Incentives, Globalization, and Redistribution*

Andreas Haufler††
University of Munich and CESifo

Carlo Perroni‡‡
University of Warwick and CESifo

March 2020

PRELIMINARY: PLEASE DO NOT QUOTE

Abstract

We offer an explanation for why taxes have become less progressive in many countries in parallel with an increase in income inequality. When performance-based compensation differentials are needed to incentivize effort, progressive income taxes lower post-tax income inequality but reduce the return to effort and thus undermine incentive contracts. Market integration can increase the spread of project returns and make contract choices more responsive to changes in the level of taxation, resulting in a lower optimum income tax rate, even when individuals are not inter-jurisdictionally mobile. Our results complement existing literature that derives falling rates of taxation from relocation responses.

KEY WORDS: Redistributive Taxation; Performance-based Contracts

JEL CLASSIFICATION: H21, F15, D63

---

*This paper was presented at conferences and workshops in Leipzig, Louvain, Munich and Osaka. We thank Philippe de Donder, Jean Hindriks, Yukihiro Nishimura, Pierre Pestieau, Dominik Sachs, Klaus Schmidt and Jonas von Wangenheim for helpful comments and suggestions, and Lucas Coronel for excellent research assistance. This paper was started when Andreas Haufler visited the University of Warwick and continued when Carlo Perroni visited the Center for Economic Studies in Munich. We wish to thank both institutions for their hospitality. Financial support by the German Research Foundation (DFG) through CRC TRR 190 is gratefully acknowledged.

††University of Munich, Seminar for Economic Policy, Akademiestr. 1, D-80799 Munich, Germany; e-mail: andreas.haufler@econ.lmu.de

‡‡University of Warwick, Gibbet Hill Road, Coventry, CV4 7AL, UK; e-mail: c.perroni@warwick.ac.uk

1 Introduction

Starting from the 1980s, the distribution of income has progressively become more unequal and more concentrated. Economic research has established this fact for international samples with a focus on the rise of top incomes (Atkinson et al., 2011), but also for the entire income distribution in specific countries, such as the United States (Piketty et al., 2018).[1] And yet, tax-transfer systems have not become more redistributive – on the contrary, in a number of countries they have actually become less so. Egger et al. (2019) have recently shown that, since the mid-1990s, economic globalization has resulted in a higher labor tax burden on the middle classes of OECD countries and a reduced labor tax burden for the top one percent of earners.[2] Detailed evidence for the United States also suggests that changes in U.S. tax policy since the 1980s have increased income inequality, raising the income share of the top 20% income earners (Bargain et al., 2015).

From an optimal taxation perspective, the observation of greater inequality going hand-in-hand with less redistribution through taxation poses a puzzle, and one that is of major policy importance. Standard theories of optimal taxation would predict an increase in tax progressivity in response to an increase in inequality (Saez, 2001; Slemrod and Bakija, 2001). Therefore, the combination of higher levels of inequality with less progressive income taxes can only be explained if the higher inequality is accompanied by an increased elasticity of the tax base. Accordingly, one prominent argument in the literature is that the tax base has become more elastic at the upper end of the income distribution because high income earners have become more inter-jurisdictionally mobile (e.g. Lehmann et al., 2014).

The international mobility of individuals remains limited in most OECD countries, however, and this includes individuals at the top of the income distribution. One reason for this limited mobility is the fact that many high-earning occupations involve

[1] Detailed international evidence on the development of various measures of income inequality are collected by the OECD (http://www.oecd.org/social/income-distribution-database.htm) and in the World Wealth and Income Database (Facundo et al., 2017).

[2] According to Egger et al., (2019), this reverses the pattern from the 1980s and early 1990s, when globalization has led to an increased progressivity of the income tax. Similar results have been found by Immervoll and Richardson (2011) in a comprehensive study of redistribution policy among OECD countries that considers both tax and benefit systems.

jurisdiction-specific investments (e.g., the provision of legal services).[3] At the same time international trade shocks, such as the 'China shock', have proven to have a powerful impact on the distribution of employment and income, contributing to the polarization of wage returns (Autor et al., 2013; 2015).

In this paper we therefore advance an alternative, and complementary, explanation for the simultaneous observation of increased income inequality and less progressive income taxation. Our argument is based on the internationalization of product markets, as opposed to labor markets, which increases competition and raises the output risks of firms and their workers (Rodrik, 1998; Autor et al., 2017). We study how this shock impacts on incentive-compatible contracts created by the private sector in response to informational frictions, and how this in turn affects optimal tax policy. We show that higher output risk induced by globalization makes incentive contracts more responsive to redistributive taxation, and will lead optimizing governments to reduce the income tax rate, and hence the (indirect) progressivity of the tax, for a wide range of redistributive objectives.

Incentive contracts play an important role in all OECD countries, particularly at the top of the income distribution. In the UK, for example, annual bonus payments accounted for 6.2% of total annual pay for the whole economy in 2017, and 24.2% in the financial and insurance industry. From 2000-2017 bonus income (Office for National Statistics, 2017), with rising bankers' bonuses accounting for two-thirds of the increase in the share of the top 1% in the period 1999-2011 (Bell and Van Reenen, 2014). Focusing on a panel of more than 3,000 US employees during the period 1976-1998, Lemieux et al. (2009) find that almost 40% of workers received some form of performance pay.[4] The prevalence of incentive contracts also means that the distribution of income is, at least in part, shaped by the stochastic structure of such contracts rather than by just

[3]Battisti et al. (2018, Table 1) collect immigration shares for 20 developed countries and find an average immigration share in the labor force of 17%, with wide variations across countries. The composition of the immigrant workforce also differs substantially between countries, with the share of low-skilled immigrants in the total immigrant population being above 80% in Germany, Italy and Austria, but only slightly above 50% in Ireland, Denmark, or Australia (Table C.1).

[4]A more general overview of incentive pay contracts is given in Bryson et al. (2012). They show (Figures 1 and 2) that the share of private sector employees with an incentive pay contract has risen substantially over time in most OECD countries, and is highest in the Scandinavian countries (around 30% in 2005) and in the U.S. (over 40%).

ability differentials.[5] There is also evidence that the increased competition in product markets that follows from international market integration is associated with steeper pay incentives (Guadalupe and Cuñat, 2009).

In this paper, we describe a model of second-best contracting where incomplete insurance is required to elicit effort in the presence of moral hazard. Risk-averse workers with heterogeneous abilities choose between an incentive contract and a less efficient fixed-wage contract that does not induce effort. In equilibrium, higher productivity workers select into performance-based contracts, whereas less productive workers choose fixed-wage contracts. Economic globalization is associated with higher-powered incentives within incentive contracts, which influences the sorting of productivity types between fixed-wage and incentive contracts. The dimension of globalization we focus on in our analysis is product market integration, characterizing the latter as an increase in the spread in the market returns to effort (potentially also accompanied by a positive productivity shift).[6]

We then turn to examining how globalization affects the choice of an optimal indirectly-progressive income tax rate. Our main finding is that economic globalization, while leading to a higher inequality of ex-post, pre-tax wage incomes, will *reduce* the optimal rate of redistributive income tax. Under a uniform distribution of ability types, this result holds irrespective of whether governments pursue a Rawlsian or a utilitarian objective. The fundamental reason for this result is that the increase in the spread of market returns that follows from globalization makes contract selection choices more responsive to tax increases, an effect that dominates any effect that the higher risk faced by individuals might entail for the value of redistribution.

The model thus predicts that, if economic globalization is the driving force behind the changes we observe, we can expect higher inequality to be associated with lower redistribution. But there is no causal relationship linking the two; they are simply concurrent outcomes that stem from a common underlying cause.

Our study is related to several strands of literature. In a national context, our anal-

[5] Abraham et al. (2017) find that incentive contracts to overcome moral hazard contribute about 10% of the observed US wage inequality.

[6] See Rodrik (1998) for an early and influential analysis that characterizes economic globalization as increasing the volatility of incomes. Our approach is also consistent with recent findings of increased concentration in product markets, which are dominated by internationally operating 'superstar firms' (Autor et al., 2017).

ysis is related to the large literature on the optimal progressivity of income taxes (see Diamond and Saez, 2011, for an overview) and on redistributive income taxation in the presence of risky wage earnings (see Boadway and Sato, 2015, for a recent synthesis). This literature, however, typically does not account for the role of incentive contracts in the determination of individual pay, and how these contracts respond to taxes. If earnings volatility is incorporated at all in these models, it is assumed to be exogenous. The same applies to the dynamic optimal taxation literature that links the optimal taxation of wage incomes to the taxation of capital (e.g. Golosov and Tsyvinski, 2007; Abraham et al. 2016).[7] In contrast, the literature on incentive contracts in labor markets has stressed the contribution of performance-based pay to rising wage inequality (Lemieux et al., 2009; Abraham et al., 2017). Our paper connects these two literatures.

The effects of economic globalization on optimal income taxation have so far been studied mainly for environments where high-income earners are internationally mobile (Simula and Trannoy, 2010; Bierbrauer et al., 2013; Lehmann et al., 2014; Tóbias, 2016).[8] These models generally find that economic integration leads to a 'race to the bottom' with respect to redistributive income taxes, where the level of taxation that remains feasible with mobile high-income earners depends on the shape of mobility costs for high income earners. Here we focus instead on the global competition in product markets and its effects on performance-based contracts as an alternative channel through which economic globalization influences tax policy choices.

The analysis proceeds as follows. Section 2 sets up our model of incentive-compatible contracting and studies the sorting of heterogeneous worker types into performance-based versus fixed-wage contracts. Section 3 analyzes how the equilibrium structure of contracts is affected by higher taxes and by a changes in the variance of output realizations. Section 4 turns to redistributive taxation and analyzes the effects of market integration on the optimal redistributive tax rate. Section 5 concludes.

[7]There are, however, some formal analogies between our results and those of Abraham et al. (2016), who show in a different context that optimal labor income taxes become less progressive when the ability to tax savings is limited.

[8]Empirical contributions to this literature include Kleven et al. (2014), who estimate the migration elasticity of foreigners from a Danish preferential foreigner tax scheme for top income earners. and Agrawal and Foremny (2019), who study the relocation of high-income households in response to tax changes in Spain.

2 Performance-based contracts

2.1 Preferences, technologies, and incentive-compatible contracts

We consider an economy populated by a unit mass of individuals who are ex-ante identical (at time $\tau = 0$) and are risk averse, with von Neumann-Morgenstern utility, $u(x)$, from consumption, x, satisfying $u'(x) > 0$ and $u''(x) < 0$. For tractability, the rest of our analysis will focus on the logarithmic case $u(x) = \ln x$, which exhibits falling absolute risk aversion and constant relative risk aversion.

At time $\tau = 1$ each individual is assigned a productivity type, α, drawn from a continuous distribution with positive support $[0, \overline{\alpha}]$ and cumulative density function $F(\alpha)$. As the number of individuals is large, $F(\alpha)$ also coincides with the c.d.f. of the ex-post distribution of productivity types in the population. Without loss of generality, we assume $\overline{\alpha} = 1$.

Production takes place through risky projects that are run by risk-neutral firms operating under conditions of perfect competition. Each project involves a single individual, hired by a firm as a worker. A project that employs an individual of productivity type α yields output

$$\frac{\alpha}{\pi}, \quad \pi \in (0,1)\,, \tag{1}$$

if successful and yields zero output if unsuccessful (irrespective of the individual's productivity type). The probability, $\gamma(e)$, of the project being successful depends on the individual's chosen level of effort, $e \in \{0,1\}$, as follows:

$$\gamma(e) = e\,\pi + (1-e)\,\eta\pi\,, \tag{2}$$

with $\eta \in (0,1)$. Hence, the probability of success is π with positive effort but it is only $\eta\pi < \pi$ if no effort is exerted. From (1), the expected output from a worker of productivity type α (gross of effort costs) is then

$$\gamma(e)\,\frac{\alpha}{\pi} = e\,\alpha + (1-e)\,\eta\alpha\,; \tag{3}$$

i.e. it is equal to the individual productivity type, α, if the individual exerts positive effort, but is reduced to $\eta\alpha < \alpha$ if the individual exerts no effort.

