1)}{\left(w^*_{\mu}\right)'\left(\left(w^*_{\mu}\right)^{-1}(x)\right)}\right. \\
&\left. - \frac{X(\left(w^*_{\mu}\right)^{-1}(x),x).\left(\left(w^*_{\mu}\right)''\left(\left(w^*_{\mu}\right)^{-1}(x)\right)\right)}{\left(\left(w^*_{\mu}\right)'\left(\left(w^*_{\mu}\right)^{-1}(x)\right)\right)^3}\right)\frac{\left(w^*_{\mu}\right)'\left(\left(w^*_{\mu}\right)^{-1}(x)\right)}{X(\left(w^*_{\mu}\right)^{-1}(x),x)} \\
&= \frac{\nabla X\left(\left(w^*_{\mu}\right)^{-1}(x),x\right).(\frac{1}{\left(w^*_{\mu}\right)'(\left(w^*_{\mu}\right)^{-1}(x))},1)}{X(\left(w^*_{\mu}\right)^{-1}(x),x)} - \frac{\left(w^*_{\mu}\right)''\left(\left(w^*_{\mu}\right)^{-1}(x)\right)}{\left(\left(w^*_{\mu}\right)'\left(\left(w^*_{\mu}\right)^{-1}(x)\right)\right)^2}
\end{aligned}$$

and similarly for $\frac{d}{dx}\left(\frac{X(\left(w^*_{\mu^0}\right)^{-1}(x),x)}{\left(w^*_{\mu^0}\right)'\left(\left(w^*_{\mu^0}\right)^{-1}(x)\right)}\right)\frac{\left(w^*_{\mu^0}\right)'\left(\left(w^*_{\mu^0}\right)^{-1}(x)\right)}{X(\left(w^*_{\mu^0}\right)^{-1}(x),x)}$. Since $x \in J$, $\left(w^*_{\mu}\right)^{-1}(x) < \left(w^*_{\mu^0}\right)^{-1}(x)$. By (20), at any skill level $f \in \left[\left(w^*_{\mu^0}\right)^{-1}\left(\min w^*_{\mu^0}\right), \left(w^*_{\mu}\right)^{-1}\left(\max w^*_{\mu}\right)\right]$, $\left(w^*_{\mu^0}\right)'(f) = \frac{1}{f}.\left(w^*_{\mu^0}(f)+1\right)$ and:

$$\left(w^*_{\mu}\right)'(f) = \frac{1}{f}.\left(w^*_{\mu}(f)+1\right).\frac{1}{\Gamma(f)}$$

where $\Gamma(f) = 1 - \frac{u_j'(w_\mu^*(f)-\theta)}{\bar{\xi}'(\frac{w_\mu^*(f)+1}{f})} f \geq 1$. Hence $\frac{\left(w_{\mu 0}^*\right)''(f)}{\left(w_{\mu 0}^*\right)'(f)} = 0$ and:

$$
\begin{aligned}
\left(w_\mu^*\right)''(f) &= \frac{\left(w_\mu^*\right)'(f)}{f\Gamma(f)} - \frac{(w_\mu^*(f)+1).\frac{d}{df}\left(f\Gamma(f)\right)}{\left(f\Gamma(f)\right)^2} \\
\frac{\left(w_\mu^*\right)''(f)}{\left(\left(w_\mu^*\right)'(f)\right)^2} &= \frac{1}{f\Gamma(f)\left(w_\mu^*\right)'(f)} - \frac{\frac{d}{df}\left(f\Gamma(f)\right)}{f\Gamma(f)\left(w_\mu^*\right)'(f)} \\
&= \frac{1-\Gamma(f)-f\Gamma'(f)}{f\Gamma(f)\left(w_\mu^*\right)'(f)} \\
&= -\frac{\Gamma(f)-1}{\left(w_\mu^*\right)'(f)\Gamma(f)}\left(\frac{1}{f} + \frac{\Gamma'(f)}{\Gamma(f)-1}\right) \\
&= -\frac{\Gamma(f)-1}{\left(w_\mu^*\right)'(f)\Gamma(f)}\left(\frac{1}{f} + \frac{\frac{d}{df}\left(\frac{u_j'(w_\mu^*(f)-\theta)}{\bar{\xi}'(\frac{w_\mu^*(f)+1}{f})}\right) f + \frac{u_j'(w_\mu^*(f)-\theta)}{\bar{\xi}'(\frac{w_\mu^*(f)+1}{f})}}{\frac{u_j'(w_\mu^*(f)-\theta)}{\bar{\xi}'(\frac{w_\mu^*(f)+1}{f})} f}\right) \\
&= -\frac{\Gamma(f)-1}{\left(w_\mu^*\right)'(f)\Gamma(f)}\left(\frac{2}{f} + \frac{\frac{d}{df}\left(\frac{u_j'(w_\mu^*(f)-\theta)}{\bar{\xi}'(\frac{w_\mu^*(f)+1}{f})}\right)}{\frac{u_j'(w_\mu^*(f)-\theta)}{\bar{\xi}'(\frac{w_\mu^*(f)+1}{f})}}\right)
\end{aligned}
$$

Note moreover that, under specification (1) considered in the body of the paper, $-\frac{u_j'(w_\mu^*(f)-\theta)}{\bar{\xi}'(\frac{w_\mu^*(f)+1}{f})} = \eta$, so $\frac{\left(w_\mu^*\right)''(f)}{\left(\left(w_\mu^*\right)'(f)\right)^2} = -\frac{\Gamma(f)-1}{\left(w_\mu^*\right)'(f)\Gamma(f)} \cdot \frac{2}{f}$.

