cesifo WORKING PAPERS

8550
2020

September 2020

Taxing Mobile and Overconfident Top Earners

Andreas Haufler, Yukihiro Nishimura

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp
An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 8550*

Taxing Mobile and Overconfident Top Earners

Abstract

We set up a simple model of tax competition for mobile, highly-skilled and overconfident managers. Firms endogenously choose the compensation scheme for managers, which consists of a fixed wage and a bonus payment in the high state. Managers are overconfident about the probability of the high state and hence of receiving the bonus, whereas firms and governments are not. In this setting we show that overconfidence (i) unambiguously increases the bonus component in the managers' compensation package and (ii) it reduces the bonus tax rate that governments set in the non-cooperative tax equilibrium. Hence overconfidence can contribute to explaining both the increasing role of bonus contracts and the fall in marginal tax rates for high-income earners.

JEL-Codes: H200, H870, G280.

Keywords: overconfidence, bonus taxes, tax competition, migration.

Andreas Haufler
Seminar for Economic Policy
University of Munich
Akademiestr. 1/II
Germany – 80799 Munich
Andreas.Haufler@econ.lmu.de

Yukihiro Nishimura
Graduate School of Economics
Osaka University / Japan
ynishimu@econ.osaka-u.ac.jp

September 2020
Paper presented at the Belgian-Japanese Public Finance Workshop in Louvain-la-Neuve and at the 2020 online conference of the IIPF. We thank David Agrawal, Traviss Cassidy, Daniel Gietl, Bernhard Kassner, Mohammed Mardan and Pierre Pestieau for helpful comments. For financial support we thank the German Research Foundation through CRC TRR 190 (Haufler) and MEXT/JSPS KAKENHI JP18H00866 (Nishimura). The paper was started when Yukihiro Nishimura visited the University of Munich. He wishes to thank the members of CES for their hospitality.

1 Introduction

During the last decades, top income earners have been able to increase their share in total national income in most countries.[1] An important contributor to this increased income concentration is the prevalence of bonus payments and other forms of incentive pay. For example, Bell and Van Reenen (2014) show that the top percentile of income earners in the United Kingdom received 35% of their total pay as bonus income in 2008, and the bonus share was even 44% in the financial sector. Similarly, Lemieux et al. (2009) find for the United States that performance pay accounts for most of the increase in wage inequality above the 80th percentile during the period 1976-1998.[2]

At the same time, top income tax rates have been reduced in many countries. Egger et al. (2019) have shown for the OECD countries that income tax systems have become less progressive since the mid-1990s. They explain this development with the increasing international mobility of high income earners. Several countries specifically try to attract international top income earners by means of further tax cuts that are only available to foreign residents (Kleven et al., 2020). Theoretical work has shown that the international mobility of top income earners reduces the optimal progressivity of income tax schedules in the countries competing for the high-skilled, mobile population (Lehmann et al., 2014). These effects are confirmed in empirical studies demonstrating that foreign residents respond to tax incentives with an elasticity that is far larger than the response of domestic residents (Kleven et al., 2014; Akcigit et al., 2016).

In this paper we introduce overconfidence as a behavioral trait of mobile top earners and ask how this affects the tax competition for them. Specifically we focus on overestimation as the most common form of overconfidence.[3] The psychology literature has shown that many individuals overestimate their own abilities and talents, as well as the probabilities of advantageous events (Taylor and Brown, 1988). This behavioral pat-

[1] See Atkinson et al. (2011) for an international comparison and Piketty et al. (2018) for a detailed study of income distribution in the United States, based on national accounts data.

[2] More generally, Bryson et al. (2012, Figures 1 and 2) show 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%) and in the U.S. (over 40%).

[3] Moore and Healy (2008) distinguish three notions of overconfidence: overestimation, overplacement (relating to comparisons with others) and overprecision. While overprecision relates to the margin of error in stochastic decisions, overestimation relates to the probability with which (typically positive) outcomes occur.

tern is particularly pronounced among high-income individuals, who have experienced success in their previous career and attribute this success largely, or even exclusively, to their own abilities (Gervais and Odean, 2001). Empirical research has convincingly shown that company CEOs and top managers exhibit systematic and persistent patterns of overconfident behavior (Malmendier and Tate, 2015; Ho et al., 2016).[4]

We cast our analysis in a framework where mobile and overconfident managers hold an incentive contract and receive a compensation that consists of both a fixed wage and a bonus in case of success. The tax we consider can therefore be conveniently interpreted as a bonus tax, but it can also be seen more generally as a higher tax rate (or surtax) on top incomes.[5] We then analyze how overconfidence of top managers affects both the compensation structure of the firms employing them, and the bonus tax rate that governments competing for the mobile high-skilled levy in the non-cooperative tax equilibrium.

Our first main finding is that overconfidence increases the share of the bonus component in total manager compensation, relative to the fixed wage rate. This higher prevalence of bonuses makes the ex-post distribution of incomes more unequal and it increases the tax base for the bonus tax. Nevertheless, we find that higher levels of overconfidence *reduce* revenue-maximizing bonus tax rates in the symmetric Nash equilibrium.

The intuition for this result is that overconfident managers overestimate the likelihood that they will receive the bonus, and hence pay the bonus tax. This increases their migration elasticity, and rational governments have to factor in this behavioral response. Since the true likelihood that the bonus is paid and the bonus tax is indeed collected is lower than anticipated by managers, overconfidence makes the bonus tax a less attractive instrument from the perspective of governments. Hence, overconfidence contributes to explaining the increased reliance of firms on bonuses and other forms of incentive pay, the high tax elasticity of highly skilled migrants, and the fall in top income tax rates faced by them.

The falling tax rate on bonus incomes does not imply, however, that overconfidence

[4]The dominant approach to empirically identify managerial overconfidence is the failure of top managers to diversify their personal investment portfolio by selling stock options of the manager's own firm that are "in the money" at the earliest possible date (Malmendier and Tate, 2005).

[5]Many countries incorporate bonuses in the tax base for the general income tax. The United States is an exception as it taxes bonuses as 'supplementary income' under a separate tax schedule. This schedule has fewer tax brackets than the general income tax schedule, but the top tax rates are aligned.

is detrimental for the government's tax revenue collections. In fact, our second main result shows that managerial overconfidence raises bonus tax revenues in equilibrium, despite the falling bonus tax rate. This is because overconfidence shifts the manager's compensation towards the taxed bonus component of his pay, and this increase in the bonus tax base dominates the effects of a lower tax rate. In fact, governments can "exploit" the overconfidence of managers in equilibrium, in the sense that the rewards to the overly high effort caused by the manager's overconfidence partly accrue to the government as tax revenue.

While our benchmark model considers only a single tax on bonuses, we show that our main results are robust to introducing a two-tier income tax where the fixed wage income is taxed at a positive, exogenous rate and bonus income is subject to an additional tax or, alternatively, to a higher marginal tax rate. In this setting it remains true that the extra tax on bonus income is reduced by overconfidence, and tax progressivity accordingly falls. In the second extension, we show that additionally incorporating shareholder overconfidence offsets some of the effects that are caused by the overconfidence of mobile managers. Nevertheless, the result that overconfidence leads to more aggressive tax competition and lower equilibrium tax rates survives also in this setting.

Our paper contributes to the literature on the optimal taxation of mobile high-income individuals (Simula and Trannoy, 2010; Bierbrauer et al., 2013; Lehmann et al., 2014; Wilson, 2015). Most of this literature considers non-linear income taxation and focuses on marginal migration costs as a determinant of optimal marginal tax rates.[6] This literature shows that marginal tax rates generally fall as a consequence of international mobility, and that marginal tax rates on top earners may even become negative. We simplify the tax schedule by focusing on a linear bonus tax, and add overly optimistic beliefs of the high-income migrants.

We also contribute to the recent literature on optimal income taxation with behavioral agents (Bernheim and Taubinsky, 2018; Farhi and Gabaix, 2020).[7] This literature has so far focused mainly on issues of tax salience and misperceptions of the tax schedule (Taubinsky and Rees-Jones, 2018; Rees-Jones and Taubinsky, 2020). By incorporating

[6]The empirical literature on tax-induced mobility confirms that high-income earners respond to higher taxes by relocating to other regions or countries (Schmidheiny and Slotwinski, 2018; Agrawal and Foremny, 2019).

[7]See Moore and Slemrod (2019) for a recent generalization of optimal tax systems with behavioral agents that incorporates non-rate policy instruments.

overconfidence, we introduce instead individual misperceptions of their own *abilities*, which is a major theme in behavioral finance (Malmendier and Tate, 2015). Moreover we take the behavioral public economics literature to an international setting where countries compete in attracting mobile and overconfident managers.

Our analysis of bonus payments and their taxation draws on the framework of Besley and Ghatak (2013). Gietl and Haufler (2018) have extended this setting to analyze international competition in bonus taxes when governments bail out failing banks. Gietl and Kassner (2020) introduce overconfidence, but analyze bonus taxation in a closed economy. In this paper, we use a simpler framework without default risks and governments bailouts, which is therefore not specific to the banking sector. We combine, however, the international mobility of managers with their overconfidence.

In the following, Section 2 describes our model, analyzing in turn the effort and migration decisions of high-skilled managers and the firm's choice of the managerial compensation structure. Section 3 then addresses the international competition in bonus taxes and derives our main results. Section 4 extends our benchmark model by considering a two-tier income tax in which all incomes are taxed at a positive, exogenous rate and an endogenous surcharge is levied on bonus income. Section 5 analyzes the case where the shareholders of firms are also overconfident, along with the managers. Section 6 concludes.

2 The model

2.1 The basic setup

We structure our analysis as a sequential four-stage game. In the first stage, governments non-cooperatively choose their bonus taxes, anticipating the responses of both firms and their managers to these taxes. In the second stage, firms choose their profit-maximizing remuneration scheme, consisting of a fixed wage and a bonus payment, and taking as given the bonus taxes that governments have set in the first stage. In the third stage, managers decide in which country to work, on the basis of the remuneration schemes offered to them by firms, as well as the taxes levied by their potential country of residence. Finally, in the fourth stage, managers choose their level of effort provision in the country (and hence firm) of their choice. We thus model a strict hierarchy of decisions where optimizing governments behave as first movers towards firms, whereas firms behave as first movers vis-à-vis managers.

The framework of our analysis is a region of two symmetric open economies $i \in \{1, 2\}$, which are small in the world market. In each of the two countries, there is a representative firm of variable size, where firm size corresponds to the number of identical divisions within the firm. Running a division requires the specific knowledge of a firm manager, which is the limiting resource in our model. Each firm employs exactly one manager per division and the number of managers a firm hires equals the number of its divisions. Hence each firm tries to attract internationally mobile managers in order to increase the number of its divisions, and hence profits.

Each division of a firm in country i has a total amount of fixed assets equal to one, which is lent in the world market. Lending operations are risky. We assume that there are two possible returns for each of the identical divisions of a firm, which can be high (h) or low (l). The division realizes a high return Y^h with probability $p^h > 0$ and a low return $Y^l < Y^h$ with probability $p^l = 1 - p^h$. Even the low return Y^l is sufficient for the firm to pay all its obligations, and to avoid default. Returns are fixed from the firm's perspective, as outputs are sold in a large world market. Hence, the representative firms in both countries produce with constant returns to scale.

