POSTGRADO ECONOMÍA Y NEGOCIOS
UNIVERSIDAD DE CHILE

"Optimal Income Taxation under Informal Consumption"

TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN ECONOMÍA

Alumno: Bastián Castro Nofal
Profesor Guía: Pablo Gutiérrez Cubillos

Santiago, mayo de 2024

Optimal Income Taxation under Informal Consumption*

Bastián Castro Nofal†

May 24, 2024

Abstract

This paper explores the optimal structure of the income tax schedule in the United States, considering factors such as private insurance and tax avoidance. The analysis reveals that tax avoidance mechanisms contribute to the observed regressivity at the top of the income distribution. Wealthier households engage in more avoidance, reducing their labor income taxes. The introduction of tax avoidance also alters the optimal income tax progressivity, with an omniscient planner allowing for increased progressivity, while a myopic planner suggests minimal changes, unaware of avoidance's impact on aggregate production.

*We express our gratitude to Pablo Gutiérrez for guiding this master's thesis and to Daniel Reck for providing the data that facilitated certain aspects of this paper. Additionally, we appreciate the valuable comments from Javier Cortés Orihuela.

†Universidad de Chile.

Agradecimientos

Esta tesis no podría haber sido posible sin la ayuda de muchas personas que han contribuido a que me convierta en quien soy. A continuación, quiero expresar mis agradecimientos a todos ellos.

A mis padres, Ivonne y Carlos, por darme la vida, las condiciones materiales para estudiar y los valores necesarios para sobrellevar ambas cosas con altura de miras. A mi hermano, Carlos Daniel, por ser un confidente y amigo, y por guiarme en las dificultades psicológicas y morales de la vida moderna.

A Luis Alberto (Q.E.P.D.), por ser la persona clave en mi decisión de estudiar Economía, por creer en esta disciplina como un instrumento para mejorar la vida humana y por haberme transmitido esa convicción. Algún día le pagaré todos los cafés y tortas que le debo.

A mi profesor guía, el Profesor Pablo Gutiérrez, por confiar en mí e introducirme a la pasión por el conocimiento. Hasta antes de entrar en su oficina en el décimonoveno piso, había acumulado cinco años de crisis vocacionales, hoy definitivamente resueltas. También agradezco a Javier Cortés, quien ha contribuido significativamente a esta definición vital, con su agudeza intelectual y su inmenso compañerismo. Dudo que existan muchos espacios en este vasto universo que mezclen horizontalidad y rigor como ellos lo logran. Asimismo, agradezco a los profesores de mi comité, Alejandro Micco y Damián Vergara. Es raro tener la oportunidad de ver a alguien que, estando a años luz en escala intelectual, se interese genuinamente por el proyecto que se le presenta.