But then, under specification (1):

$$
\begin{aligned}
&\frac{d}{dx}\left(\frac{X(\left(w_\mu^*\right)^{-1}(x),x)}{\left(w_\mu^*\right)'\left(\left(w_\mu^*\right)^{-1}(x)\right)}\right)\frac{\left(w_\mu^*\right)'\left(\left(w_\mu^*\right)^{-1}(x)\right)}{X(\left(w_\mu^*\right)^{-1}(x),x)}\\
=&\frac{\nabla X\left(\left(w_\mu^*\right)^{-1}(x),x\right).(\frac{1}{\left(w_\mu^*\right)'(\left(w_\mu^*\right)^{-1}(x))},1)}{X(\left(w_\mu^*\right)^{-1}(x),x)}-\frac{\left(w_\mu^*\right)''\left(\left(w_\mu^*\right)^{-1}(x)\right)}{\left(\left(w_\mu^*\right)'\left(\left(w_\mu^*\right)^{-1}(x)\right)\right)^2}\\
=&\frac{\nabla X\left(\left(w_\mu^*\right)^{-1}(x),x\right).(\frac{\left(w_\mu^*\right)^{-1}(x)}{1+x}\Gamma(x),1)}{X\left(\left(w_\mu^*\right)^{-1}(x),x\right)}+\frac{\Gamma(\left(w_\mu^*\right)^{-1}(x))-1}{\left(w_\mu^*\right)'(\left(w_\mu^*\right)^{-1}(x))\Gamma(\left(w_\mu^*\right)^{-1}(x))}.\frac{2}{\left(w_\mu^*\right)^{-1}(x)}\\
=&\frac{\nabla X\left(\left(w_\mu^*\right)^{-1}(x),x\right).(\frac{\left(w_\mu^*\right)^{-1}(x)}{1+x},1)}{X\left(\left(w_\mu^*\right)^{-1}(x),x\right)}+\frac{\frac{\partial X}{\partial f}\left(\left(w_\mu^*\right)^{-1}(x),x\right)\frac{\left(w_\mu^*\right)^{-1}(x)}{1+x}\left(\Gamma(\left(w_\mu^*\right)^{-1}(x))-1\right)}{X\left(\left(w_\mu^*\right)^{-1}(x),x\right)}\\
&+\frac{\Gamma(\left(w_\mu^*\right)^{-1}(x))-1}{\left(w_\mu^*\right)'(\left(w_\mu^*\right)^{-1}(x))\Gamma(\left(w_\mu^*\right)^{-1}(x))}.\frac{2}{\left(w_\mu^*\right)^{-1}(x)}\\
=&\frac{\nabla X\left(\left(w_\mu^*\right)^{-1}(x),x\right).(\frac{\left(w_\mu^*\right)^{-1}(x)}{1+x},1)}{X\left(\left(w_\mu^*\right)^{-1}(x),x\right)}\\
&+\frac{\Gamma(\left(w_\mu^*\right)^{-1}(x))-1}{\left(w_\mu^*\right)'(\left(w_\mu^*\right)^{-1}(x))\Gamma(\left(w_\mu^*\right)^{-1}(x))}\left(\frac{\frac{\partial X}{\partial f}\left(\left(w_\mu^*\right)^{-1}(x),x\right)}{X\left(\left(w_\mu^*\right)^{-1}(x),x\right)}+\frac{2}{\left(w_\mu^*\right)^{-1}(x)}\right)\\
\leq&\frac{\nabla X\left(\left(w_\mu^*\right)^{-1}(x),x\right).(\frac{\left(w_\mu^*\right)^{-1}(x)}{1+x},1)}{X\left(\left(w_\mu^*\right)^{-1}(x),x\right)}\\
\leq&\frac{\nabla X\left(\left(w_{\mu^0}^*\right)^{-1}(x),x\right).(\frac{\left(w_{\mu^0}^*\right)^{-1}(x)}{1+x},1)}{X\left(\left(w_{\mu^0}^*\right)^{-1}(x),x\right)}\\
=&\frac{d}{dx}\left(\frac{X(\left(w_{\mu^0}^*\right)^{-1}(x),x)}{\left(w_{\mu^0}^*\right)'\left(\left(w_{\mu^0}^*\right)^{-1}(x)\right)}\right)\frac{\left(w_{\mu^0}^*\right)'\left(\left(w_{\mu^0}^*\right)^{-1}(x)\right)}{X(\left(w_{\mu^0}^*\right)^{-1}(x),x)}
\end{aligned}
$$

where the first inequality follows from the first clause of Assumption 1 (lower bound on the productivity-elasticity of labour), and the second from the second clause (property (5)). Hence $\frac{d}{dx}\left(\frac{X(\left(w_\mu^*\right)^{-1}(x),x)}{X(\left(w_{\mu^0}^*\right)^{-1}(x),x)}.\frac{\left(w_{\mu^0}^*\right)'\left(\left(w_{\mu^0}^*\right)^{-1}(x)\right)}{\left(w_\mu^*\right)'\left(\left(w_\mu^*\right)^{-1}(x)\right)}\right)\leq 0$ for all $x\in J$, as required. □