We employ a standard principal-agent problem between the firms' shareholders and their managers. Managers have private effort costs and thus choose lower effort than would be optimal from the perspective of shareholders. Bonuses thus serve as a second-best instrument for the firm to solve this principal-agent problem.[8]

In each country i, managers choose an effort level e_i and all managers behave in identical ways with respect to their effort choice. We assume that the probability that the firm receives the high return Y^h is a linear function of each manager's effort level:[9]

$$p_i^h = \beta e_i, \qquad p_i^l = 1 - \beta e_i, \qquad \beta, p_i^h, p_i^l > 0. \tag{1}$$

In our model, firm managers overestimate their skills and thus overestimate the return to their effort. We denote parameters as perceived by an overconfident manager with a hat ($\hat{}$). The probabilities of the high and the low state, as perceived by managers,

[8] There is an active discussion of whether existing incentive contracts are compatible with modern shareholder value theories (Edmans and Gabaix, 2016), or reflect "rent extraction" by CEOs (Piketty et al., 2014). Our discussion restricts attention to the productivity enhancing effects of incentive contracts.

[9] Since managers are identical, manager-specific indices are omitted from all variables.

are then given by

$$\hat{p}_i^h = (1+\theta)\beta e_i, \qquad \hat{p}_i^l = 1 - (1+\theta)\beta e_i. \tag{2}$$

The parameter $\theta > 0$ in eq. (2) measures the degree of managers' overconfidence. Overconfidence thus leads each manager to overestimate the likelihood of the high state, $\hat{p}^h > p^h$, and hence to overestimate the expected reward to his effort. We confine the level of overconfidence to the range $0 \leq \theta < 1$.[10] In the following, changes in θ will be at the heart of our analysis of how overconfidence affects the firms' compensation schemes for their managers, as well as the governments' optimal bonus taxes.

2.2 Managers' effort choice

We proceed by backward induction to obtain the subgame perfect Nash equilibrium. In Stage 4, risk-neutral managers choose their effort levels. For analytical tractability, we assume that the cost of effort provision is quadratic and given by $c(e) = \eta e^2/2$. Due to these private costs, managers will not exert enough effort from the point of view of firm owners. Effort decisions are not observable. However, firm owners can mitigate the principal agent problem by a bonus payment z_i in the high return state, which occurs with a higher probability $p^h(e)$ when the manager's effort level e is increased.

Managers located in country i maximize their *perceived* location-specific utility $\hat{u}_i$, which is the excess of expected bonus payments and the fixed wage w_i over the private costs of effort.[11] Using (2) gives

$$\hat{u}_i = \hat{p}_i{}^h z_i + w_i - c(e_i) = (1+\theta)\beta e_i z_i + w_i - \frac{\eta e_i^2}{2}. \tag{3}$$

Maximizing (3) with respect to the managers' choice variable e_i yields

$$e_i = \frac{(1+\theta)\beta z_i}{\eta}\ . \tag{4}$$

Hence each manager's effort level e_i increases in the bonus payment z_i and in his level of overconfidence θ. The latter occurs because an overconfident manager overestimates

[10] The upper bound placed on θ is needed to ensure that the firms' optimal bonus choice is well-defined; see Section 2.4 below. The constraint that $\hat{p}^h < 1$ must hold in (2) will then be met by placing an upper bound on the parameter β.

[11] The perceived utility is also labelled *decision utility* in the behavioral economics literature. In contrast, the utility evaluated at the true success probability p_i^h is known as *experience utility* (e.g. Farhi and Gabaix, 2020).

the likelihood that the high state will occur, and hence the expected return to his effort. In contrast, the fixed wage w_i does not affect managers' optimal effort.

Substituting the managers' effort decision (4) into (1), we can derive the *true* equilibrium probabilities of the states h and l:

$$p_i^{h*} = \beta e_i = (1+\theta)\left[\frac{\beta^2}{\eta}\right] z_i \equiv (1+\theta)\gamma z_i, \tag{5a}$$

$$p_i^{l*} = 1 - \beta e_i = 1 - (1+\theta)\gamma z_i. \tag{5b}$$

In eq. (5a), we have introduced the parameter $\gamma > 0$ to summarize the marginal effect of the bonus payment on the probability of a high return. This effect is multiplied by the overconfidence factor $(1+\theta)$. Therefore, the *true* success probability is rising in the overconfidence of managers, because of the higher effort level that is induced by the bonus when managers overestimate their return to effort.

The managers' *perceived* success probability is then $\hat{p}_i^h = (1+\theta)p_i^{h*} = (1+\theta)^2\gamma z_i$. Substituting this along with (4) and (5a) into (3) gives the location-specific *perceived* utility of a manager working in country i. This perceived utility is increased by both a higher bonus and a higher fixed wage:

$$\hat{u}_i^* = (1+\theta)^2 \frac{\gamma z_i^2}{2} + w_i. \tag{6}$$

Again, an overconfident manager ($\theta > 0$) overvalues the influence of the bonus payment z_i on his utility in country i, as he overestimates his success probability.

2.3 Managers' migration decision

In Stage 3 managers take the bonuses z_i and fixed wages w_i as given and choose whether to work in country 1 or in country 2. Managers maximize their gross utility, which consists of the location-specific utility in (6), and the non-monetary attachment to a particular country. There are a total of $2\bar{N}$ managers in the region, which are all employed in one of the two countries. Hence $N_1 + N_2 = 2\bar{N}$, where N_i is the number of managers working in country i in equilibrium.

Managers differ in their country preferences. More precisely, managers are of type m, where m is the relative attachment to country 1 and we assume that m is distributed uniformly along $[-\bar{N}, +\bar{N}]$. Other things equal, all managers with $m > 0$ prefer to work in country 1, whereas managers with $m < 0$ prefer to work in country 2.

The utility weight of the location preference parameter m is given by the constant $a > 0$. The gross utility U_i of a manager of type m in country i is then

$$U_1(z_1, w_1, m) = \hat{u}_1^*(z_1, w_1) + am, \qquad U_2(z_2, w_2) = \hat{u}_2^*(z_2, w_2). \tag{7}$$

All managers choose to work in the country that gives them the higher gross utility. We characterize the manager that is just indifferent between working in country 1 or in country 2 by the critical location preference m^c. Equating U_1 and U_2 in (7) and using (6), we derive m^c as a function of differences in bonus payments and fixed wages between the two countries:

$$m^c = \frac{1}{a}\left[\frac{\gamma}{2}(1+\theta)^2(z_2^2 - z_1^2) + (w_2 - w_1)\right]. \tag{8}$$

Managers with $m \in [m^c, \bar{N}]$ work in country 1 and managers with $m \in [-\bar{N}, m^c]$ work in country 2. Using (8) then determines the number of managers in country i as a function of the differences in bonus payments and wages:

$$N_i = \bar{N} + \frac{1}{a}\left[\frac{\gamma}{2}(1+\theta)^2(z_i^2 - z_j^2) + (w_i - w_j)\right] \qquad \forall\, i, j \in \{1, 2\},\ i \neq j. \tag{9}$$

The larger is the bonus of country i, relative to that of country j, the more managers will work in country i in equilibrium. The same holds for the fixed wage. As managers value the bonus payment z_i in each country by their *perceived* probability of the high state $\hat{p}^h$, the effect of bonuses (z_1, z_2) on the location decision of managers is increased by managers' overconfidence, whereas the effect of the fixed wage is not.

2.4 Firms' compensation choices

In Stage 2, we turn to the remuneration decisions made by the owners of the representative, scalable firm in each country. The representative firm in country i sets the bonus $z_i \geq 0$ and the fixed wage $w_i \geq 0$ to maximize its expected after-tax profits. Both bonuses and fixed wage payments are constrained to be non-negative. Since all divisions are equal, total profits Π_i are obtained by multiplying the profits of a representative division, π_i^D, with the number of divisions, which equals the number of managers N_i. Firms assess their expected division profits using the true probabilities of the high and the low state (p_i^{h*}, p_i^{l*}). The bonus is paid only in state h, and its gross costs to the firm are increased by the bonus tax t_i. Therefore, total after-tax expected profits are

$$\Pi_i = N_i \pi_i^D = N_i \left\{ p_i^{h*}[Y^h - z_i(1+t_i)] + p_i^{l*} Y^l - w_i \right\}. \tag{10}$$

The firm maximizes its profits in (10) with respect to the bonus z_i, taking account of the managers' migration decision (9) and their equilibrium effort levels, which determine probabilities by (5a)–(5b). This gives

$$\frac{\partial \Pi_i}{\partial z_i} = (1+\theta)^2 \frac{\gamma z_i}{a}\, \pi_i^D + N_i (1+\theta)\gamma \left[(Y^h - Y^l) - 2 z_i (1+t_i) \right] = 0. \tag{11}$$

The first effect in eq. (11) gives the effect of the bonus on the number of firm divisions. This effect is unambiguously positive. As we discussed in (9) above, it incorporates the managers' migration decision that is based on their *perceived* success probability $\hat{p}^{h*} = (1+\theta)^2 \gamma z_i$. The second effect in eq. (11), which describes the marginal effect of the bonus on the profits of a representative division, must therefore be negative in the firm's optimum. For this second effect, the *true* success probability $p^{h*} = (1+\theta)\gamma z_i$ is relevant, as this determines the probability with which the bonus is actually paid.

Maximizing firm profits in (10) with respect to the fixed wage gives

$$\frac{\partial \Pi_i}{\partial w_i} = \frac{\pi_i^D}{a} - N_i \leq 0 \quad \forall\, i. \tag{12}$$

The first-order condition (12) holds with equality, if and only if the fixed wage is positive in the firm's optimum; otherwise the fixed wage is zero. In line with existing manager compensation schemes, we assume that $w_i > 0$ holds in the firm's optimum. This implies $\pi_i^D / a = N_i$. Substituting this into (11) we can simplify the first-order condition for the optimal bonus payment. The optimal bonus can then be expressed as:[12]

$$z_i^* = \frac{\Omega}{1 - \theta + 2t_i} \quad \forall\, i, \qquad \Omega \equiv Y^h - Y^l. \tag{13}$$

Hence the equilibrium bonus is rising in the output gap between the high and the low state, Ω, and it is falling in the bonus tax rate t_i. Moreover, for a given tax rate t_i, the bonus payment is unambiguously increasing in the managers' degree of overconfidence θ. Intuitively, since managers overestimate the probability of the good state h, they also overestimate the probability of receiving the bonus. This makes the bonus an attractive form of compensation from the perspective of the firm, which holds rational expectations about the likelihood of the high state.

[12]To ensure positive bonuses, the denominator in eq. (13) must be positive. This will be guaranteed if $\theta < 1$ and $t_i \geq 0$. This is the reason for placing an upper bound on θ (cf. footnote 10). The second-order condition for the firm's optimization problem with respect to z_i and w_i will also hold if $\theta < 1$. The proof is available from the authors upon request.

In Appendix A.1, we use equations (12) and (13) for both countries to derive the optimal fixed wage, as a function of both countries' tax rates. This is given by

$$w_i^* = Y^l - a\bar{N} + \frac{\gamma(1+\theta)\Omega^2}{6}\left[\frac{1}{(1+2t_j-\theta)} + \frac{4t_i - 5\theta - 1}{(1+2t_i-\theta)^2}\right]. \tag{14}$$

The second term in (14) shows that an increase in manager mobility (a decrease in a) increases the fixed wage. In contrast, the parameter a does not appear in the optimal bonus equation (13). This shows that the fixed wage is the firms' marginal instrument to attract mobile workers in our analysis, whereas the bonus payment is used to incentivize effort. The overall effects of the overconfidence parameter θ on the fixed wage will be analyzed once we have solved for the equilibrium bonus tax rate t_i.

