

Social Identity, Education and Tax Policy*

Thomas Aronsson, Stefanie Heidrich, and Magnus Wikström

Department of Economics, Umeå School of Business and Economics,
Umeå University, SE – 901 87 Umeå, Sweden

July 2014

Abstract

This paper analyzes the implications of social identity and self-categorization in the context of optimal redistributive income taxation. A two-type model is supplemented by an assumption that individuals select themselves into social categories, in which norms are formed and education effort choices partly depend on these norms. Optimal tax policy is analyzed under two different assumptions about the social objective function: a welfarist objective based on consumer preferences and a paternalist objective that does not reflect the consumer preference for social identity. We show how the welfarist government implements a tax policy to internalize the externalities arising from social norms, while the paternalist government uses tax policy to make individuals behave as if their preferences for social identity were absent.

JEL classification: D03, H21, I21, Z13

Keywords: Optimal income taxation, education, social identity, self-categorization.

* The authors would like to thank David Granlund and Ronald Wendner for helpful comments and suggestions. Research grants from the Swedish Research Council (dnr 421-2010-1420) are also gratefully acknowledged.

1. Introduction

In psychology and sociology, social identity theory has long been used to explain human behavior (see, e.g., Tajfel and Turner, 1979; Hogg, 2006). Social identity can be defined in terms of how a person's sense of self depends on the group (or groups) which the person associates with (e.g., social reference groups such as family, colleagues, friends, social class, etc.). Akerlof and Kranton (2000) show that social identity theory is useful in the context of economics, and might be suitable for analyzing a variety of issues where standard economic theory is more or less silent. Yet, although certain aspects of social identity theory - such as how various forms of social interaction may influence consumer choices - have been analyzed in previous studies, it is fair to say that most of our understanding of how public policy affects economic behavior and welfare originates from models where social identity plays no role at all.

The present paper examines tax policy implications of social identity in the context of educational choices. In economics, education is usually described as an investment that pays off in the future through higher wages. As such, the interesting tradeoff when choosing effort is that between leisure (or consumption) at present and increased productivity in the future.¹ However, if people self-select into social categories where certain types of behavior are desirable, e.g., due to category-specific norms, the incentives underlying effort choices may differ substantially from those that follow from standard economic investment-models. Our study departs from a model of educational choice and social identity presented in Akerlof and Kranton (2002), where study effort depends on such category-specific norms. To be more specific, they introduce an identity component into the individual's utility function, such that individuals differ in terms of how close their preferences or attributes are to the ideal prescribed by different social identity groups. Also, since the identity component depends on the behavior of other members in the same social reference group, externalities play an important role in the economics of identity. In other words, deviating too much from the prescribed behavior may both imply a drop in one's own utility and (positive or negative) changes in the utility of other members in the same social identity group.² The overall purpose

¹ Becker (1975) and Willis and Rosen (1979).

² These externalities also contribute in explaining the peer-group effects on study achievement discussed in the literature on educational outcomes (e.g., Hanushek, 1971; Wolfe, 1977; Evans, Oates and Schwab, 1992; Sacerdote, 2000).

of the present paper is to analyze the implications of social identity and self-categorization for optimal redistributive income taxation. Moreover, since social deployment is not necessarily desirable from society's viewpoint, we also aim at comparing welfarist and paternalist approaches to such policy.

Literature in other areas of economics shows that social norms are important for individual choices as well as for policy outcomes. Social norms (or customs) may be persistent despite that they lead to lower intrinsic utility for individuals if disobedience is associated with lost reputation (Akerlof, 1980), and may even lead individuals to conform in terms of behavior (Bernheim, 1994).³ Based on a political economy model Lindbeck, Nyberg, and Weibull (1999) examine redistributive tax-transfer policies under an employment norm that “one should live off one's own work”, and assume that the perceived cost to the individual of deviating from this norm decreases with the share of benefit recipients in society. Among other things, they find that the economy can end up in a low-tax equilibrium supported by the employed or high-tax equilibrium supported by transfer recipients. Lindbeck, Nyberg, and Weibull (2003) use a similar model to examine the implications of social insurance when the preference for leisure varies among individuals. They show that endogenous social norms may lead voters to choose less generous benefits than otherwise, thereby counteracting the free-rider problem, and also that a temporary unemployment shock may result in a persistent increase in the number of beneficiaries.⁴ Another line of research on social norms and economic behavior refers to interdependent behavior in labor supply choices (e.g., Blomquist, 1993; Aronsson, Blomquist and Sacklen, 1999), showing that norms give rise to feedback effects on labor supply of clear practical relevance for assessing the effects that taxes have on work hours.

However, there is surprisingly little research on the implications of social norms for optimal taxation.⁵ An exception is the study by Aronsson and Sjögren (2010) analyzing optimal redistributive taxation in an economy characterized by two social norms in the labor market: a work hours norm implying that individuals perceive a cost of deviating too much from the

³ See also Bruvold and Nyborg (2004), who analyze a model where norm-adherence is connected to self-image.

⁴ See Lindbeck (1995) for more informal discussions of social norms and economic behavior.

⁵ The importance of social interaction for optimal taxation has been analyzed in other contexts; in particular, in economies where consumers are concerned with their relative consumption; see, e.g., Boskin and Sheshinski (1978), Oswald (1983), Tuomala (1990), Ljunqvist and Uhlig (2000), Dupor and Liu (2003), Aronsson and Johansson-Stenman (2008, 2010, 2014), Wendner and Goulder (2008), and Eckerstorfer and Wendner (2013).

choices made by other people (i.e., interdependent labor supply choices), and a participation norm emanating from the argument that one should earn one's living from work, meaning that they combine the two labor market norms discussed above. Our study differs from theirs in at least three important ways. First, and foremost, we are concerned with the implications of social identity and norms in the context of education choices; not social norms in the labor market. Second, we consider a broader spectrum of social objectives (not just conventional welfare functions that fully reflect individual preferences) by recognizing that policy makers may not necessarily agree with individual preferences for social identity. One reason is that choices of social identity may, to some extent, reflect family characteristics, and policy makers may not want factors correlated with family characteristics to affect outcomes later in life; such influences run counter to the notion of equality of opportunity. This suggests to us that paternalist objectives are particularly interesting to examine in this context. Third, since the education choice is fundamentally intertemporal, we use a two-period model to analyze how tax policy can be used to implement a socially desirable outcome.

Our paper is also related to literature on education and optimal taxation (e.g., Boadway et al., 1996; Bovenberg and Jacobs, 2005; Guo and Krause, 2013; Jacobs, 2013), which deals with a variety of aspects of redistributive education policy. Yet, none of these studies addresses the policy implications of social identity and social norms. Our intention is to bridge this gap by introducing a social identity component into the education choice and then examine the implications for optimal redistributive taxation.

We consider a model with two productivity-types, where individual productivity is private information, along the lines of Stern (1982) and Stiglitz (1982). Such a framework is often used in theoretical literature on optimal redistributive taxation and enables us to integrate corrective and redistributive tax policy in a relatively simple way. Each individual lives for two periods; attains education in the first and earns labor income in the second. This model is here extended to accommodate social identity by allowing each individual, when young, to select into one of two social groups, which differ with respect to the prescribed study effort. As indicated above, an arguably important question is whether the government is welfarist in the traditional sense of accepting that preferences for social identity may influence effort choices, or whether it implements a paternalist policy to make individual behave as if these preferences were absent. We consider both these possibilities by comparing the education

policy of a welfarist government with that of a paternalist government, which does not share the preferences for social identity. Furthermore, the social norm is itself an endogenous variable, and our assumptions about norm formation are based on research in social psychology emphasizing that norms typically reflect more extreme attitudes within groups than just group-specific mean values which, in turn, further contributes to polarization between groups. This will be discussed in greater detail below.

In Section 2, the model of individual choice is described along with the decision-problem faced by the (welfarist and paternalist) government. The optimal tax policy is analyzed in Sections 3 and 4. Section 5 summarizes and concludes.

2. The Model

Consider an economy comprising two productivity-types, l and h , where type l will be referred to as the “low-productivity type” and type h the “high-productivity type”. This is interpretable to mean that type h has higher innate ability than type l . We also distinguish between two social identity groups, which differ with respect to the prescribed effort (through a group-specific effort norm) during the education period of the individual’s life. In the following, we just refer to these groups as the “low-effort” (L) and “high-effort” (H) group, respectively. The pre-tax wage rate facing an individual of productivity-type i in social identity group j depends on both innate ability and study effort and is given by $w_j^i = \theta^i e_j^i$, where θ^i is a measure of innate ability such that $\theta^h > \theta^l$, while e_j^i reflects the effort level during education. There are N^i individuals of productivity-type i , among which $n_L^i \in [0, N^i]$ belongs to social identity group L and $n_H^i = N^i - n_L^i$ to social identity group H .