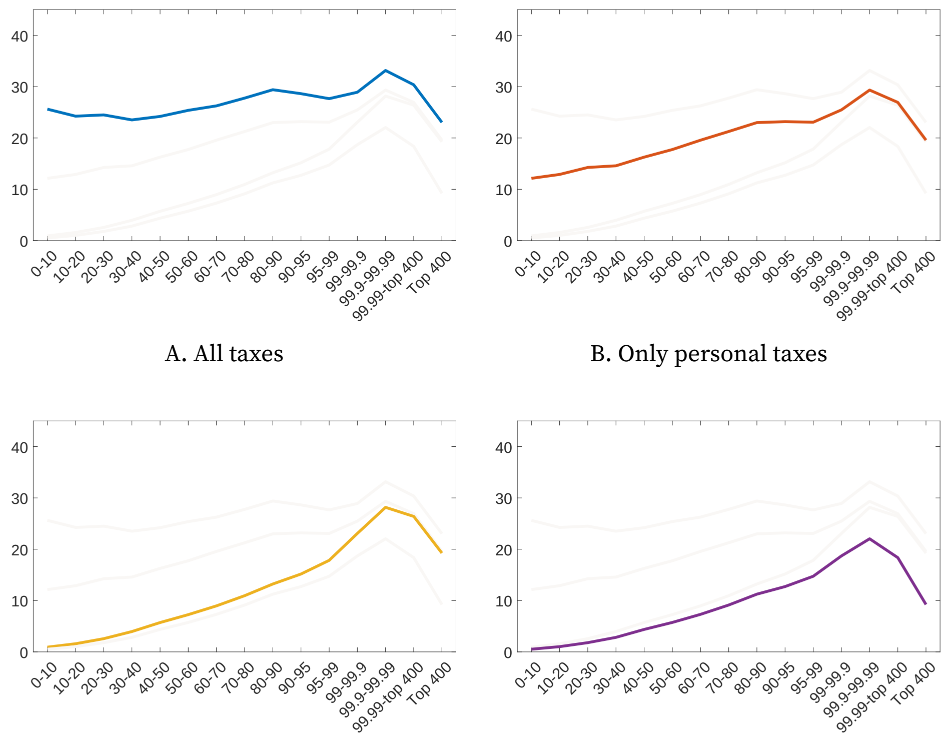
A mis amigos. A Patricio, por haber cultivado una amistad de más de 10 años en la cual he aprendido que hay más vida fuera de lo que unas cuantas ecuaciones puedan resumir. A quienes compartieron gran parte de esta etapa conmigo, mis amigos de larga data y mis compañeros de la "generación dorada" del magíster en Economía. Pensé en muchas formas de redactar esta sección, pero todas implicaban alguna forma de jerarquización, así que aquí van los nombres aleatorizados, que no es sino la forma más democrática, aunque no la más óptima: Sebastián Chacón, Bruno, Sebastián Hernández, Denisse, Pedro, Aoi, Álvaro, Ignacio Arancibia, Felipe Carrera, Sebastián Flores, Sebastián Núñez, Guillermo, Ítalo, Javier Martínez, Leonardo, Catalina, Isabel Arias, Cristóbal, Diego, Jorge, Ivar, Ricardo, Gabriela Denis, Gabriela Jaque, Ignacio Fuentes, Arturo, Miguel, Marco, Nicolás Leiva, Vicente, Martín, Camilo e Ivo. A todos ellos, gracias por haberme acompañado sentimental y académicamente, por soportar mis desvaríos vocacionales, y por las tremendas aventuras y desventuras, comedias y tragedias. Gracias por soportar durante tanto tiempo esta amalgama de miedo y convicción, cobardía y coraje que soy.

A aquellos con quienes he ido perdiendo el contacto, como Joaquín, Benjamín, Isabel Ramírez, Sergio, Nicolás Molina, Macarena, Fabián, Nataly, Felipe Muñoz y Carolina. Sepan que los guardo en mi corazón con inmenso cariño y les pido perdón por ser la persona ambivalente que he sido siempre. Volverá a ocurrir.

1. Introduction

Taxes, as the old mantra goes, are the price we pay for a civilized society. Consequently, the discourse on how to structure them is as ancient as civilization itself. In contemporary economic thought, the discussion surrounding income taxes and their progressivity—whether those with higher incomes should contribute proportionally more—traces back to Smith (1776). He proposed that individuals should pay in accordance with their abilities and the benefits derived from state protection. Despite the longstanding nature of this debate, progressive income taxes only gained prevalence in the early 20th century, with top marginal rates exceeding 70% (see, for instance, Piketty et al., 2014; Piketty, 2020), later tapering to around 40%, a trend associated with the observed increase in inequality (Atkinson et al., 2011; Zucman, 2019).

Figure 1. Average tax rates by pre-tax income group in 2018 (% of pre-tax income)



A. All taxes

B. Only personal taxes

C. Personal taxes without payroll taxes

D. Only Income taxes

Source: Saez and Zucman (2019). Panel A includes all taxes paid by the individual: sales, income, corporate, property, estate, and payroll taxes. Panel B excludes sales, property, and estate taxes; Panel C excludes what Panel B excludes and also payroll taxes, and Panel D excludes what Panel C excludes and also corporate taxes. Each panel highlights a different average tax rate measure, with the other three measures in the background in grey.

Despite the prevalence of progressive income tax rates, high-income households globally often demonstrate lower average tax rates than their low and middle-class counterparts when considering the comprehensive spectrum of income taxes paid (see, for example, Saez and Zucman (2019) and Figure 1 for the U.S.; Piketty (2020) for France; Milligan (2022) for Canada).