3 International competition in bonus taxes

3.1 The tax equilibrium

In Stage 1, the two symmetric governments choose their optimal bonus taxes t_i. Governments are assumed to maximize their tax revenue W_i. Revenue maximization is the natural government objective in our framework. It corresponds to maximizing the transfer to the local resident population, and hence local welfare, when we make the standard assumption that governments do not include footloose managers in their welfare objective.[13] Like firms, governments form an unbiased expectation about the probabilities of the different states. Hence they maximize

$$W_i = t_i p_i^{h*} z_i N_i. \tag{15}$$

In the first step, we determine the response of the equilibrium number of managers in country i when this country raises its bonus tax. Differentiating (9) with respect to t_i and taking account of (13) and (14) leads to

$$\frac{\partial N_i}{\partial t_i} = \frac{1}{a}\left[\frac{\gamma(1+\theta)^2}{2}\frac{\partial(z_i^2)}{\partial t_i} + \frac{\partial w_i}{\partial t_i} - \frac{\partial w_j}{\partial t_i}\right] = \frac{-\gamma(1+\theta)\Omega^2}{3a(1-\theta+2t_i)^2} < 0. \tag{16}$$

From (16) we see that the outflow of managers in response to a higher bonus tax unambiguously rises in the managers' overconfidence level θ. This is because the managers'

[13] Note also that firm profits are not affected by bonus taxes in our benchmark model, as we argue in more detail in Section 3.2 below.

migration decision is based on their overly high expectation of receiving the bonus, combined with their correct anticipation that the tax will reduce the bonus payment.

In the second step we maximize tax revenue in (15) with respect to t_i, using the migration response (16), along with the effects of the tax on the success probability p_i^{h*} in (5a) and on the bonus payment z_i in (13). This gives

$$\frac{\partial W_i}{\partial t_i} = \frac{z_i^2 \gamma(1+\theta)}{a} \left[aN_i \left\{ 1 - \frac{4t_i}{1-\theta+2t_i} \right\} - \frac{t_i \gamma (1+\theta)\Omega^2}{3(1-\theta+2t_i)^2} \right]. \tag{17}$$

In eq. (17), the first term in the squared bracket gives the change in tax revenues for a given number of managers. This term is in turn composed of the mechanical effect of a tax increase at an unchanged tax base, and the fall in the expected tax base induced by the higher bonus tax (resulting from both the lower bonus payment and the reduced success probability of the project). The negative second term in the squared bracket gives the loss in the tax base that results from the outmigration of managers in response to the tax.

Evaluated at $t_i = 0$, the derivative $\partial W_i / \partial t_i$ is positive in eq. (17). Also, continuity of $W_i(t_i, t_j)$ is guaranteed in our setting because all the relevant functions are continuous in t_i and t_j. Finally, Appendix A.2 derives the second-order condition for the governments' optimal choice of bonus taxes, and shows that this is fulfilled. Hence, a Nash equilibrium with positive bonus tax rates in both countries must exist in our model.

In the following, we focus on a symmetric Nash equilibrium. Setting (17) equal to zero at $t_i = t_j$, which in turn implies $N_i = \bar{N}$, yields a quadratic equation in t. Solving gives an explicit expression for the unique Nash equilibrium tax rate in both countries:

$$t_i^* = -\frac{\gamma\,\Omega^2\,(1+\theta)}{24\bar{N}a} + \sqrt{\left(\frac{\gamma\Omega^2\,(1+\theta)}{24\bar{N}a}\right)^2 + \frac{1}{4}\,(1-\theta)^2} > 0. \tag{18}$$

The simple structure of (18) allows us to unambiguously sign the derivative of the Nash equilibrium tax rate with respect to the overconfidence parameter θ:

$$\frac{\partial t_i^*}{\partial \theta} = \frac{1}{24\bar{N}a} \left(-\Omega^2 \gamma + \frac{\gamma^2 \Omega^4 (1+\theta) - 144(1-\theta) a^2 \bar{N}^2}{\sqrt{\gamma^2 \Omega^4 (1+\theta)^2 + 144(1-\theta)^2 a^2 \bar{N}^2}} \right) < 0. \tag{19}$$

Equation (19) shows that overconfidence of mobile managers reduces the optimal tax rate that both governments choose in the Nash equilibrium. Note that this result occurs even though the bonus tax base is increasing in the overconfidence level θ for any given tax rate t_i^* [eq. (13)]. However, this effect is overcompensated by a twofold increase in the elasticity with which the bonus tax base responds to taxation. First, even in the absence of manager migration, firm's bonus payments respond more elastically to the bonus tax when θ increases. This is seen from the positive first effect in the squared bracket of the first-order condition (17), whose negative component is rising in θ. Moreover, a higher level of overconfidence leads to a larger migration response of managers to the bonus tax rate, as is seen from the negative second effect in the squared bracket of (17).

Intuitively, the negative effect of overconfidence on the governments' optimal bonus tax rate arises because revenue-maximizing governments have to incorporate the higher migration elasticity with which overconfident managers respond to the bonus tax. At the same time, a rational government will calculate the expected bonus tax revenue only on the basis of the true success probability. Therefore, a higher overconfidence level raises the *elasticity* of the bonus tax base more quickly than it raises the bonus tax base itself. This makes the bonus tax a less attractive instrument from the government's perspective.[14]

Given that the optimal bonus tax rate is falling in the overconfidence level θ, we can now determine how the firms' compensation schemes for managers change as a result of overconfidence. Inserting this result in the bonus payment (13) shows that the positive direct effect that an increase in θ has on the equilibrium bonus payment z_i increases further due to the indirect effect that results from the reduced bonus tax rate. Therefore, equilibrium bonus payments unambiguously increase when managers become more overconfident.

Next we derive the equilibrium change in the fixed wage payment of firms that results from overconfidence. Evaluating the fixed wage expression (14) at the common Nash equilibrium tax rate $t_i^* = t_j^*$ in (18) and differentiating gives

$$\frac{\partial w_i^*}{\partial \theta} = \gamma \Omega^2 \frac{(2t_i - 3\theta - 1)}{(1 + 2t_i - \theta)^3} \left[1 + t_i - (1 + \theta) \frac{\partial t_i^*}{\partial \theta} \right] < 0. \tag{20}$$

This is negative because $2t_i^* - 3\theta - 1 < 0$ holds in the Nash equilibrium [see eq. (A.6) in Appendix A.2]. Also, the term in the squared bracket is unambiguously positive since $\partial t_i^* / \partial \theta < 0$ from (19). Therefore the fixed wage declines when managers become more

[14]Note the difference to a welfare-maximizing government in a closed economy, which would *increase* the bonus tax in order to correct the managers' distorted beliefs (Gerritsen, 2016). In contrast, governments in our model maximize tax revenues under conditions of international tax competition, in order to transfer the revenue to internationally immobile domestic residents.

overconfident, and the compensation structure unambiguously shifts towards more incentive pay. Our results are summarized in:

Proposition 1 *Increased overconfidence of mobile managers (a rise in θ) leads to:*
(i) firms choosing a compensation structure with higher bonus payments and lower fixed wages;
(ii) internationally mobile managers responding more elastically to bonus taxation;
(iii) governments choosing lower bonus taxes in the non-cooperative equilibrium, despite the increased bonus tax base.

3.2 Welfare effects of overconfidence

We can now determine in more detail how overconfidence affects the objective function of each agent in equilibrium. Firms' profits are unaffected by changes in θ in our model. This is seen from (12) when the fixed wage is part of the compensation package and (12) thus holds with equality. The firms' competition for mobile managers will, in equilibrium, fix division profits π_i^D at a level that depends only on the managers' international mobility, as measured by the parameter a. Total profits in equilibrium are then fixed by $\Pi_i^* = \pi_i^D \bar{N} = a\bar{N}^2$.

This leaves managers' utility and the government's tax revenue as endogenous variables that are affected by overconfidence. We adopt a paternalistic perspective and first ask how managers' *actual* utility $u_i^* = p_i^h z_i + w_i - \eta e_i^2/2$ is affected, in equilibrium, by changes in θ. Substituting (4), (5a), (13) and (14) we get

$$u_i^* = \gamma\, z_i^2 \left[1 + \theta - \frac{(1+\theta)^2}{2}\right] + w_i = Y^l - a\bar{N} + \frac{\gamma(1+\theta)\Omega^2[1 + 2t_i - 3\theta]}{2(1-\theta+2t_i)^2} \qquad (21)$$

Overconfidence distorts the manager's effort decision in the direction of 'excessive' effort, when evaluated at the actual success probability. At the same time, the higher effort caused by the manager's overconfidence relaxes the moral hazard constraint and increases the aggregate surplus from employing the manager.

Differentiating (21) with respect to θ, taking account of the endogeneity of t_i^* using (19) and using (A.7) in the appendix gives

$$\begin{aligned} \frac{\partial u_i^*}{\partial \theta} &= \frac{\gamma\Omega^2(2t_i^2 + t_i - 5t_i\theta - 4\theta)}{(1-\theta+2t_i)^3} - \frac{\gamma\Omega^2(1+\theta)(1+2t_i-5\theta)}{(1-\theta+2t_i)^3}\,\frac{\partial t_i^*}{\partial \theta} \\ &= \frac{4\gamma\Omega^2[(1-3\theta)t_i - \theta(1-\theta)]}{(1-\theta+2t_i)[(1-\theta)^2+4t_i^2]}. \end{aligned} \qquad (22)$$

Eq. (22) shows that managers' utility is hump-shaped in their overconfidence level θ. For $\theta = 0$, managers' utility is rising in θ, but higher initial levels of θ monotonously reduce the numerator in (22). The range of θ for which $\partial u_i^*/\partial\theta > 0$ holds can be further specified when (22) is evaluated at (18). This gives

$$\frac{\partial u_i^*}{\partial \theta} > 0 \iff \frac{\gamma\Omega^2}{a\bar{N}} < \frac{3(1-5\theta)(1-\theta)^2}{\theta(1+\theta)(1-3\theta)}. \tag{23}$$

A necessary condition for (23) to hold is that the numerator on the right-hand side of the inequality is positive. This defines a maximum level of overconfidence $\theta_1 < 0.2$ which is compatible with $\partial u_i^*/\partial\theta > 0$.[15]

Intuitively, for low levels of overconfidence, the extra effort supplied by the manager has a first-order effect on relaxing the moral hazard constraint and increasing the aggregate surplus from employing the manager. In contrast, the distortion of the manager's effort is only of second order when the initial level of overconfidence is very low. Therefore, the increase in the manager's income more than compensates for the higher effort cost. This effect is further strengthened by the fall in the bonus tax rate. For high levels of θ, however, the distortion of the manager's effort decision arising from overconfidence will cause a first-order welfare loss when measured by his actual utility, and this effect is reinforced by the falling bonus tax rate. For overconfidence levels above the critical value $\theta_1 (< 0.2)$, an increase in overconfidence therefore reduces the manager's true utility.

We next turn to the government's equilibrium level of tax revenues. Overconfidence has counteracting effects on total tax collections, because it increases the bonus tax base, but reduces the equilibrium tax rate under tax competition (see Proposition 1). The total effect of θ on tax revenues is derived in Appendix A.3 and given by

$$\frac{\partial W_i^*}{\partial \theta} = \frac{8t_i^2\gamma\Omega^2\bar{N}}{(1-\theta+2t_i)^2[(1-\theta)^2+4t_i^2]} > 0. \tag{24}$$

Hence, in the Nash equilibrium, the enlarged tax base dominates the effects of more aggressive tax rate competition between countries. To understand this result, recall that there are offsetting effects on the equilibrium bonus tax rate that follow from the higher tax base and the higher elasticity of the tax base. Combined with the unambiguous increase in the tax base, this leads to higher tax revenue collections in each of the two competing countries as a result of higher managerial overconfidence.

[15] Note that $\theta < 1/3$ is a necessary condition for $\partial u_i^*/\partial\theta > 0$ from (22). Therefore, the denominator on the right-hand side of the inequality in (23) is always positive in the relevant range.

