

# Commodity taxes and taste heterogeneity<sup>\*</sup>

Stéphane Gauthier<sup>†</sup>  
PSE and University of Paris 1

Fanny Henriet<sup>‡</sup>  
PSE-CNRS

September 19, 2017

## Abstract

We study optimal linear commodity taxes in the presence of non-linear income taxes when agents differ in skills and tastes for consumption. We show that optimal commodity taxes are partly determined by a many-person Ramsey rule when there is taste heterogeneity within income classes. The usual role of commodity taxes in relaxing incentive constraints explains the remaining part of these taxes when there is taste heterogeneity between income classes. We quantify these two parts using French consumption microdata and find that commodities taxes are only shaped by many-person Ramsey considerations.

**JEL classification numbers:** H21, D12, D82.

**Keywords:** taste heterogeneity, commodity taxes, income taxation, social valuations.

---

<sup>\*</sup>We thank Andrew Clark, Gabrielle Fack, Firouz Gahvari, Guy Laroque, Arnaud Lefranc and Catherine Sofer, and participants to the 2015 IEB workshop on economics of taxation, Paris 1 University, and the 2016 PET congress for comments. We also are very grateful to the editor Ayse Imrohroglu of this journal and two anonymous referees who provided us insightful feedback on a previous version of our paper.

<sup>†</sup>ENS, 48 bd Jourdan, 75014 Paris, France; +33(0)180521782; [stephane.gauthier@univ-paris1.fr](mailto:stephane.gauthier@univ-paris1.fr); webpage: <https://www.parisschoolofeconomics.eu/en/gauthier-stephane>

<sup>‡</sup>ENS, 48 bd Jourdan, 75014 Paris, France; [fanny.henriet@gmail.com](mailto:fanny.henriet@gmail.com).

# 1 Introduction

One common way of easing the fiscal burden on those in need is to make necessities tax-free or tax them at a lower rate than luxuries. Whether commodity taxes should be used in this way as part of the progressivity of the overall tax system is the subject to a continued debate in public finance. In a second-best world, in the case where social tastes for redistribution are ordered from the rich to the poor, the many-person Ramsey rule derived by Diamond and Mirrlees [12] indeed recommends to set taxes involving a lower discouragement of the demand for the goods more heavily consumed by the poor. However, this recommendation only applies when the government cannot use non-linear income taxation. The analysis of Atkinson and Stiglitz [2] and its generalization by Mirrlees [20] demonstrated that, if the redistribution from rich to poor can be effected via direct income taxes and transfers, the only role devoted to commodity taxation is to relax the incentive constraints implied by imperfect information about taxpayers. If, as is often assumed, labor skill and effort are private information to taxpayers, redistribution from high-skill (rich) to low-skill (poor) individuals requires high taxes on necessities when these taxes deter the high-skilled from reducing their labor effort. The low-skilled suffer from high taxes on necessities but they gain from the greater scope for income redistribution due to the relaxation of incentive constraints.

In the existing literature there is suspicion that the role of commodity taxes delineated by Atkinson and Stiglitz [2] and Mirrlees [20] is bound up with the restrictive modeling assumption that labor skill is the only dimension in which agents differ. The early contributions of Cremer and Gahvari [8] and [9] indeed provide examples where non-linear commodity taxes might play a redistributive role when there is heterogeneity in consumption tastes, as well as labor skill, although non-linear income taxation is also available. In practice, however, a non-linear commodity tax can only be applied to very specific goods. Most commodity taxes instead rely on total sales, e.g., Value-Added or sales taxes, which precludes imposition of non-linear taxes: if the tax rates were to change with the quantity purchased, buyers could avoid higher taxes by grouping or splitting their transactions. Our paper analyzes the respective roles of indirect linear (rather than non-linear) commodity taxation and direct non-linear income taxation when agents differ in terms of both labor skill and consumption tastes.

Much effort has recently been expended on the optimal shape of linear commodity taxes when individuals differ in two private-information dimensions: see in particular Cremer, Pestieau and Rochet [10] and [11], Saez [23], Blomquist and Christiansen [3], Kaplow [19], Diamond and Spinnewijn [13] or Golosov, Troshkin, Tsyvinski and Weinzierl [15]. This literature shows that the role of commodity taxes in relaxing incentive constraints continues to apply in this more general framework.<sup>1</sup>

---

<sup>1</sup>A notable exception is Saez [23] whose methodology does not explicitly refer to incentives. The exact link between his results and the role played by incentive constraints highlighted by Cremer, Pestieau and

Our paper highlights that this narrow role of commodity taxation depends on whether different types of agents can pay different income taxes. In reality this is open to question: unobserved heterogeneity given the income and other observable characteristics is a pervasive concern in the empirical literature, and should prevent the government from designing taxes relying on a comprehensive set of individual characteristics that are relevant in the tax design, e.g., those characteristics contributing to the formation of consumption tastes. A useful theoretical modeling that can be viewed as a shortcut to account for the presence of remaining unobserved heterogeneity given the observables obtains by considering a limited number of (publicly observed) occupations or jobs relative to the total number of different agents' (taste) types, as in Diamond and Spinnewijn [13]. Some agents with different tastes then are necessarily grouped together into the same income class. Income taxation can be used to address heterogeneity between income classes, but it is obviously no help for achieving a finer redistribution within income classes: this role is taken up by commodity taxation.

Our main result is to show that, in the presence of a limited number of occupations, optimal commodity taxes are partly shaped by a version of the many-person Ramsey rule. That is, commodity taxes are determined by the relationship between consumption and consumers' social valuations within each income class. This result provides, in sharp contrast with the existing literature, a justification for taxing more heavily the goods that are consumed in a greater proportion by agents whose social valuation is low. It is derived from a new rule for commodity taxes that decomposes the optimal discouragement into the sum of two terms: a many-person Ramsey component and a remaining part dedicated to relaxing incentive constraints. The many-person Ramsey component is driven by heterogeneity within income classes. It recommends to discourage consumption of those agents with low social valuations, even in the presence of a general non-linear income tax, therefore reinforcing some progressivity in the tax system.

An empirical illustration on French data shows that consumption taxes over the period 2010-11 are only governed by many-person Ramsey considerations, rather than incentives. We identify two taste clusters relying on household budget shares: the first one comprises old low-education individuals, often retired and living in medium-size cities while the second cluster consists of younger urban active individuals. Referring to the classification into poor, middle and rich income classes, we find taste heterogeneity in each income class. Relying on an assumption of optimal taxes we propose a new empirical test for the detection of the relevant incentive constraints and compute the social valuations of taste clusters for every income classes. This exercise extends the revealed social preferences approach started by Christiansen and Jansen [5] to the case where incentive considerations implied by nonlinear income taxation matter. We find evidence of redistribution across the two clusters at the bottom of the income distribution, but not among richer households. Redistribution among

---

Rochet [10] is unclear in the existing literature (see, e.g., footnote 8 in Saez [23]). Our paper shows how one can alter Cremer, Pestieau and Rochet [10] to obtain results whose spirit is closer to Saez [23].

75 the poor appears to favor the former taste group, which consists of older low-education individuals. This pattern of redistribution echoes the public debate about intergenerational equity in France suggesting that the overall tax system could benefit older households. Our results indeed point in this same direction but narrow the horizontal equity concerns to the poor income class.

80 The paper is organized as follows. Sections 2 and 3 describe the setup with agents differing in two dimensions and a limited number of occupations, and Section 4 provides our optimal rule of commodity taxes. Section 5 is devoted to the empirical illustration on French consumption microdata that quantifies the relative importance of the many-person Ramsey component of commodity taxes.