This pattern aligns with empirical evidence indicating that high-income earners engage in greater tax avoidance. It also corresponds with the ongoing discussion asserting that the wealth of the affluent is primarily human capital rather than financial, quantified through capital gains and dividend incomes facilitated by pass-through mechanisms (Smith et al., 2019).

This paper explores the optimal income tax policy in a Ramsey (1927) context considering the presence of tax avoidance. We present a framework that addresses the discussion in both positive and normative terms: firstly, can avoidance mechanisms account for the observed regressive average tax rates at the top? Secondly, what are the implications for optimal income taxation with avoidance mechanisms? Would taxes become more or less progressive?

Our model features heterogeneous productivity agents and partial private insurance, building on the tradition of Heathcote et al. (2017) and Heathcote and Tsujiyama (2021). We innovate by incorporating individual avoidance mechanisms related to those developed by Feldstein (1999) and Piketty et al. (2014) —an aspect hitherto unexplored in macroeconomic modeling-, albeit manifested through informal or firm consumption strategies. Furthermore, we conceptualize the trade-off between formal and informal consumption resulting from tax avoidance as a relation of relative prices.

Two natural hypotheses emerge: firstly, a model incorporating tax avoidance could elucidate the observed regressivity at the top. However, since most tax systems exhibit progressivity, this implies a greater prevalence of avoidance opportunities among higher-income households. Secondly, it is plausible that optimal tax rates may be lower than those documented in related literature, considering that avoidance diminishes the effectiveness of tax schedules. A calibrated version of the model for the U.S. economy allows us to assess and confirm these hypotheses. The results indicate that an omniscient planner, cognizant of tax avoidance but unable to eliminate it, still augments the progressivity of the tax schedule, leading to increased taxes for the high-income households, though to a lesser extent than suggested in previous literature. Conversely, a myopic planner unaware of avoidance makes considerably smaller increases in progressivity or none at all.

Related Literature. Our paper contributes to various strands of literature. Firstly, it aligns with the Optimal Income Taxation literature, which explores the optimal structure of tax schedules through the characterization of income and labor supply elasticities. This tradition is exemplified by works such as Mirrlees (1971), Diamond (1998), and Saez (2001). In our study, we depart from the non-parametric tax schedules typically assumed in this literature, opting instead for simpler yet reality-related parametric schedules.

Second, our model aligns with the dynamic stochastic general equilibrium (DSGE) literature, which employs models to simulate economic environments and assess the effects of various taxes on production and welfare. Works such as Trabandt and Uhlig (2011), Farhi and Werning (2013), Stantcheva (2017), Holter et al. (2019), and Kindermann and Krueger (2022) are notable in this tradition. We particularly draw on the tradition introduced by Heathcote et al. (2017) and Heathcote and Tsujiyama (2021), utilizing a parametric "HSV" tax system characterized by parameters determining the degree of progressivity. Our contribution lies in incorporating tax avoidance mechanisms, fundamentally altering the results of optimal income schedules chosen by a planner, and highlighting the importance of information available to them.

Furthermore, our paper connects with a literature exploring theoretical tax avoidance mechanisms, as seen in works such as Allingham and Sandmo (1972), Mayshar (1991), Feldstein (1999), Piketty and Saez (2013), and Piketty et al. (2014). We extend this literature by providing insights, following the tradition of Clotfelter (1983), that elucidate how tax avoidance can result from informal consumption within a firm, involving price mechanisms and incorporating macroeconomic dynamics.

Finally, our paper is positioned within the literature on labor versus capital-driven inequality, as investigated by Piketty and Saez (2003), Piketty (2014), and Smith et al. (2019). This body of work posits that the surge in inequality is predominantly attributable to the escalating remuneration of executives and the accumulation of human capital at the upper percentiles of the income distribution, frequently quantified as capital gains and dividend incomes due to pass-through mechanisms. Our contribution involves elucidating the theoretical mechanisms underpinning this phenomenon and scrutinizing its implications for empirical tax rates and optimal tax policy. We heavily draw on evidence of high tax avoidance at the top, as documented by Johns and Slemrod (2010), Saez and Zucman (2019) and Guyton et al. (2021). Importantly, we demonstrate that tax avoidance can elucidate the observed regressivity at the top.