Although social groups may be characterized along several dimensions, the only distinction we focus on here is study effort. As indicated above, we assume that membership in social identity group L prescribes less study effort than membership in social identity group H , ceteris paribus. The main motivation for this approach is simplicity: study effort is a decision-variable in our model. Our set up is, nevertheless, supported by research in sociology, which

indicates great influence of peers and social groups on educational aspirations.⁶ This is found to be true in general for countries that do not sort pupils into different school tracks at an early age (which is the case for the US and also for the Nordic countries).⁷

We further assume that sorting into these two groups is determined by innate study preference. Akerlof and Kranton (2002) use three different social groups in their description of high school students (leading crowd, nerds, and burnouts). The mechanism leading pupils to sort into different groups, however, is rather ad hoc in their paper. “Looks” for example might in practice not be a reliable predictor for one’s choice of social group and there is little if not no evidence in the literature to be found of such effects. Thus, in our model, we base the sorting mechanism on empirical research that points to strong transmission of educational attainment from parents to their children.⁸ The sorting into identity group L or H may, therefore, reflect that family background affects young individuals’ selection into social identity groups, i.e., the importance attached to study achievement in the individual’s environment, as well as other attributes that the individual would like to be associated with.

Consumers

By ranking the individuals of each ability-type on the basis of preferences for social identity-group, from the person with the highest to the person with the lowest preference for social identity-group L (or, equivalently, from the person with the lowest to the person with the highest preference for social identity-group H), the life-time utility function facing individual $k \in [1, N^i]$ of productivity-type i in social identity group j can be written as

$$U_j^{k,i} = u(c_j^i, x_j^i, e_j^i) + I_j^{k,i} \quad (1)$$

where c denotes consumption when young, x consumption when old, and e denotes education effort. We assume that the function $u(\cdot)$ is increasing in its first and second arguments, decreasing in the third, and strictly quasi-concave. Following Akerlof and Kranton (2002), the second term on the right hand side of equation (1) represents the identity utility component defined as

⁶ See Austin and Draper (1984) and Wentzel and Caldwell (1997).

⁷ See Buchmann and Dalton (2002).

⁸ See, e.g., Ermisch and Francesconi (2001), Anger and Heineck (2010) and Björklund and Jäntti (2012).

$$I_L^{k,i} = I_L - \beta k - \frac{1}{2}(e_L^i - \bar{e}_L)^2 \quad (2a)$$

$$I_H^{k,i} = I_H - \beta(N^i - k) - \frac{1}{2}(e_H^i - \bar{e}_H)^2 \quad (2b)$$

if the individual belongs to social identity group L and H , respectively. The term I_j represents a fixed component of the payoff associated with social identity j . Equations (2a) and (2b) presuppose that individuals of a certain productivity-type differ according to their preferences for social identity, such that individual $k=1$ has the strongest preference for being part of social identity group L and weakest preference for social identity group H , and so on. As such, the parameter β reflects how the payoff differs between individuals. The final component is a perceived cost of deviating from the behavior prescribed by the social group, where \bar{e}_j is interpretable as an identity-group-specific norm for study effort.

We abstract from any initial wealth in what follows. Adding an exogenous initial income or inherited wealth component would not affect any of the qualitative results below, which means that we refrain from such extensions here. Therefore, an individual of productivity-type i in social identity group j , faces the following life-time budget constraint:

$$-s_j^i = c_j^i \quad (3a)$$

$$s_j^i + w_j^j - T_j^i = x_j^i \quad (3b)$$

where s denotes savings. The variable $T_j^i = T(w_j^j, s_j^i)$ denotes a tax payment (positive or negative), and $T(\cdot)$ is a tax payment function. Without loss of generality, the interest rate is set to zero, and tax payments are made based on earnings and savings. Equations (3a) and (3b) mean that the individual finances his/her studies by borrowing against the future income.⁹ When employed in the second period, the individual supplies one unit of labor inelastically. Each individual is small relative to the economy as a whole and acts as an atomistic agent in the sense of treating the identity-group-specific norms, i.e., \bar{e}_j for $j=L, H$, as exogenous. Therefore, and conditional on being part of social identity group j , the individual behaves as if he/she chooses e_j^i and s_j^i to maximize utility given in equation (1) subject to the budget constraint in equations (3a) and (3b). The first order conditions are

⁹ This is a simple way of addressing “study loans”, which are used in many countries.

$$u_{j,e}^i - (e_j^i - \bar{e}_j) + u_{j,x}^i \theta^i (1 - T_{j,w}^i) = 0 \quad (4a)$$

$$-u_{j,c}^i + u_{j,x}^i (1 - T_{j,s}^i) = 0 \quad (4b)$$

in which the second subscript attached to the utility function denotes partial derivative, i.e., $u_{j,e}^i = \partial u(c_j^i, x_j^i, e_j^i) / \partial e_j^i$, $u_{j,c}^i = \partial u(c_j^i, x_j^i, e_j^i) / \partial c_j^i$ and $u_{j,x}^i = \partial u(c_j^i, x_j^i, e_j^i) / \partial x_j^i$, while $T_{j,w}^i$ and $T_{j,s}^i$ denote marginal income and savings taxes. The choice of social identity group is then based on utility comparisons between regimes L and H .

Finally, we assume that equations (2a) and (2b) are such that, for each productivity-type, there is a marginal individual, who is precisely indifferent between the two social identity groups. For productivity-type i , this means that

$$u(c_L^i, x_L^i, e_L^i) + I_L - \beta n_L^i - \frac{1}{2}(e_L^i - \bar{e}_L)^2 = u(c_H^i, x_H^i, e_H^i) + I_H - \beta(N^i - n_L^i) - \frac{1}{2}(e_H^i - \bar{e}_H)^2. \quad (5)$$

Equation (5) implicitly defines the number of members of social identity group L of productivity-type i , n_L^i , as a function of variables characterizing both social identity groups.

The process of norm formation is important both for a welfarist government (which attempts to internalize the externalities that the social norms give rise to) and paternalist government (which would like the individuals to behave as if they were not concerned with social identity). Therefore, to analyze the decision-problem faced by each such government, we must specify how the effort norms are determined. Theoretical literature dealing with the policy implications of social comparisons often assumes that people compare their own choices with an in-group mean value¹⁰. However, research within the social identity literature suggests that group norms may be more extreme than those based on group-specific averages. Tajfel (1959) suggested that categorization accentuates similarities within groups and differences between groups. In the minimal group studies, Tajfel et al. (1971) found that a maximum difference strategy (vs. an out-group) had significant influence, thereby acting towards a positive social identity. Moscovici and Zavalloni (1969) studied group discussion to consensus and found a polarization of responses compared to individual attitudes. These early studies, as well as more recent research,¹¹ thus suggest that group attitudes tend to polarize

¹⁰ This is the case in much of the literature on optimal taxation in models with relative consumption concerns referred to in the introduction.

¹¹ See, e.g., McGarty et al. (1992), Hogg and Reid (2006).

rather than depolarize; for example, that marginal members in a group are not as liked as more central members and therefore not as likely to influence the group (Hogg and Reid, 2006).

To capture the bias away from the in-group averages towards the extremes in a simple way, we consider a modal value comparison by assuming that the majority of members in social identity group L is of the lower productivity-type, and the majority of members in social identity group H is of the higher productivity-type, such that

$$\bar{e}_L = e_L^l \text{ and } \bar{e}_H = e_H^h. \quad (6)$$

To ensure that e_L^l and e_H^h represent extreme effort choices, we also assume that the inequalities $e_L^l < e_L^h, e_H^l < e_H^h$ hold.¹² Equations (5) and (6) then imply that the number of individuals in social identity group L of productivity-type l and h , respectively, can be written as follows:

$$n_L^l = n_L^l (c_L^l, x_L^l, e_L^l, c_H^l, x_H^l, e_H^l, \bar{e}_H^+) \quad (7a)$$

$$n_L^h = n_L^h (c_L^h, x_L^h, e_L^h, c_H^h, x_H^h, e_H^h, \bar{e}_L^+), \quad (7b)$$

where the sign above each argument denotes partial derivative. The corresponding number of individuals in social identity group H can then be analyzed simply by recalling that $n_H^i = N^i - n_L^i$ for $i=l, h$.