## 85 2 General setup

We consider a closed economy endowed with a population of agents who differ in terms of their labor skill  $i$  ( $i = 1, \dots, I$ ) and consumption tastes  $j$  ( $j = 1, \dots, J$ ). There are  $n_{ij}$  type  $ij$  agents and the total population size is normalized to 1. The preferences of a type- $ij$  agent are represented by the continuous utility function

$$U_i(V_j(x), y), \tag{1}$$

90 where  $x \in \mathbb{R}^+$  is a bundle of consumption goods,  $y$  is a non-negative real number that stands for pre-tax labor income, and  $V_j$  is a sub-utility function referring to consumption. The functions  $U_i$  and  $V_j$  satisfy the standard monotonicity and convexity properties. The class of utility functions (1) is identical to the one considered by Cremer, Pestieau and Rochet [10]. It generalizes the cornerstone analysis of Atkinson and Stiglitz [2] where all the agents  
95 have the same consumption tastes,  $V_j(x) = V(x)$  for every bundle  $x$ . The separability assumption embodied in (1) implies that the marginal rates of substitution between any two consumption goods are independent of both labor ability and pre-tax labor income. The rates of substitution may however vary with consumption and they possibly differ across agents. Separability is a convenient way to distinguish clearly taste heterogeneity  
100 from other dimensions of unobservable individual heterogeneity that are represented by the index  $i$ .<sup>2</sup> Although non-separability in individual preferences sounds more plausible, the empirical evidence is mixed. The spirit of our results should go through in the situation where labor supply does not have a significant impact on the rates of substitution across goods.<sup>3</sup>

---

<sup>2</sup>The interpretation of the index  $i$  as being labor skill and  $y$  taxable labor income is in line with the literature following Atkinson and Stiglitz [2] and Cremer, Pestieau and Rochet [10]. Still it can be viewed as a more general proxy for individual heterogeneity unrelated to consumption tastes. With this last interpretation  $y$  represents taxable income.

<sup>3</sup>In their recent synthesis for the Mirrlees review Crawford, Keen and Smith [7] reject weak separability between labor and consumption but they note on page 288 that ‘this finding needs to be interpreted with

105 The government observes individual income and the aggregate consumption of each good. Labor skill and consumption tastes remain private information to the agents. The government can use linear consumption taxes and non-linear income taxes. Consumption taxes are given by  $q-p$ , where  $q$  and  $p$  respectively are the vectors of consumer and producer prices. An agent with pre-tax income of  $y$  has post-tax income of  $R(y)$ .

110 Income heterogeneity appears via the income classes  $k$  ( $k = 1, \dots, K$ ). Following Diamond and Spinnewijn [13] an income class may be viewed as some job or occupation. All the agents in class  $k$  earn the same pre-tax income of  $y_k$  and post-tax income of  $R_k = R(y_k)$ , independently of their labor skill. For illustrative purpose one can think of an occupation as a given predetermined task, e.g., build a house, that yields some payment independently  
 115 of the effort spent to perform it: some certainly need work harder than others to complete the task. The total number  $K$  is exogenously given and it can take any value between 1 and  $IJ$ . When  $K < IJ$  there remains unobserved individual heterogeneity controlling for the income level.<sup>4</sup> The diversity of offered tasks, and thus possible payments, then cannot match workers' diversity.

120 Given  $(y, R)$  a type- $ij$  agent chooses a consumption bundle  $\xi_j(q, R)$  maximizing  $V_j(x)$  subject to the budget constraint  $q \cdot x \leq R$ . Separability between consumption and labor in (1) makes consumption choices independent of labor ability. The agent obtains conditional indirect subutility  $V_j(\xi_j(q, R)) \equiv V_j(q, R)$  from consumption, with a slight abuse of notation. A type- $ij$  agent in class  $k$  has utility  $U_i(V_j(q, R_k), y_k)$  and thus self-selects into  
 125 this class if and only if

$$U_i(V_j(q, R_k), y_k) \geq U_i(V_j(q, R_{k'}), y_{k'}) \quad (2)$$

for all  $k'$ .

A tax system is defined by a vector  $q$  of consumer prices, an income tax profile  $(y_k, R_k)$  and an allocation rule  $(\mu_{ijk})$ , where  $\mu_{ijk}$  equals 1 if  $ij$  agents are assigned to class  $k$ , and 0 otherwise. This satisfies incentive compatibility if (2) holds for each type  $ij$  and class  $k$   
 130 such that  $\mu_{ijk} = 1$ . It is feasible when

$$\sum_{jk} n_{jk} [(q - p) \cdot \xi_j(q, R_k) + (y_k - R_k)] \geq 0, \quad (3)$$

---

caution.<sup>7</sup>

<sup>4</sup>The precise circumstances yielding a certain number  $K$  of different income levels in the economy are irrelevant to our analysis. An alternative equivalent interpretation involves self-selection of agents into  $K \leq IJ$  income classes in a economy where there are at least  $IJ$  classes. Agents with different characteristics then earn the same before-tax income. This bunching interpretation would require minor changes in the argument developed in Section 3 to deal with incentive constraints in the presence of empty income classes.

where

$$n_{jk} \equiv \sum_i n_{ij} \mu_{ijk}$$

is the number of taste  $j$  agents in class  $k$ .<sup>5</sup>

A tax system is socially optimal when it maximizes some social objective subject to the incentive and feasibility constraints (2) and (3). In what follows, the social objective is assumed to be Paretian, i.e., it is strictly increasing with agents' utilities.<sup>6</sup>

### 135 3 A sub-program

Consider some 'reference' tax system  $((\bar{\mu}_{ijk}), \bar{q}, (\bar{y}_k, \bar{R}_k))$  satisfying (2) and (3) for every type  $ij$  and income class  $k$  such that  $\bar{\mu}_{ijk} = 1$ . We are looking for necessary conditions for the reference tax system to be socially optimal.

We do not solve the whole problem for a socially optimal tax system. Instead we derive 140 conditions for  $(\bar{q}, (\bar{R}_k))$  to be optimal conditionally to  $((\bar{\mu}_{ijk}), (\bar{y}_k))$ . The allocation  $(\bar{\mu}_{ijk})$  of individuals to income classes and the before-tax income profile  $(\bar{y}_k)$  are arbitrarily given and possibly suboptimal. Hence the tax system  $((\bar{\mu}_{ijk}), \bar{q}, (\bar{y}_k, \bar{R}_k))$  is not necessarily socially optimal.

We appeal to the standard tax reform methodology. The reforms under scrutiny involve 145 tax systems  $((\bar{\mu}_{ijk}), q, (\bar{y}_k, R_k))$  satisfying incentive requirements (2) and feasibility (3) for every type  $ij$  and income class  $k$  such that  $\bar{\mu}_{ijk} = 1$ .

We suppose that the economy is locally non-satiated at  $((\bar{\mu}_{ijk}), \bar{q}, (\bar{y}_k, \bar{R}_k))$ : a small amount of additional resources would allow the government to achieve a weak Pareto improvement without violating the incentive-compatibility requirements. Under this assumption, there is no  $((\bar{\mu}_{ijk}), q, (\bar{y}_k, R_k))$  improving locally upon  $((\bar{\mu}_{ijk}), \bar{q}, (\bar{y}_k, \bar{R}_k))$  if, given 150  $((\bar{\mu}_{ijk}), (\bar{y}_k))$ ,  $(q, (R_k))$  locally maximizes the collected taxes subject to (2) for all  $ijk$  such that  $\bar{\mu}_{ijk} = 1$ , and

$$U_i(V_j(q, R_k), \bar{y}_k) \geq U_i(V_j(\bar{q}, \bar{R}_k), \bar{y}_k) \quad (4)$$

for all  $ijk$  such that  $\bar{\mu}_{ijk} = 1$ . Inequalities (4) ensure that no agent suffers from the reform.

To deal with incentive requirements (2), consider some type  $ij$  and two income classes 155  $k$  and  $k'$  such that  $\bar{\mu}_{ijk'} = 1$  (note that type  $ij$  agents are assumed to be assigned to income class  $k'$  in the reference situation). In the reference situation, the incentive constraint

$$U_i(V_j(\bar{q}, \bar{R}_{k'}), \bar{y}_{k'}) \geq U_i(V_j(\bar{q}, \bar{R}_k), \bar{y}_k) \quad (5)$$

is satisfied. If the inequality (5) is strict, then any tax system where  $(\bar{q}, (R_k))$  is close

---

<sup>5</sup>Throughout the paper, the letters  $i, j$  and  $k$  respectively represent labor skill, consumption taste, and income class. For instance  $n_{ij}$  is the number of agents with labor skill  $i$  and taste  $j$ , while  $n_{jk}$  is the number of agents with taste  $j$  in income class  $k$ .