The paper is organized as follows: Section 2 presents our theoretical model. Section 3 provides the calibration and results of the model, assuming homogeneous tax avoidance. In contrast, Section 4 performs the same analysis with heterogeneous tax avoidance. Section 5 introduces alternative models and deliberates on the stability of the original findings. Finally, Section 6 synthesizes the results and offers a brief conclusion.

2. Model

2.1. A Tractable Macro Model

In this paper, we adopt the alternative specification employed by Heathcote and Tsujiyama (2021), featuring autonomous agents who engage in the purchase of private insurance within decentralized financial markets[1]. The population consists of a unit mass of individuals, each characterized by varying labor productivity denoted as w. This productivity is determined by two orthogonal idiosyncratic components: $\alpha \in \mathcal{A} \subseteq \mathbb{R}$ represents shocks that cannot be privately insured and are interpreted as fixed effects drawn prior to agents entering the economy, encompassing factors such as ability, education, cultural capital, etc.; and $\varepsilon \in \mathcal{E} \subseteq \mathbb{R}$ represents

[1]We comment on the macrodynamics of this model in Appendix B.2.

shocks that can be perfectly privately insured, interpreted as life-cycle shocks. Neither α nor ε is observable by the tax authority.

The budget constraint for an agent with a specific value of α is expressed as:

$$\int \mathbf{B}(\alpha, \varepsilon)\mathbf{Q}(\varepsilon)d\varepsilon = 0, \tag{1}$$

where $\mathbf{B}(\alpha, \varepsilon)$ denotes the quantity of insurance claims purchased, which pays a unit of consumption only if the drawn shock corresponds to $\varepsilon \in \mathcal{E}$, and $\mathbf{Q}(\varepsilon)$ represents the price of the bundle of claims. Actuarially fair prices imply that $\mathbf{Q}(E) = \int_E dF(\varepsilon)$, where E is a subset of all shocks.

Taxation is levied at the individual level and encompasses earnings along with insurance payments. This implies that the individual's budget constraints can be expressed as:

$$c(\alpha, \varepsilon) = y(\alpha, \varepsilon) - T(y(\alpha, \varepsilon)). \tag{2}$$

Here, $y(\alpha, \varepsilon)$ represents labor income, $T(y(\alpha, \varepsilon))$ denotes net tax revenues at income level y, and $c(\alpha, \varepsilon)$ signifies consumption. The individual's income before taxes and transfers is then defined by:

$$y(\alpha, \varepsilon) = \exp(\alpha + \varepsilon)h(\alpha, \varepsilon) + B(\alpha, \varepsilon), \tag{3}$$

where $h(\alpha, \varepsilon)$ measures the hours worked by the agent. Agents share identical preferences for consumption and work effort, which take a separable form:

$$u(c, h) = \log(c) - \frac{h^{1+\sigma}}{1+\sigma}$$

where $\sigma > 0$, providing a Frisch elasticity of labor supply of $1/\sigma$. The aggregate output in the economy is the aggregate effective labor supply, divided between private consumption and a non-valued public good G. Consequently, the resource constraint of the economy is given by:

$$\int\int c(\alpha, \varepsilon)dF_\alpha(\alpha)dF_\varepsilon(\varepsilon) + G = \int\int \exp(\alpha + \varepsilon)h(\alpha, \varepsilon)dF_\alpha(\alpha)dF_\varepsilon(\varepsilon) \tag{4}$$

The individual agent's optimization problem is to choose $c(\alpha, \varepsilon)$, $h(\alpha, \varepsilon)$, and $B(\alpha, \varepsilon)$ to maximize:

$$\max_{\{c(\alpha,\varepsilon),h(\alpha,\varepsilon),B(\alpha,\varepsilon)\}} \int \left[\log c(\alpha, \varepsilon) - \frac{h(\alpha, \varepsilon)^{1+\sigma}}{1+\sigma}\right] dF_\varepsilon(\varepsilon) \tag{5}$$

subject to equations (1), (2) and (3).