Finally, the aggregate welfare changes resulting from overconfidence can be obtained by adding up (22) and (24); see eq. (A.11) in Appendix A.3. Evaluating at t_i^* in (18) gives

$$\frac{\partial(W_i^* + N_i u_i^*)}{\partial\theta} > 0 \iff \frac{\gamma\Omega^2}{a\bar{N}} < \frac{3(1-3\theta)}{\theta}. \tag{25}$$

Condition (25) is unambiguously fulfilled at $\theta = 0$, and it cannot be fulfilled for $\theta \geq 1/3$. More generally, it defines a critical level $\theta_2 < 1/3$ at which aggregate welfare begins to fall in θ.[16] The overall efficiency effects resulting from overconfidence are therefore hump-shaped in the overconfidence level θ. As in our discussion of (22), the effect of relaxing the moral hazard constraint dominates for low levels of θ, whereas the distortion in the manager's effort decision dominates when θ is sufficiently high.

We summarize these results in:

Proposition 2 *Increased overconfidence of mobile managers (a rise in θ):*
(i) raises managers' actual utility for $\theta < \theta_1 < 1/5$, but reduces it for $\theta > \theta_1$;
(ii) increases equilibrium tax revenues for all initial levels of θ;
(iii) raises total welfare for $\theta < \theta_2 < 1/3$, but reduces it for $\theta > \theta_2$.

Proposition 2 shows that, in addition to the overall efficiency changes resulting from overconfidence, there is a redistributive effect from the manager to the government in equilibrium. This redistributive effect arises because the manager's overconfidence leads to a higher reliance on the taxed bonus component of his pay, whereas the untaxed fixed wage component is reduced. In this sense the manager's overconfidence is "exploited" by the government in our model. This differs from the existing literature in closed economies, where a higher bonus share allows the firm to meet the participation constraint of an overconfident manager at a lower expected total salary (Gervais et al., 2011; De la Rosa, 2011). In our model, firms will not find it profitable to reduce total manager compensation, due to the international competition for scarce and mobile managers. As a result, the gains from the managers' misconception are passed on to governments in the form of higher tax revenue.

We conclude this section by emphasizing that employing an overconfident manager does not have any negative effects on the level of firm profits in our benchmark model. In Appendix A.4 we analyze the case where overconfidence leads managers to take fewer precautions, leading to higher costs for the firm in the case of adverse events. We model

[16] It is easily checked that, in the relevant range of $\theta < 1/3$, it is always true that $\theta_2 > \theta_1$.

this as a loss to the firm arising in the low state that is proportional to the manager's degree of overconfidence. Hence, in the low state, the return to the firm is given by $Y^l - \kappa\theta$, where $\kappa \geq 0$. We show that this additional effect *increases* the firm's optimal bonus payment in equilibrium, and also that it strengthens the downward effect on the equilibrium tax rate. Therefore, our results in Proposition 1 are even strengthened by this extension. Intuitively, the loss that overconfidence creates in the low state makes the firm even more willing to incentivize the manager's effort, in order to increase the likelihood that the high state with Y^h occurs. A further effect is that the loss in the low state will lead to a reduction in the manager's fixed wage, thus adding a direct, negative effect of overconfidence on the manager's equilibrium utility. Therefore, if the loss parameter κ is sufficiently large, the managers' utility may now be a falling function of θ for all levels of overconfidence.

4 Two-tier income taxation

In our benchmark model, we have considered only a tax on bonus payments, but have left untaxed the fixed salary earned by managers in each country. This has direct implications for tax revenue when firms change the mix of bonus versus fixed wage compensation. In this section we show that qualitatively the same results are obtained when we allow both the fixed wage and the bonus to be taxed, but the bonus is taxed at a higher rate under a general and progressive income tax.

It is well-known from the optimal income tax literature that replicating the directly progressive tax schedules that exist in most countries through a two-tier income tax with two endogenously chosen tax rates is extremely difficult, even if redistributive motives between individuals are explicitly accounted for.[17] In the following, we therefore treat the general income tax rate τ_i, which applies to all incomes, as exogenous. We then analyze a two-tier tax system with an additional, endogenous surtax t_i that applies only to bonuses. Given that all managers in our symmetric model earn the same compensation, this is equivalent to a directly progressive two-tier income tax system where the fixed wage falls in the lower ('general') income tax bracket with a tax rate

[17] Slemrod et al. (1994) have shown that no directly progressive income tax schedule will result in settings where standard assumptions are made with respect to social welfare functions and the distribution of abilities. Conversely, Apps et al. (2014) study the conditions under which two-tier income tax systems are directly progressive, but these schemes are not easily tractable analytically.

of τ_i, and the bonus payment in case of success falls in the higher bracket with tax rate $(\tau_i + t_i)$. We derive the optimal bonus tax in this extended setting, and ask how it is affected by overconfidence for any given level of the lower-bracket tax rate τ_i. Our approach therefore allows to directly study the implications that overconfidence has for the progressivity of this two-tier income tax.

With two-tier taxation of the manager's remuneration, the after-tax profits of the firm [eq. (10)] change to

$$\tilde{\Pi}_i = \tilde{N}_i \left\{ \tilde{p}_i^{h*}[Y^h - \tilde{z}_i(1 + \tilde{t}_i + \tau_i)] + \tilde{p}_i^{l*} Y^l - \tilde{w}_i(1 + \tau_i) \right\}, \tag{26}$$

where the tilde symbol (˜) is used for endogenous variables in this extension.

Using (5a) and (5b) and the assumption that the optimal fixed wage paid by the firm remains positive, the optimal bonus payment is now

$$\tilde{z}_i^* = \frac{\Omega}{(1-\theta)(1+\tau_i) + 2\tilde{t}_i} \quad \forall\, i. \tag{27}$$

In comparison to the benchmark case [eq. (13)], the larger denominator in (27) reflects the additional taxation of the bonus by the general income tax τ_i. The optimal fixed wage in this setting is given in Appendix B.1 [eq. (B.3)].

Tax revenue for government i is now

$$\tilde{W}_i = \tilde{N}_i[\tilde{p}_i^{h*}(\tilde{t}_i + \tau_i)\tilde{z}_i + \tau_i \tilde{w}_i] \qquad \forall\, i. \tag{28}$$

Appendix B.2 derives the optimal bonus tax in this extended setting. This is given by

$$\tilde{t}_i^* = \frac{B + \sqrt{B^2 - AC}}{A}, \tag{29}$$

$$\begin{aligned}
A &= 8\left[3 + \tau_i\left(7 + 4\tau_i + Y_a^l\right)\right] > 0, \\
B &= -4\tau_i(1-\theta)(1+\tau_i)\left[4(1+\tau_i) + Y_a^l\right] - (1+\theta)(1+2\tau_i)\gamma\Omega_a^2 < 0, \\
C &= 2(1+\tau_i)(1-\theta)\left\{(1-\theta)(1+\tau_i)(Y_a^l\tau_i + 4\tau_i^2 + \tau_i - 3) + \tau_i\,(1+\theta)\gamma\Omega_a^2\right\},
\end{aligned}$$

where $Y_a^l \equiv Y^l/(a\bar{N})$ and $\Omega_a^2 \equiv \Omega^2/(a\bar{N})$.

We can infer from the structure of (29), together with $A > 0$ and $B < 0$ for all $\theta \in (0,1)$, that the bonus tax will be positive if and only if the term C is negative. In this case it will thus be optimal to tax bonus income at a higher rate than income from the fixed wage. Inspection of the term C shows that this will unambiguously be the case when $\tau_i = 0$, in which case we return to our benchmark setting.[18] As τ_i is continuously raised, this tends to reduce the optimal bonus tax $\tilde{t}_i^*$. Therefore, $\tilde{t}_i^*$ will be lower than the corresponding bonus tax rate in the benchmark case, t_i^*, for all $\tau_i > 0$ and $\theta \in [0, 1]$. At some point the general income tax rate will reach a critical level, denoted by τ_i^{crit}, at which the term C turns positive and hence $\tilde{t}_i^* < 0$. Our following analysis focuses on the parameter range $\tau_i \in [0, \tau_i^{crit}]$ for which the bonus surtax $\tilde{t}_i^*$ is positive, and generates additional tax revenue.

Our main interest is in the question of how an increase in the overconfidence level of managers affects the optimal bonus tax under this extension. This is summarized in:

Proposition 3 *When all incomes are taxed at an exogenous rate $\tau_i \in [0, \tau_i^{crit}]$, and the optimal bonus tax rate $\tilde{t}_i^*$ is positive in the symmetric, non-cooperative equilibrium, then increased overconfidence of mobile managers (a rise in θ) leads to:*
(i) higher bonus payments and lower fixed wages by firms;
(ii) managers responding more elastically to bonus taxation;
(iii) lower optimal bonus taxes $\tilde{t}_i^*$ in both countries.

Proof: See Appendix B3.

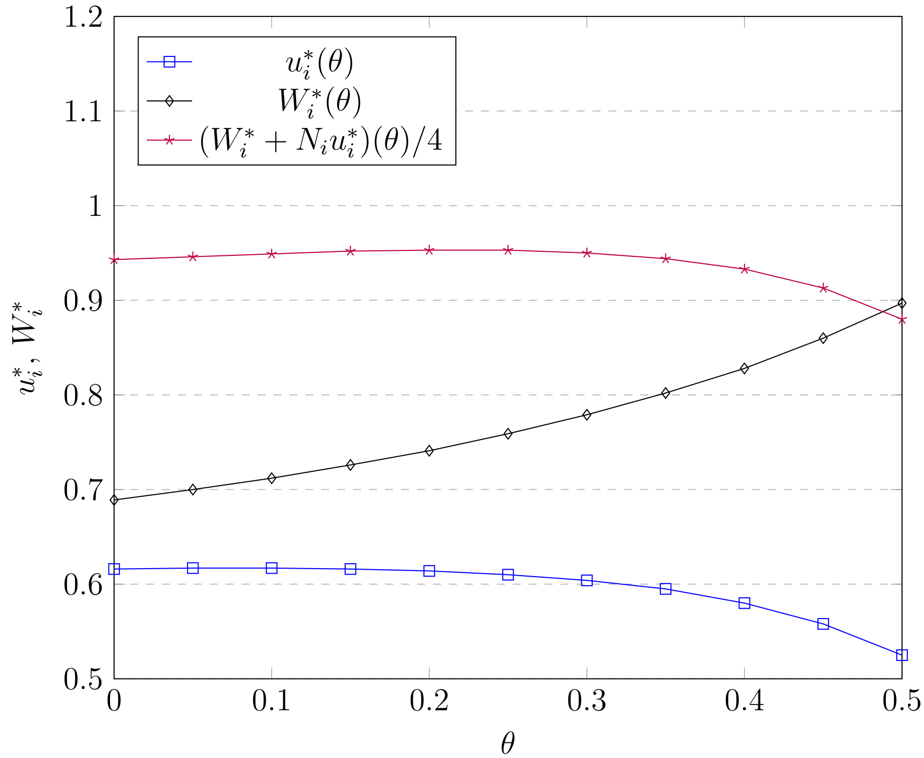
Proposition 3 states that our results from the benchmark case (Proposition 1) carry over to an extended setting with a positive general rate of income taxation, whenever the endogenous bonus surtax is positive in the optimum. In this case the intuition from our benchmark model carries over as well. The higher bonus payments by firms will increase the bonus tax base. However, the higher overconfidence of managers simultaneously increases their migration response, making the bonus tax base respond more elastically to a higher bonus tax rate. In equilibrium, the effect of the higher tax base elasticity dominates, and the optimal bonus tax must fall. To put it differently, a directly progressive income tax system will become less progressive in the government's optimum, when the overconfidence level of mobile managers increases.