<sup>6</sup>See Remark 2 in Section 5.3 for a discussion of the Rawlsian case.

enough to  $(q, (\bar{R}_k))$  preserves incentive compatibility. If, however, the incentive constraint (5) holds at equality in the reference situation,

$$U_i(V_j(\bar{q}, \bar{R}_{k'}), \bar{y}_{k'}) = U_i(V_j(\bar{q}, \bar{R}_k), \bar{y}_k), \quad (6)$$

160 then incentive compatibility requirements in the new tax system  $((\bar{\mu}_{ijk}), q, (\bar{y}_k, R_k))$  may be violated. For such a type  $ij$  and these two income classes  $k$  and  $k'$  the alternative tax system preserves incentive compatibility if and only if

$$U_i(V_j(q, R_{k'}), \bar{y}_{k'}) \geq U_i(V_j(q, R_k), \bar{y}_k). \quad (7)$$

Since  $U_i(V_j(q, R_{k'}), \bar{y}_{k'}) \geq U_i(V_j(\bar{q}, \bar{R}_{k'}), \bar{y}_{k'})$  by (4), a sufficient condition for (7) is

$$U_i(V_j(\bar{q}, \bar{R}_{k'}), \bar{y}_{k'}) \geq U_i(V_j(q, R_k), \bar{y}_k). \quad (8)$$

Hence, for this type  $ij$  agent (with  $\bar{\mu}_{ijk'} = 1$ ) and this income class  $k$ , a sufficient condition 165 for the inequality (8) to be met is, using (6),

$$U_i(V_j(q, R_k), \bar{y}_k) \leq U_i(V_j(\bar{q}, \bar{R}_k), \bar{y}_k). \quad (9)$$

We are now in a position to derive necessary conditions for conditional optimality of  $(\bar{q}, (\bar{R}_k))$ . Let  $\bar{V}_{jk} \equiv V_j(\bar{q}, \bar{R}_k)$  stand for the sub-utility of a type- $ij$  agent under the reference tax system. Under non-satiation, the reference tax system is optimal only if, given the allocation  $(\bar{\mu}_{ijk})$  and pre-tax income levels  $(\bar{y}_k)$ , the tax tools  $(q, (R_k)) = (\bar{q}, (\bar{R}_k))$  maximize collected taxes

$$\sum_{jk} n_{jk} [(q - p) \cdot \xi_j(q, R_k) + (\bar{y}_k - R_k)]$$

subject to

$$V_j(q, R_k) \geq \bar{V}_{jk} \quad (10)$$

for all  $ijk$  such that  $\bar{\mu}_{ijk} = 1$ , and

$$V_j(q, R_k) \leq \bar{V}_{jk} \quad (11)$$

for all  $ijk$  such that  $U_i(\bar{V}_{jk}, \bar{y}_k) = U_i(\bar{V}_{jk'}, \bar{y}_{k'})$  and  $\bar{\mu}_{ijk'} = 1$  for some  $k' \neq k$ .

The inequalities (10) and (11) respectively obtain from (4) and (9), using the fact that 170  $U_i$  is increasing in  $V_j$ . The choice of commodity taxes and after-tax incomes in the original program with two dimensions of individual heterogeneity can thus be made referring to the above sub-program with only one dimension of individual heterogeneity: only taste heterogeneity is relevant in (10) and (11). This simplification is made possible thanks to the separability assumption embodied in (1), which allows us to get rid of heterogeneity in 175 labor skill  $i$  in the treatment of incentive constraints.

We solve this problem using the Lagrangian approach. For this approach to be valid we require the qualification of the active constraints in the reference situation, i.e., the matrix whose  $jk$ -th row is  $\nabla V_j(\bar{q}, \bar{R}_k)$  has rank  $JK$ . Then there exist  $JK$  Lagrange multipliers  $n_{jk}\lambda_{jk} - \tilde{n}_{jk}\gamma_{jk}$  such that the reference tax system is optimal only if  $(\bar{q}, (\bar{R}_k))$  is a local extremum of the Lagrangian function

$$\sum_{jk} n_{jk} [(q - p) \cdot \xi_j(q, R_k) - R_k] + \sum_{jk} (n_{jk}\lambda_{jk} - \tilde{n}_{jk}\gamma_{jk}) [V_j(q, R_k) - \bar{V}_{jk}],$$

where

$$n_j \equiv \sum_k n_{jk} \quad \text{and} \quad \tilde{n}_{jk} \equiv \sum_i \sum_{k' \neq k} n_{ij} \bar{\mu}_{ijk'} \mathbb{1} [U_i(\bar{V}_{jk'}, \bar{y}_{k'}) = U_i(\bar{V}_{jk}, \bar{y}_k)]$$

are respectively the number of taste  $j$  agents, and the number of taste  $j$  agents assigned to class  $k' \neq k$  who contemplate switching to  $k$ .

In what follows we set  $n_{jk}\lambda_{jk} \geq 0$  and  $\tilde{n}_{jk}\gamma_{jk} = 0$  when the multiplier  $n_{jk}\lambda_{jk} - \tilde{n}_{jk}\gamma_{jk}$  is non-negative, and  $\tilde{n}_{jk}\gamma_{jk} > 0$  and  $n_{jk}\lambda_{jk} = 0$  otherwise. This convention makes it clear that for any given  $jk$  the only relevant (binding) constraint is (10) if  $n_{jk}\lambda_{jk} \geq 0$ , and (11) if  $\tilde{n}_{jk}\gamma_{jk} > 0$ . Type  $jk$  agents are ‘envied’ (or mimicked) when  $\tilde{n}_{jk}\gamma_{jk} > 0$ .

## 4 A Many-Person Ramsey Rule

The first-order condition for  $q^h$  to be optimal is, using  $\alpha_{jk}$  for the marginal utility of income of a taste  $j$  agent in class  $k$ ,

$$\sum_{jk} n_{jk} \left( \xi_{jk}^h + \sum_{\ell} t^{\ell} \frac{\partial \xi_{jk}^{\ell}}{\partial q^h} \right) - \sum_{jk} (n_{jk}\lambda_{jk} - \tilde{n}_{jk}\gamma_{jk}) \alpha_{jk} \xi_{jk}^h = 0,$$

where all of the variables are evaluated at the reference tax system.

The Lagrange multiplier  $\lambda_{jk}$  associated with the constraint (10) reflects the social desire to increase the utility of a  $jk$  person. One can therefore measure the marginal social valuation of a 1 unit income transfer toward a  $jk$  person by

$$\beta_{jk} = \sum_{\ell} t^{\ell} \frac{\partial \xi_{jk}^{\ell}}{\partial R} + \lambda_{jk} \alpha_{jk}, \tag{12}$$

which we call the ‘intrinsic’ social valuation of one taste  $j$  agent assigned to class  $k$ .

The same income transfer also benefits taste  $j$  agents allocated to another class  $k' \neq k$  when they switch to  $k$ . The impact on the objective of the resulting tightening of the associated incentive constraint is measured by

$$\tilde{\beta}_{jk} = \gamma_{jk} \alpha_{jk} \tag{13}$$

where  $\gamma_{jk}$  is the Lagrange multiplier associated with the sufficient condition (11) for incentive constraints to be met.

As a result the total social value of a one-unit income transfer toward each taste  $j$  agent in class  $k$ , which we call the ‘consolidated’ social valuation, is

$$b_{jk} = n_{jk}\beta_{jk} - \tilde{n}_{jk}\tilde{\beta}_{jk}. \quad (14)$$

195 The consolidated social valuation is equal to the intrinsic social valuation of taste  $j$  agents in class  $k$  net of the (non-negative) incentive correction  $\tilde{n}_{jk}\tilde{\beta}_{jk}$ .