Remark 2. This is the only part of the proof of Theorem 2 that relies on the utility specification (1). Versions of this result hold for the general specification (17), with the same conclusion but subtly different assumptions. In particular,

it is clear from the calculations above that the desired inequality holds in the general case whenever $\left|\frac{\frac{\partial X}{\partial s}(f,x)}{X(f,x)}\right| \geq \frac{2}{f} + \frac{\frac{d}{df}\left(\frac{u_j'(w_\mu^*(f)-\theta)}{\bar{\xi}'(\frac{w_\mu^*(f)+1}{f})}\right)}{\frac{u_j'(w_\mu^*(f)-\theta)}{\bar{\xi}'(\frac{w_\mu^*(f)+1}{f})}} = \frac{2}{f} + \frac{\frac{d}{df}\left(u_j'(w_\mu^*(f)-\theta)\right)}{u_j'(w_\mu^*(f)-\theta)} - \frac{\frac{d}{df}\left(\bar{\xi}'(\frac{w_\mu^*(f)+1}{f})\right)}{\bar{\xi}'(\frac{w_\mu^*(f)+1}{f})}$. So, for instance, the result holds under the same assumption for the skill-elasticity of labour (first clause of Assumption 1) whenever the disutility for inequality function u_j is concave (for all types j), and utility for money is linear. Observations like this (or others involving strengthenings of Assumption 1) can be used to provide generalisations of Theorem 2 to the utility specification (17).

Recall that $G_{w_\mu^*}$ is the cumulative distribution for $g_{w_\mu^*}$ (and similarly for $G_{w_{\mu^0}^*}$), i.e. $G_{w_\mu^*}(x) = 0$ for $x < 1$, $G_{w_\mu^*}(1) = \frac{1}{2}$ and $G_{w_\mu^*}(x) = \frac{1}{2} + \frac{1}{2}\int_1^x \frac{X(\left(w_\mu^*\right)^{-1}(y),y)}{\left(w_\mu^*\right)'\left(\left(w_\mu^*\right)^{-1}(y)\right)\int_{\underline{f}}^{\overline{f}} X(f,w_\mu^*(f))df}dy$ for $x > 1$. Clearly $G_{w_\mu^*}$ is differentiable on $(\min w_\mu^*, \max w_\mu^*)$; moreover, since, as noted in Appendix A.2, $X(\left(w_\mu^*\right)^{-1}\left(\min w_\mu^*\right), \min w_\mu^*) = 0$, $G_{w_\mu^*}$ is differentiable (with derivative zero) for all $x \leq \min w_\mu^*$. Similarly, $G_{w_{\mu^0}^*}$ is differentiable on $[0, \max w_{\mu^0}^*)$.

As a point of notation for the following proofs, define $H_\mu, H_{\mu^0} : (1,\infty) \to [0,1]$ by: $H_\mu(x) = \int_1^x \frac{X(\left(w_\mu^*\right)^{-1}(y),y)}{\left(w_\mu^*\right)'\left(\left(w_\mu^*\right)^{-1}(y)\right)\int_{\underline{f}}^{\overline{f}} X(f,w_\mu^*(f))df}dy$ and $H_{\mu^0}(x) = \int_1^x \frac{X(\left(w_{\mu^0}^*\right)^{-1}(y),y)}{\left(w_{\mu^0}^*\right)'\left(\left(w_{\mu^0}^*\right)^{-1}(y)\right)\int_{\underline{f}}^{\overline{f}} X(f,w_{\mu^0}^*(f))df}dy$.

Lemma 4. $G_{w_\mu^*}(x) > G_{w_{\mu^0}^*}(x)$ for all $x \in \left(w_\mu^*(\hat{f}), \max w_\mu^*\right]$ where $\hat{f}$ is as in Lemma 1.

Proof of Claim 4. Clearly it suffices to show that $H_\mu(x) > H_{\mu^0}(x)$ for all $x \in \left(w_\mu^*(\hat{f}), \max w_\mu^*\right]$. Since $w_\mu^*(f) < w_{\mu^0}^*(f)$ for $f > \hat{f}$, $\left(w_\mu^*\right)^{-1}(x) \geq \left(w_{\mu^0}^*\right)^{-1}(x)$ for $x \geq w_\mu^*(\hat{f})$. Moreover, since $\frac{\partial X}{\partial x} > 0$ wherever X is non-zero, $X(f, w_\mu^*(f)) < X(f, w_{\mu^0}^*(f))$ for all $f > \hat{f}$ with $X(f, w_{\mu^0}^*(f)) > 0$. Hence, for $x \in \left(w_\mu^*(\hat{f}), \max w_\mu^*\right]$:

$$
\begin{aligned}
1 - H_\mu(x) &= \int_x^\infty \frac{X(\left(w_\mu^*\right)^{-1}(y), y)}{\left(w_\mu^*\right)'\left(\left(w_\mu^*\right)^{-1}(y)\right)\int_{\underline{f}}^{\overline{f}} X(f, w_\mu^*(f))df} dy \\
&= \int_{\left(w_\mu^*\right)^{-1}(x)}^{\overline{f}} \frac{X(f, w_\mu^*(f))}{\int_{\underline{f}}^{\overline{f}} X(f, w_\mu^*(f))df} df \\
&\leq \int_{\left(w_{\mu^0}^*\right)^{-1}(x)}^{\overline{f}} \frac{X(f, w_\mu^*(f))}{\int_{\underline{f}}^{\overline{f}} X(f, w_\mu^*(f))df} df \\
&< \int_{\left(w_{\mu^0}^*\right)^{-1}(x)}^{\overline{f}} \frac{X(f, w_{\mu^0}^*(f))}{\int_{\underline{f}}^{\overline{f}} X(f, w_{\mu^0}^*(f))df} df \\
&= \int_x^\infty \frac{X(\left(w_{\mu^0}^*\right)^{-1}(y), y)}{\left(w_{\mu^0}^*\right)'\left(\left(w_{\mu^0}^*\right)^{-1}(y)\right)\int_{\underline{f}}^{\overline{f}} X(f, w_{\mu^0}^*(f))df} dy \\
&= 1 - H_{\mu^0}(x)
\end{aligned}
$$

where the first inequality follows from the first observation, and the second from the second observation and Lemma 2. This establishes the result. □