Social Decision-Problem

The government (or social planner) is assumed to observe income and savings at the individual level, whereas individual productivity is private information. A nonlinear tax attached to earnings and savings means that the government can implement any desired combination of consumption and effort (subject to informational limitations, see below). Therefore, we follow convention in the literature on optimal nonlinear taxation and write the public decision-problem as a direct decision-problem, where the government directly decides upon consumption and effort. The optimal tax policy through which this desired resource allocation can be implemented in a decentralized economy is derived by comparing the first order conditions of the public decision-problem with the individuals' first order conditions for

¹² This always holds if the utility function is quasi-linear in the second period consumption.

education effort and savings. The government is also assumed to recognize how the social norms are determined, i.e., according to equations (6), and treat them as endogenous.¹³

We consider a Pareto efficient policy, where the government maximizes utility for one sub-group, e.g., low-productivity individuals belonging to social identity group L , subject to minimum utility restrictions for all other sub-groups. The only difference in preferences between individuals of the same productivity-type arises through the identity utility defined in equations (2a) and (2b), according to which individuals differ in their preferences for social identity groups L and H . Therefore, and conditional on group-choice, we can suppress constant terms and write the utility of productivity-type i in social identity group j as

$$U_j^i = u(c_j^i, x_j^i, e_j^i) - \frac{1}{2}(e_j^i - \bar{e}_j)^2, \quad (8a)$$

which is the utility component of any such i - j individual that a welfarist government may directly affect through tax policy. A paternalist government, on the other hand, does not share the preferences for social identity; instead, this government wants each individual to behave as if the life-time utility takes the following form for any individual of productivity-type i and social identity group j :

$$V_j^i = u(c_j^i, x_j^i, e_j^i). \quad (8b)$$

We consider the conventional case where the government wants to redistribute from high-productivity to low-productivity individuals. Therefore, since productivity is private information, the government must prevent high-productivity individuals to mimic low-productivity individuals. This is accomplished by introducing the self-selection constraints

$$U_j^h \geq u\left(c_j^l, x_j^l, \frac{\theta^l}{\theta^h} e_j^l\right) - \frac{1}{2}\left(\frac{\theta^l}{\theta^h} e_j^l - \bar{e}_j\right)^2 = \hat{U}_j^h \text{ for } j=L, H. \quad (9)$$

If (9) is satisfied, none of the individuals of productivity-type h has an incentive to become a mimicker, irrespective of the strength of the preference for belonging to social identity group j . The left hand side of the weak inequality is the utility of the true high-productivity individual, while the right hand side is the utility of the mimicker for whom the hat symbol is attached to the utility function. The variable $\theta^l / \theta^h < 1$ measures the relative productivity,

¹³ For purposes of comparison, we will also briefly discuss the case where the government treats the identity-group-specific norms as exogenous (see Section 3 below).

meaning that $\hat{e}_j^h = (\theta^l / \theta^h) e_j^l < e_j^l$ denotes the mimicker's education effort: although the mimicker earns as much income and consumes the same amount as a low-productivity individual, the mimicker is more productive and, therefore, needs to exert less effort.

To focus on tax policy, we abstract from public expenditures on education. Although seemingly restrictive, it is not important for our understanding of optimal taxation whether or not public education also (in addition to innate productivity and effort) affects the labor earnings in the second period of the individuals' lives.¹⁴ Therefore, by using the government's budget constraint, $\sum_i [n_L^i T(w_L^i, s_L^i) + (N^i - n_L^i) T(w_H^i, s_H^i)] = 0$, together with the private budget constraints, the economy's resource constraint becomes

$$\sum_i [n_L^i (w_L^i - c_L^i - x_L^i) + (N^i - n_L^i) (w_H^i - c_H^i - x_H^i)] = 0. \quad (10)$$

The resource constraint means that income is used for private consumption.

The decision-problem for the welfarist government can then be written as (if we assume that the government attempts to maximize the utility of the low-ability type of social identity group L)

$$\begin{aligned} & \underset{c_L^l, x_L^l, e_L^l, c_L^h, x_L^h, e_L^h, c_H^l, x_H^l, e_H^l, c_H^h, x_H^h, e_H^h}{Max} U_L^l \\ \text{s.t. } & (i) U_H^l \geq \bar{U}_H^l, U_L^h \geq \bar{U}_L^h, U_H^h \geq \bar{U}_H^h, \\ & (ii) U_L^h \geq \hat{U}_L^h, U_H^h \geq \hat{U}_H^h, \\ & (iii) \sum_i [n_L^i (w_L^i - c_L^i - x_L^i) + (N^i - n_L^i) (w_H^i - c_H^i - x_H^i)] = 0 \text{ for } i = l, h, \\ & (iv) \text{ equations (6) and (7).} \end{aligned}$$

where \bar{U}_H^l , \bar{U}_L^h and \bar{U}_H^h are minimum utility restrictions. The corresponding decision-problem of the paternalist government becomes

$$\begin{aligned} & \underset{c_L^l, x_L^l, e_L^l, c_L^h, x_L^h, e_L^h, c_H^l, x_H^l, e_H^l, c_H^h, x_H^h, e_H^h}{Max} V_L^l \\ \text{s.t. } & (i) V_H^l \geq \bar{V}_H^l, V_L^h \geq \bar{V}_L^h, V_H^h \geq \bar{V}_H^h, \\ & (ii) U_L^h \geq \hat{U}_L^h, U_H^h \geq \hat{U}_H^h, \\ & (iii) \sum_i [n_L^i (w_L^i - c_L^i - x_L^i) + (N^i - n_L^i) (w_H^i - c_H^i - x_H^i)] = 0 \text{ for } i = l, h, \\ & (iv) \text{ equations (6) and (7).} \end{aligned}$$

¹⁴ One way of extending the model is to assume that the government raises a net revenue, g , which is spent on education to increase individual skills, such that $w_j^i = \theta^i e_j^i h(g)$, where $h'(g) > 0$ and $h(0) = 1$. This will not change the qualitative results derived below.

The only difference between these two decision-problems is that the welfarist government recognizes the consumer preferences for social identity and aims at internalizing the externalities that this social interaction gives rise to, whereas the paternalist government bases its objective on the intrinsic part of the consumers' utility functions (as represented by the function $u(\cdot)$ in equation (1)). Note also that irrespective of whether the government is welfarist or paternalist, the self-selection constraints (constraints (ii) in the decision-problems characterized above) are always based on the actual consumer objectives: the reason is, of course, that these constraints are used to counteract mimicking and must, therefore, reflect the incentives faced by the consumers.

3. Optimal Taxation

This section begins with a presentation of the marginal income tax rates implemented by a welfarist government, and then continues with the corresponding marginal tax policy implemented by the paternalist government.

Throughout, we focus on income taxes and do not present any results for marginal savings taxes. The reason is that the social identity choices made by the consumers directly affect the optimal marginal income tax rates, while they have no direct influence on the policy incentives underlying the marginal savings taxes. As such, in a first best setting where the self-selection constraints do not bind, neither the welfarist nor the paternalist government would use marginal savings taxes. If the self-selection constraints bind, the marginal savings tax would still be zero for high-productivity individuals, while it would be positive (negative) for low-productivity individuals depending on whether low-ability individuals have a stronger (weaker) preference for early consumption compared to the corresponding mimicker. These results are well understood from earlier research and will not be further discussed here (see, e.g., Brett, 1997).

3.1 Welfarist Policy

The Lagrangean of the public-decision problem facing a welfarist government can be written as

$$\begin{aligned} \mathcal{L} = & U_L^l + \delta_L^h (U_L^h - \bar{U}_L^h) + \sum_{i=l}^h \delta_H^i (U_H^i - \bar{U}_H^i) + \sum_{j=L}^H \lambda_j (U_j^h - \hat{U}_j^h) \\ & + \gamma \sum_i \left[n_L^i (w_L^i - c_L^i - x_L^i) + (N^i - n_L^i) (w_H^i - c_H^i - x_H^i) \right] \end{aligned} \quad (11)$$

where δ_j^i , λ_j and γ are Lagrange multipliers. The first order conditions are presented in the Appendix.

First Best Taxation

To simplify the presentation, consider first the special case where individual productivity is observable. In terms of the model set out above, this special case means that the self-selection constraints become redundant and $\lambda_L = \lambda_H = 0$. As such, it also provides a suitable starting point: since the government can redistribute through productivity-specific lump-sum taxes, the only reason for distorting the education choice is to correct for externalities. Therefore, the optimal marginal income tax rates will solely reflect (i) the welfare contributions of the social norms, and (ii) how each productivity-identity group affects these norms through effort choices. The welfare contribution of each social norm can be derived by differentiating the Lagrangean in equation (11) with respect to \bar{e}_L and \bar{e}_H , respectively, as follows:

$$\frac{\partial \mathcal{L}}{\partial \bar{e}_L} = \delta_L^h (e_L^h - \bar{e}_L) + \gamma \frac{\partial n_L^h}{\partial \bar{e}_L} (G_L^h - G_H^h) \quad (12a)$$

$$\frac{\partial \mathcal{L}}{\partial \bar{e}_H} = \delta_H^l (e_H^l - \bar{e}_H) + \gamma \frac{\partial n_L^l}{\partial \bar{e}_H} (G_L^l - G_H^l) \quad (12b)$$

where $G_j^i = w_j^i - c_j^i - x_j^i$ is the net contribution to public revenue by productivity-type i in social identity group j , while equations (7a) and (7b) imply $\partial n_L^h / \partial \bar{e}_L > 0$ and $\partial n_L^l / \partial \bar{e}_H > 0$ based on our earlier assumptions. We base most of our interpretations below on the additional (and reasonable) assumption that for each productivity-type, individuals in social identity group H contribute more to the tax revenue than individuals in social identity group L , such that $G_H^i > G_L^i$ for $i=l, h$.