Appealing to the Slutsky properties, the first-order condition in  $q^h$  can be rewritten as

$$\sum_{\ell} t^{\ell} \frac{\partial \hat{\xi}^h}{\partial q^{\ell}} = -\xi^h + \sum_{jk} b_{jk} \xi_{jk}^h \quad (15)$$

where

$$\xi^h \equiv \sum_{jk} n_{jk} \xi_{jk}^h$$

represents the aggregate demand for good  $h$ , and  $\hat{\xi}^h$  is the aggregate compensated demand for this good. Using the first-order condition in the post-tax income chosen for class  $k$ ,

$$\sum_j b_{jk} = \sum_j n_{jk} \equiv n_k, \quad (16)$$

200 the first-order condition in  $q^h$  finally yields a version of the many-person Ramsey rule that is set out in Proposition 1.

**Proposition 1.** *Consider some given (possibly sub-optimal) allocation rule  $(\bar{\mu}_{ijk})$  and some given (possibly sub-optimal) profile  $(\bar{y}_k)$  of before-tax income. Optimal commodity taxes are such that, for every consumption good  $h$ ,*

$$\sum_{\ell} t^{\ell} \frac{\partial \hat{\xi}^h}{\partial q^{\ell}} = \sum_k n_k \Phi_k^h \quad (17)$$

where

$$\Phi_k^h \equiv \sum_j \frac{b_{jk}}{n_k} \xi_{jk}^h - \sum_j \frac{n_{jk}}{n_k} \xi_{jk}^h = \text{cov} \left( \frac{b_{jk}}{n_{jk}}, \xi_{jk}^h \right).$$

205 The covariance  $\Phi_k^h$  is positive when the agents in class  $k$  who have high consolidated social valuations like good  $h$ . Rule (17) closely resembles the standard many-person Ramsey rule, with  $\Phi_k^h$  being the distributive factor for good  $h$ . In the typical textbook formulation, however, intrinsic social valuations replace the consolidated social valuations in the expression of the distributive factors. The consolidated social valuation of an agent under-

210 estimates her intrinsic social valuation when an income transfer toward this agent tightens  
incentive constraints: her situation gets more desirable to some agents in other income  
classes. This may help explain why the empirical literature following Ahmad and Stern [1]  
has sometimes produced negative estimates of social values consistent with the many-person  
Ramsey rule. The reason would be that this literature in fact has produced estimates of  
215 the consolidated social valuations, rather than the intrinsic valuations, which explains the  
discrepancy.

If some taste  $j$  agents in class  $k' \neq k$  contemplate switching to  $k$ , the covariance  $\Phi_k^h$   
involves the hypothetical consumption levels that  $jk'$  agents would have in class  $k$ . In order  
to disentangle effective (observed) consumption from the hypothetical consumption of the  
mimickers we require a more suitable form of (17). Note that

$$\Phi_k^h = \text{cov}(\beta_{jk}, \xi_{jk}^h) - \text{cov}\left(\frac{\tilde{n}_{jk}}{n_{jk}}\tilde{\beta}_{jk}, \xi_{jk}^h\right),$$

where

$$\text{cov}(\beta_{jk}, \xi_{jk}^h) \equiv \phi_k^h = \sum_j \frac{n_{jk}}{n_k} \beta_{jk} \xi_{jk}^h - \sum_j \frac{n_{jk}}{n_k} \beta_{jk} \sum_j \frac{n_{jk}}{n_k} \xi_{jk}^h \quad (18)$$

fits the standard concept of distributive factor of good  $h$  (here this is calculated within  
class  $k$ ) relying on intrinsic social valuations. Similarly,

$$\text{cov}\left(\frac{\tilde{n}_{jk}}{n_{jk}}\tilde{\beta}_{jk}, \xi_{jk}^h\right) = \sum_j \frac{\tilde{n}_{jk}}{n_k} \tilde{\beta}_{jk} (\xi_{jk}^h - \xi_k^h) \quad (19)$$

where

$$\xi_k^h \equiv \sum_j \frac{n_{jk}}{n_k} \xi_{jk}^h$$

220 is the average (observed) demand for good  $h$  in class  $k$ . A type- $ij$  agent in class  $k' \neq k$   
would have the same consumption  $\xi_{jk}^h$  as the taste  $j$  agents initially assigned to class  $k$ .  
The covariance in (19) is thus a weighted sum of the differences between the hypothetical  
consumption of good  $h$  by agents who contemplate switching to class  $k$  and the actual  
consumption of this same good by agents who are actually allocated to class  $k$  in the  
225 reference situation. This is the ‘incremental demand of the mimickers’ that appears in  
Guesnerie [17] and Cremer, Pestieau and Rochet [10].

Proposition 1 can then be rewritten as follows:

**Proposition 2.** *Consider some given (possibly sub-optimal) allocation rule  $(\bar{\mu}_{ijk})$  and  
some given (possibly sub-optimal) profile  $(\bar{y}_k)$  of before-tax income. Optimal commodity*

230 *taxes are such that, for every consumption good  $h$ ,*

$$\sum_{\ell} t^{\ell} \frac{\partial \hat{\xi}^h}{\partial q^{\ell}} = \sum_k n_k \phi_k^h - \sum_{jk} \tilde{n}_{jk} \tilde{\beta}_{jk} (\xi_{jk}^h - \xi_k^h), \quad (20)$$

where  $\phi_k^h$  is the within class  $k$  distributive factor for good  $h$ .

Proposition 2 is the main result of our paper. It yields the rules obtained in the previous literature as special cases.

1. Incentive considerations are irrelevant in the single class  $K = 1$  case, i.e.  $\tilde{n}_{jk} \tilde{\beta}_{jk} = 0$  for all  $jk$ . The income tax then degenerates to a uniform lump-sum tax and the rule given in Proposition 2 can be simplified to

$$\sum_{\ell} t^{\ell} \frac{\partial \hat{\xi}^h}{\partial q^{\ell}} = \sum_k n_k \phi_k^h = \phi_1^h.$$

235 This coincides with the familiar many-person Ramsey rule (when a uniform income tax is allowed). The discouragement of the compensated demand for good  $h$  should equal the distributive factor of this good. Indirect taxation is useful as long as there is taste heterogeneity in the (assumed unique) income class.

2. When there are  $K = IJ$  different classes,  $\phi_k^h = 0$  for all  $h$  and  $k$ , and the optimal rule for indirect taxation given in (20) becomes

$$\sum_{\ell} t^{\ell} \frac{\partial \hat{\xi}^h}{\partial q^{\ell}} = - \sum_{jk} \tilde{n}_{jk} \tilde{\beta}_{jk} (\xi_{jk}^h - \xi_k^h). \quad (21)$$

240 Now the discouragement only relies on the incremental net demand of the mimickers, as in Mirrlees [20], Guesnerie [17], and Cremer, Pestieau and Rochet [10] and [11]. The formula in equation (21) is of the same kind as those derived by Saez [23] from his Assumption 3. It also applies in the special case where tastes are perfectly correlated with skill, as in Golosov, Troshkin, Tsyvinski and Weinzierl [15]. Indirect taxation is  
 245 useless in the homogeneous-taste case,  $V_j(x) = V(x)$  for all  $j$ , since then  $\xi_{jk}^h = \xi_k^h$  for all  $j$ , which is Atkinson and Stiglitz [2] theorem. It is also useless in the absence of incentive problems ( $\tilde{n}_{jk} \tilde{\beta}_{jk} = 0$  for all  $jk$ ). Otherwise consumption of a given good should be discouraged when agents who contemplate switching to  $k$  have a higher consumption of this good than the agents who are actually assigned to class  $k$ . The  
 250 role of indirect taxes is to relax the incentive constraints, which is possible only if agents in different income classes also have different tastes, i.e., in the presence of heterogeneity across income classes.

The considerations resulting from these two polar cases come together in a surprisingly simple manner in the more general configuration where unobserved heterogeneity given the income matters ( $1 \leq K < IJ$ ). Optimal discouragement then equals the sum of the distributive factors over the different income classes and a measure of the excess consumption of mimickers.