Effort is costly: the individual incurs a private utility cost ce, $c > 0$, irrespective of her productivity type. For the remainder of our discussion, we shall assume that, if the individual could secure an income equal to her expected output in all states, then exerting positive effort would be always preferred to exerting no effort. This implies $\ln \alpha - c > \ln \eta\alpha$, which in turn requires

$$c < -\ln \eta \;\equiv \bar{c}\,. \tag{4}$$

A worker's productivity type is publicly observable, and so is output. Effort, however, is not observable (or, if it is observable, it is not verifiable), which implies that wage payments cannot be directly conditioned on effort. Contracting between firms and workers thus runs against a fundamental moral hazard problem, which can only be (partially) addressed by an incentive-compatible contract that induces workers to exert effort by conditioning the wage payment on output. This implies a *performance-based* contract prescribing two different gross-of-tax wage levels: a higher wage, w_H, paid if the project is successful, and a lower wage, w_L, if the project is unsuccessful. The lower wage level can be thought of as a base wage, with the higher wage level adding a bonus, $w_H - w_L$, to the base wage contingent on success.

Income is taxed at rate t, and revenues from the tax are distributed equally amongst the population in lump-sum fashion, i.e. through a uniform transfer that individuals take as exogenous. As the population has unit mass, the uniform transfer equals total tax revenue, g. This is an *indirectly* progressive, linear income tax scheme that collects comparatively more tax revenues from higher earners while redistributing all proceeds on an equal per-capita basis. It is also an anonymous scheme, in the sense that the tax an individual pays and the transfer she receives are not conditioned on her type (whether or not that can be observed by the tax planner) but only depends on her income.

Given this tax/transfer scheme, if an individual receives a gross wage income equal to w, her net-of-tax wage income is $w^N = (1-t)\,w$ and her total disposable income is $(1-t)\,w + g$, which must equal consumption, x. A worker's expected utility from such a contract, if she chooses to exert positive effort, is thus

$$\pi \,\ln\big((1-t)\,w_H + g\big) + (1-\pi)\,\ln\big((1-t)\,w_L + g\big) - c \;\equiv EU^E. \tag{5a}$$

If zero effort is chosen, expected utility from the incentive contract is instead[9]

$$\eta\pi \,\ln\big((1-t)\,w_H + g\big) + (1-\eta\pi)\,\ln\big((1-t)\,w_L + g\big) \;\equiv EU^N. \tag{5b}$$

[9]Note that workers' utility in this no-effort incentive contract will generally differ from utility in the (no-effort) fixed wage contract, as given in (11) below.

In order to induce positive effort from workers, performance-based contracts must satisfy the incentive-compatibility constraint $EU^E \geq EU^N$. Competitive firms will choose the contract that involves the lowest expected wage cost to them and still induces workers to exert effort. Hence, $EU^E = EU^N$ must hold in a competitive equilibrium. Equating expected utilities in (5a) and (5b) gives

$$\frac{(1-t)\, w_H + g}{(1-t)\, w_L + g} = e^{\frac{c}{\pi\,(1-\eta)}} \;\equiv \Gamma > 1\,. \tag{6}$$

Condition (6) determines the equilibrium spread between the high wage, w_H, and the low wage, w_L, that just induces effort. The expected level of worker remuneration is then pinned down by the free entry and exit of risk-neutral firms under perfect competition, which implies that expected profits for a firm offering a performance-based contract to a worker of productivity type α (who will exert positive effort in equilibrium) must be zero:

$$\alpha - \big(\pi\, w_H + (1-\pi)\, w_L\big) \;\equiv E\Pi^E = 0\,. \tag{7}$$

The zero-profit condition (7) and the incentive-compatibility constraint (6) jointly determine the equilibrium levels of gross-of-tax wages in each of the two states, as functions of the productivity type, α, and of the tax/transfer scheme, (t, g):

$$w_L = \frac{\alpha - \pi\,(\Gamma-1)\, g/(1-t)}{1+\pi\,(\Gamma-1)}\,, \tag{8a}$$

$$w_H = \frac{\alpha\Gamma + (1-\pi)\,(\Gamma-1)\, g/(1-t)}{1+\pi\,(\Gamma-1)}\,, \tag{8b}$$

where Γ has been introduced in (6). From (8a) and (8b) we see that both state-contingent wage levels are rising in the worker's productivity level, α.

By (7), in a performance-based contract each worker must receive her full productivity, α, as her *expected* gross wage. However, in order to induce effort, this needs to be delivered in the form of a lottery that leaves the worker exposed to income risk: since $\Gamma > 1$, the high wage, w_H, is always above the expected productivity, α, of a worker who exerts effort, whereas the low wage, w_L, is below the worker's expected productivity. Given that workers are risk-averse and firms are risk-neutral (and therefore capable, in principle, to insure workers at no cost), such an outcome is inefficient.

Firms can, alternatively, offer workers fixed-wage contracts that fully insure the worker, paying the same wage level, w_F, in all contingencies, whether or not the project is successful. But then the firm must fully anticipate that, absent any incentive for workers to exert positive effort, workers of any productivity type will choose to exert zero effort. From (3), expected output will thus only be $\eta\alpha$. Under conditions of perfect competition, expected firm profits under this contract must also be zero:

$$\eta\alpha - w_F \;\equiv E\Pi^F = 0\,, \tag{9}$$

which implies a fixed pre-tax wage rate

$$w_F = \eta\alpha\,. \tag{10}$$

Expected utility for a worker of type α in a fixed-wage contract will thus equal

$$\ln\big((1-t)\,\eta\alpha + g\big) \;\equiv EU^F. \tag{11}$$

By (4), exerting positive effort is productively efficient. So, a fixed-wage contract removes the welfare cost associated with income risk but replaces it with a welfare cost that comes from inefficient production choices.

2.2 Contract choice and the distribution of income

A performance-based contract will only be selected if it yields a higher expected utility than a fixed-wage contract does, i.e. if $EU^E > EU^F$. Since EU^E must equal EU^N in equilibrium, we can directly compare EU^N to EU^F.

Replacing the expression for w_L from (8a) into (5b), using (11), and differentiating $EU^N - EU^F$ with respect to α, we obtain

$$\frac{\partial(EU^N - EU^F)}{\partial\alpha} = \frac{(1-t)(1-\eta)\,g}{\big((1-t)\,\alpha + g\big)\big((1-t)\,\eta\alpha + g\big)} > 0\,. \tag{12}$$

Therefore, if $EU^N \geq EU^F$ for a productivity type α', this must also be true for productivity types $\alpha \geq \alpha'$; and if $EU^N < EU^F$ for a productivity type α'', this must also be true for productivity types $\alpha < \alpha''$. Equating EU^N with EU^F, and solving for α, we obtain

$$\alpha = \frac{g}{(1-t)\,K} \equiv \hat{\alpha}\,, \qquad K \equiv -\frac{\eta\,\big(1+\pi\,(\Gamma-1)\big) - \Gamma^{\eta\pi}}{1+\pi\,(\Gamma-1)\; - \Gamma^{\eta\pi}}\,, \tag{13}$$

where K is positive for $\eta < 1$ and c sufficiently low.[10] Because of (12), the critical value $\hat{\alpha}$ partitions the range of productivity types into two intervals: all workers with

[10]See the proof of Proposition 1.

a productivity level $\alpha < \hat{\alpha}$ will choose the fixed-wage contract and exert zero effort, and workers with a productivity level $\alpha \geq \hat{\alpha}$ will choose a performance-based contract and exert positive effort.

A relevant constraint for performance-based contracts is that the low wage, w_L, cannot become negative. Lemma 1 shows that this constraint is always met in our model:

Lemma 1. *For all productivity types, $\alpha > \hat{\alpha}$, that select into performance-based contracts, the low wage level, w_L, will be strictly positive.*

PROOF*:* See Appendix B.1.

We summarize the properties of the equilibrium contract structure in the following result:

Proposition 1. *For any given combination of tax and transfer, and for effort cost, c, lying below a positive threshold $\tilde{c} < \bar{c}$, there exists a productivity level, $\hat{\alpha} \in (0, 1)$ such that:*

- *individuals with productivity types $\alpha \geq \hat{\alpha}$ will select into performance-based contracts with state-contingent wage levels, w_H and w_L, as given in* (8a)-(8b)*, and will exert positive effort;*

- *individuals of productivity types $\alpha < \hat{\alpha}$ will receive a fixed wage $w_F = \eta\alpha$ and will exert no effort.*

PROOF*:* See Appendix B.2.

Our result that performance-based contracts are concentrated among high-ability individuals is aligned with the empirical evidence. Lemieux et al. (2009, Table 1), for example, document for a sample of 3,050 U.S. employees that hourly earnings were 30% higher in performance-pay jobs, as compared to non-performance-pay jobs. Similarly, Bryson et al. (2012, Tables 4 and 5) find a statistically significant positive relationship between high education and the incidence of performance pay in a broad sample of developed countries.

Realized ex-post, pre-tax wages will then be as follows: for each productivity type $\alpha < \hat{\alpha}$, all individuals of that type will receive a wage w_F, whereas for each productivity type $\alpha \geq \hat{\alpha}$, a fraction π of individuals of that type will receive a wage w_H and a fraction $1 - \pi$ of individuals of that type will receive a wage w_L. Let $w_F^{-1}(w)$,

$w_L^{-1}(w)$ and $w_H^{-1}(w)$ denote the inverse mappings from wage realizations to productivity types, α, corresponding respectively to (10), (8a) and (8b):

$$w_F^{-1}(w) = w/\eta\,; \tag{14}$$

$$w_L^{-1}(w) = w + \pi\,(\Gamma - 1)\,\big(w + g/(1-t)\big)\,; \tag{15}$$

$$w_H^{-1}(w) = \frac{w_L^{-1}(w) - (\Gamma - 1)\,g/(1-t)}{\Gamma}\,. \tag{16}$$

The distribution of realized, pre-tax wage income in the economy will thus arise as a mixture of three component distributions:

(i) a distribution with c.d.f. $F\big(w_F^{-1}(w)\big)/F(\hat{\alpha})$ over the support $[0, \hat{\alpha}/\eta]$ and mixture weight $F(\hat{\alpha})$;

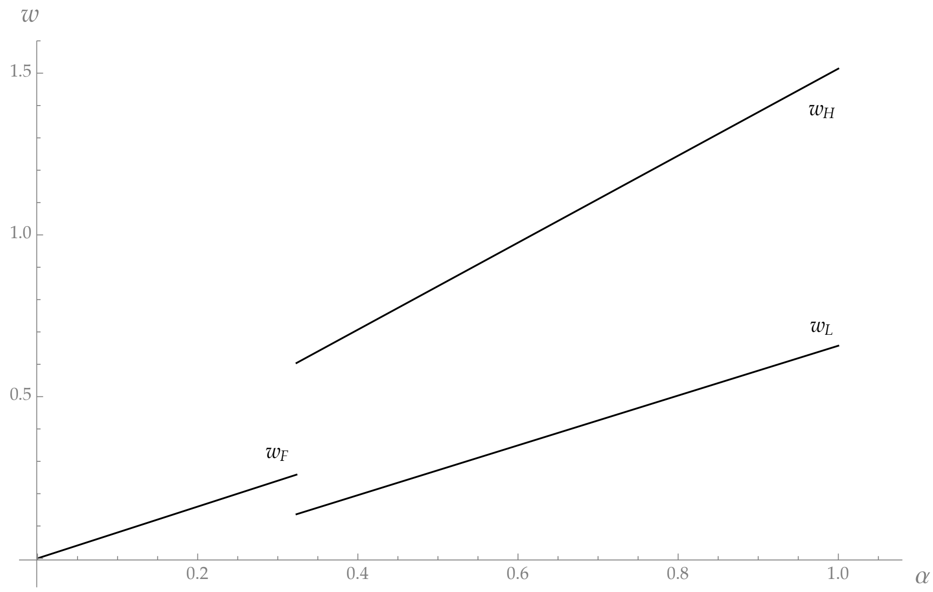
(ii) a distribution with c.d.f. $F\big(w_L^{-1}(w)\big)/\big(1 - F(\hat{\alpha})\big)$ over the support $(w_L(\hat{\alpha}), w_L(1)]$ and mixture weight $(1-\pi)\,\big(1 - F(\hat{\alpha})\big)$.

(iii) a distribution with c.d.f. $F\big(w_H^{-1}(w)\big)/\big(1 - F(\hat{\alpha})\big)$ over the support $(w_H(\hat{\alpha}), w_H(1)]$ and mixture weight $\pi\,\big(1 - F(\hat{\alpha})\big)$.