We have derived the following result:

Proposition 1. *In a first best setting where $\lambda_L = \lambda_H = 0$, the marginal income tax policy implemented by a welfarist government can be characterized as*

$$T_{L,w}^l = -\frac{\partial \mathcal{L}}{\partial \bar{e}_L} \frac{1}{\gamma n_L^l \theta^l}, \quad T_{L,w}^h = 0, \quad T_{H,w}^l = 0, \quad \text{and} \quad T_{H,w}^h = -\frac{\partial \mathcal{L}}{\partial \bar{e}_H} \frac{1}{\gamma n_H^h \theta^h},$$

where $\partial \mathcal{L} / \partial \bar{e}_L$ and $\partial \mathcal{L} / \partial \bar{e}_H$ are given by equations (12).

Proof: see the Appendix.

Note that it is only the low-productivity type in social identity group L and high-productivity type in social identity group H that generate externalities. As such, it is only the education choices in these two sub-groups that will be distorted in the first best optimum. Given the assumptions set out above, we have $e_H^l < \bar{e}_H$. Therefore, if $G_H^l > G_L^l$, it follows that $\partial \mathcal{L} / \partial \bar{e}_H < 0$, since an increase in \bar{e}_H leads to lower utility for low-productivity individuals in social identity group H and to lower tax revenue through an increase in the number of high-productivity individuals that select into social identity group L . This means that $T_{H,w}^h > 0$, suggesting that the incentives faced by a welfarist government to internalize externalities generate an element of tax progression. However, $\partial \mathcal{L} / \partial \bar{e}_L$ can be either positive or negative because an increase in \bar{e}_L leads to higher utility for high-productivity individuals in social reference group L (since $e_L^h > \bar{e}_L$ by our earlier assumptions) and to lower tax revenue (if $G_H^h > G_L^h$). As a consequence, $T_{L,w}^l$ may be either positive or negative at the optimum depending on which effect dominates.

Finally, notice that the only reason for a welfarist government to distort the individual's education choice in a first best setting is to influence \bar{e}_L and \bar{e}_H . This is seen from the following corollary to Proposition 1, which characterizes the marginal tax policy that would follow in the special case where the government treats \bar{e}_L and \bar{e}_H as exogenous:

Corollary 1. *In a first best setting, and if the welfarist government treats \bar{e}_L and \bar{e}_H as exogenous, the marginal tax rates on earnings are zero, i.e.,*

$$T_{w_L}^l = T_{w_H}^l = T_{w_L}^h = T_{w_H}^h = 0.$$

The interesting thing to note here is that Corollary 1 applies despite that the selection into social identity groups is endogenous. In other words, a welfarist government that treats the identity-group-specific norms as exogenous has no incentive to influence the selection into social identity groups by taxing earnings. As we will see below, this result does not apply for a paternalist government, which would like the individuals to behave as if social identity were of no concerns for them.

Second Best Taxation

With Proposition 1 at our disposal, we are now ready to examine the implications of social identity choices for optimal second best taxation. To shorten the notation, let

$$MRS_{j,ex}^i = \frac{u_{j,e}^i - (e_j^i - \bar{e}_j)}{u_{j,x}^i} \quad \text{and} \quad M\hat{R}S_{j,ex}^h = \frac{\left(\hat{u}_{j,e}^h - \left(\frac{\theta^l}{\theta^h} e_j^l - \bar{e}_j \right) \right)}{\hat{u}_{j,x}^l} \quad (13)$$

denote the marginal rate of substitution between effort and second period consumption for productivity-type i and the mimicker, respectively, in social identity group j . With binding self-selection constraints, the partial welfare effect of an increase in \bar{e}_L and \bar{e}_H , respectively, extends to read

$$\frac{\partial \mathcal{L}}{\partial \bar{e}_L} = \delta_L^h (e_L^h - \bar{e}_L) + \lambda_L \left(e_L^h - \frac{\theta^l}{\theta^h} e_L^l \right) + \gamma \frac{\partial n_L^h}{\partial \bar{e}_L} (G_L^h - G_H^h) \quad (14a)$$

$$\frac{\partial \mathcal{L}}{\partial \bar{e}_H} = \delta_H^l (e_H^l - \bar{e}_H) + \lambda_H \left(e_H^h - \frac{\theta^l}{\theta^h} e_H^l \right) + \gamma \frac{\partial n_L^l}{\partial \bar{e}_H} (G_L^l - G_H^l). \quad (14b)$$

The difference between equations (12) and (14) is the second term on the right hand side of equations (14a) and (14b): each such component is positive, meaning that an increase in each social norm contributes to relax one of the self-selection constraints. This is so because an increase in \bar{e}_L leads to lower utility (through greater effort) for the mimicker in social identity group L , ceteris paribus, while an increase in \bar{e}_H contributes to increase the distance between the effort norm and the mimicker's effort in social identity group H . The second best optimal tax policy is presented in Proposition 2:

Proposition 2. *In a second best setting where the self-selection constraints bind, the marginal income tax policy implemented by a welfarist government can be characterized as*

$$T_{L,w}^l = -\frac{\lambda_L \hat{u}_{L,x}}{\gamma n_L^l \theta^l} \left(MRS_{L,ex}^l - \frac{\theta^l}{\theta^h} MR\hat{S}_{L,ex}^h \right) - \frac{\partial \mathcal{L}}{\partial \bar{e}_L} \frac{1}{\gamma n_L^l \theta^l},$$

$$T_{L,w}^h = 0,$$

$$T_{H,w}^l = -\frac{\lambda_H \hat{u}_{H,x}}{\gamma n_H^l \theta^l} \left(MRS_{H,ex}^l - \frac{\theta^l}{\theta^h} MR\hat{S}_{H,ex}^h \right)$$

$$T_{H,w}^h = -\frac{\partial \mathcal{L}}{\partial \bar{e}_H} \frac{1}{\gamma n_H^h \theta^h}$$

where $\partial \mathcal{L} / \partial \bar{e}_L$ and $\partial \mathcal{L} / \partial \bar{e}_H$ are given by equations (14).

Proof: see the Appendix.

There are two important differences between the tax formulas in Proposition 2 and the corresponding first best policy analyzed in Proposition 1. First, the component

$$-\frac{\lambda_j \hat{u}_{j,x}}{\gamma n_j^l \theta^l} \left(MRS_{j,ex}^l - \frac{\theta^l}{\theta^h} MR\hat{S}_{j,ex}^h \right)$$

in the marginal income tax rate of the low-productivity type in each social identity group is positive, and interpretable as the marginal tax rate that would be implemented for this low-productivity type if \bar{e}_L and \bar{e}_H were exogenous to the government. This tax incentive serves to make mimicking unattractive by exploiting that each low-productivity type and corresponding mimicker differ from one another with respect to the marginal value of leisure. The corresponding component for high-productivity individuals is zero here (because the relative productivity, θ^l / θ^h , is constant). As such, this mechanism is well known from earlier studies (e.g., Stiglitz, 1982).¹⁵ Second, increases in \bar{e}_L and \bar{e}_H contribute to relax the self-selection constraints as explained above, which can be seen from the second term on the right hand side of the expressions for $\partial \mathcal{L} / \partial \bar{e}_L$ and $\partial \mathcal{L} / \partial \bar{e}_H$ given in equations (14): this

¹⁵ Note that this component would be present also in the absence of any preference for social identity, and works in the direction of regressive taxation, in the sense that the marginal tax rates decline with productivity, *ceteris paribus*. Guo and Krause (2013) use a model with two productivity-types different from ours, where the individual both supplies work hours and faces direct expenditures on education, and where the policy instruments are labor income taxes and taxes on education expenditure. Although they find that the income tax is regressive in the sense described above, they also find that the optimal education policy is progressive in the sense that the marginal education tax is negative (i.e., a marginal subsidy) for the low-productivity type and zero for the high-productivity type.

mechanism works to reduce the marginal tax rates facing the externality generating consumers.

3.2 Paternalist Policy

Turning to the paternalist government's decision-problem, the Lagrangean is now given by

$$\begin{aligned} \mathcal{L} = & V_L^l + \delta_L^h (V_L^h - \bar{V}_L^h) + \sum_{i=l}^h \delta_H^i (V_H^i - \bar{V}_H^i) + \sum_{j=L}^H \lambda_j (U_j^h - \hat{U}_j^h) \\ & + \gamma \sum_i [n_L^i (w_L^i - c_L^i - x_L^i) + (N^i - n_L^i)(w_H^i - c_H^i - x_H^i)] \end{aligned} \quad (15)$$

As explained above, the paternalist government does not share the individual preferences for social identity; therefore, V_j^i is interpretable as the utility function that the government would like productivity-type i in social identity group j to have (which, in turn, coincides with the intrinsic component of the individual's utility function given by the function $u(\cdot)$ in equation (1)). The first order conditions are presented in the Appendix.