In each configuration income taxes address agent heterogeneity across income classes. Indirect taxes are useful under two kinds of circumstances. First, they help address equity as soon as individuals allocated to the same income class have different consumption tastes, as they then allow the tax authority to manage within income class taste heterogeneity. Second, they ensure the relaxation of the incentive constraints when individuals who are allocated to different income classes have different tastes, i.e. when there is between-class taste heterogeneity.

Proposition 2 shows that indirect taxation is useless in the absence of taste heterogeneity ( $J = 1$ ) since then  $\phi_k^h$  is zero for all income class  $k$  and consumption good  $h$ . Therefore, under the separability assumption embodied in the utility function (1), the Atkinson and Stiglitz theorem still holds when the before-tax income is sub-optimal and when it does not reflect individual heterogeneity in labor characteristics of the taxpayers ( $K < I$ ).

When taste heterogeneity matters ( $J > 1$ ) the considerations coming from the many-person Ramsey rule partly affect optimal indirect taxes. Still we cannot conclude that indirect taxes should always serve to reinforce the progressivity of the tax system, with luxuries being more heavily taxed than necessities. In particular there is pressure for high taxes on necessities in the redistributive case where agents from the higher income classes have lower social values and like necessities more than do the less well-off (when endowed with lower income, they would consume more necessities than do the poor).

## 5 An illustration on data from France

One appealing feature of the tax rule given in Proposition 2 is to involve sufficient statistics that disentangle clearly redistribution within and between income classes. A version of the many-person Ramsey rule applies to within income class redistribution while incentives instead relate to heterogeneity between income classes. The rule given in Proposition 2 bears on relatively weak theoretical requirements: it does not rely on a special formulation for the (Paretian) social welfare function and it applies given any arbitrary (possibly sub-optimal) occupational choices and pre-tax income distribution. It seems therefore well suited to meet the data. In this section we use the French household expenditures survey ‘Budget de Famille’ for year 2011 to obtain an order of magnitude for the various social valuations used in the theoretical part and eventually assess how much the many-person Ramsey considerations are relevant in the design of consumption taxes in France.

## 5.1 Taste heterogeneity

290 The ‘Budget de Famille’ survey provides us with physical quantities and household expenditures on consumption items disaggregated at the 5-digit COICOP international classification.<sup>7</sup> Expenditures are aggregated into 8 broad categories consistent with the 2-digit level of COICOP grouping: Food (01), Alcoholic beverages, tobacco and narcotics (02), Clothing and footwear (03), Furnishings, household equipment and routine household maintenance  
295 (05), Transport (07), Leisure, Communication and Culture (08 and 09), Restaurants and hotels (11) and Other goods and services (12). For each broad category we compute a Stone price index from the unit values (ratios of expenditures over quantities) of the items in the category.

We consider three income classes in accordance with the usual decomposition into poor  
300 (which consist of the bottom three deciles of the per consumption unit income distribution), middle (the next four deciles) and rich (the remaining three top deciles).

In order to identify taste groups we cluster the household budget shares into  $J = 2$  groups. In a first step each point (a vector of 8 budget shares) is randomly assigned to a cluster and we compute the geometric center of each cluster. In a second step, we reassign  
305 points to the cluster whose step 1 center is the closest and we compute the new geometric center of each cluster. The reassignment is repeated until within cluster variation cannot be reduced any further. The within cluster variation is calculated as the sum of the euclidean distance between the points and their respective cluster centers. Table 4 shows that the algorithm succeeds very well in reducing within cluster variation of Food. It also performs  
310 well for Transport and Restaurants and hotels. These are the three categories that display the greatest difference of average budget shares across the two clusters. Taste 1 households food expenditure is more than twice the amount spend by Taste 2 on food items. For the remaining categories the within cluster variance of budget shares has about the same magnitude as the variance of budget shares. Table 4 also shows that the two clusters face  
315 similar prices and Table 5 highlights that for every income class the average income is about identical across clusters. This suggests that food behavior differences across clusters can be plausibly attributed to taste heterogeneity.

Tables 1 provides summary statistics about the two taste groups. Food consumers (Taste 1) are older, less educated and they live in medium-sized cities. The importance  
320 of age for shaping food consumption is well documented for France in the recent sociology literature, e.g., Plessz and Gojard [22]. Taste 2 households instead work in intermediate or high skill professions in densely populated cities.

The tax rates on the various categories of goods for year 2011 collected by the Insti-

---

<sup>7</sup>Our initial sample consists of the 10,342 surveyed households living in continental France. From this sample we keep observations such that the family head is between 18 and 80 years old and is not self-employed. To fit our theoretical setup we neutralize household size heterogeneity by dividing household expenditures by the number of consumption units within the household. All the results thus apply to single (childless) equivalent adults.

Table 1: SUMMARY STATISTICS

	Taste 1	Taste 2
Age <sup>1</sup>	54.33	46.20
Nb of persons	2.30	2.20
Nb of children	0.67	0.65
Education		
College diploma	11.0	21.7
Upper secondary education	20.5	29.7
Vocational training certificates	44.7	36.1
No diploma	23.8	12.5
Professional category		
Farmers	0.1	0.1
Businessmen, craftsmen, shopkeepers	2.5	2.9
Others inactive	6.5	6.0
Senior executives and higher skill professions	7.1	16.1
Employees	13.7	16.1
Workers	16.7	16.7
Technicians and associate (intermediate) professionals	12.4	20.0
Retirees	40.8	22.0
Population in the area		
Below 20,000	18.6	19.0
20,000-200,000	42.6	39.9
Above 200,000	38.8	41.2

Note 1: Age of the head of household (in years).

tut des Politiques Publiques are reported in Table 4. Food is subject to low taxes, which  
 325 suggests that food consumers (Taste 1) could be socially favored. From Proposition 2 we  
 know however that redistributive statements must be treated with caution: such statements  
 are relevant only if a significant impact of within income class heterogeneity is found in  
 consumption taxes. To assess the respective contributions of within and between income  
 classes heterogeneity, we shall proceed in two steps. First we exploit Proposition 1 to com-  
 330 pute the consolidated social valuations that best fit the tax rule (17). In a second step, we  
 provide a simple new test that detects binding incentive constraints and gives the incentive  
 corrections. The contributions of within and between income class heterogeneity then can  
 be obtained from (14) by splitting consolidated social valuations into intrinsic social valua-  
 tions and incentive corrections, which eventually allows us decompose the discouragement  
 335 in line with Proposition 2.

## 5.2 Consolidated social valuations

We start by recovering the consolidated social valuations from the first-order conditions  
 (15) and (16) for optimal taxes. Referring to ad valorem tax rates ( $t_{\text{val}}^h$ ), obtained by  
 equalizing  $t^h/q^h$  and  $t_{\text{val}}^h/(1+t_{\text{val}}^h)$ , the first-order condition (15) associated with category  
 340  $h$  is

$$\sum_{\ell} \frac{t_{\text{val}}^{\ell}}{1+t_{\text{val}}^{\ell}} \hat{\varepsilon}^{h\ell} = -1 + \sum_{jk} b_{jk} \frac{q^h \xi_{jk}^h}{q^h \xi^h}, \quad (22)$$

where the aggregate elasticity

$$\hat{\varepsilon}^{h\ell} = \sum_{jk} n_{jk} \frac{q^h \xi_{jk}^h}{q^h \xi^h} \varepsilon_{jk}^{h\ell}$$

is a weighted sum of the elasticities  $\hat{\varepsilon}_{jk}^{h\ell}$  of  $jk$  household compensated demand for category  
 $h$  with respect to price  $q_{\ell}$ . The weights in the sum are the ratios of  $jk$  expenditures  $n_{jk} q^h \xi_{jk}^h$   
 over total expenditures  $q^h \xi^h$  for category  $h$ . The elasticities  $\hat{\varepsilon}_{jk}^{h\ell}$  obtain from the estimation  
 of taste  $j$  demand functions, assuming that they obey an AIDS specification, and computed  
 345 at the average income of class  $k$ .<sup>8</sup>