It remains to discuss whether the welfare effects from our benchmark model carry over to this extended setting. The two-tier income tax framework is too complex to derive the welfare effects of overconfidence analytically. It is straightforward, however, to carry out a numerical analysis. Figure 1 provides a representative example of the results.[19]

Figure 1 shows that the effects of overconfidence on the manager's utility and on tax revenue are analogous to those discussed for our benchmark model (Proposition 2).

[18] This can easily be checked by setting $\tau_i = 0$ in expressions A–C. In this case the optimal tax

Figure 1: Welfare effects of changes in θ under two-tier income taxation



The manager's welfare is hump-shaped in θ, and it becomes a decreasing function of θ at a critical value of $\theta_1 \approx 0.08$. Tax revenue is monotonously rising in θ. Total welfare is again hump-shaped in θ, with a critical level $\theta_2 \approx 0.22$ at which it becomes a decreasing function of θ. These results change only little when the basic income tax rate τ_i is varied exogenously. In sum, we can therefore conclude that all the results from our benchmark model remain qualitatively unchanged under a two-tier income tax with an exogenously set base rate of $\tau \in [0, \tau_i^{crit}]$.

5 Overconfident shareholders and managers

A further interesting extension of our benchmark model is to analyze the case where the firms' shareholders are also overconfident, in addition to their managers. The literature on behavioral corporate finance has pointed to a number of examples where

in (29) collapses to eq. (18) in Section 3.

[19]The numerical analysis is based on the following parameter vales: $Y^l = 1$, $\Omega = 0.5$, $\bar{N} = 5$; $a = 0.05$, $\gamma = 0.5$, $\tau = 0.2$. In Figure 1, we show $(W_i^* + N_i u_i^*)/4$, in order to depict total welfare in the same graph as its components.

investor sentiments deviate from rational expectations, and these behavioral biases can go hand in hand with the biases of the top management. In general, investor sentiments are found to be more cyclical than those of top managers or CEOs, with periods of overoptimism alternating with periods of excessive pessimism.[20] In our treatment here, we focus on a period where investor sentiment, like that of managers, is overoptimistic.

To isolate the effects of this model extension, we return to a setting where only bonuses are taxed, but the fixed wage is not ($\tau_i = 0$). To simplify notation, we further assume that the degree of overconfidence is the same for the firms' managers and their shareholders. Therefore, shareholders assess their expected division profits using the overestimated success probability $\hat{p}_i^{h*} = (1+\theta)^2\gamma z_i$. Denoting variables under overconfident expectations of shareholders by a 'hat' symbol, investors' perceived total after-tax expected profits are then given by

$$\hat{\Pi}_i = \hat{N}_i \left\{(1+\theta)^2\gamma\hat{z}_i[Y^h - \hat{z}_i(1+\hat{t}_i)] + [1-(1+\theta)^2\gamma\hat{z}_i]Y^l - \hat{w}_i\right\}. \tag{30}$$

In comparison to our benchmark case [eq. (10)], shareholders now overestimate both the expected firm profits and the expected bonus payments to managers. The first-order condition with respect to the bonus payment z_i is:

$$\frac{\partial\hat{\Pi}_i}{\partial\hat{z}_i} = (1+\theta)^2\frac{\gamma\hat{z}_i}{a}\,\hat{\pi}_i^D + \hat{N}_i(1+\theta)^2\gamma\left[\Omega - 2\hat{z}_i(1+\hat{t}_i)\right] = 0. \tag{31}$$

In (31), the positive first term is similar to our previous analysis [eq. (11)], as the migration response of the manager depends only on this agent's overconfident expectations. However, the negative second term is now also multiplied by the factor $(1+\theta)^2$, as shareholders expect to make higher bonus payments, relative to the case where their expectations are unbiased.

The effects of a higher fixed wage on firm profits are the same as in the benchmark model [eq. (12)]. We assume again that $\hat{w}_i > 0$ holds in the firm's optimum so that $\hat{\pi}_i^D/a = \hat{N}_i$. The optimal bonus payment is then

$$\hat{z}_i^* = \frac{\Omega}{1+2\hat{t}_i} \quad \forall\, i. \tag{32}$$

Eq. (32) is independent of the parameter θ, in contrast to the case where only managers are overconfident [eq. (13)]. Hence, shareholders no longer make an inflated bonus payment to managers. The reason for this is straightforward. Since shareholders overestimate the probability of the good state h in the same way as managers do, they

[20] See the surveys in Baker et. al. (2007), and Malmendier and Tate (2015, pp. 55-57 and Figure 2).

also expect to pay the bonus more often. Therefore, shareholders can no longer exploit the manager's biased beliefs by offering a compensation package with a higher bonus component. In this sense, the overconfidence of the manager and the overoptimism of the firms' shareholders offset each other in determining the optimal bonus payment.

In Appendix C.1, we use the equilibrium bonus payment (32) to derive the optimal fixed wage. This is given by

$$\hat{w}_i^* = Y^l - a\bar{N} + \frac{\gamma(1+\theta)^2\Omega^2}{6}\left[\frac{1}{(1+2\hat{t}_j)} + \frac{4\hat{t}_i - 1}{(1+2\hat{t}_i)^2}\right]. \tag{33}$$

We show in Appendix C.1 that $\hat{w}_i^*$ in (33) exceeds w_i^* in (14) for any $\theta > 0$. This is because shareholders overestimate the additional expected profit derived from opening an extra division, and therefore compete more aggressively for mobile managers by means of the fixed wage. In sum, the compensation pattern for managers now features a higher fixed wage, but a lower bonus payment, in comparison to the case where only managers are overconfident.

In Stage 1, the two symmetric governments choose their optimal bonus taxes $\hat{t}_i$. Governments continue to form unbiased expectations about the probabilities of the different states and maximize revenues as given in (15). Using (32) and (33) gives

$$\frac{\partial \hat{W}_i}{\partial \hat{t}_i} = \frac{\hat{z}_i^2\gamma(1+\theta)}{a}\left[a\hat{N}_i\left\{1 - \frac{4\hat{t}_i}{1+2\hat{t}_i}\right\} - \frac{\hat{t}_i\gamma(1+\theta)^2\Omega^2}{3(1+2\hat{t}_i)^2}\right], \tag{34}$$

which is similar in structure to eq. (17), and the interpretation given there applies.[21] Note, however, that the migration response [the negative second term in the squared bracket of (34)] is less strong now (i.e., less negative in response to a rise in $\hat{t}_i$), as compared to the case where only managers are overconfident [eq. (17)].[22] This is because the bonus component in executive compensation is now lower, and hence a bonus tax increase affects mobile managers less.

The unique Nash equilibrium tax rate in both countries can again be expressed in closed form:

$$\hat{t}_i^* = -\frac{\gamma\,\Omega^2\,(1+\theta)^2}{24\bar{N}a} + \sqrt{\left(\frac{\gamma\Omega^2\,(1+\theta)^2}{24\bar{N}a}\right)^2 + \frac{1}{4}} > 0. \tag{35}$$

In Appendix C.4, we show that the optimal bonus tax rate in (35) is higher as compared to the case where shareholders behave rationally [eq. (18)]. This is a direct consequence

[21] Appendix C.2 proves that the second-order condition for the optimal tax rate is met.

[22] See Appendix C.3 for the formal proof. Note that the migration response is still stronger than in the case where all agents behave rationally.

of the less elastic response of manager migration in response to the bonus tax. In sum, tax competition for mobile and overconfident managers becomes less aggressive when shareholders are equally overconfident, and therefore reduce the bonus component in managerial pay.

However, the joint overconfidence of managers and shareholders still depresses equilibrium tax rates, relative to a scenario where all agents are rational. This is seen by differentiating $\hat{t}_i^*$ with respect to θ:

$$\frac{\partial \hat{t}_i^*}{\partial \theta} = \frac{1}{24\bar{N}a}\left(-2\Omega^2\gamma(1+\theta) + \frac{2\gamma^2\Omega^4(1+\theta)^3}{\sqrt{\gamma^2\Omega^4(1+\theta)^4 + 144a^2\bar{N}^2}}\right) < 0, \tag{36}$$

which is unambiguously negative. Hence, the result that overconfidence of mobile managers reduces optimal bonus taxes stays intact in a setting where shareholders are also overconfident. We summarize our results in this section in:

Proposition 4 *When shareholders and mobile managers are equally overconfident, the following holds:*
(*i*) *the bonus component in managerial pay is independent of investors' and managers' joint overconfidence level θ;*
(*ii*) *equilibrium bonus taxes are lower than in the case where all agents behave rationally, but are higher than in the case where only managers are overconfident.*

Our analysis in this section thus shows, somewhat counterintuitively, that adding investor overconfidence does not necessarily reinforce the effects of managerial overconfidence, but may even partly offset them. The critical effect in our model is that the overconfidence of shareholders will *reduce* bonus payments to managers, other things being equal, because overconfident shareholders can no longer exploit the biased expectations of managers. This dampens, but it does not fully neutralize, the effect that the overconfidence of mobile managers has on tax competition between countries.

6 Conclusion

Our analysis has shown that overconfidence as a behavioral trait of high-skilled managers can contribute to explaining several important developments that have been observed over the past decades. On the one hand, it offers an additional rationalization for the increased use of incentive pay contracts in most OECD countries. Higher shares of incentive pay are in turn an obvious driver of wage income inequality. On the other hand, overconfidence in connection with incentive pay provides a mechanism that increases the migration elasticity of the mobile highly skilled. This adds to the explanation for the rather high migration elasticities that have been found empirically for this segment of the labor market (Kleven et al., 2020). The high migration elasticities in turn offer a rationale for the fall in top income tax rates that has been observed since the mid-1990s (Egger et al., 2019).

Our results have been derived in a simplified setting where general wage income remained either untaxed, or it was taxed at an exogenously given tax rate. A first extension of our analysis would therefore be to study the effects of overconfidence under a more general system of optimal non-linear income taxation. A further extension would be to study optimal contracts by firms when only some part of their managers are overconfident, and this is private information. We leave these and other extensions to future research.

Appendix

A. Appendix to Section 3

A.1 Derivation of equation (14)

Firm i chooses the bonus z_i and the fixed wage w_i, which depends on z_j and w_j. Hence the system of first-order conditions in (12) is interdependent, and given by

$$\begin{aligned}\frac{\partial \Pi_i}{\partial w_i} &= \frac{1}{a}\left\{p_i^{h*}[Y^h - z_i(1+t_i)] + p_i^{l*}(Y^l - \kappa\theta) - w_i\right\} \\ &\quad - \bar{N} - \frac{1}{a}\left[\frac{\gamma}{2}(1+\theta)^2(z_i^2 - z_j^2) + w_i - w_j\right] = 0, \qquad \text{(A.1)} \\ \frac{\partial \Pi_j}{\partial w_j} &= \frac{1}{a}\left\{p_j^{h*}[Y^h - z_j(1+t_j)] + p_i^{l*}(Y^l - \kappa\theta) - w_j\right\} \\ &\quad - \bar{N} - \frac{1}{a}\left[\frac{\gamma}{2}(1+\theta)^2(z_j^2 - z_i^2) + w_j - w_i\right] = 0. \qquad \text{(A.2)}\end{aligned}$$

Substituting in the equilibrium bonuses z_i and z_j from (13) and the equilibrium probabilities from (5a)–(5b) yields

$$\begin{aligned} w_i &= \begin{pmatrix} \frac{2}{3} & \frac{1}{3} \end{pmatrix} \begin{pmatrix} p_i^{h*}(\Omega - z_i(1+t_i)) + Y^l - \kappa\theta - a\bar{N} - \frac{\gamma}{2}(1+\theta)^2(z_i^2 - z_j^2) \\ p_j^{h*}(\Omega - z_j(1+t_j)) + Y^l - \kappa\theta - a\bar{N} - \frac{\gamma}{2}(1+\theta)^2(z_j^2 - z_i^2) \end{pmatrix} \\ &= Y^l - \kappa\theta - a\bar{N} + \frac{1}{3}\frac{\gamma(1+\theta)\Omega^2}{(1+2t_j-\theta)}\left(1 - \frac{1+t_j}{1+2t_j-\theta}\right) + \frac{2}{3}\frac{\gamma(1+\theta)\Omega^2}{(1+2t_i-\theta)}\left(1 - \frac{1+t_i}{1+2t_i-\theta}\right) \\ &\quad + \frac{\gamma(1+\theta)^2\Omega^2}{6}\left(\frac{1}{(1+2t_j-\theta)^2} - \frac{1}{(1+2t_i-\theta)^2}\right). \qquad \text{(A.3)}\end{aligned}$$

Simplifying (A.3) leads to eq. (14) in the main text.