Proposition 4. *Under Assumption 1, $g_{w_\mu^{*0}}$ dominates $g_{w_\mu^*}$ in the sense of first-order stochastic dominance, i.e. $G_{w_\mu^*}(x) \geq G_{w_{\mu^0}^*}(x)$ for all x. Moreover, the inequality is strict for all $x \in G^{-1}_{w_\mu^*}((0.5, 1])$.*[37]

Proof of Proposition 4. Given the formula for $G_{w_\mu^*}$, it suffices to show that $H_\mu(x) \geq H_{\mu^0}(x)$ for all x, with strict inequality for all $x \in \left(\min\left\{\min w_\mu^*, \min w_{\mu^0}^*\right\}, \max\left\{\max w_\mu^*, \max w_{\mu^0}^*\right\}\right) = \left(\min w_\mu^*, \max w_{\mu^0}^*\right)$. Since $\max w_\mu^* < \max w_{\mu^0}^*$, there exists an interval $(\underline{x}, \max w_{\mu^0}^*)$ for some $\underline{x} < \max w_\mu^*$, with $H_\mu(x) > H_{\mu^0}(x)$ for all $x \in (\underline{x}, \max w_{\mu^0}^*)$. Moreover, since $\min w_\mu^* < \min w_{\mu^0}^*$, there exists an interval $(\min w_\mu^*, \overline{x})$ for some $\overline{x} > \min w_{\mu^0}^*$ with $H_\mu(x) > H_{\mu^0}(x)$ for all $x \in (\min w_\mu^*, \overline{x})$. We now show that there exists no $x \in \left(\min w_{\mu^0}^*, \max w_\mu^*\right)$ such that $H_\mu(x) < H_{\mu^0}(x)$. For reductio, assume that there is such a point. By the previous observations, there exists a point where H_μ crosses H_{μ^0} from above, and another higher point where H_μ crosses H_{μ^0} from below: $y = \inf\left\{x : H_\mu(x) < H_{\mu^0}(x)\right\}$ and $y' = \sup\left\{x : H_\mu(x) < H_{\mu^0}(x)\right\}$, respectively, are such points. Clearly $y < y'$. By their definition, $H_\mu(y) = H_{\mu^0}(y)$, $H_\mu(y') = H_{\mu^0}(y')$, $\frac{dH_\mu}{dx}(y) < \frac{dH_{\mu^0}}{dx}(y)$ and $\frac{dH_\mu}{dx}(y') > \frac{dH_{\mu^0}}{dx}(y')$, and thus:

[37] The inverse distribution G^{-1} is defined in footnote 15.

$$\frac{X(\left(w_\mu^*\right)^{-1}(y),y)}{\left(w_\mu^*\right)'\left(\left(w_\mu^*\right)^{-1}(y)\right)\int_{\underline{f}}^{\overline{f}}X(f,w_\mu^*(f))df} < \frac{X(\left(w_{\mu^0}^*\right)^{-1}(y),y)}{\left(w_{\mu^0}^*\right)'\left(\left(w_{\mu^0}^*\right)^{-1}(y)\right)\int_{\underline{f}}^{\overline{f}}X(f,w_{\mu^0}^*(f))df}$$

$$\frac{X(\left(w_\mu^*\right)^{-1}(y'),y')}{\left(w_\mu^*\right)'\left(\left(w_\mu^*\right)^{-1}(y')\right)\int_{\underline{f}}^{\overline{f}}X(f,w_\mu^*(f))df} > \frac{X(\left(w_{\mu^0}^*\right)^{-1}(y'),y')}{\left(w_{\mu^0}^*\right)'\left(\left(w_{\mu^0}^*\right)^{-1}(y')\right)\int_{\underline{f}}^{\overline{f}}X(f,w_{\mu^0}^*(f))df}$$

Claim 4. There exists $\min w_\mu^* < \hat{x} \leq w_\mu^*(\hat{f})$ such that $\frac{dH_\mu}{dx}(x) > \frac{dH_{\mu^0}}{dx}(x)$ for all $\min w_\mu^* < x < \hat{x}$ and $\frac{dH_\mu}{dx}(x) \leq \frac{dH_{\mu^0}}{dx}(x)$ for all $\hat{x} < x < w_\mu^*(\hat{f})$.

Proof of Claim 4. By the previous observations, $\frac{dH_{\mu^0}}{dx}(x) = 0 < \frac{dH_\mu}{dx}(x)$ for all $x \in (\min w_\mu^*, \min w_{\mu^0}^*]$. Since $w_\mu^*(f) > w_{\mu^0}^*(f)$ for $f < \hat{f}$ by Lemma 1, $\left(w_\mu^*\right)^{-1}(x) < \left(w_{\mu^0}^*\right)^{-1}(x)$ for $x \in \left(\min w_{\mu^0}, w_\mu^*(\hat{f})\right)$. It follows from Lemma 3 that $\frac{X(\left(w_\mu^*\right)^{-1}(x),x)}{X(\left(w_{\mu^0}^*\right)^{-1}(x),x)} . \frac{w_{\mu^0}'\left(\left(w_{\mu^0}^*\right)^{-1}(x)\right)}{w_\mu'\left(\left(w_\mu^*\right)^{-1}(x)\right)}$ is monotonically decreasing on $x \in \left(\min w_{\mu^0}^*, w_\mu^*(\hat{f})\right)$. It follows from the fact that $\frac{dH_{\mu^0}}{dx}(x) = 0 < \frac{dH_\mu}{dx}(x)$ for all $x \in (\min w_\mu^*, \min w_{\mu^0}^*]$ that