Replacing into (22) the profile ( $b_{1k}$ ) obtained from (16),

$$b_{1k} = \sum_j n_{jk} - b_{2k} \equiv n_k - b_{2k} \quad (23)$$

for all  $k$ , one can finally compute the values ( $b_{2k}^*$ ) of the consolidated social valuations ( $b_{2k}$ )

---

<sup>8</sup>Table 4 and 5 gives information about price indexes and income for each taste group. Taste groups  
 face about the same prices for every consumption category, and they have about the same income whatever  
 the income class  $k$  is. Details about the demand estimation are given in Appendix B.

that minimize

$$\sum_h \left( \sum_\ell \frac{t^\ell}{q^\ell} \hat{\varepsilon}^{h\ell} + 1 - \sum_k n_k \frac{q^h \xi_{1k}^h}{q^h \xi^h} - \sum_k b_{2k} \frac{q^h \xi_{2k}^h - q^h \xi_{1k}^h}{q^h \xi^h} \right)^2. \quad (24)$$

The corresponding estimates ( $b_{1k}^*$ ) of ( $b_{1k}$ ) satisfy (23), where ( $b_{2k}$ ) is replaced with ( $b_{2k}^*$ ) for every  $k$ . We disregard in (24) the first-order conditions associated with Alcoholic beverages, tobacco and narcotics and Transport since these two categories are likely to be taxed for reasons other than efficiency and equity, e.g., public health or environmental considerations. In (24) the index  $h$  thus covers  $8 - 2 = 6$  categories of goods while  $\ell$  still covers all the 8 categories to allow for possible substitution with the two categories Alcoholic beverages, tobacco and narcotics and Transport.

Table 2: SOCIAL VALUATIONS

Taste group ( $j$ )	Taste 1			Taste 2		
Income class ( $k$ )	Poor	Middle	Rich	Poor	Middle	Rich
Percentage of equivalent adults ( $n_{jk}$ )	(15)	(18)	(9)	(21)	(17)	(20)
Consolidated valuation ( $b_{jk}^*/n_{jk}$ )	1.134	0.979	1.050	0.882	1.018	0.978
Incentive statistics ( $s_{jk}^*$ )	-0.147	-0.146	-0.079	-0.117	-0.17	-0.154
Intrinsic valuation ( $\beta_{jk}^*/n_{jk}$ )	1.134	0.979	1.050	0.882	1.018	0.978

The first row of Table 2 reports the profile ( $b_{jk}^*/n_{jk}$ ). The ratio  $b_{jk}^*/n_{jk}$  measures the gross social value of a one euro income transfer toward one equivalent adult of group  $jk$ . Such a transfer costs 1 euro, and so it is socially profitable if  $b_{jk}^*/n_{jk}$  is greater than 1. We find variations in the consolidated social valuations across tastes among the poor only. Social valuations are about identical across tastes among richer adults. Redistribution among the poor favors Taste 1 at the expense of Taste 2: the net social gain of a 1 euro income transfer toward a poor Taste 1 is 0.13 euro while the same transfer toward a poor Taste 2 has net cost of 0.12 euro. Redistribution through consumption taxes in France thus tends to play as a transfer from young educated non-food poor consumers to the remaining older less educated food poor consumers.

REMARK 1. If, as we have assumed, the government chooses the taxes maximizing social welfare, then (24) computed at ( $b_{jk}^*$ ) should be close to 0. We find a value of 0.036 at this point. We have reproduced the same exercise 1, 000 times starting from tax rates randomly drawn between 0 and 1, i.e., we have computed consolidated social valuations satisfying (23) that minimize (24) for such arbitrary taxes. From random taxes the average value of (24) equal to 0.48, more than 10 times higher than 0.036. This makes plausible the

hypothesis that the government indeed optimizes for the various taxes in France.  $\square$

### 5.3 Incentives and intrinsic social valuations

By (14) a low consolidated social valuation may reflect a low intrinsic social valuation of  
 375 group  $jk$  or some incentive issues involving this group being mimicked. To disentangle intrinsic social valuations from incentive corrections we need to identify the relevant incentive constraints (those  $jk$  such that  $\tilde{n}_{jk}\tilde{\beta}_{jk} > 0$ ). Our identification strategy starts from (14) which, using (12) and (13), rewrites

$$n_{jk} \sum_{\ell} t^{\ell} \frac{\partial \xi_{jk}^{\ell}}{\partial R^k} - b_{jk} = \tilde{n}_{jk}\tilde{\beta}_{jk} - n_{jk}\alpha_{jk}\lambda_{jk}. \quad (25)$$

The right-hand side of (25) equals  $\tilde{n}_{jk}\tilde{\beta}_{jk}$  if (10) is slack and (11) is binding. It equals  
 380  $-n_{jk}\alpha_{jk}\lambda_{jk}$  if (10) is binding and (11) is slack. One can consequently detect the presence of asymmetric information at the disaggregated taste  $\times$  income class level by considering the right-hand side of (25), with  $b_{jk}^*$  replacing  $b_{jk}$ ,

$$s_{jk}^* = n_{jk} \sum_{\ell} t^{\ell} \frac{\partial \xi_{jk}^{\ell}}{\partial R^k} - b_{jk}^*. \quad (26)$$

A positive  $s_{jk}^*$  reveals that  $jk$  agents are envied and provides an estimate  $\tilde{n}_{jk}^*\tilde{\beta}_{jk}^*$  for  $\tilde{n}_{jk}\tilde{\beta}_{jk}$ .  
 385 Otherwise, the incentive correction  $\tilde{n}_{jk}^*\tilde{\beta}_{jk}^*$  is 0 and intrinsic and consolidated social valuations coincide.

The second row of Table 2 shows that the values of the statistics are always negative. It follows that all the incentive corrections are 0, and so the intrinsic and consolidated social valuations coincide. The intrinsic social valuations are reported in the third row of Table 2.

390 We conclude that commodity taxation in France only reflects many-person Ramsey considerations. Actually within income class heterogeneity does not matter for middle and upper income classes: the intrinsic valuations of the two taste groups are about equal for every such income classes. But taste heterogeneity matters among the poor: the French government favors Taste 1 poor. Taste 1 are older, less educated and spend more on food  
 395 items than Taste 2. This finding is reminiscent of the public debate about intergenerational equity arguing that the French tax system could be detrimental to the younger generations (see for instance the administrative report [14] or Pisani-Ferry [21] for a recent synthesis). In a report published at the moment when the Budget de Famille survey was conducted, the Conseil des Prélèments Obligatoires [6] advocated for putting the emphasis on intergenerational concerns and stated on page 46 that ‘in terms of intergenerational redistribution,  
 400 it appears that the commodity taxation currently involves an instantaneous transfer to the

benefit of households and individuals over 65 years of age.’ The results in Table 2 clearly accord with this view. They however yield a picture of the French situation where this diagnosis does not apply to every income classes, but instead is narrowed to the less well-off part of the population. The consumption patterns of the various groups of households point toward the crucial role played by low food taxes to favor Taste 1 poor.