The partial welfare effects of increases in the social norms, i.e., \bar{e}_L and \bar{e}_H , can now be written as

$$\frac{\partial \mathcal{L}}{\partial \bar{e}_L} = \lambda_L \left(e_L^h - \frac{\theta^l}{\theta^h} e_L^l \right) + \gamma \frac{\partial n_L^h}{\partial \bar{e}_L} (G_L^h - G_H^h) \quad (16a)$$

$$\frac{\partial \mathcal{L}}{\partial \bar{e}_H} = \lambda_H \left(e_H^h - \frac{\theta^l}{\theta^h} e_H^l \right) + \gamma \frac{\partial n_L^l}{\partial \bar{e}_H} (G_L^l - G_H^l). \quad (16b)$$

Compared to the welfarist model, equations (16a) and (16b) imply that the welfare effect of an increase in each such norm is decomposed into two (instead of three) components, i.e., the first term on the right hand side of equation (14a) and (14b), respectively, is absent here since the paternalist government does not share the consumer preference for social identity. The remaining terms are identical to their counterparts in the welfarist case. As explained above, the first term on the right hand side of equation (16a) and (16b), respectively, is positive, since an increase in \bar{e}_L makes mimicking less attractive in social identity group L by necessitating greater effort of the mimicker, whereas an increase in \bar{e}_H makes mimicking less attractive in social identity group H due to lower identity utility for the mimicker. The second term on the

right hand side of equation (16a) and (16b), respectively, is negative if $G_L^i < G_H^i$ and positive otherwise for $i=l, h$.

First Best Policy

As in subsection 3.1, we begin by considering the special case where the self-selection constraints do not bind, i.e., where $\lambda_L = \lambda_H = 0$. Therefore, if $G_L^i < G_H^i$ for $i=l, h$, equations (16a) and (16b) imply $\partial\mathcal{L}/\partial\bar{e}_L < 0$ and $\partial\mathcal{L}/\partial\bar{e}_H < 0$. Proposition 3 describes the first best policy of a paternalist government:

Proposition 3. *In a first best setting where $\lambda_L = \lambda_H = 0$, the marginal income tax rates implemented by a paternalist government can be written as*

$$\begin{aligned} T_{L,w}^l &= -\frac{1}{\gamma\theta^l n_L^l} \frac{\partial\mathcal{L}}{\partial\bar{e}_L} = -\frac{1}{\theta^l n_L^l} \frac{\partial n_L^h}{\partial\bar{e}_L} (G_L^h - G_H^h) \\ T_{L,w}^h &= -\frac{1}{\theta^h} \frac{(e_L^h - \bar{e}_L)}{u_{L,x}^h} + \frac{1}{\theta^h n_L^h} (G_L^h - G_H^h) \frac{(e_L^h - \bar{e}_L)}{2\beta} \\ T_{H,w}^l &= -\frac{1}{\theta^l} \frac{(e_H^l - \bar{e}_H)}{u_{H,x}^l} - \frac{1}{\theta^l n_H^l} (G_L^l - G_H^l) \frac{(e_H^l - \bar{e}_H)}{2\beta} \\ T_{H,w}^h &= -\frac{1}{\gamma n_H^h \theta^h} \frac{\partial\mathcal{L}}{\partial\bar{e}_H} = -\frac{1}{\theta^h n_H^h} \frac{\partial n_L^l}{\partial\bar{e}_H} (G_L^l - G_H^l). \end{aligned}$$

Proof: see the Appendix.

Recall from Proposition 1 that the only reason for a welfarist government to distort the effort choice in a first best world is to influence \bar{e}_L and \bar{e}_H (implying a positive marginal tax for the high-productivity type in social identity group H , while the marginal tax imposed on the low-productivity type in social identity group L could be either positive or negative). This policy incentive is present here as well; yet, in modified form, since the paternalist government does not share the consumer preference for social identity. Accordingly, if $G_L^i < G_H^i$ for $i=l, h$, the first best tax policy of the paternalist government implies $T_{L,w}^l > 0$ and $T_{H,w}^h > 0$. The intuition is that a decrease in \bar{e}_L increases the number of high-productivity individuals in social

identity group H , and a decrease in \bar{e}_H leads to an increase in the number of low-productivity individuals in social identity group H , ceteris paribus, which contribute to increased tax revenue if $G_L^i < G_H^i$. Conversely, if $G_L^i > G_H^i$, the paternalist policy implies $T_{L,w}^l < 0$ and $T_{H,w}^h < 0$; let be that this outcome seems unlikely.

However, contrary to the welfarist government, a paternalist government imposes non-zero marginal income taxes also on the high-productivity type in social identity group L and low-productivity type in social identity group H . As can be seen from the second and third formulas in the proposition, there are two reasons for this. The first term on the right hand side of the formula for $T_{L,w}^h$ is negative (since $e_L^h > \bar{e}_L$) and represents a pure paternalist motive for subsidizing the income of the high-productivity type in social identity group L . The intuition is that such a marginal subsidy counteracts the incentive for this agent to choose less effort in response to the effort norm. By analogy, the first term on the right hand side of the formula for $T_{H,w}^l$ is positive (since $e_H^l < \bar{e}_H$) and constitutes a pure paternalist motive for marginal income taxation of the low-productivity type in social identity group H , who would otherwise exert too much effort in response to the effort norm.

The second reason for imposing non-zero marginal income taxes on the h - L and l - H individuals is captured by the second term on the right hand side of the formulas for $T_{L,w}^h$ and $T_{H,w}^l$ in Proposition 3. When the government does not share the individual preferences for social identity, the individuals' selection into social identity groups may influence the marginal tax policy: these effects are absent under a welfarist government, which is seen from Corollary 1. To be more specific, there is a paternalist motive for influencing the tax revenue through the selection into social identity groups. We show in the Appendix that these components are derived from the following expressions

$$\gamma(G_L^i - G_H^i) \left[\frac{\partial n_L^i}{\partial e_L^i} - \frac{u_{L,e}^i}{u_{L,x}^i} \frac{\partial n_L^i}{\partial e_L^i} \right] = -\gamma(G_L^i - G_H^i) \frac{(e_L^i - \bar{e}_L)}{2\beta} \quad (17a)$$

$$\gamma(G_L^i - G_H^i) \left[\frac{\partial n_L^i}{\partial e_H^i} - \frac{u_{H,e}^i}{u_{H,x}^i} \frac{\partial n_L^i}{\partial e_H^i} \right] = \gamma(G_L^i - G_H^i) \frac{(e_H^i - \bar{e}_H)}{2\beta} \quad (17b)$$

where the right hand side follows from the comparative statics properties of equations (7a) and (7b). Note that the right hand side of equation (17a) is zero for productivity-type l and non-zero for productivity-type h , while the right hand side of equation (17b) is zero for productivity-type h and non-zero for productivity-type l . The intuition is that a compensated increase in e_L^h (where the utility-compensation is based on the preferences of the paternalist government) leads to lower identity utility for h - L individuals and, therefore, a decrease in the number of individuals of productivity-type h in social identity group L . Similarly, a compensated increase in e_H^l leads to that fewer individuals of productivity-type l choose social identity group L (due to that the discrepancy between e_H^l and \bar{e}_H decreases). In turn, this leads to increased tax revenue if $G_L^i < G_H^i$ for $i=l, h$, which constitutes an incentive to subsidize income at the margin for the high-productivity type in social identity group L and the low-productivity type in social identity group H (the second term on the right hand side of each such tax formula is negative).

We summarize the qualitative implications of Proposition 3 as follows:

Corollary 2. *In a first best setting where $\lambda_L = \lambda_H = 0$, and if $G_L^i < G_H^i$ for $i=l, h$, the optimal tax policy implemented by a paternalist government satisfies $T_{L,w}^l > 0$, $T_{L,w}^h < 0$ and $T_{H,w}^h > 0$, while $T_{H,w}^l$ can be either positive or negative at the optimum.*

Second Best Policy

We now turn to the second best model, where the self-selection constraints bind. To shorten the notation, let

$$PRS_{j,ex}^i = \frac{u_{j,e}^i}{u_{j,x}^i} \text{ and } \hat{PRS}_{j,ex}^h = \frac{\hat{u}_{j,e}^h}{\hat{u}_{j,x}^h} \quad (18)$$

represent the marginal rate of substitution between effort and second period consumption for productivity-type i and the mimicker, respectively, in social identity group j based on the preferences imposed on them by the paternalist government. As such, these differ from the individuals' own marginal rates of substitution presented in equations (13). Also, to suppress policy incentives already explained, we use $T_{j,w}^{i,FB}$ to denote the first best marginal tax formula

for productivity-type i in social identity group j as defined in Proposition 3, although here evaluated in the second best allocation. The optimal marginal tax policy is characterized in Proposition 4.