$$\left\{ x \in (\min w_\mu^*, w_\mu^*(\hat{f})) : \begin{array}{c} \frac{X(\left(w_\mu^*\right)^{-1}(x),x)}{w_\mu'\left(\left(w_\mu^*\right)^{-1}(x)\right)\int_1^\infty X(\left(w_\mu^*\right)^{-1}(y),y)} \\ > \frac{X(\left(w_{\mu^0}^*\right)^{-1}(x),x)}{w_{\mu^0}'\left(\left(w_{\mu^0}^*\right)^{-1}(x)\right)\int_1^\infty X(\left(w_{\mu^0}^*\right)^{-1}(y),y)} \end{array} \right\} \neq \emptyset$$

Letting

$$\hat{x} = \sup \left\{ x \in (\min w_\mu^*, w_\mu^*(\hat{f})) : \begin{array}{c} \frac{X(\left(w_\mu^*\right)^{-1}(x),x)}{w_\mu'\left(\left(w_\mu^*\right)^{-1}(x)\right)\int_1^\infty X(\left(w_\mu^*\right)^{-1}(y),y)} \\ > \frac{X(\left(w_{\mu^0}^*\right)^{-1}(x),x)}{w_{\mu^0}'\left(\left(w_{\mu^0}^*\right)^{-1}(x)\right)\int_1^\infty X(\left(w_{\mu^0}^*\right)^{-1}(y),y)} \end{array} \right\}$$

and noting that $\hat{x} \geq \min w_{\mu^0}^*$, it follows from the previous observations that $\frac{dH_\mu}{dx}(x) > \frac{dH_{\mu^0}}{dx}(x)$ for all $\min w_\mu^* < x < \hat{x}$ and $\frac{dH_\mu}{dx}(x) \leq \frac{dH_{\mu^0}}{dx}(x)$ for all $\hat{x} < x < w_\mu^*(\hat{f})$, as required.

□

Lemma 4 implies that $y < y' < w_\mu^*(\hat{f})$. However Claim 4 contradicts the existence of such y and y'. So $H_\mu(x) \geq H_{\mu^0}(x)$ for all $x \in (1, \infty)$. Moreover, the claims taken together imply that $H_\mu(x) > H_{\mu^0}(x)$ for all $w_\mu^*(\hat{f}) \leq x \leq \max w_{\mu^0}^*$ and that there exists $\min w_\mu^* < \hat{x} < w_\mu^*(\hat{f})$ such that $\frac{dH_\mu}{dx}(x) > \frac{dH_{\mu^0}}{dx}(x)$ for all $\min w_\mu^* < x < \hat{x}$ and $\frac{dH_\mu}{dx}(x) \leq \frac{dH_{\mu^0}}{dx}(x)$ for all $\hat{x} < x < w_\mu^*(\hat{f})$. It follows that $H_\mu(x) > H_{\mu^0}(x)$ for all $x \in (\min w_\mu^*, \max w_{\mu^0}^*)$, as required.

□

We now complete the proof of Theorem 2. Let $a' = \frac{1}{2} - \frac{1}{2}\int_{\hat{f}}^{\overline{f}} \frac{X(f,w^*_\mu(f))}{\int_{\underline{f}}^{\overline{f}} X(f,w^*_\mu(f))df} df$, where $\hat{f}$ is as in Lemma 1. By the definition of $G_{w^*_\mu}$ and $G_{w^*_{\mu^0}}$, for every $0 < b \leq 0.5$, $G^{-1}_{w^*_\mu}(b) = G^{-1}_{w^*_{\mu^0}}(b)$; by Lemma 4, $G^{-1}_{w^*_\mu}(1-a) < G^{-1}_{w^*_{\mu^0}}(1-a)$ for every $0 \leq a \leq a'$. Moreover, by Proposition 4, $G^{-1}_{w^*_\mu}(1-a) < G^{-1}_{w^*_{\mu^0}}(1-a)$ for every $0 \leq a < 0.5$, under Assumption 1. It follows that $\iota(g_{w^*_\mu}) < \iota(g_{w^*_{\mu^0}})$ for every $\iota^{quant}_{a,b}$ with $0 \leq b \leq 0.5$ and $0 \leq a \leq a'$; and moreover this inequality holds for every ι in the quantile family under Assumption 1.

Now consider the share family. For every $0 < b \leq 0.5$, $\int_0^b G^{-1}_{w^*_\mu}(\tau)d\tau = b = \int_0^b G^{-1}_{w^*_{\mu^0}}(\tau)d\tau$. It follows from Lemma 4, that $\int_{1-a}^1 G^{-1}_{w^*_\mu}(\tau)d\tau < \int_{1-a}^1 G^{-1}_{w^*_{\mu^0}}(\tau)d\tau$ for every $0 \leq a \leq a'$, and from Proposition 4 that this holds for all $0 \leq a < 0.5$ under Assumption 1. It follows that $\iota(g_{w^*_\mu}) < \iota(g_{w^*_{\mu^0}})$ for every $\iota^{share}_{a,b}$ with $0 \leq b \leq 0.5$ and $0 \leq a \leq a'$; and moreover this inequality holds for every ι in the share family under Assumption 1.