REMARK 2. A Rawlsian objective is inconsistent with the assumption that actual taxes are optimal since social valuations are positive for all agents. However it could be that the objective fits Rawlsian criteria but actual taxes are not optimally chosen, leaving aside Remark 1 above. Since no relevant incentive constraints are detected in the current situation, one can analyze the impact of small enough tax changes in the Rawlsian case abstracting from incentive considerations.<sup>9</sup> In the current situation Taste 2 poor ( $k = P$ ) households have the lowest AIDS sub-utility from consumption. Given  $((\bar{\mu}_{ijk}), (\bar{y}_k))$ , if the government only values the utility of these agents, there exists  $(q, (R_k))$  that improves upon  $(\bar{q}, (\bar{R}_k))$  while satisfying incentive (2) and feasibility (3) requirements. Feasible tax reforms satisfy

$$\sum_{\ell} \left( \xi^{\ell} + \sum_h t^h \frac{\partial \xi^h}{\partial q^{\ell}} \right) dq^{\ell} - \sum_k n_k dR_k \leq 0.$$

The best reforms are such that this constraint binds. Given  $((\bar{\mu}_{ijk}), (\bar{y}_k))$  we therefore consider budget neutral reforms  $(dt^{\ell}, dR_P)$ ,

$$\left( \xi^{\ell} + \sum_h t^h \frac{\partial \xi^h}{\partial q^{\ell}} \right) \frac{dq^{\ell}}{dR_P} - n_P = 0. \quad (27)$$

To the first-order, this reform yields a change in Taste 2 poor sub-utility  $V_2(q, R_P)$  equal to

$$\alpha_{2P} \left( -\xi_{2P}^{\ell} \frac{dq^{\ell}}{dR_P} + n_{2P} \right) dR_P.$$

Using (27), and switching to the variables used in the empirical illustration, the term into brackets rewrites

$$1 - \frac{n_P \frac{q^{\ell} \xi_{2P}^{\ell}}{q^{\ell} \xi^{\ell}}}{1 + \sum_{jk} n_{jk} \sum_h \frac{t_{\text{val}}^h}{1 + t_{\text{val}}^h} \varepsilon_{jk}^{hl} \frac{q^h \xi_{jk}^h}{q^{\ell} \xi^{\ell}}}. \quad (28)$$

The expression in (28) measures the monetary change in social welfare for a Rawlsian government implied by a 1 euro increase in the after-tax income of a poor Taste 2 household

<sup>9</sup>Recent tax reforms exercises include Golosov, Tsyvinski and Werquin [16] and Jacquet and Lehmann [18].

financed by a tax on good  $\ell$ . Table 3 reports these monetary gains for each good  $\ell$ .

Table 3: IMPACT OF TAX REFORMS ON A RAWLSIAN SOCIAL WELFARE

Good $\ell$	Alco	Cloth	Leis	Food	Furn	Othe	Rest	Tran
(28)	0.722 <sup>1</sup>	0.765	0.778	0.920	0.914	0.795	0.860	0.655

Note 1: A one-unit additional income transfer toward a poor Taste 2 adult financed by a tax on Alcohol yields a (net) monetary change in social welfare equal to 0.722 euro.

The gain is always positive, which indicates that the actual consumption taxes are too low from a Rawlsian perspective: an additional income transfer toward the poor financed by higher consumption taxes would be socially profitable. The most profitable among such reforms would be to raise the tax on Food, which indeed is relatively less consumed by Taste 2. Table 3 also allows us to identify the best consumption tax reform, i.e., a reform maintaining the income taxes at their current levels. Consider a change in the tax on good  $h$  for an amount that would yield one additional unit of after-tax income to a poor Taste 2 equivalent adult. Suppose that this reform is financed by a change in the tax on good  $\ell$ . The resulting change in social welfare is proportional to the difference between the terms in (28) for goods  $h$  and  $\ell$ . The best consumption tax reform involves a lower tax on Transport (absent from externality issues), which is relatively more consumed by Taste 2, financed by a higher tax on Food, yielding a positive net social gain of  $0.920 - 0.655 = 0.265$  euro.  $\square$

## 6 Conclusion

The existing literature that examines optimal linear commodity taxes when taxpayers differ according to labor skill and consumption tastes concludes that these taxes are only used to relax incentive constraints if the government can also use nonlinear income taxation. Indeed, with this rich set of tax instruments, the government deals with redistribution concerns using direct taxes and transfers. Our results show how the presence of unobserved individual heterogeneity given the (publicly observed) individual income alters this conclusion: commodity taxes then matter for redistribution given the income, implying that goods consumed by socially favored agents tend to be taxed less heavily at the optimum. This conclusion follows from a new commodity tax rule that disentangles redistribution within and between income classes, and where redistribution within income classes is governed by a variant of the familiar many-person Ramsey rule. An illustration on French data shows the extent of within income class heterogeneity and suggests a low impact of incentives in the design of commodity taxes.

Our empirical analysis is subject to several caveats:

- 445
450
455
460
 • The necessary conditions for optimal consumption taxes derived in Proposition 2 suppose that consumption taxes only obey to equity and efficiency considerations. It is clear that taxation of categories such as Transport, Health or Education also involve externality and public good considerations. We also abstract from other aspects, e.g., administrative, lobbying or cross-border adjustments and employment, that are important in the actual process for designing consumption taxes. Our theoretical approach should be flexible enough to deal with such considerations. However our approach heavily relies on the separability assumption that demand for consumption goods depends on post-tax income but not on labor effort itself. It seems difficult to disentangle tastes from other dimensions of individual heterogeneity in the absence of separability. These are areas for further work.
- Our methodology deals with incentive constraints in the presence of a greater number of dimensions of individual heterogeneity than usually considered in the literature by deriving a sub-program where only taste heterogeneity matters. This program involves sufficient, but not necessary, conditions for the incentive constraints to be met. As a result, we do not necessarily detect all the binding incentive constraints. A future line of research would go toward a better understanding of the pattern of the relevant incentive constraints disaggregated at the level of agents' types needed in models with asymmetric information. The strategy used in the empirical illustration, based on qualification requirements in Lagrangian analysis, might be a step in this direction.

## References

- [1] Ahmad, E. and N. Stern, 1984, The theory of reform and Indian indirect taxes, *Journal of Public Economics* 25, 259-298.
- 465
 [2] Atkinson, A. and J. Stiglitz, 1976, The design of tax structure: Direct versus indirect taxation, *Journal of Public Economics* 6, 55-75.
- [3] Blomquist, S. and V. Christiansen, 2008, Taxation and heterogeneous preferences, *Finanz Archiv* 64, 218-244.
- 470
 [4] Chiappori, P.A. and B. Salanié, 2000, Testing for asymmetric information in insurance markets, *Journal of Political Economy* 108, 56-78.
- [5] Christiansen, V. and E.V. Jansen, 1978, Implicit social preferences in the Norwegian system of indirect taxation, *Journal of Public Economics* 10, 217-245.
- [6] Conseil des Prélèvements Obligatoires, 2008, La répartition des prélèvements obligatoires entre générations et la question de l'équité intergénérationnelle, avail-

- 475 able in French at: [www.ladocumentationfrancaise.fr/var/storage/rapports-publics/084000692.pdf](http://www.ladocumentationfrancaise.fr/var/storage/rapports-publics/084000692.pdf)
- [7] Crawford, I., M. Keen and S. Smith, 2010, Value Added Tax and Excises, chapter 4 in: Adam, S. et al., Dimensions of Tax Design, Oxford University Press.
- [8] Cremer, H. and F. Gahvari, 1998, On Optimal Taxation of Housing, *Journal of Urban Economics* 43, 315-335.
- 480 [9] Cremer, H. and F. Gahvari, 2002, Nonlinear pricing, redistribution and optimal tax policy, *Journal of Public Economic Theory* 4, 139-161.
- [10] Cremer, H., P. Pestieau and J.C. Rochet, 2001, Direct versus indirect taxation: the design of the tax structure revisited, *International Economic Review* 42, 781-799.
- 485 [11] Cremer, H., P. Pestieau and J.C. Rochet, 2003, Capital income taxation when inherited wealth is not observable, *Journal of Public Economics* 87, 2475-2490.
- [12] Diamond, P. and J. Mirrlees, 1971, Optimal taxation and public production II: tax rules, *American Economic Review* 61, 261-78.
- [13] Diamond, P. and J. Spinnewijn, 2011, Capital income taxes with heterogeneous discount rates, *American Economic Journal: Economic Policy* 3, 52-76.
- 490 [14] France Stratégie, 2016, Jeunesse, vieillissement : quelles politiques ?, note Enjeux available at <http://www.strategie.gouv.fr/sites/strategie.gouv.fr/files/atoms/files/17-27-jeunesse-vieillissement2.pdf>
- [15] Golosov, M., M. Troshkin, A. Tsyvinski, and M. Weinzierl, 2013, Preference heterogeneity and optimal commodity taxation, *Journal of Public Economics* 97, 160-175.
- 495 [16] Golosov, M., A. Tsyvinski and N. Werquin, 2014, A variational approach to the analysis of tax systems, NBER Working Paper 20780.
- [17] Guesnerie, R., 1995, A Contribution to the Pure Theory of Taxation, *Econometric Society Monographs*, Cambridge University Press.
- 500 [18] Jacquet, L. and E. Lehmann, 2017, Optimal income taxation with composition effects, CEPR Discussion Paper 12276.
- [19] Kaplow, L., 2008, Optimal policy with heterogeneous preferences, *B. E. J. Econ. Analysis and Policy* 8, Article 40, 1-28.
- [20] Mirrlees, J., 1976, Optimal tax theory: a synthesis, *Journal of Public Economics* 6, 327-358.
- 505

[21] Pisani-Ferry, J., 2016, The Politics of Young and Old, France Stratégie, available at <http://www.strategie.gouv.fr/point-de-vue/politics-young-and-old>.