Proposition 4. *In a second best resource allocation where the self-selection constraints bind, the tax policy implemented by a paternalist government satisfies*

$$T_{L,w}^l = T_{L,w}^{l,FB} - \frac{\lambda_L}{\gamma n_L^l \theta^l} \left(e_L^h - \frac{\theta^l}{\theta^h} e_L^l \right) - \frac{\lambda_L \hat{u}_{L,x}}{\gamma n_L^l \theta^l} \Delta_L - \frac{\lambda_L}{\gamma n_L^l \theta^h} \left(\frac{\theta^l}{\theta^h} e_L^l - \bar{e}_L \right)$$

$$T_{L,w}^h = T_{L,w}^{h,FB} + \frac{\lambda_L}{\gamma n_L^h \theta^h} \left(e_L^h - \bar{e}_L \right)$$

$$T_{H,w}^l = T_{H,w}^{l,FB} - \frac{\lambda_H \hat{u}_{H,x}}{\gamma n_H^l \theta^l} \Delta_H - \frac{\lambda_H}{\gamma n_H^l \theta^h} \left(\frac{\theta^l}{\theta^h} e_H^l - \bar{e}_H \right)$$

$$T_{H,w}^h = T_{H,w}^{h,FB} - \frac{\lambda_H}{\gamma n_H^h \theta^h} \left(e_H^h - \frac{\theta^l}{\theta^h} e_H^l \right)$$

where $\Delta_j = PRS_{j,ex}^l - \frac{\theta^l}{\theta^h} PRS_{j,ex}^h$ for $j=L,H$.

Proof: see the Appendix.

Three additional components arise here compared to the first best policy rules characterized in Proposition 3. First, there is a direct incentive to make mimicking less attractive though marginal taxation of the low-productivity type captured by

$$-\frac{\lambda_j \hat{u}_{j,x}}{\gamma n_j^l \theta^l} \Delta_j - \frac{\lambda_j}{\gamma n_j^l \theta^h} \left(\frac{\theta^l}{\theta^h} e_j^l - \bar{e}_j \right) > 0 \text{ for } j=L,H$$

in the first and third tax formulas in Proposition 4. Therefore, this effect works to increase the marginal tax rates faced by low-productivity individuals (compared to the marginal tax rates under full information). The interpretation is that a decrease in e_j^l relaxes the self-selection constraint both because $\Delta_j < 0$ (as in conventional models of optimal income taxation) and by increasing the distance between the mimicker's effort and the effort norm. Second, there is an incentive to relax the self-selection constraint though marginal income taxation of the high-productivity type in social identity group L , which works to offset the subsidy result derived under first best conditions (see Corollary 2). This effect is captured by the second term on the

right hand side of the formula for $T_{L,w}^h$ (which is positive since $e_L^h > \bar{e}_L$), implying that the marginal tax facing h - L individuals can be either positive or negative here. The intuition is that a decrease in e_L^h leads to higher identity utility (and, therefore, counteracts the incentive to become a mimicker) for h - L individuals by reducing distance between e_L^h and \bar{e}_L , ceteris paribus. There is no corresponding effect in the marginal income tax implemented for h - H individuals where $e_H^h = \bar{e}_H$.

The third additional component reflects a desire to relax the self-selection constraints through changes in the two social norms, \bar{e}_L and \bar{e}_H : this will only affect the marginal income tax rates of the low-productivity type in social identity group L and high-productivity type in social identity group H . Since the government may relax the self-selection constraint by increasing the social norms, there is an incentive to subsidize l - L and h - H individuals as reflected by the second term on the right hand side in their tax formulas.

4. Summary and Conclusions

In this paper, we have used an education framework to study the implications of self-categorization for optimal tax policy. We motivate our study mainly by the observation that intergenerational social mobility is often considered to be lower than is optimal from society's point of view. In the model it is assumed that individuals differ in two respects. First, as in a standard optimal taxation framework, we assume that individuals are of two types, a high- and a low-productivity type reflecting innate ability. Second, we introduce an element of social inertia by assuming that individuals differ with respect to their preferences for the social group they want to be associated with. We exemplify by assuming that there are two social identity groups in which norms regarding study effort are formed. Individuals will then self-select into one of the groups depending on ability and group preference, and they will make educational (effort) choices based on their ability and the group norm.

Norm formation is endogenous and assumed to be based on modal values. We base this assumption on social psychology findings that norms appear to be more polarized than group averages. Two versions of the social decision problem are then considered; one in which a welfarist government attempts to internalize the externalities that the social comparisons give

rise to, and the other where a paternalist government disregards the value individuals place on self-categorization and tries to induce individuals to behave as if they were not concerned with social identity. Under each of these assumptions, taxation is analyzed in a first best scenario as well as a second best framework where individual productivity cannot be observed.

In the full information scenario, the welfarist government uses nonlinear taxation to correct for externalities generated by the decisive (norm forming) types, i.e., high-productivity individuals in the high-effort group and low-productivity individuals in the low-effort group, respectively. Marginal tax rates are non-zero for these groups, and it is shown that the marginal tax rate for the high-productivity type in the high-effort group is typically positive, meaning that there is an element of tax progression in the optimal tax formula. In the second best solution, we show, among other things, that increases in the effort norms helps in relaxing the self-selection constraints, which means that the optimal policy contains one factor that works in the direction of reducing the marginal tax rates of the externality generating individuals.

In contrast to the welfarist government, a paternalist government has a motive to correct effort choices to “undo” the effects of self-categorization. This means that there is an element in the tax formula such that high-productivity individuals choosing the low-effort group are subsidized at the margin, and vice versa, low productivity individuals in the high-effort group are taxed at the margin. We also show that the earnings of both the low-productivity type in the low-effort group and the high-productivity type in the high-effort group are tax at the margin in a first best optimum. Another reason for imposing non-zero marginal tax rates is that a paternalist government wants to influence the tax revenue through the selection into social groups.

Let us end by briefly discussing two possible directions of future research. First, our paper solely focuses on taxation, which means that we have neglected the role of public expenditure. While earlier research shows that different types of public expenditure are useful instruments for redistribution, less is known about the role that such expenditure may play in connection to social norms and, in particular, how corrective and redistributive aspects of public expenditure interact in this context. For instance, will concerns for social identity

among consumers motivate higher or lower public investments in education, and how would welfarist and paternalist governments differ with respect to such investments? Second, a model with more than two periods would allow us to relate effort choices (and possibly also the social identity component underlying such choices) to time-inconsistent preferences for immediate gratification. If individuals later in life regret their lack of study effort, at least a paternalist government may have incentives to use taxation and public expenditure to correct these behavioral mistakes. We hope to address these and other questions in future research.

Appendix

Social First Order Conditions under Welfarism

Use $G_j^i = w_j^i - c_j^i - x_j^i$, $\hat{u}_j^h = u^h(c_j^l, x_j^l, \frac{\theta^l}{\theta^h} e_j^l)$, $n_L^i = n^i$ and $n_H^i = \bar{n}^i - n^i$. We have

$$\delta_j^l u_{j,c}^l - \lambda_j \hat{u}_{j,c}^h + \gamma \left[-n_j^l + \frac{\partial n^l}{\partial c_j^l} (G_L^l - G_H^l) \right] = 0 \quad (\text{A1})$$

$$\delta_j^l u_{j,x}^l - \lambda_j \hat{u}_{j,x}^h + \gamma \left[-n_j^l + \frac{\partial n^l}{\partial x_j^l} (G_L^l - G_H^l) \right] = 0 \quad (\text{A2})$$

$$\begin{aligned} \delta_j^l [u_{j,e}^l - (e_j^l - \bar{e}_j)] - \lambda_j \frac{\theta^l}{\theta^h} \left[\hat{u}_{j,e}^h - \left(\frac{\theta^l}{\theta^h} e_j^l - \bar{e}_j \right) \right] \\ + \gamma \left[n_j^l \theta^l + \frac{\partial n^l}{\partial e_j^l} (G_L^l - G_H^l) \right] + \frac{\partial \mathcal{L}}{\partial \bar{e}_j} \frac{\partial \bar{e}_j}{\partial e_j^l} = 0 \end{aligned} \quad (\text{A3})$$

$$(\delta_j^h + \lambda_j) u_{j,c}^h + \gamma \left[-n_j^h + \frac{\partial n^h}{\partial c_j^h} (G_L^h - G_H^h) \right] = 0 \quad (\text{A4})$$

$$(\delta_j^h + \lambda_j) u_{j,x}^h + \gamma \left[-n_j^h + \frac{\partial n^h}{\partial x_j^h} (G_L^h - G_H^h) \right] = 0 \quad (\text{A5})$$

$$(\delta_j^h + \lambda_j) [u_{j,e}^h - (e_j^h - \bar{e}_j)] + \gamma \left[n_j^h \theta^h + \frac{\partial n^h}{\partial e_j^h} (G_L^h - G_H^h) \right] + \frac{\partial \mathcal{L}}{\partial \bar{e}_j} \frac{\partial \bar{e}_j}{\partial e_j^h} = 0 \quad (\text{A6})$$

for $j=L, H$, and $\delta_L^l = 1$.