A.5 Proofs of results in Section 3.4

Proof of Theorem 3. We prove Theorem 3 under the general specification (17). Let (c^*, q^*) be the allocation arising from an equilibrium in this market, let p^* be the equilibrium price vector (giving a price of 1 to the numéraire, and the equilibrium prices to the other commodities, i.e. the good at various inequality levels) and w^* the equilibrium wage schedule. Suppose, for reductio, that (c, q) is a feasible allocation Pareto dominating (c^*, q^*)—it yields higher utility for all consumers and strictly higher utility for a set of consumers of strictly positive measure. It follows that c and c^* differ on a set of strictly positive measure.

For each consumer k, by the utility maximisation (point 2. in the definition of equilibrium), if $U_k(c(k)) > U_k(c^*(k))$, then the budget constraint is not respected: $p^*(c_1(k)) + c_2(k) > \hat{n}_k$. Moreover, since the utility function (17) is strictly increasing in n, and so is locally non-satiated, if $U_k(c(k)) \geq U_k(c^*(k))$, then $p^*(c_1(k)) + c_2(k) \geq \hat{n}_k$. So, under the allocation (c, q):

$$\int_P p^*(c_1(k)) + c_2(k)dk > \int_P \hat{n}_k dk \tag{38}$$

At prices p^*, any firm producing the good by hiring the H-type worker at skill level f for wage i makes profits $p^*(i).f - (i+1)$. Since (c, q) satisfies the feasibility constraints, the total profits under p^* are strictly positive: for r generating q as specified in Section 3.4,

$$\begin{aligned}\int_{\underline{f}}^{\overline{f}} \left(p^*(r_2(f)).f - (r_2(f)+1)\right) r_1(f)df &= \int_P p^*(c_1(k)) + c_2(k)dk - \int_P \hat{n}_k dk \\ &> 0\end{aligned}$$

where the equality follows from (14) and (15), and the inequality from (38). Consider $\bar{r}$ that maximises the total profits across all firms; since $\bar{r}_1(f) =$

$X(f, \bar{r}_2(f))$, $\bar{r}_2$ maximises $\int_{\underline{f}}^{\overline{f}} (p^*(\bar{r}_2(f)).f - (\bar{r}_2(f)+1))\, X(f, \bar{r}_2(f))df$. By the Euler-Lagrange equation for this optimisation problem, there is a constant C with $(p^*(\bar{r}_2(f)).f - (\bar{r}_2(f)+1))\, X(f, \bar{r}_2(f)) = C$ for all f for which $\bar{r_2}(f) > 0$; moreover, since (c, q) results in strictly positive profits, $C > 0$. Since $p^*(i) = \frac{i+1}{w^{*-1}(i)}$ (Appendix A.1), it follows that:

$$\frac{f}{w^{*-1}(\bar{r}_2(f))} - 1 = \frac{C}{(\bar{r}_2(f)+1)X(f, \bar{r}_2(f))}$$

for all $f \in [\underline{f}, \overline{f}]$ for which $\bar{r_2}(f) > 0$, whence $f > w^{*-1}(\bar{r}_2(f))$ for all such f. So, for all f for which $\bar{r_2}(f) > 0$, $w^*(f) > r_2(f)$, whence, by the fact $\frac{\partial X}{\partial x} > 0$ wherever X is non-zero,

$$N = \int_{\underline{f}}^{\overline{f}} fX(f, w^*(f))df > \int_{\underline{f}}^{\overline{f}} fX(f, r_2(f))df$$

contradicting the market clearing conditions (14). Hence there exists no such $\bar{r}$, contradicting the assumption that (c, q) is a feasible allocation, and thus establishing the result. □

Proof of Proposition 1. For ease of presentation, we work with utility specification (1); a similar argument holds for the specification (17). Let $w^*_{\mu^0} = Cf - 1$, and let $\underline{f_K}$ be as defined in (11)—i.e. it is the lowest productivity level employed under $w^*_{\mu^0}$. By assumption $\underline{f_K} > \underline{f}$; by the observations in Appendix A.2, $\underline{f_K} = \inf\{f : X(f, Cf-1) > 0\}$. We consider the case where $w^*_{\mu^0}\left(\underline{f_K}\right) > \theta$; similar arguments yield the result for the other case. Under $w^*_{\mu^0}$, the utility of a consumer k of type j serviced by workers of type f is $\hat{n_k}+v_k+\eta_j\theta-C-\eta_j(Cf-1)$. Since this is decreasing in f, it suffices to show the result under any allocation (consistent with $w^*_{\mu^0}$) where more inequality averse consumers receive less unequal goods: i.e. whenever consumers of types $j > j'$ are serviced by f and f' respectively, $f > f'$. Let (c_0, q_0) be any such allocation. Let $\bar{k}$ be the largest $j < K$ with $\mu_j \neq 0$. By the assumption on the number of inequality neutral consumers under μ, there exists an interval $[f_{\bar{k}}, f'_{\bar{k}}]$ containing skill levels servicing consumers of type $\bar{k}$ under (c_0, q_0) and $w^*_{\mu^0}$ such that $Cf - 1 > \theta$ for all $f \in [f_{\bar{k}}, f'_{\bar{k}}]$. Without loss of generality, we can assume that $f_{\bar{k}} > \underline{f_K}$; if this is not the case, we can replace it with any $f_{\bar{k}} < f'_{\bar{k}}$ for which it is. Take any $D > 0$ with $C + \eta_{\bar{k}}(C\underline{f_K} - 1) < D < C + \eta_{\bar{k}}(Cf_{\bar{k}} - 1)$; clearly such D exists. By the former inequality, $\frac{Df_K - 1}{\eta_{\bar{k}}\underline{f_K}+1} > C\underline{f_K} - 1$, so, by the continuity of X and the fact that $\frac{\partial X}{\partial x} > 0$ wherever it is positive, $X(f, \frac{Df_K-1}{\eta_{\bar{k}}\underline{f_K}+1}) > 0$. Hence $f_D = \inf\left\{f : X(f, \frac{Df-1}{\eta_{\bar{k}}f+1}) > 0,\ \frac{Df-1}{\eta_{\bar{k}}f+1} > \theta\right\} < \underline{f_K}$, and $\int_{f_D}^{f_K} fX(f, \frac{Df-1}{\eta_{\bar{k}}f+1})df > 0$. Hence there exist $f_1 \geq f_{\bar{k}}$ and $f_2 < \underline{f_K}$ such that $\int_{f_1}^{f'_{\bar{k}}} fX(f, Cf-1)df = \int_{f_D}^{f_2} fX(f, \frac{Df-1}{\eta_{\bar{k}}f+1})df > 0$. Take any such f_1, f_2, define the wage schedule w by