[22] Plessz, M. and S. Gojard, 2013, Do processed vegetables reduce the socio-economic differences in vegetable purchases? A study in France, *European Journal of Public Health*, 23, 747-752.

510

[23] Saez, E., 2002, The desirability of commodity taxation under non-linear income taxation and heterogeneous tastes, *Journal of Public Economics* 83, 217-230.

# A Consumption

We use ad valorem tax rates for year 2011 collected by the French Institut des Politiques Publiques at the 5-digit COICOP level. If different 5-digit COICOP goods within the same 2-digit category (the level of aggregation that we consider in the main text) are taxed at different rates, then the tax rate of the 2-digit category is set at the average tax rate weighted by the expenditures for the 5-digit items in this category. Tax rates computed for 2-digit COICOP categories are reported in the first column of Table 4. This same table reports the average (Stone) price indexes faced by the two taste groups, as well as budget share information. Food budget share has the largest total variance and most of it is imputable to between taste group heterogeneity. Information on consumption patterns and income at the taste  $\times$  income are given in Table 5.

Table 4: TASTE GROUP CONSUMPTION PATTERNS

	Tax rates (%)	Price indexes <sup>1</sup>		Budget share			
		Taste 1	Taste 2	Mean		Variance	
				Taste 1	Taste 2	Total	Within/Total
Food	5.9	3.00	3.02	0.37	0.15	0.022	0.367
Alco	66.4	5.45	5.37	0.05	0.04	0.006	1.000
Clot	19.6	15.76	15.63	0.06	0.08	0.006	0.973
Furn	19.4	2.69	2.71	0.04	0.04	0.003	0.995
Tran	100.2	2.06	2.08	0.09	0.18	0.015	0.866
Cult & Comm	11.6	3.85	3.94	0.14	0.19	0.011	0.939
Rest	5.5	10.41	10.26	0.05	0.12	0.009	0.904
Othe	19.6	2.70	2.77	0.19	0.21	0.016	0.993

Note 1: Stone price indexes

# B AIDS

The AIDS assumes that the budget share of good  $\ell$  for a household  $g$  in taste group  $j$  whose after-tax income (total expenditures) is  $R_g$  obeys

$$\frac{q^\ell \xi_h^\ell}{R_g} = \delta_j^\ell + \zeta_j f_g + \sum_{\ell'} \gamma_j^{\ell\ell'} \log q^{\ell'} + \beta_j^\ell \log \left( \frac{r_g}{Q_j} \right),$$

where  $r_g$  is the after-tax adult equivalent income of household  $g$  (that is,  $R_g$  divided by the number of consumption units in the household) and  $Q_j$  is the true taste  $j$  consumer price index associated with the AIDS.

The categories of goods are purged from durable items, which we consider as fixed expenditures. The expenditures for categories Housing, water, electricity, gas and other

Table 5: CLASS CONSUMPTION PATTERNS

$j$	Taste 1			Taste 2		
	Poor <sup>1</sup>	Middle <sup>1</sup>	Rich <sup>1</sup>	Poor <sup>1</sup>	Middle <sup>1</sup>	Rich <sup>1</sup>
$n_{jk}$	0.15	0.18	0.09	0.21	0.17	0.20
Food	0.40	0.38	0.35	0.12	0.16	0.15
Alco	0.05	0.05	0.05	0.05	0.05	0.04
Clot	0.06	0.06	0.06	0.09	0.08	0.08
Furn	0.03	0.04	0.06	0.02	0.03	0.05
Tran	0.07	0.09	0.10	0.17	0.19	0.17
Cult & Comm	0.14	0.14	0.14	0.21	0.18	0.18
Rest	0.03	0.05	0.07	0.08	0.11	0.13
Othe	0.22	0.20	0.17	0.26	0.21	0.20
Income	5414	11034	21601	5711	10850	19193

Note 1: Poor: [422, 8120); Middle: [8120, 14200); Rich: [14200, 83412].

fuels (04), Health (06) and Education (10) are also considered as fixed. The French survey does not provide imputed rents to landlords, but it reports housing loans and various housing insurance contracts subscribed by landlords in two original categories 13 and 14 that are absent from the COICOP classification. We take into account these expenditures in order to treat tenants and landlords symmetrically. Finally price indexes for Health (06) and Education (10), which are two categories comprising mostly publicly provided items in France, cannot be accurately computed. For this reason we also treat these expenditures as fixed. Hence the vector  $f_g$  comprises 6 variables that correspond to fixed adult equivalent expenditures, namely household  $g$  expenditures for durables and COICOP categories 04, 06, 10, and the Budget de Famille categories 13 and 14 that relate to landlord expenditures.

The demand functions are estimated for each taste group using the `micEconAids` package developed for R, imposing the usual AIDS restrictions on parameters that ensure that budget shares sum to 1, and that homogeneity and symmetry properties of demand are satisfied. For the sake of saving space Tables 6 and 7 in this appendix report the estimates of the AIDS elasticities computed at the average prices and income of each taste group. However, to recover the consolidated social valuations from (24) and the incentive corrections from (26), we have considered the finest level of taste  $\times$  income.

Table 6: COMPENSATED AND INCOME ELASTICITIES FOR TASTE 1

	pFoo	pAlc	pClo	pFur	pTra	pLei	pRes	pOth	Income
qFoo	-0.665	0.073	0.084	0.028	0.105	0.144	0.073	0.159	0.902
qAlc	0.470	-0.813	0.049	0.050	-0.049	0.111	-0.066	0.248	1.134
qClo	0.437	0.040	-0.998	0.008	0.121	0.205	-0.018	0.205	1.105
qFur	0.270	0.076	0.015	-0.934	0.033	0.145	0.162	0.233	1.469
qTra	0.386	-0.028	0.085	0.012	-0.919	0.149	0.118	0.197	1.307
qLei	0.354	0.042	0.097	0.036	0.099	-0.837	0.069	0.138	0.858
qRes	0.402	-0.057	-0.019	0.091	0.177	0.155	-0.966	0.216	1.403
qOth	0.321	0.078	0.080	0.048	0.108	0.113	0.079	-0.826	0.818

Table 7: COMPENSATED AND INCOME ELASTICITIES FOR TASTE 2

	pFoo	pAlc	pClo	pFur	pTra	pLei	pRes	pOth	Income
qFoo	-0.976	0.066	0.153	0.035	0.190	0.241	0.053	0.237	1.195
qAlc	0.182	-0.816	0.076	-0.070	0.219	0.085	0.038	0.285	1.045
qClo	0.219	0.039	-0.880	-0.007	0.090	0.213	0.123	0.202	1.012
qFur	0.171	-0.122	-0.023	-0.845	0.192	0.238	0.091	0.298	1.457
qTra	0.138	0.058	0.045	0.029	-0.763	0.168	0.115	0.211	1.127
qLei	0.167	0.021	0.103	0.034	0.160	-0.766	0.124	0.157	0.745
qRes	0.059	0.015	0.096	0.021	0.175	0.200	-0.824	0.258	1.297
qOth	0.150	0.065	0.090	0.039	0.184	0.144	0.147	-0.819	0.755