In all proofs below, we use that the norms are given by $\bar{e}_L = e_L^l$ and $\bar{e}_H = e_H^h$, meaning that the partial welfare effects of increases in these norms will only be part of the social first order conditions for effort of the low-productivity type in social identity group L and high-productivity type in social identity group H , respectively.

Proof of Proposition 1

In the special case where the self-selection constraints are not binding, we have $\lambda_L = \lambda_H = 0$ in equations (A1)-(A6). Consider first the marginal income tax rate implemented for the low-productivity type in social identity group L . Combining equations (A2) and (A3) while using equations (13) give

$$(MRS_{L,ex}^l + \theta^l)\gamma n_j^l = \gamma(G_L^l - G_H^l) \left[MRS_{L,ex}^l \frac{\partial n_L^l}{\partial x_L^l} - \frac{\partial n_L^l}{\partial e_L^l} \right] - \frac{\partial \mathcal{L}}{\partial \bar{e}_L}. \quad (\text{A7})$$

Making use of the private first order condition from equation (4a) we can replace $MRS_{L,ex}^l + \theta^l$ with $\theta^l T_{L,w}^l$ in equation (A7). Finally, using equations (7) to derive

$$\frac{\partial n_L^l}{\partial x_L^l} = \frac{1}{2\beta} \frac{\partial u_L^l}{\partial x_L^l} \quad (\text{A8a})$$

$$\frac{\partial n_L^l}{\partial e_L^l} = \frac{1}{2\beta} \left(\frac{\partial u_L^l}{\partial e_L^l} - (e_L^l - \bar{e}_L) \right) \quad (\text{A8b})$$

we can see that the two terms within square brackets in equation (A7) cancel out, which means that we arrive at the expression for $T_{L,w}^l$ in the proposition. The marginal income tax rate facing productivity-type h in social identity group H can be derived analogously by using equations (A5) and (A6). Following again the same approach, the formulas for the optimal marginal tax rates for productivity-type l in social identity group H and productivity-type h in social identity group L will reduce to zero, since they do not contain the partial welfare effect of any of the two social norms. ■

Proof of Proposition 2

In a second best setting where $\lambda_j \neq 0$ for $j=L, H$, we see from the social first order conditions that only the low-productivity type will be additionally distorted in its effort choice compared to the first best. The equation analogous to equation (A7) for type l individuals in social identity group L then becomes

$$\begin{aligned}
(MRS_{L,ex}^l + \theta^l) \gamma n_j^l &= \gamma (G_L^l - G_H^l) \left[MRS_{L,ex}^l \frac{\partial n_L^l}{\partial x_L^l} - \frac{\partial n_L^l}{\partial e_L^l} \right] \\
&\quad - MRS_{L,ex}^l \lambda_L \hat{u}_{L,x}^h + \lambda_L \frac{\theta^l}{\theta^h} \left[\hat{u}_{L,e}^h - \left(\frac{\theta^l}{\theta^h} e_L^l - \bar{e}_L \right) \right] - \frac{\partial \mathcal{L}}{\partial \bar{e}_L}
\end{aligned} \tag{A9}$$

Using the definition of $\hat{MRS}_{j,ex}^h$ from equation (13), and $MRS_{L,ex}^l + \theta^l = \theta^l T_{L,w}^l$, we arrive at the marginal tax formula for l - L individuals in the proposition. The corresponding tax formulas for other combinations of productivity and social identity groups are derived in the same general way. ■

Social First Order Conditions under Paternalism

$$\delta_j^l u_{j,c}^l - \lambda_j \hat{u}_{j,c}^h + \gamma \left[-n_j^l + \frac{\partial n^l}{\partial c_j^l} (G_L^l - G_H^l) \right] = 0 \tag{A10}$$

$$\delta_j^l u_{j,x}^l - \lambda_j \hat{u}_{j,x}^h + \gamma \left[-n_j^l + \frac{\partial n^l}{\partial x_j^l} (G_L^l - G_H^l) \right] = 0 \tag{A11}$$

$$\delta_j^l u_{j,e}^l - \lambda_j \frac{\theta^l}{\theta^h} \left[\hat{u}_{j,e}^h - \left(\frac{\theta^l}{\theta^h} e_j^l - \bar{e}_j \right) \right] + \gamma \left[n_j^l \theta^l + \frac{\partial n^l}{\partial e_j^l} (G_L^l - G_H^l) \right] + \frac{\partial \mathcal{L}}{\partial \bar{e}_j} \frac{\partial \bar{e}_j}{\partial e_j^l} = 0 \tag{A12}$$

$$(\delta_j^h + \lambda_j) u_{j,c}^h + \gamma \left[-n_j^h + \frac{\partial n^h}{\partial c_j^h} (G_L^h - G_H^h) \right] = 0 \tag{A13}$$

$$(\delta_j^h + \lambda_j) u_{j,x}^h + \gamma \left[-n_j^h + \frac{\partial n^h}{\partial x_j^h} (G_L^h - G_H^h) \right] = 0 \tag{A14}$$

$$(\delta_j^h + \lambda_j) u_{j,e}^h - \lambda_j (e_j^h - \bar{e}_j) + \gamma \left[n_j^h \theta^h + \frac{\partial n^h}{\partial e_j^h} (G_L^h - G_H^h) \right] + \frac{\partial \mathcal{L}}{\partial \bar{e}_j} \frac{\partial \bar{e}_j}{\partial e_j^h} = 0 \tag{A15}$$

for $j=L, H$, and $\delta_L^l = 1$.

Proof of Proposition 3

Consider productivity-types l and h in social identity groups group L . By using equations (A11) and (A12) together with the definition of marginal rate of substitution in equations (18), we have

$$(PRS_{L,ex}^l + \theta^l) \gamma n_L^l = \left(PRS_{L,ex}^l \frac{\partial n_L^l}{\partial x_L^l} - \frac{\partial n_L^l}{\partial e_L^l} \right) \gamma (G_L^h - G_H^h) - \frac{\partial \mathcal{L}}{\partial \bar{e}_L}. \tag{A16}$$

Similarly, combining equations (A14) and (A15) gives

$$\left(PRS_{L,ex}^h + \theta^h \right) \gamma n_L^h = \left(PRS_{L,ex}^h \frac{\partial n_L^h}{\partial x_L^h} - \frac{\partial n_L^h}{\partial e_L^h} \right) \gamma (G_L^h - G_H^h). \quad (A17)$$

Note that $\partial \bar{e}_L / \partial e_L^h = \partial \bar{e}_H / \partial e_L^h = 0$, which explains why there are no partial welfare effects of social norms in equation (A17).

Now, due to the difference between the *MRS* and *PRS* measures, we can use the private first order condition for effort to derive

$$PRS_{j,ex}^i + \theta^i = \theta^i T_{j,w}^i + \frac{(e_j^i - \bar{e}_j)}{u_{j,x}^i}. \quad (A18)$$

Substituting into equations (A16) and (A17) gives

$$\theta^l T_{L,w}^l \gamma n_L^l = \left(PRS_{L,ex}^l \frac{\partial n_L^l}{\partial x_L^l} - \frac{\partial n_L^l}{\partial e_L^l} \right) \gamma (G_L^h - G_H^h) - \frac{\partial \mathcal{L}}{\partial \bar{e}_L} - \frac{(e_L^l - \bar{e}_L)}{u_{L,x}^l} \quad (A19)$$

$$\theta^h T_{L,w}^h \gamma n_L^h = \left(PRS_{L,ex}^h \frac{\partial n_L^h}{\partial x_L^h} - \frac{\partial n_L^h}{\partial e_L^h} \right) \gamma (G_L^h - G_H^h) - \frac{(e_L^h - \bar{e}_L)}{u_{L,x}^h}. \quad (A20)$$

Note that equations (7a) and (7b) imply

$$PRS_{L,ex}^l \frac{\partial n_L^l}{\partial x_L^l} - \frac{\partial n_L^l}{\partial e_L^l} = 0 \quad (A21)$$

$$PRS_{L,ex}^h \frac{\partial n_L^h}{\partial x_L^h} - \frac{\partial n_L^h}{\partial e_L^h} = \frac{1}{2\beta} (e_L^h - \bar{e}_L). \quad (A22)$$

The right hand side of equation (A22) is due to the discrepancy between the *MRS* and *PRS* measures; this discrepancy vanishes in equation (A21), since $\bar{e}_L = e_L^l$. Substituting equations (A21) and (A22) into equations (A19) and (A20) and using $\bar{e}_L = e_L^l$ gives the expressions for $T_{L,w}^l$ and $T_{L,w}^h$ in the proposition. The marginal income tax rates implemented for productivity-types l and h in social identity group H can be derived analogously. ■