A2. Second-order condition for optimal bonus taxes

Differentiating the first-order condition for bonus taxes (17) with respect to t_i gives

$$\begin{aligned}\frac{\partial^2 W_i}{\partial t_i^2} &= \frac{\partial z_i^2}{\partial t_i}\frac{\gamma(1+\theta)}{a}\left[aN_i\left\{1 - \frac{4t_i}{1-\theta+2t_i}\right\} - \frac{t_i\gamma(1+\theta)\Omega^2}{3(1-\theta+2t_i)^2}\right] \\ &\quad + \frac{z_i^2\gamma(1+\theta)}{a}\Big[-a\frac{\gamma(1+\theta)\Omega^2}{3a(1-\theta+2t_i)^2}\left\{1 - t_i\frac{4}{1-\theta+2t_i}\right\} - aN_i\frac{4(1-\theta)}{(1-\theta+2t_i)^2} \\ &\quad - \frac{\gamma(1+\theta)\Omega^2}{3(1-\theta+2t_i)^2} + \frac{4t_i\gamma(1+\theta)\Omega^2}{3(1-\theta+2t_i)^3}\Big]. \qquad \text{(A.4)}\end{aligned}$$

At $t_i = t_i^*$ [eq. (18)], the first term on the right-hand side is zero. The remaining terms are rearranged to yield:

$$\frac{\partial^2 W_i}{\partial t_i^2} = \frac{z_i^2 \gamma(1+\theta)}{a}\Big[- aN_i \frac{4(1-\theta)}{(1-\theta+2t_i)^2} + \frac{\Omega^2(-1+2t_i+\theta)2\gamma(1+\theta)}{3(1-\theta+2t_i)^3}\Big]. \quad \text{(A.5)}$$

A sufficient condition for (A.5) to be negative is $-1+\theta+2t_i < 0$. Substituting the equilibrium tax rate from (18) gives

$$-1+\theta+2t_i = -\left[\frac{\gamma\,\Omega^2(1+\theta)}{12\bar{N}a} + (1-\theta)\right] + \sqrt{\left(\frac{\gamma\Omega^2\,(1+\theta)}{12\bar{N}a}\right)^2 + \,(1-\theta)^2} < 0, \text{(A.6)}$$

which demonstrates that the second-order condition for an optimum is fullfilled. □.

A3. Derivation of eqs. (24) and (25)

All derivatives include $\partial t_i^*/\partial\theta$. Evaluating this at the equilibrium tax rate $t_i = t_i^*$ yields

$$\left.\frac{\partial t_i^*}{\partial \theta}\right|_{t_i=t_i^*} = \frac{-t_i[2(1+\theta)(1-\theta)+(1-\theta)^2-4t_i^2]}{(1+\theta)[(1-\theta)^2+4t_i^2]}\,. \quad \text{(A.7)}$$

From (5a) and (13), tax revenue is $W_i^* = p_i^{h*} t_i z_i N_i = t_i(1+\theta)\gamma z_i^2 = \dfrac{t_i\,(1+\theta)\,\gamma\Omega^2}{(1-\theta+2\,t_i)^2} N_i$. Differentiating and dividing by W_i^* gives

$$\frac{1}{W_i^*}\frac{\partial W_i^*}{\partial\theta} = \frac{3+\theta+2t_i}{(1+\theta)(1-\theta+2t_i)} + \frac{1-\theta-2\,t_i}{t_i(1-\theta+2t_i)}\frac{\partial t_i^*}{\partial\theta}\,. \quad \text{(A.8)}$$

Substituting (A.7) in (A.8) and simplifying terms gives

$$\frac{1}{W_i^*}\frac{\partial W_i^*}{\partial\theta} = \frac{8t_i}{(1+\theta)[(1-\theta)^2+4t_i^2]} > 0. \quad \text{(A.9)}$$

Multiplying (A.9) by W_i^* gives (24) in the main text.

Total welfare in the economy is, using (21):

$$W_i^* + N_i u_i^* = \left(Y^l - a\bar{N} + \frac{\gamma(1+\theta)\Omega^2[1+2t_i-3\theta]}{2(1-\theta+2t_i)^2} + \frac{t_i\,(1+\theta)\,\gamma\Omega^2}{(1-\theta+2\,t_i)^2}\right) N_i. \quad \text{(A.10)}$$

Differentiating by θ, taking account of $N_i = \bar{N}$ in the symmetric equilibrium and using (A.7) gives,

$$\begin{aligned}\frac{1}{\gamma\Omega^2\bar{N}}\frac{\partial(W_i^*+u_i^*N_i)}{\partial\theta} &= \frac{4(1+t_i)(t_i-\theta)}{(1+\theta)(1-\theta+2t_i)^3} - \frac{4(1+\theta)(t_i-\theta)}{(1-\theta+2t_i)^3}\frac{\partial t_i^*}{\partial\theta} \\ &= \frac{4(t_i-\theta)}{(1-\theta+2t_i)[(1-\theta)^2+4t_i^2]}\,. \end{aligned} \quad \text{(A.11)}$$

Evaluating (A.11) at t_i^* using (18) yields (25) in the main text.

A4. Loss in the low state caused by overconfidence

We extend our benchmark model by assuming that, in the low state, the return to the firm is given by $Y^l - \kappa\theta$, where $\kappa \geq 0$. The firm's after-tax expected profits are then

$$\Pi_i = N_i \pi_i^D = N_i \left\{ p_i^{h*}[Y^h - z_i(1+t_i)] + p_i^{l*}(Y^l - \kappa\theta) - w_i \right\}. \tag{A.12}$$

Differentiating with respect to z_i and assuming that (12) holds with equality yields

$$z_i^* = \frac{\Omega^+}{1-\theta+2t_i} \quad \forall\, i, \qquad \Omega^+ \equiv Y^h - (Y^l - \kappa\theta), \tag{A.13}$$

which is *higher* than the optimal bonus in the benchmark model [eq. (13)]. The loss that overconfidence creates in the low states adds an extra reason to firms to incentivize managers to increase their effort in order to reach the high state.

The structure of the optimal tax rate in (18) is unchanged by this extension, but Ω is replaced by Ω^+ in (A.13). Therefore, the derivative with respect to θ changes to

$$\frac{\partial t_i^*}{\partial \theta} = \frac{1}{24\bar{N}a}\left\{ -\Omega^+\gamma[\Omega^+ + 2(1+\theta)\kappa] + \frac{\gamma^2(\Omega^+)^3(1+\theta)[\Omega^+ + 2(1+\theta)\kappa] - 144(1-\theta)a^2\bar{N}^2}{\sqrt{\gamma^2(\Omega^+)^4(1+\theta)^2 + 144(1-\theta)^2a^2\bar{N}^2}} \right\} < 0. \tag{A.14}$$

The sum of the terms added by the parameter κ in (A.14) is negative. Hence, a given increase in overconfidence lowers the equilibrium tax rate by more when overconfidence causes a loss to the firm in the low state. This is a direct consequence of the higher equilibrium bonus that results from this extension.

In a symmetric tax equilibrium, the manager's fixed wage is now given by

$$w_i^* = (Y^l - \kappa\theta) - a\bar{N} + \frac{\gamma(1+\theta)(Y^h - Y^l + \kappa\theta)^2(t_i - \theta)}{(1+2t_i-\theta)^2}, \tag{A.15}$$

so that

$$\frac{\partial w_i^*}{\partial \theta} = \gamma(\Omega^+)^2\,\frac{(2t_i - 3\theta - 1)}{(1+2t_i-\theta)^3}\left[1 + t_i - (1+\theta)\frac{\partial t_i^*}{\partial \theta}\right] + \frac{\kappa[-(1+2t_i-\theta) + p^{h*}(2t_i - 2\theta)]}{(1+2t_i-\theta)} < 0. \tag{A.16}$$

The first term in (A.16) is negative from (A.6) [cf. eq. (20) in the main text]. The additional effects arising from the loss in the low state, appearing in the negative second term, reinforce the reduction of the fixed wage.

The managers' *actual* utility u_i^*, evaluated at the true success probability p^{h*} is

$$u_i^* = Y^l - \kappa\theta - a\bar{N} + \frac{\gamma(1+\theta)(\Omega^+)^2(1+2t_i-3\theta)}{2(1-\theta+2t_i)^2}. \tag{A.17}$$

Differentiating with respect to θ and taking account of the endogeneity of t_i gives

$$\begin{aligned}\frac{\partial u_i^*}{\partial \theta} &= \frac{\gamma(\Omega^+)^2(t_i + 2t_i^2 - 5\theta t_i - 4\theta)}{(1-\theta+2t_i)^3} - \frac{\gamma(\Omega^+)^2(1+\theta)(1+2t_i-5\theta)}{(1-\theta+2t_i)^3}\frac{\partial t_i^*}{\partial \theta} \\ &+ \frac{\kappa[-(1+2t_i-\theta) + p^{h*}(1+2t_i-3\theta)]}{(1-\theta+2t_i)}. \end{aligned} \tag{A.18}$$

The first two terms in (A.18) are analogous to (22) in the main text. The additional effects arising from the loss in the low state are collected in the negative third term, with κ as a multiplier. Clearly, if κ is sufficiently large, then (A.18) is negative for any level of θ.