Now, we introduce the tax function $T(y)$, commonly known as "HSV" after its reintroduction by Heathcote et al. (2017). However, its specification traces back to Feldstein (1969), and

Persson (1983) and Benabou (2000) were the first to incorporate it into the context of dynamic macroeconomic models with heterogeneous agents. This function is defined as:

$$T(y) = y - \lambda y^{1-\tau}. \tag{6}$$

Two notable properties arise from this expression. Firstly, the equation implies a log-linear relationship between pre-government and disposable earnings, denoted as $\tilde{y}_i$, given that $\tilde{y}_i = \lambda y_i^{1-\tau}$. As demonstrated by Heathcote et al. (2017), this log-linear relationship is empirically suitable. In this case, the parameter $1-\tau$ measures the elasticity of post-tax to pre-tax income. Secondly, the progressivity of a tax system can be characterized using the parameter τ, where $\tau > 0$ indicates a progressive system, $\tau < 0$ signifies a regressive one, and $\tau = 0$ represents a flat tax system.

Solving the first-order conditions with the HSV tax function results in the following equilibrium allocations for consumption, hours, and individual earnings:

$$c(\alpha) = \lambda(1-\tau)^{\frac{1-\tau}{1+\sigma}} \left\{ \mathbf{E}\left[\exp(\varepsilon)^{\frac{1+\sigma}{\sigma}} \right] \right\}^{\frac{\sigma(1-\tau)}{1+\sigma}} \exp((1-\tau)\alpha), \tag{7}$$

$$h(\varepsilon) = (1-\tau)^{\frac{1}{1+\sigma}} \left\{ \mathbf{E}\left[\exp(\varepsilon)^{\frac{1+\sigma}{\sigma}} \right] \right\}^{\frac{-1}{1+\sigma}} \exp\left(\frac{1}{\sigma}\varepsilon\right), \tag{8}$$

$$y(\alpha) = (1-\tau)^{\frac{1}{1+\sigma}} \left\{ \mathbf{E}\left[\exp(\varepsilon)^{\frac{1+\sigma}{\sigma}} \right] \right\}^{\frac{\sigma}{1+\sigma}} \exp(\alpha). \tag{9}$$

Here, $y(\alpha)$ represents the end-of-period family income, the only type of income observed by the tax authority, and is defined as $y(\alpha) = \int \exp(\alpha + \varepsilon) h(\alpha, \varepsilon) dF_\varepsilon(\varepsilon)$.

Finally, we turn to describing the "Ramsey planner", which will be utilized throughout the rest of the paper. This planner, as outlined by Ramsey (1927), refers to an entity that selects the optimal tax function within a specified parametric class denoted as $\mathcal{T}$. For the HSV class, $\mathcal{T} = \{T : \mathbb{R}_+ \to \mathbb{R} | T(y) = y - \lambda y^{1-\tau} \text{ for } y \in \mathbb{R}_+, \lambda \in \mathbb{R}_+, \tau \in [-1, 1]\}$. Thus, the Ramsey problem involves maximizing social welfare by choosing a tax progressivity parameter while ensuring that allocations form a competitive equilibrium:

$$\max_{T \in \mathcal{T}} \int W(\alpha) \int u(c(\alpha, \varepsilon), h(\alpha, \varepsilon), e(\alpha, \varepsilon)) dF_\alpha(\alpha) dF_\varepsilon(\varepsilon) \tag{10}$$

subject to (4) and to (7) and (8) being solutions to the family problem. Here, $W(\alpha)$ is a Pareto weight function which we determine later.

2.2. A Simple Model of Informal Consumption

In this section, we provide intuition for our results using a simplified microeconomic model, drawing on the ideas of tax avoidance literature but presenting a different perspective on evasion costs and mechanisms.

2.2.1. Literature review

Theoretical contributions in the field of tax avoidance comprise a spectrum of models that conceptualize these mechanisms as strategic decisions undertaken by households. The most clear example is the seminal model of Allingham and Sandmo (1972) which delved into the domain of individual tax avoidance, with a particular focus on risk aversion within the framework of expected utility. The model posits that a taxpayer may be inclined to report a taxable income below its actual value, where the deterrence mechanism against income tax evasion in this model is contingent upon a fixed probability of detecting any understatement of taxable income by the planner, accompanied by a proportional penalty, in addition to settling the genuine tax liability. Expanding on this groundwork, Mayshar (1991) and then Feldstein (1999) advanced a more traditional model of utility maximization subject to budget constraints but retaining the same underlying characteristics: a portion of income remains concealed from tax authorities through evasion incurring a cost. Both contributions contribute significantly to the concept of "tax technology," which determine the taxes paid as a function of policies selected by the authority.