$$w(f) = \begin{cases} w^*_{\mu^0}(f) & f \in [\underline{f}, \overline{f}] \setminus \left([f_1, f'_{\bar{k}}] \cup [f_D, f_2]\right) \\ \frac{Df-1}{\eta_{\bar{k}}f+1} & f \in [f_D, f_2] \\ 0 & f \in [f_1, f'_{\bar{k}}] \end{cases}$$

and let c be the allocation which coincides with c_0 except for consumers serviced by workers of skill level in $[f_1, f'_{\bar{k}}]$ under $w^*_{\mu^0}$, who are serviced instead by workers of skill level $[f_D, f_2]$ under w. (I.e. under c, such consumers receive goods of inequality $\frac{Df-1}{\eta_{\bar{k}}f+1}$ and numéraire $\hat{n_k} - \frac{D+\eta_{\bar{k}}}{\eta_{\bar{k}}f+1}$, for $f \in [f_D, f_2]$.) Letting q be the production allocation generated by w under (16), (c, q) satisfies (14) and (15), and hence is a feasible allocation, by construction. Moreover, the only consumers receiving a different allocation under (c, q) and (c_0, q_0) are those serviced by $f \in [f_D, f_2]$ and $f \in [f_1, f'_{\bar{k}}]$ respectively; since, by the previous inequality of D, the utility obtained under (c_0, q_0), $\hat{n_k} + v_k + \eta_j\theta - C - \eta_j(Cf-1) < \hat{n_k} + v_k + \eta_j\theta - D$, the utility obtained under (c, q), these consumers have strictly higher utility under (c, q). By (14) and the fact that $\int_{f_D}^{f_2} fX(f, \frac{Df-1}{\eta_{\bar{k}}f+1})df > 0$, this is a set of positive measure, so (c, q) is a feasible allocation Pareto dominating (c_0, q_0), as required. □

References

Acemoglu, D. and Autor, D. (2011). Chapter 12 - Skills, Tasks and Technologies: Implications for Employment and Earnings. In Card, D. and Ashenfelter, O., editors, *Handbook of Labor Economics*, volume 4, pages 1043–1171. Elsevier.

AFL-CIO (2020). Company Pay Ratios. `https://aflcio.org/paywatch/company-pay-ratios`.

Almås, I., Cappelen, A. W., Sørensen, E. Ø., and Tungodden, B. (2010). Fairness and the Development of Inequality Acceptance. *Science*, 328(5982):1176–1178.

Almås, I., Cappelen, A. W., and Tungodden, B. (2020). Cutthroat Capitalism versus Cuddly Socialism: Are Americans More Meritocratic and Efficiency-Seeking than Scandinavians? *Journal of Political Economy*, 128(5):705551.

Alvaredo, F. (2018). *The World Inequality Report*. Harvard University Press.

Andorfer, V. A. and Liebe, U. (2012). Research on Fair Trade Consumption—A Review. *Journal of Business Ethics*, 106(4):415–435.

Antunez, K. and Papuchon, A. (2019). Les français plus sensibles aux inégalités de revenus et plus attachés au maintien des prestations sociales. synthèse des résultats du baromètre d'opinion 2018 de la drees. Technical Report 35. `https://drees.solidarites-sante.gouv.fr/IMG/pdf/dd35.pdf`.

Atkinson, A. B., Piketty, T., and Saez, E. (2011). Top Incomes in the Long Run of History. *Journal of Economic Literature*, 49(1):3–71.

Baksi, S. and Bose, P. (2007). Credence goods, efficient labelling policies, and regulatory enforcement. *Environmental and Resource Economics*, 37(2):411–430.

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., and Vohs, K. D. (2001). Bad is stronger than good. *Review of general psychology*, 5(4):323–370.

Baumol, W. J. and Oates, W. E. (1988). *The Theory of Environmental Policy*. Cambridge university press.

Card, D., Cardoso, A. R., Heining, J., and Kline, P. (2018). Firms and labor market inequality: Evidence and some theory. *Journal of Labor Economics*, 36(S1):S13–S70.

Card, D., Mas, A., Moretti, E., and Saez, E. (2012). Inequality at work: The effect of peer salaries on job satisfaction. *American Economic Review*, 102(6):2981–3003.

Chakravarty, S. R. (2009). Inequality, polarization and poverty. *Advances in distributional analysis. New York*.

Clark, A. E. and d'Ambrosio, C. (2015). Attitudes to income inequality: Experimental and survey evidence. In *Handbook of Income Distribution*, volume 2, pages 1147–1208. Elsevier.