Figures 1 and 2 provide an illustration. Figure 1 gives the wage realizations by productivity type. For the given parameters, the critical value, $\hat{\alpha}$, is approximately 0.32. Individuals just below this type receive a fixed wage of about 0.25, whereas individuals marginally above $\hat{\alpha}$ receive a high wage of 0.6 with probability 0.4, and a low wage of 0.12 with probability 0.6. Figure 2 plots the distribution of pre-tax wages for the same example. Note that the distribution of productivity types is assumed to be uniform, but the resulting distribution of pre-tax wages is not, with a large mass of low earners at the bottom of the distribution and a fat tail of high earners.

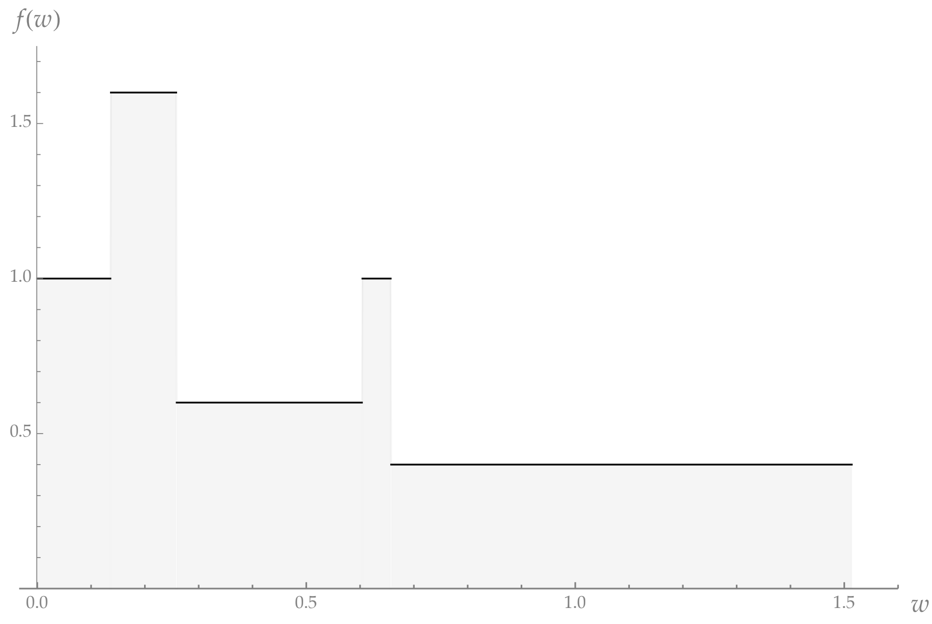
3 Taxation, market integration, and the equilibrium structure of contracts

We next examine how changes in the economic environment that are exogenous to agents' decisions, namely a change in the tax/redistribution scheme and a change in output volatility caused by market integration, are reflected in the equilibrium structure of contracts and the distribution of pre-tax wages. In the next section we will build on these results to examine how the choice of an optimal redistributive tax is affected by an increase in output risk.



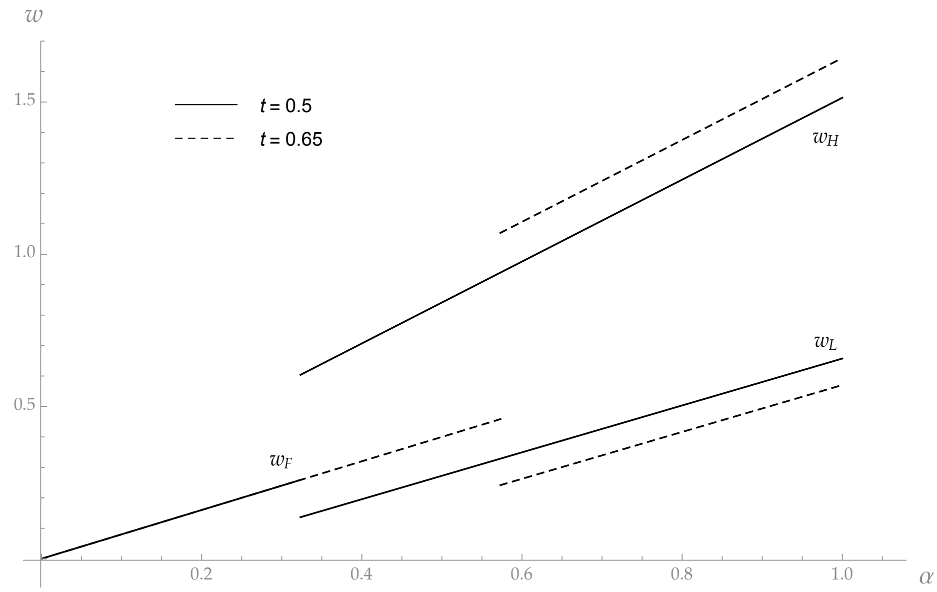
With $\pi = 2/5$, $\eta = 4/5$, $c = \bar{c}/5$, $t = 0.5$, $g \approx 0.24$ (from the government budget constraint, assuming a uniform distribution of productivity types, $F(\alpha) = \alpha$)

Figure 1: Wage realizations by productivity type



With $\pi = 2/5$, $\eta = 4/5$, $c = \bar{c}/5$, $t = 0.5$, $g \approx 0.24$, $F(\alpha) = \alpha$

Figure 2: Distribution of pre-tax wages



With $\pi = 2/5$, $\eta = 4/5$, $c = \bar{c}/5$, $F(\alpha) = \alpha$

Figure 3: Contract responses to tax changes

3.1 Taxation and contract choice

With $K > 0$, it is straightforward to establish that

$$\frac{\partial \hat{\alpha}}{\partial t} = \frac{g}{(1-t)^2 K} = \frac{\hat{\alpha}}{(1-t)} > 0, \qquad \frac{\partial \hat{\alpha}}{\partial g} = \frac{1}{(1-t)K} = \frac{\hat{\alpha}}{g} > 0. \tag{17}$$

An increase in the tax, t, for $g > 0$, raises the critical productivity level $\hat{\alpha}$ below which workers select low-output, fixed-wage contracts. An increase in the transfer, g, also raises $\hat{\alpha}$.

The transfer, g, and the tax rate, t, are linked through the government budget constraint:

$$g = t\,Q, \qquad Q = \eta \int_0^{\hat{\alpha}(t,g)} \alpha \,\mathrm{d}F(\alpha) + \int_{\hat{\alpha}(t,g)}^1 \alpha \,\mathrm{d}F(\alpha), \tag{18}$$

where we have written $\hat{\alpha}(t, g)$ to highlight the dependence of $\hat{\alpha}$ on t and g. The tax base, Q, consists of the expected output of all workers, which are employed either under a fixed-wage contract without effort (the first term in Q), or under a performance-based contract with effort (the second term).

From (17), $\hat{\alpha}$ is increasing in both t and g and the analysis of a change in t must incorporate the induced change in the transfer g. It can never be optimal, however, to select a tax rate at a level such that the total derivative of g with respect to t, $\mathrm{d}g/\mathrm{d}t$ is negative. Then, in the relevant range of rationalizable tax rate choices, t and g must be positively related. Therefore, an increase in the tax will unambiguously raise $\hat{\alpha}$, i.e. $\mathrm{d}\hat{\alpha}/\mathrm{d}t > 0$, and thus lower the total income tax base in (18). This is summarized in:

Proposition 2. *Consider two tax levels t' and $t'' > t'$, with $\hat{\alpha}'' \in (0,1)$ denoting a productivity type that is indifferent between the two form of contracts under $t = t''$. Then the range of productivity types that select into performance based contracts is narrower under t'' than under t', and the tax base is also smaller.*

Figure 3 incorporates the same parameterization as in Figures 1 and 2, but with a higher tax rate, $t = 0.65$, up from $t = 0.5$.[11] There are now more productivity types selecting into fixed-wage contracts (about 57% of them, up from 32%), resulting in a reduction in the tax base (by an amount $f(\alpha)(1-\eta)\alpha$ for each productivity type that switches from one contract form to the other). Those productivity types that do select into performance-based contracts now face a larger variation in pre-tax income. This is because contracts respond to offset the tax change by increasing the pre-tax wage differential, $w_H - w_L$, in order to still induce effort under the higher tax rate.

3.2 Output volatility, contracts, and the distribution of income

Consider next the effects of an increase in the spread of output realizations. In our formalization, output in the favourable state for a project employing an individual of productivity α and exerting positive effort is α/π whereas output in the unfavourable state is zero. A decrease in the parameter π thus amounts to a mean-preserving spread in the returns to effort, which leaves expected output unchanged at α while raising output risk.

Such a change can be interpreted as being associated with economic globalization. Appendix A presents a fully specified model that rationalizes this interpretation. In short, if ex-ante identical firms compete in any given market and face idiosyncratic cost shocks, there will be realizations in which the most productive firm will be able to capture the entire market, while other firms are left with a zero market share. Market

[11] Note that in both cases, the maximum average tax rate is much lower, peaking at about 34% for $\alpha = 1$ with $t = 0.5$ and at about 46% for $\alpha = 1$ with $t = 0.65$.

integration, gives each firm access to a larger combined market with more competitors. This increases the amount of revenues a firm can obtain if it manages to capture the larger market, but it makes such an event comparatively less likely because of the larger number of competitors. The question we want to address then is how greater product market integration, which we model here simply as a decrease in π, affects the equilibrium choice of contracts.

From (13), the critical value $\hat{\alpha}$ is decreasing in K, which in turn can be shown to be increasing in π (see (B.1) in the appendix). Intuitively, a lower success probability π raises the risk involved in a performance-based contract, making this contract less attractive to risk-averse workers, in comparison to a fixed-wage contract. From (8b), we can also see that the high wage level, w_H, in a performance-based contract is rising when π falls. It immediately follows from this that a decrease in π must unambiguously increase the fraction of total income that accrues to a subset of high earners in favorable realizations. We summarize in:

Proposition 3. *For a given t, consider a reduction in the success probability of all projects from π' to $\pi'' < \pi'$, with $\hat{\alpha}'' \in (0, 1)$ denoting a productivity type that is indifferent between the two forms of contracts under $\pi = \pi''$. Then:*

i. the range of productivity types that select into performance based contracts is narrower under π'' than under π', and the tax base is also smaller;

ii. the fraction of pre-tax income accruing to income recipients above a certain percentile threshold, p'', in the distribution of ex-post, pre-tax income is greater under π'' than under π'.

PROOF*:* See Appendix B.3

In qualitative terms, the effect of a mean-preserving spread in the returns to effort associated with greater product market integration (a fall in π), will be analogous to that of an increase in t: it will reduce the range of productivity types that select into incentive contracts and raise the dispersion of pre-tax wages within incentive contracts. This model implication is consistent with the empirical evidence that economic globalization increases wage inequality.[12]

[12]See Dreher and Gaston (2008) for empirical evidence that a broad-based measure of globalization

Figure 4 incorporates the same parameterization as in Figures 1 and 2, but with a lower probability of success, $\pi = 0.3$, down from $\pi = 0.4$. The critical productivity level below which individuals select into a fixed-wage contract has risen to about 0.49. Figure 5 shows how the distribution of pre-tax wages has become more concentrated. For $\pi = 0.3$, bonus recipients are those above the 85th percentile of the income distribution ($\hat{\alpha} \approx 0.49 \Rightarrow \pi(1 - F(\hat{\alpha})) \approx 0.15$). The fraction of total pre-tax income accruing to them is approximately 46%. The corresponding figure for earners above the 85th percentile of the distribution when $\pi = 0.4$ is lower, at approximately 36%. Approximate values of the Gini coefficient for $\pi = 0.4$ and $\pi = 0.3$ are respectively 0.39 and 0.45.