Recent literature has extended its focus to incorporate sheltered-income decisions within the context of optimal labor income taxation. Piketty and Saez (2013) employ a model where the agent can shield income, incurring a convex disutility in doing so. The authors demonstrate that this scenario creates a fiscal externality, allowing the government to enhance efficiency and tax capacity by closing tax avoidance opportunities. Consequently, optimal tax rates are contingent not solely on the actual elasticity of labor supply but also on a total elasticity that accounts for both real effects and sheltering—an aspect integrated into the computation of the ETI. In an extension of this framework, Piketty et al. (2014) presented a more nuanced model that integrates sheltered income and the dynamics of compensation bargaining.

While these micro models excel at elucidating the intricacies of tax avoidance in a direct and sophisticated manner, they often involve tax schedules that may not be closely aligned with the actual tax structures in reality. Additionally, some of the nuances introduced by heterogeneous agents with an income distribution may be overlooked in these models.

Much less has been written about tax avoidance as firm consumption. As an illustrative instance, Clotfelter (1983) introduced a model wherein taxable income could be diminished through business expenses, including those associated with travel and entertainment (T&E). Employing a framework wherein an input for production is subject to consumption, the author demonstrates a pronounced price sensitivity, which suggests that prevailing tax regulations distort business expenditures, ultimately employed to reduce tax liabilities. To the best of our knowledge, this marks the initial consideration of tax avoidance mechanisms intertwined with consumption within a firm.

Evidence on firm consumption's role in tax avoidance remains limited but offers valuable insights into the underlying mechanisms. Some papers show evidence that CEOs, for example, consume more perks (Malmendier and Tate, 2009), with goods such as jets (Yermack, 2006) and mansions (Liu and Yermack, 2012) being the most important examples. There is also theory that develops the idea of pet projects (Chetty and Saez, 2010) as a way for managers to make investments that are not productive for shareholders. More recently and importantly, Leite (2024) found that in Portugal, individuals who control firms shift 36% of their monthly personal expenditures through firms and 31% of their household expenditures, constituting approximately 1% of GDP. This suggests that the channel of firm consumption may be important in tax avoidance.

2.2.2. The model

Let us address the decision problem faced by an individual agent. This agent is subject to a prescribed income tax rate denoted as τ, after which they allocate their after-tax income to formal consumption (c) and informal consumption (e). Informal consumption may result from various strategies, including tax avoidance, fringe benefits, expenses associated with travel and entertainment (as developed by Clotfelter, 1983), and income reclassification from wages to capital gains, among other contributing factors. It should be regarded as an umbrella term that encompasses consumption prior to the receipt of personal income. Within this context, the individual possesses the flexibility to allocate their entire income to informal consumption, in which case they would incur no tax liability. Consequently, the agent seeks a solution to the following optimization problem:

$$\max_{\{c,e\}} U(c, e)$$
$$s.t.\ p_1 c = (1 - \tau)(y - p_2 e)$$

It is essential to emphasize that the costs associated with informal consumption are fully incorporated into the utility function. It is evident that, if c and e were considered perfect substitutes, leading to a linear utility function, the rational choice for the agent would be to allocate their entire income to informal consumption ($c = 0$), thus avoiding any tax liability. However, as pointed out by Andreoni et al. (1998), the predictions of the Allingham-Sandmo model sharply contrast with the observed high compliance levels in modern tax systems, characterized by low audit rates and relatively modest penalties. Therefore, the introduction of imperfect substitution is imperative to account for the observed phenomenon where households allocate only a minor portion of their income to tax evasion strategies. In that sense, we can simplify the model by incorporating a more implicit modeling of costs of avoidance through utility functions that allow for imperfect substitution.