Proof of Proposition 4

When the self-selection constraints bind (such that $\lambda_L \neq 0$ and $\lambda_H \neq 0$), the analogue to equation (A16) is given by

$$\begin{aligned}
\left(PRS_{L,ex}^l + \theta^l \right) \gamma n_j^l &= \left(PRS_{L,ex}^l \frac{\partial n_L^l}{\partial x_L^l} - \frac{\partial n_L^l}{\partial e_L^l} \right) \gamma \left(G_L^h - G_H^h \right) \\
&\quad - \lambda_L \hat{u}_{L,x}^l \left(PRS_{L,ex}^l - \frac{\theta^l}{\theta^h} P\hat{R}S_{L,ex}^h \right) \quad . \quad (A23) \\
&\quad - \lambda_L \frac{\theta^l}{\theta^h} \left(\frac{\theta^l}{\theta^h} e_L^l - \bar{e}_L \right) - \frac{\partial \mathcal{L}}{\partial \bar{e}_L}
\end{aligned}$$

The difference to the first best setting is the second and third term on the right hand side. Similarly, the analogue to equation (A17) becomes

$$\left(PRS_{L,ex}^h + \theta^h \right) \gamma n_j^h = \left(PRS_{L,ex}^h \frac{\partial n_L^h}{\partial x_L^h} - \frac{\partial n_L^h}{\partial e_L^h} \right) \gamma \left(G_L^h - G_H^h \right) + \lambda_j \left(e_L^h - \bar{e}_L \right) \quad (A18)$$

where the final term on the right hand side constitutes the addition due to the binding self-selection constraints. By proceeding in exactly the same way as in the proof of Proposition 3, we can derive the expressions for $T_{L,w}^l$ and $T_{L,w}^h$ in Proposition 4. Again, the marginal income tax rates implemented for productivity-types l and h in social identity group H are derived in an analogous way. ■

References

- Anger, S. and Heineck, G. (2010), Do smart parents raise smart children? The Intergenerational Transmission of Cognitive Abilities. *Journal of Population Economics*, **23** (3), 1105-1132.
- Akerlof, G. (1980), A Theory of Social Custom, of which Unemployment may be one Consequence. *Quarterly Journal of Economics* **94**, 749-775.
- Akerlof, G. and R. Kranton (2000), Economics and Identity. *The Quarterly Journal of Economics* **115**, 715-753.
- Akerlof, G. and R. Kranton (2002), Identity and Schooling: Some Lessons for the Economics of Education. *Journal of Economic Literature* **40**, 1167-1201.
- Aronsson, T., S. Blomquist, and H. Sacklen (1999), How to Identify Interdependent Behaviour in an Empirical Model of Labor Supply. *Journal of Applied Econometrics* **14**, 607-626.

- Aronsson, T. and O. Johansson-Stenman (2008), When the Joneses' Consumption Hurts: Optimal Public Good Provision and Nonlinear Income Taxation. *Journal of Public Economics* **92**, 986-997.
- Aronsson, T. and O. Johansson-Stenman (2010), Positional Concerns in an OLG Model: Optimal Labor and Capital Income Taxation. *International Economic Review* **51**, 1071-1095.
- Aronsson, T. and O. Johansson-Stenman (2014), Positional Preferences in Time and Space: Optimal Income Taxation with Dynamic Social Comparisons. *Journal of Economic Behavior and Organization* **101**, 1-23.
- Aronsson, T. and T. Sjögren (2010) Optimal Income Taxation and Social Norms in the Labor Market. *International Tax and Public Finance* **17**, 67-89.
- Austin, A. M. B., and Draper, D. C. (1984), The Relationship among Peer Acceptance, Social Impact, and Academic Achievement in Middle Childhood. *American Educational Research Journal*, **21** (3), 597-604.
- Becker, G. S. (2009), Human capital: A Theoretical and Empirical Analysis, with special reference to education. University of Chicago Press.
- Bernheim, D. (1994), A Theory of Conformity. *Journal of Political Economy* **102**, 1-35841-877.
- Björklund, A., and Jäntti, M. (2012), How Important is Family Background for Labor-economic Outcomes? *Labour Economics* **19** (4), 465-474.
- Blomquist, S. (1993), Interdependent Behavior and the Effect of Taxes. *Journal of Public Economics* **51**, 211-218.
- Boskin, M. J., and E. Sheshinski (1978), Individual Welfare Depends upon Relative Income. *Quarterly Journal of Economics* **92**, 589-601.
- Bowles, S., Gintis, H. (2000), Reciprocity, Self-Interest and the Welfare State. *Nordic Journal of Political Economy* **26**, 33-53.
- Bruvoll, A., Nyborg, K. (2004), The Cold Shiver of not Giving Enough: On the Social Costs of Recycling Campaigns. *Land Economics* **80**, 540-549.

- Buchmann, C., and Dalton, B. (2002), Interpersonal Influences and Educational Aspirations in 12 Countries: The Importance of Institutional Context. *Sociology of Education* **75** (2), 99-22.
- Dupor, B. and W. F. Liu (2003), Jealousy and Overconsumption. *American Economic Review* **93**, 423-428.
- Eckerstorfer, P. and R. Wendner (2013), Asymmetric and Non-atmospheric Consumption Externalities, and Efficient Consumption Taxation. *Journal of Public Economics* **106**, 42-56.
- Ermisch, J., and Francesconi, M. (2001), Family Matters: Impacts of Family Background on Educational Attainments. *Economica* **68** (270), 137-156.
- Guo, J-T. and A. Krause (2011), Optimal Nonlinear Income Taxation with Habit Formation". *Journal of Public Economic Theory* **13**, 463-480.
- Guo, J-T. and A. Krause (2013), Optimal Nonlinear Taxation of Income and Education Expenditures. *Oxford Economic Papers* **65**, 74-9.
- Hogg, M. A. (2006), Social Identity Theory, in: Burke P. J. (ed.), *Contemporary Social Psychology Theories*, pp. 111-135, Stanford University Press, Stanford, CA.
- Hogg, M. A. and Reid, S.A. (2006), Social Identity, Self-Categorization, and the Communication of Group Norms. *Communication Theory* **16**, 7-30.
- Lindbeck, A. (1995), Welfare State Disincentives with Endogenous Habits and Norms. *Scandinavian Journal of Economics* **97**, 477-494.
- Lindbeck, A., S. Nyberg, and J. Weibull (1999), Social Norms and Economic Incentives in the Welfare State. *Quarterly Journal of Economics* **114**, 1-35.
- Lindbeck, A., S. Nyberg, and J. Weibull (2003), Social Norms and Welfare State Dynamics. *Journal of the European Economic Association* **1**, 533-542.
- Ljungqvist, L. and H. Uhlig (2000), Tax Policy and Aggregate Demand Management Under Catching Up with the Joneses. *American Economic Review* **90**, 356-366.

- McGarty, C., Turner, J. C., Hogg, M. A., David, B. and Wetherell, M. S. (1992), Group Polarization as Conformity to the Prototypical Group Member. *British Journal of Social Psychology* **31**: 1–19.
- Moscovici, S. and Zavalloni, M. (1969), The Group as a Polarizer of Attitudes. *Journal of Personality and Social Psychology* **12**, 125–135.
- Oswald, A. (1983), Altruism, Jealousy and the Theory of Optimal Non-Linear Taxation. *Journal of Public Economics* **20**, 77–87.
- Stern, N.H. (1982), Optimum Taxation with Errors in Administration. *Journal of Public Economics* **17**, 181-211.
- Stiglitz, J. E. (1982), Self-Selection and Pareto Efficient Taxation. *Journal of Public Economics* **17**, 213-240.
- Tajfel, H. (1959), Quantitative Judgement in Social Perception. *British Journal of Psychology* **50**, 16–29.
- Tajfel, H. (1969), Cognitive Aspects of Prejudice. *Journal of Social Issues* **25**, 79–97.
- Tajfel, H., Billig, M., Bundy, R. P., and Flament, C. (1971), Social Categorization and Intergroup Behaviour. *European Journal of Social Psychology* **1**, 149–177.
- Tajfel, H., and Turner, J. C. (1979), An Integrative Theory of Intergroup Conflict. In Austin, W.G. and S. Worchel (eds.), *The Social Psychology of Intergroup Relations*, pp 33–47, Monterey, CA: Brooks/Cole.
- Tuomala, M. (1990), *Optimal Income Tax and Redistribution*, Oxford: Clarendon Press.
- Wendner, R. and L. H. Goulder (2008), Status Effects, Public Goods Provision, and the Excess Burden. *Journal of Public Economics* **92**, 1968-1985.
- Wentzel, K. R., and Caldwell, K. (1997), Friendships, Peer Acceptance, and Group Membership: Relations to Academic Achievement in Middle School. *Child Development* **68** (6), 1198-1209.
- Willis, R. J., and Rosen, S. (1979), Education and Self-Selection. *Journal of Political Economy* **87** (5), S7-S36.