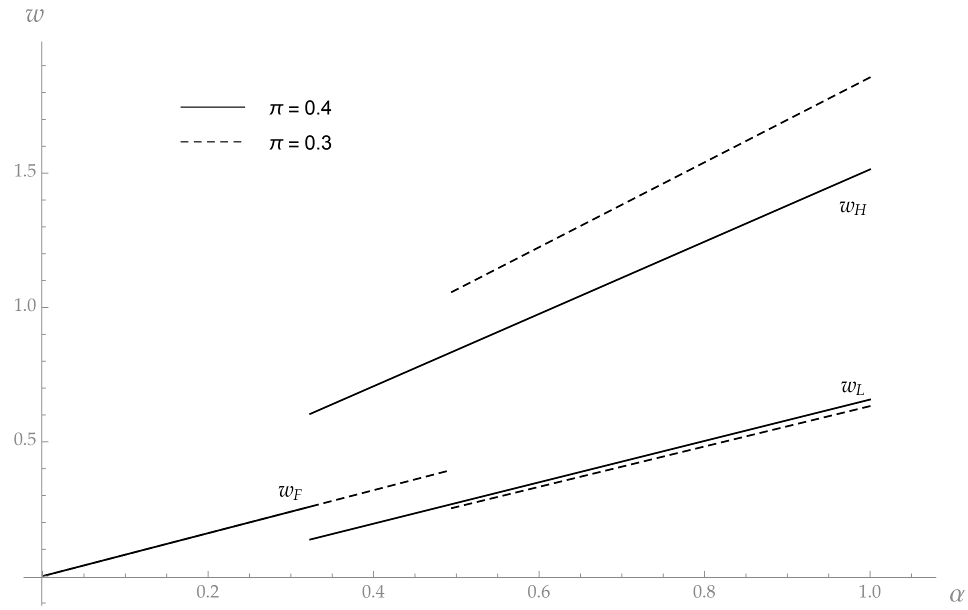
In our discussion, we have characterized the effect of economic globalization in product markets as consisting of an increase in the variance of outcomes for individual projects. However, the same conclusions continue to apply if, additionally, globalization also results in an increase in productivity.[13]

To see this, take the distribution of productivity types in our model, which has a support $[0, 1]$, to represent the 'baseline' distribution. Now suppose that globalization results in an economy-wide, type-neutral productivity shift that increases the productivity of each worker by a factor $\lambda > 1$. Further, let the cutoff productivity of the individual that is indifferent between contract types before globalization (scenario 0), for a tax rate t_0 and tax revenue g_0, be $\hat{\alpha}_0$. After globalization, the productivity of this same type is $\lambda\hat{\alpha}_0$. If this individual is also the cut-off productivity type following globalization (scenario 1), so that $\hat{\alpha}_1 = \lambda\hat{\alpha}_0$, then the model remains structurally unchanged. For an unchanged tax rate, $t_1 = t_0$, total tax revenues will then equal $g_1 = \lambda g_0$ (as the income of all workers is multiplied by λ, irrespective of contract type); and from (13) we see that the condition $\hat{\alpha}_1 = \lambda\hat{\alpha}_0 = \lambda g_0/((1-t)K)$ will indeed be met if the tax rate remains the same.[14] We argue below that there is indeed no reason for the tax

increases industrial wage inequality. Bergh and Nilsson (2010) confirm this result for a broad panel of countries. Dorn et al. (2018) have recently reexamined the link between globalization and income inequality for 140 countries over the period 1970-2014. They find a robust positive relationship between globalization and wage inequality for most countries, including China and Eastern Europe, though the effect is not significant for the most advanced economies.

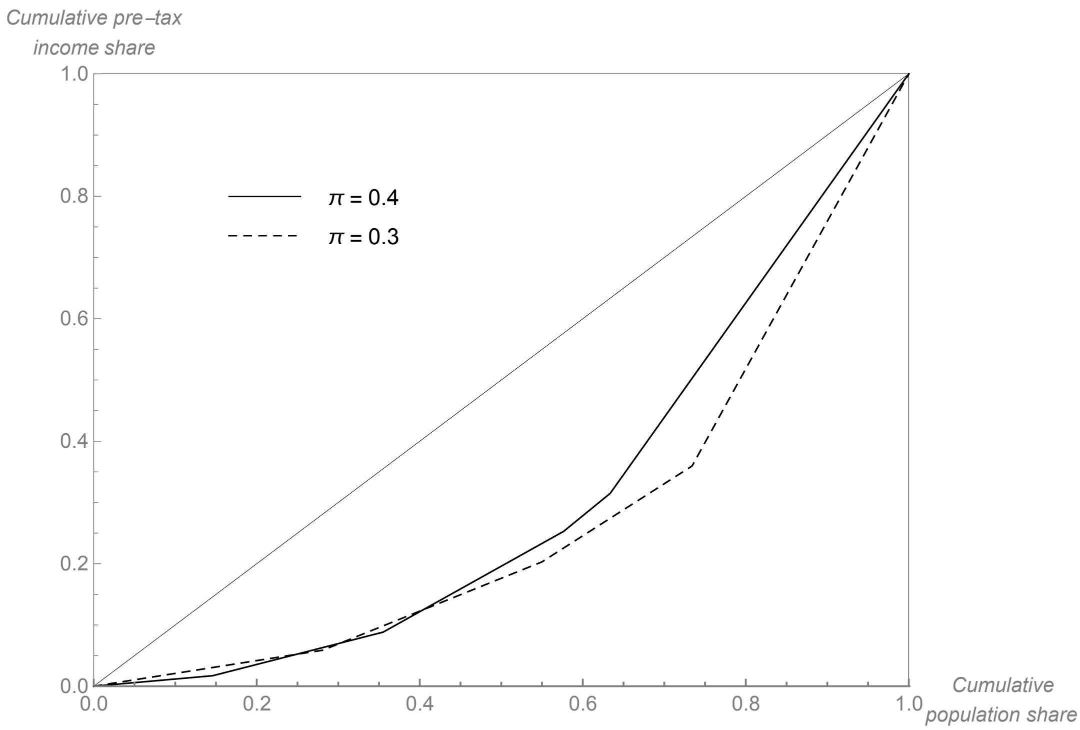
[13]The model we present in Appendix A, which provides microfoundations for our interpretation of higher output risk as being equivalent to product market integration, does predict a positive productivity shift.

[14]Intuitively, with logarithmic utility the trade-off between the fixed wage contract and the incentive



With $t = 0.5$, $\eta = 4/5$, $c = \bar{c}/5$, $F(\alpha) = \alpha$

Figure 4: Contract responses to output volatility



With $t = 0.5$, $\eta = 4/5$, $c = \bar{c}/5$, $F(\alpha) = \alpha$

Figure 5: Output volatility and income concentration

rate to change in scenario 1. It follows that our problem is fully invariant to uniform changes in the productivity of all types. Hence, all the results derived in our analysis continue to hold if, in addition to increasing the volatility of output for each worker, the globalization of product markets also leads to productivity gains.

4 Optimal redistributive taxation and economic globalization

We now turn to our main question – how the choice of redistributive taxes is affected by an increase in output risk brought about by market integration.

Since agents are risk-averse in our model and firms are risk-neutral, optimal contracts under full information (i.e. with effort being observable) would be directly conditioned on effort and would fully eliminate income risk for workers. Under asymmetric information, performance-based contracts that impose risks on risk-averse agents arise endogenously, by deliberate design, as a second-best market remedy. As a result, the linear income tax cannot play any role in mitigating this risk. However, by redistributing income between different productivity types, linear income taxes can still insure individuals against their individual productivity draw.[15]

A linear income tax will be able to reduce the variation in post-tax outcomes that comes from exogenous 'luck' in relation to agents' productivity draws, as in Varian (1980).[16] Income based taxation, however, restricts redistribution to be conditioned

contract remains unaffected when incomes under both contract types increase proportionally and effort costs in the incentive contract stay nominally unchanged.

[15]The optimal income tax literature often distinguishes between an *insurance role* of the income tax, which reduces the variance of exogenous earnings risk for a given productivity type, and a *redistributive role*, which reduces the after-tax variation in the incomes of different productivity types (see e.g. Boadway and Sato, 2015). Still, as stressed for example by Sinn (1995), the redistributive role of the income tax can alternatively be seen as providing social insurance against a variation in lifetime incomes that is unpredictable ex-ante. We follow this line of argument and interpret the redistributive role of the income tax equivalently as an insurance against individual productivity draws.

[16]Based on the assumption of exogenous income variation, Hoynes and Luttmer (2011) empirically estimate the insurance value of state tax-and-transfer programmes in the United States. They find that this insurance value rises with income, counteracting the redistributive effect of tax-and-transfer programmes.

on income realizations rather than on productivity types, and the ranking of realized earnings does not fully reflect the ranking of productivity types. By redistributing from ex-post high earners to ex-post low earners the indirectly progressive income tax achieves redistribution "on average" from high-productivity to low-productivity types, thus providing ex-ante insurance against adverse productivity draws. But it does so at the cost of interfering with contract design: a linear income tax that reduces the after-tax wage gap in the two states of the world will require a correspondingly higher before-tax wage differential for the contract to remain incentive-compatible. This, in turn, will put a break on redistribution.[17]

A change in output risk caused by economic globalization will affect the extent to which contract choices respond to taxes and will thus be reflected in the level of redistributive taxation that is optimally selected by the tax planner.

4.1 Revenue maximization

A natural starting point for studying how changes in the economic environment translate into changes in tax policy choices is the case where a government aims at maximizing tax revenues. Revenue maximization is consistent with the government pursuing redistribution under a Rawlsian social welfare function objective where only the utility level of the least well-off type enters as an argument. In an economy where the lowest productivity type receives no wage income under either contract ($\underline{\alpha} = 0$), this is the same as maximizing tax revenues and hence the per-capita transfer that can be redistributed to all individuals, including the least well-off.[18]

From (18), the first-order condition for a revenue-maximizing tax rate, t^*, is

$$\frac{\mathrm{d}g}{\mathrm{d}t^*} = Q + t^* \frac{\mathrm{d}Q}{\mathrm{d}t^*} = 0, \tag{19}$$

[17]Our following analysis is restricted to interior optima for the linear income tax rate. This requires that workers' productivity discount in a fixed wage contract is sufficiently large (i.e., η is sufficiently below unity) to prevent a confiscatory tax rate $t^* = 1$ in the government's optimum.

[18]From an ex-ante perspective, the tax planner's objective in this case does *not* coincide with the ex-ante expected utility of an individual as implied by an expectation over $u(x_{\alpha_s}) - e_\alpha c$ (with $s \in \{0, 1\}$ denoting the output state), as it would do under a utilitarian objective. Rather, it corresponds to a case where individuals exhibit ex-ante infinite risk aversion, represented by the ex-ante expected utility $\min_{\alpha,s} \{u(x_{\alpha_s}) - e_\alpha c\}$.

where

$$\frac{\mathrm{d}Q}{\mathrm{d}t} = -(1-\eta)\, f(\hat{\alpha})\, \hat{\alpha}\, \frac{\mathrm{d}\hat{\alpha}}{\mathrm{d}t}; \qquad \frac{\mathrm{d}\hat{\alpha}}{\mathrm{d}t} = \frac{\partial \hat{\alpha}}{\partial t} + \frac{\partial \hat{\alpha}}{\partial g}\, \frac{\mathrm{d}g}{\mathrm{d}t} > 0\,, \tag{20}$$

where the expression for $\partial\hat{\alpha}/\partial t > 0$ is given in (17), and $\mathrm{d}g/\mathrm{d}t$ in (20) is zero in the optimum by the envelope theorem.

In order to examine how market integration, modelled as an exogenous reduction in π, affects the optimal tax rate t^*, it is sufficient to focus on expression K in (13) – since π enters (19) and (20) only through K. From (B.1) in the appendix, K is positively related to π. Relying on the implicit function theorem and assuming the second-order condition for an optimum to hold,[19] we get, at $t = t^*$:

$$\frac{\mathrm{d}t}{\mathrm{d}K} = \frac{\mathrm{d}^2 g/(\mathrm{d}t\,\mathrm{d}K)}{\mathrm{d}^2 g/\mathrm{d}t^2} \qquad \Rightarrow \qquad \mathrm{sign}\left(\frac{\mathrm{d}t}{\mathrm{d}K}\right) = \mathrm{sign}\left(\frac{\mathrm{d}^2 g}{\mathrm{d}t\,\mathrm{d}K}\right).$$

Differentiating (19) with respect to K yields:

$$\frac{\mathrm{d}^2 g}{\mathrm{d}t\,\mathrm{d}K} = \Phi\, \left(f(\hat{\alpha}) + \frac{t}{1-t} \left(\frac{\partial f(\hat{\alpha})}{\partial \alpha}\, \hat{\alpha} + 2 f(\hat{\alpha}) \right) \right), \tag{21}$$

where $\Phi \equiv (1-\eta)\,\hat{\alpha}\, g/\left(K\left((1-t)\,K + t\,(1-\eta)\, f(\hat{\alpha})\,\hat{\alpha}\right)\right) > 0$. The expression on the right-hand side of (21) is positive if the sum of terms in parentheses is positive, i.e. if

$$\frac{\partial f(\hat{\alpha})}{\partial \alpha} \frac{\hat{\alpha}}{f(\hat{\alpha})} \equiv \xi_{f,\alpha} \geq -\frac{1+t}{t}\,. \tag{22}$$

A sufficient condition for this to be satisfied for all values of $t \in (0,1)$ is $\xi_{f,\alpha} > -2$. If this condition is met, then a *reduction* in π will lower the equilibrium tax rate t^*. This will be the case when $f(\alpha)$ is either non-decreasing, or only moderately decreasing in the productivity type α. Condition (22) will always be fulfilled, for example, under a uniform distribution of productivity types, where $f'(\hat{\alpha}) = 0$.[20]

[19] The second-order condition is met if revenue is concave in t, i.e. if $2\,\mathrm{d}Q/\mathrm{d}t + t\,\mathrm{d}^2Q/\mathrm{d}t^2 < 0$; or $-t(\mathrm{d}^2Q/dt^2)/(\mathrm{d}Q/\mathrm{d}t) < 2$. Using the fact that $\mathrm{d}g/\mathrm{d}t = 0$ at an optimum, this can be shown to require $\xi_{f,\alpha} < 1 + 2/t$. This condition is always met by a distribution of ability types such that $f'(\alpha) \leq 0$ (sufficiency).