Cowell, F. (2011). *Measuring Inequality*. Oxford University Press.

Crifo, P. and Forget, V. D. (2015). The economics of corporate social responsibility: A firm-level perspective survey. *Journal of Economic Surveys*, 29(1):112–130.

Darby, M. R. and Karni, E. (1973). Free competition and the optimal amount of fraud. *The Journal of law and economics*, 16(1):67–88.

Edmans, A. and Gabaix, X. (2016). Executive compensation: A modern primer. *Journal of Economic literature*, 54(4):1232–87.

Fehr, E. and Schmidt, K. M. (2003). Theories of fairness and reciprocity-evidence and economic applications. In Dewatripont, M., Hansen, L. P., and Turnovsky, S. J., editors, *Advances in Economics and Econometrics*, volume 1 of *Econometric Society Monographs*, pages 208–257.

Fehr, E. and Schmidt, K. M. (2006). The economics of fairness, reciprocity and altruism–experimental evidence and new theories. *Handbook of the economics of giving, altruism and reciprocity*, 1:615–691.

Gabaix, X. (2016). Power laws in economics: An introduction. *Journal of Economic Perspectives*, 30(1):185–206.

Gabaix, X. and Landier, A. (2008). Why has CEO pay increased so much? *The Quarterly Journal of Economics*, 123(1):49–100.

Hammond, P. J. (1979). Straightforward Individual Incentive Compatibility in Large Economies. *The Review of Economic Studies*, 46(2):263–282.

Hill, B. and Lloyd, T. (2021). Are People Willing to Pay for Reduced Inequality? HEC Paris Research Paper Series ECO/SCD-2020-1392, HEC Paris. `https://papers.ssrn.com/abstract=3711785`.

Huneeus, F., Kroft, K., and Lim, K. (2021). Earnings Inequality in Production Networks. Technical report, National Bureau of Economic Research.

Johansson-Stenman, O., Carlsson, F., and Daruvala, D. (2002). Measuring future grandparents' preferences for equality and relative standing. *The Economic Journal*, 112(479):362–383.

Kahneman, D., Knetsch, J. L., and Thaler, R. H. (1991). Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias. *The Journal of Economic Perspectives*, 5(1):193–206. `http://www.jstor.org/stable/1942711`.

Katz, L. F. and Autor, D. H. (1999). Chapter 26 - Changes in the Wage Structure and Earnings Inequality. In Ashenfelter, O. C. and Card, D., editors, *Handbook of Labor Economics*, volume 3, pages 1463–1555. Elsevier.

Kiatpongsan, S. and Norton, M. I. (2014). How Much (More) Should CEOs Make? a Universal Desire for More Equal Pay. *Perspectives on Psychological Science*, 9(6):587–593.

Kohut, A. (2013). Economies of emerging markets better rated during difficult times. *Pew Research Center*. https://www.pewresearch.org/wp-content/uploads/sites/2/2013/05/Pew-Global-Attitudes-Economic-Report-FINAL-May-23-20131.pdf.

Kuziemko, I., Norton, M. I., Saez, E., and Stantcheva, S. (2015). How elastic are preferences for redistribution? evidence from randomized survey experiments. *American Economic Review*, 105(4):1478–1508.

Lambert, P. (2001). *The Distribution and Redistribution of Income: Third Edition*. Manchester University Press.

Manili, J. (2021). Climate Change, Disclosure Laws, and the Supermarket Mechanism. Technical report, Bocconi.

Mas-Colell, A., Whinston, M. D., and Green, J. R. (1995). *Microeconomic Theory*. Oxford University Press, USA.

Mohan, B., Norton, M. I., and Deshpandé, R. (2015). Paying up for fair pay: Consumers prefer firms with lower CEO-to-worker pay ratios. *Harvard Business School Marketing Unit Working Paper*, (15-091).

Mohan, B., Schlager, T., Deshpandé, R., and Norton, M. I. (2018). Consumers Avoid Buying From Firms With Higher CEO-to-Worker Pay Ratios. *Journal of Consumer Psychology*, 28(2):344–352.

Mueller, H. M., Ouimet, P. P., and Simintzi, E. (2017). Wage inequality and firm growth. *American Economic Review*, 107(5):379–83.

OECD (2021). Income inequality (indicator). https://doi.org/10.1787/459aa7f1-en.

Osberg, L. and Smeeding, T. (2006). "Fair" inequality? attitudes toward pay differentials: The United States in comparative perspective. *American sociological review*, 71(3):450–473.

Pedersen, L. H., Fitzgibbons, S., and Pomorski, L. (2021). Responsible investing: The ESG-efficient frontier. *Journal of Financial Economics*, 142(2):572–597.

Piketty, T. (2013). *Le Capital Au XXIe Siècle*. Le Seuil.

Rozin, P. and Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion. *Personality and social psychology review*, 5(4):296–320.

Shaked, A. and Sutton, J. (1983). Natural Oligopolies. *Econometrica*, 51(5):1469–1483.

Song, J., Price, D. J., Guvenen, F., Bloom, N., and Von Wachter, T. (2019). Firming up inequality. *The Quarterly journal of economics*, 134(1):1–50.

Støstad, M. N. and Cowell, F. (2021). Inequality as an Externality: Consequences for Tax Design. Technical report, Paris School of Economics.

Thurow, L. C. (1971). The income distribution as a pure public good. *The Quarterly Journal of Economics*, pages 327–336.

Tirole, J. (1988). *The Theory of Industrial Organization*. MIT press.

Weitzman, M. L. (1974). Prices vs. quantities. *The review of economic studies*, 41(4):477–491.