B. Appendix to Section 4

B.1 Derivation of the fixed wage

Differentiating (26) with respect to fixed wages gives

$$\begin{aligned}\frac{\partial \tilde{\Pi}_i}{\partial \tilde{w}_i} &= \frac{1}{a}\left\{\tilde{p}_i^{h*}[Y^h - \tilde{z}_i(1+\tilde{t}_i+\tau_i)] + \tilde{p}_i^{l*}Y^l - \tilde{w}_i(1+\tau_i)\right\} \\ &- \bar{N}(1+\tau_i) - \frac{1}{a}\left[\frac{\gamma}{2}(1+\theta)^2(\tilde{z}_i^2 - \tilde{z}_j^2) + \tilde{w}_i - \tilde{w}_j\right](1+\tau_i) = 0, \end{aligned} \tag{B.1}$$

$$\begin{aligned}\frac{\partial \tilde{\Pi}_j}{\partial \tilde{w}_j} &= \frac{1}{a}\left\{\tilde{p}_j^{h*}[Y^h - \tilde{z}_j(1+\tilde{t}_j+\tau_j)] + \tilde{p}_i^{l*}Y^l - \tilde{w}_j(1+\tau_j)\right\} \\ &- \bar{N}(1+\tau_j) - \frac{1}{a}\left[\frac{\gamma}{2}(1+\theta)^2(\tilde{z}_j^2 - \tilde{z}_i^2) + \tilde{w}_j - \tilde{w}_i\right](1+\tau_j) = 0, \end{aligned} \tag{B.2}$$

where the only difference to (A.1)–(A.2) lies in the tax factors $(1+\tau_i)$ and $(1+\tau_j)$. Substituting in equilibrium bonuses from (27) and equilibrium probabilities from (5a)–(5b) and simplifying leads to

$$\begin{aligned}\tilde{w}_i^* &= \left[\frac{2}{3(1+\tau_i)} + \frac{1}{3(1+\tau_j)}\right]Y^l - a\bar{N} \\ &+ \frac{\gamma(1+\theta)\Omega^2}{6}\left[\frac{1}{(1+\tau_j)[(1-\theta)(1+\tau_j)+2\tilde{t}_j]} + \frac{4\tilde{t}_i - 5\theta(1+\tau_i) - (1+\tau_i)}{(1+\tau_i)((1-\theta)(1+\tau_i)+2\tilde{t}_i)^2}\right]. \end{aligned} \tag{B.3}$$

B.2 Derivation of optimal bonus tax

Differentiating (9) with respect to t_i, taking account of (27) and (B.3), leads to

$$\frac{\partial \tilde{N}_i}{\partial \tilde{t}_i} = \frac{1}{a}\left[\frac{\gamma(1+\theta)^2}{2}\frac{\partial(\tilde{z}_i^2)}{\partial \tilde{t}_i} + \frac{\partial \tilde{w}_i}{\partial \tilde{t}_i} - \frac{\partial \tilde{w}_j}{\partial \tilde{t}_i}\right] = \frac{-\gamma(1+\theta)\Omega^2}{3a(1+\tau_i)[(1-\theta)(1+\tau_i)+2\tilde{t}_i]^2} < 0. \tag{B.4}$$

Differentiating (28) with respect to t_i^*, using (B.4) and (B.3), gives

$$\begin{aligned}\frac{\partial \tilde{W}_i(\tilde{t}_i,\tilde{t}_j)}{\partial \tilde{t}_i} &= \tilde{N}_i\left[\tilde{z}_i^2\gamma(1+\theta) - \frac{4(1+\theta)\gamma\tilde{z}_i^2(\tilde{t}_i+\tau_i)}{(1-\theta)(1+\tau_i)+2\tilde{t}_i} + \frac{4\tau_i\gamma(1+\theta)\Omega^2(2\tau_i(1+\theta)+1+\tau_i-\tilde{t}_i)}{3(1+\tau_i)[(1-\theta)(1+\tau_i)+2\tilde{t}_i]^3}\right] \\ &- \frac{\gamma(1+\theta)\Omega^2[(1+\theta)\gamma\tilde{z}_i^2(\tilde{t}_i+\tau_i)+\tilde{w}_i\tau_i]}{3a(1+\tau_i)[(1-\theta)(1+\tau_i)+2\tilde{t}_i]^2} = 0 \quad \forall\, i \neq j. \qquad \text{(B.5)}\end{aligned}$$

We evaluate the above formula at the symmetric equilibrium where $\tilde{t}_i = \tilde{t}_j$ and $\tau_i = \tau_j$. The fixed wage (B.3) is then

$$\tilde{w}_i^* = \frac{Y^l}{1+\tau_i} - a\bar{N} + \frac{\gamma(1+\theta)\Omega^2[\tilde{t}_i - (1+\tau_i)\theta]}{(1+\tau_i)[(1-\theta)(1+\tau_i)+2\tilde{t}_i]^2}\ . \qquad \text{(B.6)}$$

Using this in (B.5) and dividing by $\bar{N}\gamma(1+\theta)\Omega^2$ gives

$$\begin{aligned}\frac{\partial \tilde{W}_i(\tilde{t}_i,\tilde{t}_j)}{\partial \tilde{t}_i}\frac{1}{\bar{N}\gamma(1+\theta)\Omega^2} &= \frac{(1-\theta)(1+\tau_i)-2\tilde{t}_i-4\tau_i}{[(1-\theta)(1+\tau_i)+2\tilde{t}_i]^3} + \frac{4\tau_i[2\theta(1+\tau_i)+1+\tau_i-\tilde{t}_i]}{3(1+\tau_i)[(1-\theta)(1+\tau_i)+2\tilde{t}_i]^3} \\ -\left[\frac{(1+\tau_i)(\tilde{t}_i+\tau_i)+\tau_i[\tilde{t}_i-(1+\tau_i)\theta]}{(1+\tau_i)[(1-\theta)(1+\tau_i)+2\tilde{t}_i]^2} + \frac{\tau_i[\frac{Y^l}{1+\tau_i}-a\bar{N}]}{\gamma(1+\theta)\Omega^2}\right] &\left[\frac{\gamma(1+\theta)\Omega^2/\bar{N}}{3a(1+\tau_i)[(1-\theta)(1+\tau_i)+2\tilde{t}_i]^2}\right] = 0. \qquad \text{(B.7)}\end{aligned}$$

Multiplying eq. (B.7) by $3[(1-\theta)(1+\tau_i)+2\tilde{t}_i]^4(1+\tau_i)^2$ and rearranging terms yields a quadratic equation for $\tilde{t}_i^*$ given by

$$-\frac{A}{2}\tilde{t}_i^2 + B\tilde{t}_i - \frac{C}{2} = 0, \qquad \text{(B.8)}$$

where the terms $A - C$ are in eq. (29) in the main text. The solution to (B.8) is then given in (29).

B.3 Proof of Proposition 3

Denote $B = B(\theta)$ and $C = C(\theta)$ as functions of θ. The optimal bonus tax $\tilde{t}_i^*$ in (29) is decreasing in θ iff:

$$B'(\theta) + \frac{2B(\theta)B'(\theta) - AC'(\theta)}{2\sqrt{(B(\theta))^2 - AC(\theta)}} < 0. \qquad \text{(B.9)}$$

For $C(\theta)$, the following can be readily confirmed:

$$C'(\theta)\frac{1-\theta}{2} = -C(\theta) + 2\tau_i(1+\tau_i)(1-\theta)\gamma\Omega_a^2 > 0. \qquad \text{(B.10)}$$

Therefore, if $B'(\theta) < 0$ (which holds, for example, if $\tau_i = 0$), then condition (B.9) is fulfilled immediately.

If $B'(\theta) > 0$, condition (B.9) holds iff $B'(\theta)[\sqrt{(B(\theta))^2 - AC(\theta)} + B(\theta)] < AC'(\theta)/2$. Using (B.10) and rearranging, this is equivalent to:

$$-(B'(\theta))^2 2\tau_i(1+\tau_i)(1-\theta)\gamma\Omega_a^2 < \frac{A}{4}(C'(\theta))^2 - C'(\theta)B'(\theta)\left[B(\theta) + \frac{1-\theta}{2}B'(\theta)\right] \quad \text{(B.11)}$$

Since $B(\theta)$ is a linear function of θ, we have $B(\theta) = (1-\theta)B(0) + \theta B(1)$ and $B'(\theta) = B(1) - B(0)$. Hence

$$B(\theta) + \frac{1-\theta}{2}B'(\theta) = \frac{1-\theta}{2}B(0) + \frac{1+\theta}{2}B(1) < 0 \text{ for all } \theta \in (0,1),$$

since $B(0) < 0$ and $B(1) < 0$. Therefore, the LHS of (B.11) is negative and the RHS of (B.11) is positive. Hence, condition (B.9) also holds for $B'(\theta) > 0$ and $\partial\tilde{t}_i^*/\partial\theta < 0$ is true for all levels of $B(\theta)$.

The equilibrium bonus payment is immediately seen to rise in θ from eq. (27). Also, evaluating the fixed wage expression (B.3) at the common Nash equilibrium tax rate $\tilde{t}_i^*$ in (29) gives

$$\frac{\partial\tilde{w}_i^*}{\partial\theta} = \gamma\Omega^2 \frac{(2\tilde{t}_i - 3\theta - 1 - 3\tau_i\theta - \tau_i)}{((1-\theta)(1+\tau_i) + 2\tilde{t}_i)^3}\left[1 + \tilde{t}_i + \tau_i - (1+\theta)\frac{\partial\tilde{t}_i^*}{\partial\theta}\right], \quad \text{(B.12)}$$

which is negative since $\tilde{t}_i^*(\tau_i,\theta) \leq \tilde{t}_i^*(0,\theta) = t_i^*(\theta)$. Finally, (B.4) shows that the outflow of managers in response to a higher bonus tax rises in θ. □

C. Appendix to Section 5

C.1 Derivation of equation (33)

We proceed in the same way as in Appendix A.1. The interdependent system of first-order conditions in (12) is now given by

$$\begin{aligned}\frac{\partial\hat{\Pi}_i}{\partial\hat{w}_i} &= \frac{1}{a}\left\{\hat{p}_i^{h*}[Y^h - \hat{z}_i(1+\hat{t}_i)] + \hat{p}_i^{l*}Y^l - \hat{w}_i\right\} \\ &\quad - \bar{N} - \frac{1}{a}\left[\frac{\gamma}{2}(1+\theta)^2(\hat{z}_i^2 - \hat{z}_j^2) + \hat{w}_i - \hat{w}_j\right] = 0, \quad \text{(C.1)} \\ \frac{\partial\hat{\Pi}_j}{\partial\hat{w}_j} &= \frac{1}{a}\left\{\hat{p}_j^{h*}[Y^h - \hat{z}_j(1+\hat{t}_j)] + \hat{p}_i^{l*}Y^l - \hat{w}_j\right\} \\ &\quad - \bar{N} - \frac{1}{a}\left[\frac{\gamma}{2}(1+\theta)^2(\hat{z}_j^2 - \hat{z}_i^2) + \hat{w}_j - \hat{w}_i\right] = 0, \quad \text{(C.2)}\end{aligned}$$

where the only difference to (A.1)–(A.2) lies in the expected success probability $\hat{p}_i^h = (1+\theta)p_i^h$. Substituting in the equilibrium bonuses $\hat{z}_i$ and $\hat{z}_j$ from (32) and the equilibrium probabilities from (5a)–(5b) yields

$$\begin{aligned}
\hat{w}_i &= \begin{pmatrix} \frac{2}{3} & \frac{1}{3} \end{pmatrix} \begin{pmatrix} \gamma(1+\theta)^2\hat{z}_i(\Omega - \hat{z}_i(1+\hat{t}_i)) + Y^l - a\bar{N} - \frac{\gamma}{2}(1+\theta)^2(\hat{z}_i^2 - \hat{z}_j^2) \\ \gamma(1+\theta)^2\hat{z}_j(\Omega - \hat{z}_j(1+\hat{t}_j)) + Y^l - a\bar{N} - \frac{\gamma}{2}(1+\theta)^2(\hat{z}_j^2 - \hat{z}_i^2) \end{pmatrix} \\
&= Y^l - a\bar{N} + \frac{1}{3}\frac{\gamma(1+\theta)^2\Omega^2}{(1+2\hat{t}_j)}\left(1 - \frac{1+\hat{t}_j}{1+2\hat{t}_j}\right) + \frac{2}{3}\frac{\gamma(1+\theta)^2\Omega^2}{(1+2\hat{t}_i)}\left(1 - \frac{1+\hat{t}_i}{1+2\hat{t}_i}\right) \\
&+ \frac{\gamma(1+\theta)^2\Omega^2}{6}\left[\frac{1}{(1+2\hat{t}_j)^2} - \frac{1}{(1+2\hat{t}_i)^2}\right]. \qquad \text{(C.3)}
\end{aligned}$$

Simplifying (C.3) leads to eq. (33) in the main text.