[20] Condition (22) refers to the distribution of productivity types and does *not* imply a condition on the distribution of realized wages. The latter results in our setting from a mixture of distributions of type-specific distributions of wage realizations. In a generalized version of our model where the type-specific distributions of wage realizations are continuous (e.g. Pareto), the shape and tail behaviour of the overall distribution of realized wages (e.g. a Pareto tail) would thus be inherited from the shape of the type-specific distributions, not from the shape of the distribution of productivity types.

A mean-preserving output spread thus causes the efficiency costs of taxation to rise, as incentive contracts respond more sensitive to taxation. Intuitively, a fall in π increases income volatility in a performance-based contract and increases the cutoff productivity level, $\hat{\alpha}$, above which workers select into this contract. At the higher level of $\hat{\alpha}$, the level of the tax base is reduced and the remaining tax base responds more strongly to taxation. Both of these effects imply that the elasticity $\varepsilon = (t/Q)(\mathrm{d}Q/\mathrm{d}t)$ with which the tax base responds to a change in the tax rate tends to increase in absolute value. This increase in the elasticity of the tax base is similar to models where the high-income workers are internationally mobile (Lehmann et al., 2014). The mechanism, however, is entirely different – there is no inter-jurisdictional factor mobility in our model.

Under revenue maximization, the optimal tax rate will therefore fall unless, at the higher level of $\hat{\alpha}$, there are substantially fewer workers whose contract choice is distorted by the tax. Since the government is only concerned with the ex-post welfare of the poorest individual, any changes in the insurance role that the indirectly progressive income tax takes in our model, are not relevant under this objective function.

We summarize our results in this section in the following proposition:

Proposition 4. *A mean-preserving spread in output realizations (a fall in π) reduces the revenue-maximizing tax rate: (i) if the distribution of productivity types is uniform; (ii) for any distribution of types that fulfills condition* (22).

PROOF*:* See Appendix B.4.

Note that if we associate globalization with a positive productivity shift, as discussed at the end of Section 3, such a productivity shift has no effect on the optimal tax under a revenue-maximizing objective: in (19), both Q and $\mathrm{d}Q/\mathrm{d}t^*$ are multiplied by the the productivity factor λ in this case, leaving the optimal tax rate unchanged.

We also emphasize that the uniform distribution of types singled out in Proposition 4 does *not* imply a uniform distribution of incomes. Rather, it generates a right tail as only some of the highest-productivity workers will be successful, and these receive high bonus payments in equilibrium. Moreover, this right tail becomes fatter as market integration increases, as fewer high-productivity workers will achieve comparatively greater success.

4.2 Utilitarian objective

A utilitarian social welfare function can be written as

$$W = \int_0^{\hat{\alpha}} EU^F(\alpha)\,\mathrm{d}F(\alpha) + \int_{\hat{\alpha}}^1 EU^E(\alpha)\,\mathrm{d}F(\alpha)\,. \tag{23}$$

The first term on the right-hand side sums over the expected utilities of workers with a fixed wage contract, whereas the second term represents workers with an incentive contract. This is a non-trivial change from the Rawlsian case (revenue maximization), because now all individuals, with different types of contracts, enter the government's objective.

Since $EU^E(\hat{\alpha}) = EU^F(\hat{\alpha})$ (by the definition of $\hat{\alpha}$), the first-order condition for a welfare-maximizing tax rate can be written as

$$\frac{\mathrm{d}W}{\mathrm{d}t} = \int_0^{\hat{\alpha}} \frac{\mathrm{d}EU^F(\alpha)}{\mathrm{d}t}\,\mathrm{d}F(\alpha) + \int_{\hat{\alpha}}^1 \frac{\mathrm{d}EU^E(\alpha)}{\mathrm{d}t}\,\mathrm{d}F(\alpha) = 0\,. \tag{24}$$

Expanding the terms, $\mathrm{d}EU^Z(\alpha)/\mathrm{d}t$, $Z \in \{E, F\}$, we can express them as (see Appendix B.5)

$$\frac{\mathrm{d}EU(\alpha)}{\mathrm{d}t} = \frac{\mathrm{d}g/\mathrm{d}t - b(\alpha)\,\alpha}{g + (1-t)\,b(\alpha)\,\alpha}\,, \tag{25}$$

with $b(\alpha) = \eta$ for $\alpha < \hat{\alpha}$ and $b(\alpha) = 1$ for $\alpha \geq \hat{\alpha}$. Just by inspecting (25) and (24), we can conclude that, since $b(\alpha) > 0$ for all α, we must have $\mathrm{d}g/\mathrm{d}t > 0$ at an optimum. Thus, under a utilitarian objective the optimal tax falls short of the level that maximizes tax revenues.[21] As noted earlier, a welfare-maximizing tax could never exceed the level at which revenues are maximized, and will only reach that level when the redistribution motive is at its strongest – under a Rawlsian objective.

To understand how a mean-preserving spread in output realizations (a reduction in π) is reflected in the optimal choice of tax, note first that all effects only flow through the induced increase in the cutoff productivity $\hat{\alpha}$. In the absence of a change in $\hat{\alpha}$, there would be no change in $\mathrm{d}g/\mathrm{d}t$ (nor in g), nor would there be any change in the $b(\alpha)$ terms that enter expression (25).[22]

[21] The Rawlsian case corresponds to $b(\underline{\alpha}) = 1$ for $\underline{\alpha} = 0$ and $b(\alpha) = 0$ for all $\alpha > 0$; (24) then becomes $(\mathrm{d}g/\mathrm{d}t)/g = 0$.

[22] If there is also a productivity shift, all terms $\mathrm{d}EU(\alpha)/\mathrm{d}t$ in (25) remain unchanged, as both the numerator and the denominator on the right-hand side are multiplied by the factor λ. Therefore, the optimal tax rate implied by the first-order condition (24) must again remain unchanged.

We can intuitively discuss how a reduction in π affects each of the terms in (25) and hence the optimality condition for the tax rate chosen by a utilitarian planner. The first effect, how a reduction in π affects $\mathrm{d}g/\mathrm{d}t$, is the same that we characterized for the revenue-maximizing case in (21). A fall in π makes the tax base respond more elastically to a tax increase, thus lowering $\mathrm{d}g/\mathrm{d}t$.

The new element under a utilitarian objective comes from a redistributive effect that is incorporated in the weights $b(\alpha)$ in the second term. Note first that the smaller is $b(\alpha)$ the closer is (25) to $(\mathrm{d}g/\mathrm{d}t)/g$, the expression that is driven to zero under a Rawlsian objective. Note further that the switch in contract form at $\hat{\alpha}$ generates a discontinuity in the weights $b(\alpha)$ in (24), which, at $\hat{\alpha}$, rise from η to 1. An increase in $\hat{\alpha}$ then mechanically makes (24) closer to the optimality condition under a Rawlsian objective, i.e. it strengthens the weight of the redistributive motive in the choice of t.

The intuition behind this redistributive effect of the tax is that there is a 'leisure' component of welfare, having a consumption value equal to $(1-\eta)\,\alpha$ and producing a utility differential equal to the effort cost c, that at the discontinuity point goes from being untaxed (below $\hat{\alpha}$) to being taxed (above $\hat{\alpha}$). Since higher-income individuals bear a comparatively greater tax burden, this adds to the redistributive effectiveness of the tax for a given $\hat{\alpha}$. If then $\hat{\alpha}$ rises following a decrease in π, the redistributive effectiveness of the tax further improves, with comparatively lower productivity types previously above the threshold now moving below the threshold and benefiting comparatively more from a higher tax. Other things bein equal, this makes a higher tax more attractive to a utilitarian planner.[23]

Thus, a mean-preserving spread in output realizations makes the tax base more tax-elastic but increases the redistributive value of the tax. Still, it is not difficult to identify scenarios where the first effect remains the dominating effect and so a decrease in π unambiguously lowers t^*. This is true, for example, if the distribution of productivity types is uniform. For $F(\alpha)=\alpha$, expression (18) simplifies to

$$g = t\,Q = \frac{t}{2}\left(1-(1-\eta)\,\hat{\alpha}^2\right), \tag{26}$$

[23]Note also the difference to a standard optimal tax problem with endogenous labor supply and no incentive contracts. In such a setting the income tax causes the tax base to shrink comparatively more for high ability types than for low ability types, which both generates distortions in labor supply decisions *and* makes the tax less effective as a redistributive device. In our setting, in contrast, the tax causes the tax base to shrink for low ability types, but not for high ability types, thus making the tax *more* effective as a redistributive device.

and the response of the tax base Q with respect to the tax rate t becomes

$$\frac{\mathrm{d}Q}{\mathrm{d}t} = -(1-\eta)\,\hat{\alpha}\,\frac{\mathrm{d}\hat{\alpha}}{\mathrm{d}t}\,. \tag{27}$$

In Appendix B.5, we derive the first-order condition for the optimal tax rate:

$$\frac{\mathrm{d}W}{\mathrm{d}t} = \frac{\Theta}{\eta}\,\frac{(1-t)\,K^2}{\mu\left((1-t)\,K+\mu\right)} - 1 = 0, \qquad \mu \equiv \sqrt{(1-t)^2\,K^2 + (1-\eta)t^2}, \tag{28}$$

where

$$\Theta \equiv \eta\,\ln\frac{g+(1-t)}{g+(1-t)\,\hat{\alpha}} + \ln\frac{g+(1-t)\,\eta\hat{\alpha}}{g} > 0. \tag{29}$$

Since $\hat{\alpha} \in (0,1)$ the fractions for which the logarithms are taken in (29) are both greater than unity. We can then proceed in the same way as in the previous section: assuming that the second-order condition for an optimum are satisfied and relying on the positive relationship between K and π in (B.1), the effect of a change in π on the optimal tax rate has the same sign as $\mathrm{d}^2W/(\mathrm{d}t\,\mathrm{d}K)$.

Appendix B.6 derives the term $\mathrm{d}^2W/(\mathrm{d}t\,\mathrm{d}K)$ in (B.17), clearly distinguishing the higher efficiency costs and the higher redistributive value of the tax. The appendix then proceeds to prove that the first effect always dominates the second. This is summarized in:

Proposition 5. *Under a utilitarian objective, and for a uniform distribution of productivity types, a mean-preserving spread in output realizations (a fall in π) reduces the welfare-maximizing tax rate.*

PROOF*:* See Appendix B.6

Proposition 5 corresponds to the result described by Proposition 4 for the Rawlsian case. With a uniform distribution of productivity types, our model therefore unambiguously predicts a *negative* effect of market integration (characterized here as a mean-preserving spread in output realizations) on the optimal redistributive income tax rate. Once more, it is worth stressing that a uniform distribution of productivity types does not produce a uniform distribution of income realizations.

5 Globalization, inequality and redistribution

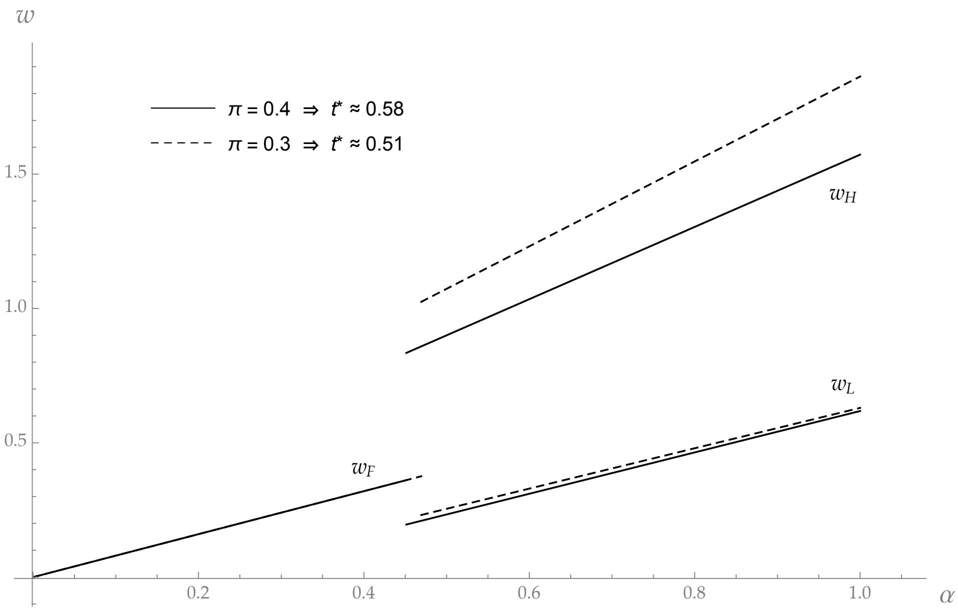
The predictions we have derived in the previous sections have direct implications for the relationship between changes in tax policy choices and changes in the distribution of income, when this relationship is mediated through a change in π. In Section 3, we have shown that economic globalization, characterized as a decrease in π, raises wage inequality and the concentration of earnings at the top of the income distribution (Proposition 3). And our results in Section 4 state that this can be accompanied by a fall in the optimal rate of redistributive taxation (Propositions 4 and 5).

Figure 6 continues the example that we have described in the previous sections, using the same parameterization but now making the level of taxation equal to its optimum under a utilitarian objective. This optimum rate is approximately 58% under the baseline value of π of 0.4. If π is reduced to 0.3, the optimal tax rate falls to approximately 51%. The combination of the exogenous change in π and the endogenous response in the choice of tax is associated with a reduction in the proportion of individuals in performance-based contracts and an increase in the concentration of after-tax earnings (Figure 7) in the upper tail of the distribution of earnings realizations. The Gini coefficient for after-tax income is 0.19 in the baseline scenario ($\pi = 0.4$, $t = 58\%$), rising to 0.21 for $\pi = 0.3$ if t remains at 58%, and rising further to 0.24 if t adjusts to the optimal rate of 51%. Pre-tax earnings also become more concentrated – approximate values of the pre-tax income Gini coefficient for $\pi = 0.4$ and $\pi = 0.3$ are respectively 0.42 and 0.45.

Thus, if we were to compare the above two scenarios, and for each we only observed the level of taxation and the distribution of earnings, but not the different values of π, we might be tempted to conclude that the higher inequality we observe in the second scenario, in terms of both pre-tax and after-tax earnings, follows from an independent change in tax policy (arising, for example, from a change in the government's distributional objective). Instead, the decrease in taxation and the increase in earnings inequality in this example both are consequences of structural changes associated with economic globalization.

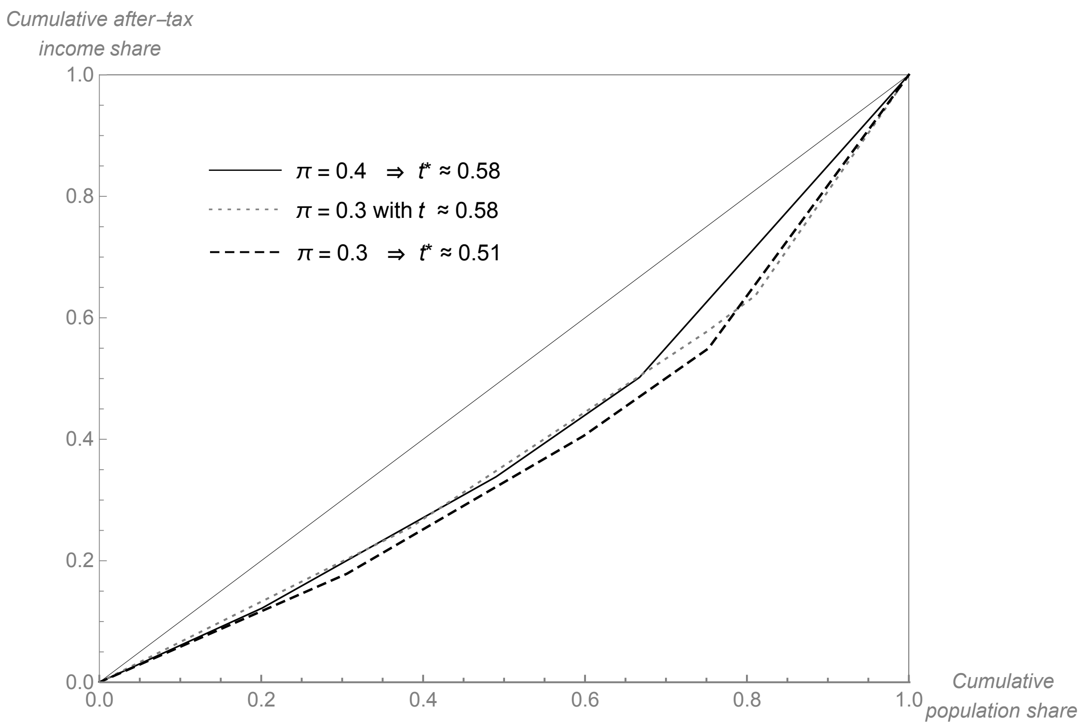
6 Conclusion

Economic globalization makes factors of production and the goods that are produced with them more inter-jurisdictionally mobile. In this paper we have focused on the latter aspect of globalization, i.e. market integration. Increased competition in broader product markets translates into an increase in the spread of returns for individual projects. This both produces an increase in earnings inequality and raises the efficiency



With $\eta = 4/5$, $c = \bar{c}/5$, $F(\alpha) = \alpha$

Figure 6: Output volatility, optimal taxes, and equilibrium contracts – utilitarian objective



With $\eta = 4/5$, $c = \bar{c}/5$, $F(\alpha) = \alpha$

Figure 7: Output volatility, optimal taxes, and after-tax income concentration – utilitarian objective

cost of redistributive taxation and lead to lower optimal levels of taxation.

The link operates through performance-based contracts, which have played an increasingly important role in developed economies in recent decades, in particular for high-income earners. Higher idiosyncratic income risk arising from globalized product markets raises the risk that performance-based contracts must impose on high productivity workers in order to overcome the moral hazard problem. This higher volatility of individual earnings makes incentive contracts both less attractive and also more sensitive to the disincentive effects caused by an indirectly progressive income tax, raising the cost of redistribution by comparison to its effectiveness.

This prediction aligns with the recent empirical evidence presented in Egger et al. (2019), who show that globalization has reduced the progressiveness of income taxes in the OECD countries since the mid-1990s. So far, the main explanation offered in the literature for the causal link between globalization and a reduced progressivity of the income tax system has been the increasing international mobility of top income earners (Simula and Trannoy, 2010; Bierbrauer et. al., 2013; Lehmann et al., 2014). Our contribution to this debate is in showing that the same pattern can also result from the increased competition in product markets that accompanies economic globalization, even when factors of production are not inter-jurisdictionally mobile.

References

Abraham, A., Alvarez-Parra, F., Forstner, S. (2017). The effects of moral hazard on wage inequality in a frictional labor market. Mimeo, European University Institute, Florence.

Abraham, A., Koehne, S., Pavoni, N. (2016). Optimal income taxation when asset taxation is limited. *Journal of Public Economics* 136, 14-29.

Agrawal, D., Foremny, D. (2019). Relocation of the rich: Migration in response to top tax rate changes from Spanish reforms. *Review of Economics and Statistics* 101, 214-232.

Autor, D., Dorn, D., Hanson, G. (2013). The China syndrome: Local labor market effects of import competition in the United States. *American Economic Review* 103, 2121-2168.

Autor, D., Dorn, D., Hanson, G. (2015). Untangling trade and technology: Evidence from local labor markets. *Economic Journal* 125, 621-646.

Autor, D., Dorn, D., Katz, L.F., Patterson, C., Van Reenen, J. (2017). The fall of the labor share and the rise of superstar firms. NBER Working Paper No. 23396. Cambridge, Mass.

Alvaredo, F., Chancel, L., Piketty, T., Saez, E., Zucman, G. (2017). Global inequality dynamics: New Findings from WID.world. *American Economic Review, Papers and Proceedings* 107, 404-409.

Atkinson, A., Piketty, T., Saez, E. (2011). Top incomes in the long run of history. *Journal of Economic Literature* 49, 3-71.

Bargain, O., Dolls, M., Immervoll, H., Neumann, D., Peichl, A., Pestel, N., Siegloch, S. (2015). Tax policy and income inequality in the United States, 1979-2007. *Economic Inquiry* 53, 1061-1085.

Battisti, M., Felbermayr, G., Peri, G., Poutvaara, P. (2018). Immigration, search and redistribution: A quantitative assessment of native welfare. *Journal of the European Economic Association* 16, 1137-1188.

Bell, B., Van Reenen, J. (2014). Bankers and their bonuses. *The Economic Journal* 124, F1-F21.

Bergh, A., Nilsson, T. (2010). Do liberalization and globalization increase income inequality? *European Journal of Political Economy* 26, 488-505.

Bierbrauer, F., Brett, C., Weymark, J. (2013). Strategic nonlinear income tax competition with perfect labor mobility. *Games and Economic Behavior* 82, 292-311.

Boadway, R., Sato, M. (2015). Optimal income taxation with risky earnings: A synthesis. *Journal of Public Economic Theory* 17, 773-801.

Bryson, A., Freeman, R., Lucifora, C., Pelizzari, M., Perotin, V. (2012). Paying for performance: incentive pay schemes and employees' financial participation. Centre for Economic Performance Discussion Paper No. 1112. London School of Economics.

Diamond, P.A., Saez, E. (2011). The case for a progressive tax: From basic research to policy recommendations. *Journal of Economic Perspectives* 25, 165-190.

Dorn, F., Fuest, C., Potrafke, N. (2018). Globalization and income inequality revisited. CESifo Working Paper No. 6859. Munich.

Dreher, A., Gaston, N. (2008). Has globalization increased inequality? *Review of International Economics* 16, 516-536.

Eaton, J., Kortum, S. (2002). Technology, geography, and trade. *Econometrica* 70, 1741-1779.

Egger, P., Nigal, S. , Strecker, N. (2019). The taxing deed of globalization. *American Economic Review* 109, 353-390.

Golosov, M., Tsyvinski, A. (2007). Optimal taxation with endogenous insurance markets. *Quarterly Journal of Economics* 122, 487-534.

Guadalupe, M., and Cuñat, V. (2009). Globalization and the provision of incentives inside the firm: The Effect of foreign competition. *Journal of Labor Economics* 27, 179-212.

Hoynes, H., Luttmer, E. (2011). The insurance value of state tax-and-transfer programmes. *Journal of Public Economics* 95, 1466-1484.

Immervoll, H., Richardson, L. (2011). Redistribution policy and inequality reduction in OECD countries. OECD Social, Employment and Migration Working Paper No. 122. Paris.

Kleven, H., Landais, C., Saez, E., Schultz, E. (2014). Migration and wage effects of taxing top earners. Evidence from the foreigners' tax scheme in Denmark. *Quarterly Journal of Economics* 129, 333-378.

Lehmann, E., Simula, L., Trannoy, A. (2014). Tax me if you can! Optimal non-linear income tax between competing governments. *Quarterly Journal of Economics* 129, 1995-2030.

Lemieux, T., MacLeod, W.B., Parent, D. (2009). Performance pay and wage inequality. *The Quarterly Journal of Economics* 124, 1-49.

Office for National Statistics (2017). *Average weekly earnings, bonus payments in Great Britain: 2017*.