Since $\hat{t}_i^* < 1/2$ follows from (35), we get

$$\begin{aligned}
\frac{\partial \hat{w}_i^*}{\partial \theta} &= \frac{\gamma\Omega^2}{6}2(1+\theta)\left[\frac{1}{(1+2\hat{t}_i)} + \frac{4\hat{t}_i - 1}{(1+2\hat{t}_i)^2}\right] + \frac{\gamma\Omega^2}{6}(1+\theta)^2\frac{\partial}{\partial\theta}\left[\frac{1}{(1+2\hat{t}_i)} + \frac{4\hat{t}_i - 1}{(1+2\hat{t}_i)^2}\right] \\
&= \frac{\gamma\Omega^2}{6}\frac{(1+\theta)^2(1-2\hat{t}_i)}{(1+2\hat{t}_i)^3}\left[\frac{\Delta - \Gamma}{(1+\theta)\left(\Gamma + 12\bar{N}a - \Delta\right)}\right]\left[\frac{\Delta - \Gamma + 12\bar{N}a}{\Delta}\right] > 0 \qquad \text{(C.4)}
\end{aligned}$$

$$\text{where} \qquad \Gamma = \gamma\,\Omega^2\,(1+\theta)^2 \qquad \text{and} \qquad \Delta \equiv \sqrt{\gamma^2\Omega^4\,(1+\theta)^4 + 144\,\bar{N}^2a^2}.$$

Since w_i^* in (14) and $\hat{w}_i^*$ in (33) are equal for $\theta = 0$, and since w_i^* is falling in θ from (20), but $\hat{w}_i^*$ is rising in θ from (C.4), it must be true that $\hat{w}_i^* > w_i^*$ holds for all $\theta > 0$.

C.2 Second-order condition of (34)

Differentiating the first-order condition for bonus taxes (34) with respect to t_i gives

$$\begin{aligned}
\frac{\partial^2 \hat{W}_i}{\partial \hat{t}_i^2} &= \frac{\partial \hat{z}_i^2}{\partial \hat{t}_i}\frac{\gamma(1+\theta)}{a}\left[aN_i\left\{1 - \frac{4\hat{t}_i}{1+2\hat{t}_i}\right\} - \frac{\hat{t}_i\gamma(1+\theta)^2\Omega^2}{3(1+2\hat{t}_i)^2}\right] \\
&+ \frac{\hat{z}_i^2\gamma(1+\theta)}{a}\Big[-a\frac{\gamma(1+\theta)^2\Omega^2}{3a(1+2\hat{t}_i)^2}\left\{1 - \hat{t}_i\frac{4}{1+2\hat{t}_i}\right\} - aN_i\frac{4}{(1+2\hat{t}_i)^2} \\
&- \frac{\gamma(1+\theta)^2\Omega^2}{3(1+2\hat{t}_i)^2} + \frac{4\hat{t}_i\gamma(1+\theta)^2\Omega^2}{3(1+2\hat{t}_i)^3}\Big]. \qquad \text{(C.5)}
\end{aligned}$$

At $\hat{t}_i = \hat{t}_i^*$ [eq. (35)], the first term on the RHS is zero. Rearranging then yields:

$$\frac{\partial^2 \hat{W}_i}{\partial \hat{t}_i^2} = \frac{\hat{z}_i^2\gamma(1+\theta)}{a}\Big[-aN_i\frac{4}{(1+2\hat{t}_i)^2} + \frac{\Omega^2(-1+2\hat{t}_i)2\gamma(1+\theta)}{3(1+2\hat{t}_i)^3}\Big], \qquad \text{(C.6)}$$

which is negative since $-1 + 2\hat{t}_i < 0$. □

C.3 Comparison of migration response

Differentiating (9) with respect to t_i and taking account of (32) and (33) leads to

$$\frac{\partial \hat{N}_i}{\partial \hat{t}_i} = \frac{1}{a}\left[\frac{\gamma(1+\theta)^2}{2}\frac{\partial(\hat{z}_i^2)}{\partial \hat{t}_i} + \frac{\partial \hat{w}_i}{\partial \hat{t}_i} - \frac{\partial \hat{w}_j}{\partial \hat{t}_i}\right] = \frac{-\gamma(1+\theta)^2\Omega^2}{3a(1+2\hat{t}_i)^2} < 0. \tag{C.7}$$

From (16) and (C.7) we have:

$$\frac{\partial \hat{N}_i(\hat{t}_i)}{\partial \hat{t}_i} - \frac{\partial N_i(\hat{t}_i)}{\partial t_i} = -\frac{\gamma\Omega^2(1+\theta)\theta(\theta^2-\theta+4\hat{t}_i^2-4\hat{t}_i\theta-1)}{3a(1+2\hat{t}_i)^2(1+2\hat{t}_i-\theta)^2}.$$

For the range of $\theta \in (0,1)$ we have $\theta^2-\theta<0$. For the relevant range of equilibrium tax rates $0<\hat{t}_i^*<1/2$ [see eq. (35)], we have $4\hat{t}_i^2-4\hat{t}_i\theta-1<4\hat{t}_i^2-1<0$. Now,

$$\frac{\partial \hat{N}_i(\hat{t}_i^*)}{\partial \hat{t}_i} - \frac{\partial N_i(t_i^*)}{\partial t_i} = \left(\frac{\partial \hat{N}_i(\hat{t}_i^*)}{\partial \hat{t}_i} - \frac{\partial N_i(\hat{t}_i^*)}{\partial t_i}\right) + \left(\frac{\partial N_i(\hat{t}_i^*)}{\partial \hat{t}_i} - \frac{\partial N_i(t_i^*)}{\partial t_i}\right). \tag{C.8}$$

We show in Appendix C.4 that $\hat{t}_i^* > t_i^*$, and (16) shows that $\partial^2 N_i(t_i)/(\partial t_i)^2 > 0$. Therefore, the sign of (C.8) is positive, implying a less strong (i.e., less negative) migration response in (C.7), as compared to (16). □

C.4 Comparison of equilibrium tax rates

From (18) and (35) we have:

$$\begin{aligned}
&\hat{t}_i^* > t_i^* \iff \\
&-\gamma\Omega^2(1+\theta)\theta + \sqrt{\gamma^2\Omega^4(1+\theta)^4 + 144\bar{N}^2a^2} > \sqrt{\gamma^2\Omega^4(1+\theta)^2 + 144\bar{N}^2a^2(1-\theta)^2} \iff \\
&\gamma^2\Omega^4(1+\theta)^3 + 72\,\bar{N}^2a^2(2-\theta) > \gamma\Omega^2(1+\theta)\sqrt{\gamma^2\Omega^4\,(1+\theta)^4 + 144\,\bar{N}^2a^2} \iff \\
&(72\bar{N}^2a^2(2-\theta))^2 + 144\gamma^2\Omega^4\bar{N}^2a^2(-\theta^2+\theta+1)(1+\theta)^2 > 0.
\end{aligned} \tag{C.9}$$

The last expression of (C.9) is positive for $\theta \in (0,1)$. □

References

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.

Akcigit, U., Baslandze, S., Stantcheva, S. (2016). Taxation and the international mobility of inventors. *American Economic Review 106*, 2930-2981.

Apps, P., Long, N.V., Rees, R. (2014). Optimal piecewise linear income taxation. *Journal of Public Economic Theory 16*, 523-545.

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

Baker, M., Ruback, R., Wurgler, J. (2007). Behavioral corporate finance: A survey. Chapter 4 in: B.E. Eckbo (ed.), Handbook of Corporate Finance: Empirical Corporate Finance, Vol.1. Elsevier/North-Holland, New York.

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

Besley, T., Ghatak, M. (2013). Bailouts and the optimal taxation of bonus pay. *American Economic Review Papers & Proceedings 103*, 163-167.

Bernheim, B., Taubinsky, D. (2018). Behavioral public economics. In: B. Bernheim, S. DellaVigna, D. Laibson (eds.) Handbook of Behavioral Economics: Applications and Foundations 1, Vol. 1, Chap. 5, 381-515.

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

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 Disc. Paper No. 1112. London School of Economics.

De la Rosa, L.E. (2011). Overconfidence and moral hazard. *Games and Economic Behavior 73*, 429-451.

Edmans, A., Gabaix, X. (2016). Executive compensation: A modern primer. *Journal of Economic Literature 54*, 1232-1287.

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

Farhi, E., Gabaix, X. (2020). Optimal taxation with behavioral agents. *American Economic Review 110*, 298-336.

Gerritsen, A. (2016). Optimal taxation when people do not maximize well-being. *Journal of Public Economics 144*, 122-139.

Gervais, S., Heaton, J.B., Odean, T. (2011). Overconfidence, compensation contracts, and capital budgeting. *Journal of Finance 66*, 1735-1777.

Gervais, S., Odean, T. (2001). Learning to be overconfident. *Review of Financial Studies 14*, 1-27.

Gietl, D., Haufler, A. (2018). Bonus taxes and international competition for bank managers. *European Economic Review 110*, 41-60.

Gietl, D., Kassner, B. (2020). Managerial overconfidence and bank bailouts. *Journal of Economic Behavior and Organization*, forthcoming.

Greve, P., Biemann, T., Ruigrok, W. (2015). Foreign executive appointments: A multilevel examination. *Journal of World Business 50*, 674-686.

Ho, P.H., Huang, C.W., Lin, C.Y., Yen, J.F. (2016). CEO overconfidence and financial crisis: Evidence from bank lending and leverage. *Journal of Financial Economics 120*, 194-209.

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.

Kleven, H., Landais, C., Munoz, M., Stantcheva. S. (2020). Taxation and migration: Evidence and policy implications. *Journal of Economic Perspectives 34(2)*, 119-142.

Lehmann, E., Simula, L., Trannoy, A. (2014). Tax me if you can! Optimal nonlinear 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.

Malmendier, U., Tate, G. (2005). CEO overconfidence and corporate investment. *Journal of Finance 60*, 2661-2700.

Malmendier, U., Tate, G. (2015). Behavioral CEOs: The role of managerial overconfidence. *Journal of Economic Perspectives 29*, 37-60.

Moore, D., Healy, P. (2008). The trouble with overconfidence. *Psychological Review 115*, 502-517.

Moore, D., Slemrod, J. (2019). Optimal tax systems with endogenous behavioral biases. SSRN Working Paper. https://ssrn.com/abstract=3401834

Office for National Statistics (2017): Average weekly earnings, bonus payments in Great Britain: 2017. www.ons.gov.uk/employmentandlabourmarket/peopleinwork/earningsandworkinghours/bulletins/averageweeklyearningsbonuspaymentsingreatbritain/2017

Piketty, T., Saez, E., Stantcheva. S. (2014). Optimal taxation of top labor incomes: A tale of three elasticities. *American Economic Journal: Economic Policy 6(1)*, 230-271.

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

Rees-Jones, A., Taubinsky, D. (2020). Measuring "schmeduling". *Review of Economic Studies*, forthcoming.

Schmidheiny, K., Slotwinki, M. (2018). Tax-induced mobility: Evidence from a foreigners' tax scheme in Switzerland. *Journal of Public Economics 167*, 293-324.

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

Slemrod, J., Yitzhaki, S., Mayshar, J., Lundolm, M. (1994). The optimal two-bracket linear income tax. *Journal of Public Economics 53*, 269-290.

Taubinsky, D., Rees-Jones, A. (2018). Attention variation and welfare: Theory and evidence from a tax salience experiment. *Review of Economic Studies 85*, 2462-2496.

Taylor, S.E., Brown, J.D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin 103*, 193-210.

Wilson, J.D. (2015). Competitive income taxation with migration. In: R. Avi-Yonah and J. Slemrod (eds), Taxation and Migration, Chapter 2. Kluwer Law International BV.