Piketty, T., Saez, E., Zucman, G. (2018). Distributional national accounts: Methods and estimates for the United States. *Quarterly Journal of Economics* 133, 553-609.

Rodrik, D. (1998). Why do more open economies have bigger governments? *Journal of Political Economy* 106, 997-1032.

Saez, E. (2001). Using elasticities to derive optimal income tax rates. *Review of Economic Studies* 68, 205-229.

Simula, L., Trannoy, A. (2010). Optimal income tax under the threat of migration by top-income earners. *Journal of Public Economics* 94, 163-173.

Sinn, H.W. (1995). A theory of the welfare state. *Scandinavian Journal of Economics* 97, 495-526.

Slemrod, J., Bakija, J. (2001). Growing inequality and decreased tax progressivity. In: K.A. Hassett and R.G. Hubbard (eds.), Inequality and Tax Policy, Chapter 6, 192-226. American Enterprise Institute, Washington, D.C..

Tóbias, A. (2016). Income redistribution in open economies. *Journal of Public Economics* 134, 19-34.

Varian, H. (1980). Redistributive taxation as social insurance. *Journal of Public Economics* 14, 49-68.

Appendix

A Market integration and revenue volatility

In this appendix, we describe a model of product market integration whose predictions can be mapped to the predictions of the model we use to derive results in the main text.

There are N symmetric economies. In each of these economies, agents are endowed with a given amount of a non-produced good that can be either consumed or used as an input into the production of consumption goods. There is a continuum of consumption good types, $i \in [0,1] \equiv I$, that can (potentially) be produced in each economy. In each economy, there is a unit mass of identical consumers with preferences

$$u(y,x) = y + \int_0^1 \min\left\{\overline{x}(i), x(i)\right\} \mathrm{d}i,$$

where y is consumption of the endowment good and $x(i)$ is consumption of product i. Consumers view products of the same type, i, that are produced in different economies as being homogeneous in consumption. Without loss of generality, we assume that products are ordered so that $i'' > i' \Rightarrow x(i'') > x(i')$. We also assume that $\overline{x}(i)$ is a continuous function of i. Given these preferences, a consumer is willing to buy up to $\overline{x}(i)$ units of product i at any price $p(i) \leq 1$.

In each economy, and for each product i, there is a finite number, $M \geq 2$, of workers/managers of type i that can be employed to oversee the production of good i. Each worker-/manager of type i is hired by a single firm – and so, for each good i, there are at most M firms (potentially) producing product i in each economy. There are otherwise no barriers to entry. A worker of type i can be either employed to produce good i or, alternatively, employed to produce one unit of the endowment good. Doing the latter always requires positive effort on the part of the worker. Each worker/manager is also a consumer. With a unit mass of workers/consumers in the economy, this implies that the mass of product types equals $1/M$.

Each of the firms that has hired a worker of type i – i.e. each firm that operates in *sector i* – can produce any quantity of product i using the endowment good as an input at a constant marginal input cost equal to $1/\phi$, where ϕ is a firm-specific productivity draw. The probability of a firm experiencing a certain productivity draw depends on whether or not the worker/manager the firm has hired chooses to exert positive effort – at a positive cost to the worker/manager. If the manager exerts positive effort, the draw is $\overline{\phi} > 1$ with probability γ and $\underline{\phi} \in [1, \overline{\phi})$ otherwise. If the worker/manager exerts no effort, the draw is $\underline{\phi}$ with certainty. Once productivity draws are realized, firms compete with each other in prices in all consumer markets to which they have access.

Consider next an autarkic economy. With M firms in sector i, and assuming that each worker/manager in each firm has exerted positive effort, with probability

$$M\gamma(1-\gamma)^{M-1} \equiv M\pi,$$

a single firm will experience a draw $\overline{\phi}$ and all other firms will experience a productivity draw $\underline{\phi}$. In this case, the firm that has experienced the favourable draw will be able, by pricing just marginally below $\overline{p} = 1/\underline{\phi}$, to keep other firms from producing, thus securing profits equal to

$$\left(1/\underline{\phi} - 1/\overline{\phi}\right) \overline{x}(i) = \mu\, \overline{x}(i) \equiv R(i).$$

In all other realizations, there will either be at least two firms experiencing $\overline{\phi}$, or all firms will experience $\underline{\phi}$; in all these cases, price competition will drive prices down to a point where $p(i) = 1/\phi$. The market will then be shared by more than one firm and all firms, whether they are active or inactive, will experience zero profits. Thus, under autarky, any given firm in sector i experiences a profit $R(i)$ with probability π and a profit equal to zero with probability $1 - \pi$. Expected interim profits with effort for any firm operating in sector i are therefore $\pi R(i) = \gamma(1-\gamma)^M \mu\, \overline{x}(i) \equiv \alpha(i)$.

We assume $\overline{x}(i) > 1/(\pi\,\mu), \forall i$. This implies $\alpha(i) > 1, \forall i$; i.e. the expected return of a worker of type i in the production of good i is greater than the alternative of employing the worker to produce one unit of the endowment good. Only interim profits, $R(i)$, are verifiable, not the productivity draw; and so payments to worker/managers can only be conditioned on $R(i)$: the worker/manager's pay is $w_L(i)$ when the firm experiences a zero interim profit realization and $w_H(i)$ when it experiences a positive interim profit realization. The compensation level $w_H(i)$ is formally structured as follows: $w_H(i) = w_L(i) + \sigma(i)R(i)$; i.e. $w_H(i)$ equals the base pay, $w_L(i)$, plus a bonus component consisting of a share, $\sigma(i)$, of the interim profits, $R(i)$. Structuring pay in this way ensures that the interim incentives of the managers when making pricing decisions (once the effort choice has been made) are aligned with those of the firm.

The ex-ante expected profits of the firm, under positive effort, will then equal $E[\Pi(i)] = \pi\left(R(i) - w_H(i)\right) - \left(1 - \pi\right) w_L(i)$. Absent barriers to entry, these will have to be zero in equilibrium. The equilibrium contract for each worker type will therefore be a contract $\left(w_L(i), w_H(i) = w_L(i) + \sigma(i)R(i)\right)$ that satisfies the incentive compatibility constraint with equality (as we detail discuss in the main text) and results in zero ex-ante expected profits for the firm.

Now suppose that the N economies become integrated. In this case, we have

$$\pi(N) = \gamma\left(1-\gamma\right)^{NM-1}, \qquad R(i,N) = N\,\mu\,\overline{x}(i),$$

with $d\pi(N)/dN < 0$ and $dR(i,N)/dN > 0$. Being the only firm in sector i to experience $\overline{\phi}$ is now less likely; but if that happens, the firm that experiences $\overline{\phi}$ earns positive profits in each of the N consumer markets.[24] Expected interim profits, gross of the expected payments

[24]This is similar to the mechanism at work in trade models where Ricardian comparative advantage results from idiosyncratic technology shocks, as in Eaton and Kortum (2002). When trade opens up, the country with the lowest production cost for any given product variety will secure the demand from all countries for that variety – a 'winner-takes-all' effect.

to worker/managers, equal

$$\pi(N)\,R(i,N) = \gamma(1-\gamma)^{NM-1} N\,\mu\,\overline{x}(i).$$

These are not invariant to a change in N, i.e. an increase in N combines a mean-preserving spread with a positive expected productivity shift that affects all productivity types uniformly. As discussed in Section 3, however, such a uniform productivity shift is of no consequence for our results.

B Derivations of expressions and proofs of results

B.1 Proof of Lemma 1

From (8a), the productivity level below which $w_L \leq 0$ in a performance-based contract is given by $\pi\,(\Gamma-1)\,g/(1-t) \equiv \alpha^0$. The expected payoff to an individual in a performance-based contract must equal EU^N in equilibrium; this is the level of expected utility associated with receiving w_H with probability $\eta\pi$ and w_L with probability $1-\eta\pi$, i.e. a lottery with expected value

$$\eta\pi\,w_H + (1-\eta\pi)\,w_L = \frac{\alpha[1+\eta\pi(\Gamma-1)] - \pi g(\Gamma-1)(1-\eta)/(1-t)}{1+\pi(\Gamma-1)} \equiv EW^N.$$

The productivity level for which $EW^N = EW^F = \eta\alpha$ is $(\Gamma-1)\,g/(1-t) \equiv \tilde{\alpha} > \alpha^0$. By risk aversion, an individual will always prefer a certain prospect to an uncertain one of equal expected value; and so, for $\alpha = \tilde{\alpha}$, it must be the case that $EU^F > EU^N$. By (12), this implies $\tilde{\alpha} < \hat{\alpha}$ and $\alpha^0 < \hat{\alpha}$ (since $\alpha^0 < \tilde{\alpha}$). Thus, productivity types that select into performance-based contracts will receive a positive wage even when unsuccessful. Since $w_F > 0$ and $w_H > w_L$, wages will be positive for all productivity types in all realizations, implying that a limited liability constraint $w \geq 0$, if present, can never be binding. □

B.2 Proof of Proposition 1

If $\eta \to 1$, then there can be no positive solution $\hat{\alpha} < 1$: the insurance motive dominates, and the fixed-wage contract becomes be preferred to the performance-based contract by all productivity types. For $\eta < 1$, the expression K is negative for c approaching $\overline{c}$ and approaches infinity for c approaching zero; and so there can always find a threshold, $\tilde{c} < \overline{c}$, such that for $c < \tilde{c}$ the condition $K > g/(1-t)$ is met and therefore $\hat{\alpha} \in (0, \overline{\alpha})$.

The rest of the results follow from (12). □

B.3 Proof of Proposition 3

Differentiating the expression for K in (13) with respect to π and letting $\Psi \equiv -(1-\eta)\,\Gamma^{(\eta\pi-1)}/\left(1+\pi(\Gamma-1)-\Gamma^{\eta\pi}\right)^2 < 0$, we obtain

$$\frac{dK}{d\pi} = \Psi\left(\Gamma\left(1-\frac{c}{(1-\eta)\,\pi}\right)-1\right) > 0\,, \tag{B.1}$$

where the sign follows from (6) and from the expression in parentheses being negative.

The expression for $\hat{\alpha}$ is $g/\big((1-t)K\big)$, which is decreasing in K and increasing in g. In turn, for a given t, g is decreasing in $\hat{\alpha}$. Consider then an increase in K accompanied by a reduction in $\hat{\alpha}$. If the increase in g caused by the fall in $\hat{\alpha}$ were large enough to offset the negative effect on $\hat{\alpha}$ of a higher K, then $\hat{\alpha}$ would rise, and so g would have to fall rather than rise: a contradiction. So, for a given t, $\hat{\alpha}$ must be decreasing in K (and thus in π), as stated in part (i) of the proposition.

For part (ii), w_H is unambiguously decreasing in π from (8b). For given t, and with $\hat{\alpha}$ rising from (13) and (B.1), a counteracting effect comes from the simultaneous reduction in the tax base, and hence in g. This effect, however, can never dominate the direct effect of a change in π. If it did, then it would not only lower w_H, but also raise w_L from (8a). This would align the pre-tax wage levels for an unchanged expected wage, and would thus make the performance-based contract unambiguously more attractive, relative to the fixed-wage contract. But then $\hat{\alpha}$ would need to fall, a contradiction to part (i) proven above. □

B.4 Proof of Proposition 4

Differentiating (19) with respect to K gives

$$\frac{\partial^2 g}{\partial t\,\partial K} = \frac{dQ}{dK} + t\,\frac{d^2Q}{dt\,dK}\,; \tag{B.2}$$

where

$$t\,\frac{d^2Q}{dt\,dK} = -t\,(1-\eta)\left(\frac{df(\hat{\alpha})}{d\alpha}\,\hat{\alpha}\,\frac{d\hat{\alpha}}{dt}\,\frac{d\hat{\alpha}}{dK} + f(\hat{\alpha})\,\frac{d\hat{\alpha}}{dt}\,\frac{d\hat{\alpha}}{dK} + f(\hat{\alpha})\,\hat{\alpha}\,\frac{d^2\hat{\alpha}}{dt\,dK}\right). \tag{B.3}$$

From the definitions of $\hat{\alpha}$ and Q in (13) and (18), we get the following derivatives

$$\frac{d\hat{\alpha}}{dt} = \frac{\partial\hat{\alpha}}{\partial t} = \frac{g}{(1-t)^2K} = \frac{\hat{\alpha}}{1-t}\,, \tag{B.4}$$

$$\frac{dQ}{dK} = -(1-\eta)\,f(\hat{\alpha})\,\hat{\alpha}\,\frac{d\hat{\alpha}}{dK}\,. \tag{B.5}$$

From (13) and (B.5), and letting $Z \equiv \big((1-t)\,K + t\,(1-\eta)\,f(\hat{\alpha})\,\hat{\alpha}\big)\,K$:

$$\frac{d\hat{\alpha}}{dK} = -\frac{g}{(1-t)\,K^2} + \frac{t}{(1-t)\,K}\,\frac{dQ}{dK} = -\frac{g}{Z}\,, \tag{B.6}$$

$$\frac{d^2\hat{\alpha}}{dt\,dK} = \frac{d\hat{\alpha}}{dK}\,\frac{1}{1-t} = -\frac{g}{(1-t)\,Z}\,. \tag{B.7}$$

Substituting (B.4)–(B.7) into (B.3) gives

$$t\,\frac{d^2Q}{dt\,dK} = \frac{t\,(1-\eta)\,\hat{\alpha}\,g}{(1-t)\,Z}\left(\frac{df(\hat{\alpha})}{d\alpha}\,\hat{\alpha} + 2f(\hat{\alpha})\right). \tag{B.8}$$

Using (B.6) to rewrite dQ/dK and substituting this along with (B.8) into (B.2) yields (21) in the main text. □

B.5 Derivation of (25) and (28)

From the utilitarian welfare function (23), the first-order condition for a welfare-maximizing t is

$$\frac{\mathrm{d}W}{\mathrm{d}t} = \int_{\underline{\alpha}}^{\hat{\alpha}} \frac{\mathrm{d}EU^F(\alpha)}{\mathrm{d}t}\,\mathrm{d}\alpha + \int_{\hat{\alpha}}^{\overline{\alpha}} \frac{\mathrm{d}EU^E(\alpha)}{\mathrm{d}t}\,\mathrm{d}\alpha - \left(EU^E(\hat{\alpha}) - EU^F(\hat{\alpha})\right)\frac{\mathrm{d}\hat{\alpha}}{\mathrm{d}t} = 0\,. \tag{B.9}$$

The last term on the LHS of (B.9) is zero because $EU^E(\hat{\alpha}) = EU^F(\hat{\alpha})$ by the definition of $\hat{\alpha}$. Expanding the remaining terms $\mathrm{d}EU^Z(\alpha)/\mathrm{d}t$, $Z \in \{E, F\}$, we get

$$\frac{\mathrm{d}EU^Z(\alpha)}{\mathrm{d}t} = \frac{\partial EU^Z(\alpha)}{\partial g}\,\frac{\mathrm{d}g}{\mathrm{d}t} + \frac{\partial EU^Z(\alpha)}{\partial t}\,, \qquad Z \in \{E, F\}\,; \tag{B.10}$$

where

$$\frac{\partial EU^E(\alpha)}{\partial g} = \frac{1}{g + (1-t)\,\alpha} > 0\,; \qquad \frac{\partial EU^F(\alpha)}{\partial g} = \frac{1}{g + (1-t)\,\eta\,\alpha} > 0\,; \tag{B.11}$$

$$\frac{\partial EU^E(\alpha)}{\partial t} = \frac{-\alpha}{g + (1-t)\,\alpha} < 0\,; \qquad \frac{\partial EU^F(\alpha)}{\partial t} = \frac{-\eta\,\alpha}{g + (1-t)\,\eta\,\alpha} < 0\,. \tag{B.12}$$

The terms in (B.12) are summarized in (25) in the main text. Since $\partial EU^Z(\alpha)/\partial g$ is positive and $\partial EU^Z(\alpha)/\partial t$ is negative, we must have $\mathrm{d}g/\mathrm{d}t > 0$ at an optimum. Using (B.11)-(B.12) gives

$$\frac{\mathrm{d}W}{\mathrm{d}t} = \frac{\Theta}{\eta}\left(\frac{g}{(1-t)} + \frac{\mathrm{d}g}{\mathrm{d}t}\right) - 1 = 0\,, \tag{B.13}$$

where Θ is given in (29) in the main text. Finally, it is possible to obtain a closed-form solution for g. Solving the balanced-budget condition (26) in conjunction with the expression for $\hat{\alpha}$ in (13) gives

$$g = \frac{t\,(1-t)\,K}{(1-t)\,K + \mu}\,, \qquad \mu \equiv \sqrt{(1-t)^2K^2 + (1-\eta)\,t^2} > 0\,. \tag{B.14}$$

Using (B.14) to compute the terms in the squared bracket in (B.13) gives (28) in the main text. □

B.6 Proof of Proposition 5

Differentiating the first-order condition (28) with respect to K gives, in a first step

$$\frac{\mathrm{d}^2W}{\mathrm{d}t\,\mathrm{d}K} = \frac{(1-t)\,K}{((1-t)K + \mu)\,\mu^2}\left(A\,\frac{\Theta}{\eta} + \frac{B}{(K+\mu)(K+1)}\right), \tag{B.15}$$

where

$$A \equiv \frac{(1-t)\,K + 2\mu}{\mu} \big(\mu - (1-t)\,K\big) > 0\,,$$

$$B \equiv K\,(K+1)\,\big((1-t)^2 K + \mu\big) - \mu\,(K+\mu)\,\big(K + (K+1)/(K+\eta)\big)\,.$$

Using the first-order condition (28) to eliminate Θ/η gives, after some manipulations,

$$\frac{\mathrm{d}^2 W}{\mathrm{d}t\,\mathrm{d}K} = \frac{(1-t)\,K}{((1-t)\,K+\mu)\mu^2} \left(\frac{((1-t)\,K + 2\mu)\,\big(\mu^2 - (1-t)^2 K^2\big)}{(1-t)\,K^2} + \frac{K\mu + (1-t)^2 (K+1)\,K^2 - \mu^2 K}{(K+\mu)(K+1)} - \frac{\mu}{K+\mu} \right). \tag{B.16}$$

Using $\mu^2 - (1-t)^2 K^2 = t^2(1-\eta)$, this can be further simplified to yield

$$\frac{\mathrm{d}^2 W}{\mathrm{d}t\,\mathrm{d}K} = \frac{(1-t)\,(1-\eta)\,K}{((1-t)\,K+\mu)\,\mu} \left(t^2 \left(\frac{1}{(K+\mu)\,K} + \frac{2}{(1-t)\,K^2} \right) - \frac{1}{(K+1)\,(K+\eta)} \right). \tag{B.17}$$

The first term in the large brackets of (B.17) is positive; this term gives the higher efficiency costs of taxation. The second term in (B.17) is negative; this term describes the increased redistributive motive of the tax.

We need to show that (B.17) is positive for all levels of K and t. Ignoring the first positive term in parentheses, a sufficient condition for this to be true is that

$$\frac{2t^2(K+1)(K+\eta) - (1-t)\,K^2}{(1-t)\,K^2(K+1)(K+\eta)} > 0 \;\Rightarrow\; (2t^2 + t - 1)\,K^2 + 2t^2(K + \eta + \eta\,K) \equiv E(K) > 0\,, \tag{B.18}$$

for all levels of K. We first evaluate $E(K)$ in (B.18) for $K \to 0$. This gives $E(K)|_{K\to 0} = 2t^2\eta > 0$, since t approaches zero less quickly than K does. This follows from the fact that $\hat{\alpha}$ in (13) must approach a maximum when $K \to 0$ and hence effort costs c are at a maximum:

$$\frac{\partial K}{\partial c} = \Psi\,\pi\,\big((1-\eta) + (\Gamma - 1)(1 - \eta\pi)\big) < 0\,, \tag{B.19}$$

where Ψ has been defined in the Proof of Proposition 2.

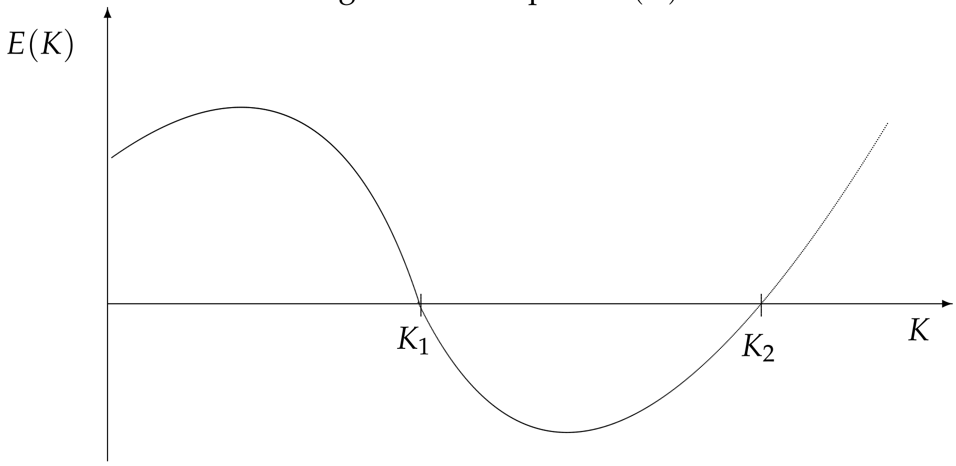
Next, we evaluate $E(K)$ for $K \to \infty$. In this case $\hat{\alpha} \to 0$ from (13) and the tax becomes a non-distortive instrument. Hence $t \to 1$, which implies $E(K)|_{K\to\infty} > 0$ from (B.18).

It remains to be shown that $E(K) > 0$ also for all intermediate values of K. To show this, we differentiate $E(K)$ in (B.18) with respect to K, giving

$$\frac{dE}{dK} = 2(2t^2 + t - 1)\,K + 2t^2(1+\eta) + \big(4t(K + \eta + \eta\,K + K^2) + K^2\big)\,\frac{\mathrm{d}t}{\mathrm{d}K}\,. \tag{B.20}$$

We first show that dE/dK is positive for both $K \to 0$ and $K \to \infty$. Evaluating (B.20) at $K \to 0$ leaves the second and third terms. These must both be positive, since $E(K)|_{K\to 0} > 0$ is a sufficient condition for $(\mathrm{d}t/\mathrm{d}K)|_{K\to 0} > 0$. Evaluating (B.20) at $K \to \infty$, the first term must be positive, as $t \to 1$ from the argument above. Moreover, the third term in (B.20) is positive given that $E|_{K\to\infty} > 0$ is a sufficient condition for $(\mathrm{d}t/\mathrm{d}K)|_{K\to\infty} > 0$. It follows that $E(K) < 0$ can

Figure A.1: Graph of $E(K)$



only occur for some intermediate range of K when the graph of $E(K)$ has two local extrema, as shown in Figure 8.

We proceed by contradiction and assume that this is indeed the case. Then the $E(K)$ curve must cross the $K = 0$ axis twice, at levels K_1 and K_2. We rewrite $E(K)$ in (B.18) as

$$E(K) = K\left((2t^2 + t - 1)\,K + 2t^2(1 + \eta) + \frac{2t^2\eta}{K}\right) . \tag{B.21}$$

Assume that (B.21) is zero at $K = K_1$, implying that the sum of terms in the large parentheses is zero. Moreover, by assumption, $E(K)$ must be negative between K_1 and K_2. Hence $t|_{K_2}$ must be smaller than $t|_{K_1}$.[25] But then a comparison of the three terms in the large parentheses of (B.21) shows that these are all smaller at K_2 than at K_1. The first term must be negative if $E(K_1) = 0$ is to hold, and it is smaller (i.e., more negative) at K_2, since $K_2 > K_1$ and $t|_{K_2} < t|_{K_1}$. Moreover, under these conditions the positive second and third terms are both unambiguously smaller at K_2 then they are at K_1. Hence when $E(K_1) = 0$, there cannot be a value $K_2 > K_1$ for which $E(K_2) = 0$ holds. But then the graph of $E(K)$ cannot be as depicted in Figure 8. By contradiction, this implies that $E(K) > 0$ must hold for any level of K. Since $E(K) > 0$ is sufficient for $\mathrm{d}t/\mathrm{d}K > 0$, this proves the result. □

[25]To be precise, $E(K) < 0$ is only a necessary, but not a sufficient, condition for t to fall between K_1 and K_2. But if t does not fall between K_1 and K_2, then it cannot fall for any level of K, because $E(K) > 0$ holds for $K < K_1$ and for $K > K_2$. The result would then follow immediately.