Let us examine the scenario in which the prices for formal and informal consumption are

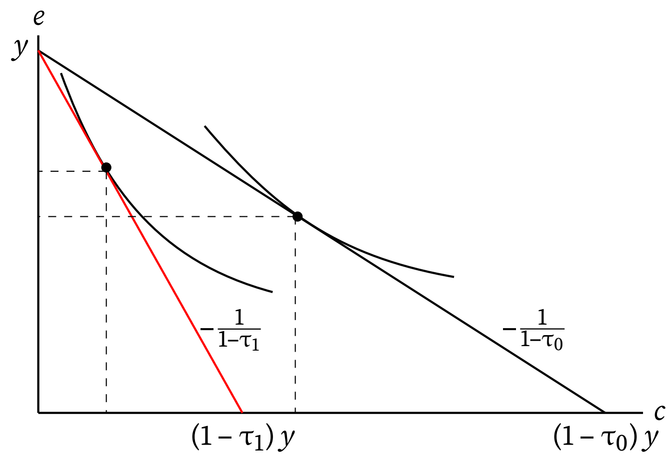


Figure 2. Increase in the tax as a change in relative prices.

equal ($p_1 = p_2$), and imperfect substitution prevails—a condition alone sufficient to guarantee an interior solution for both allocations, as previously discussed. As we can see in Figure 2, it becomes evident that an increase in the income tax rate can be interpreted as a rise in the relative prices of formal and informal consumption. Consequently, a higher income tax rate renders formal consumption more costly relative to informal consumption. It follows naturally that, under a progressive tax schedule, households are inclined to engage in a greater degree of informal consumption, as their price for formal consumption is higher than that faced by lower-income households.

This becomes more evident when considering the HSV function, where $c = \lambda(y - e)^{1-\tau}$. As depicted in Figure 3, an increase in household income induces a substitution effect, making tax avoidance relatively cheaper compared to formal consumption. Higher income not only expands the potential for consumption but also enhances the capacity for tax evasion to a greater extent than consumption.

The conclusion is straightforward: households with higher incomes will exhibit a preference for informal over formal consumption. This tendency becomes more pronounced in the presence of higher tax rates, even when faced with equivalent costs of avoidance as those experienced by less affluent households.

2.2.3. Comparison

We now compare our model with the conventional approach in tax avoidance models. Let us consider two models: a version of our own micro model assuming Cobb-Douglas utility for consumption and evasion, and a simplified and modified version of the model developed by Piketty et al. (2014). In Piketty's model, ordinary taxable income z is the difference between real income y and sheltered income e, where $z = y - e$ is taxed at normal income tax rates, and e is taxed at a lower rate denoted as π. The objective is to maximize a quasi-linear utility

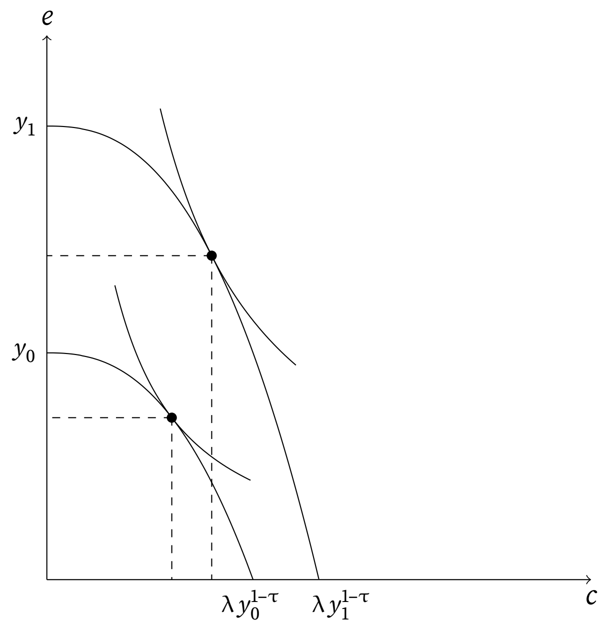


Figure 3. Increase in the agent's income as a substitution effect.

$u(c, e) = c - d(e)$, subject to the budget constraint $c = y - T(y - e) - \pi e$, where $d(e)$ is a cost of sheltering income that increases and is convex in e.

Let us assume a HSV tax function in both models, and a certain disutility function of sheltering income that is similar in structure to the disutility of work, so that the first problem is given by:

$$\max_{\{c_O, e_O\}} \log c_O + \omega \log e_O$$
$$\text{s.t. } c_O = \lambda(y_O - e_O)^{1-\tau},$$

and the second by:

$$\max_{\{c_P, e_P\}} c_P - \frac{e_P^{1+\omega}}{1+\omega}$$
$$\text{s.t. } c_P = \lambda(y_P - e_P)^{1-\tau} + (1-\pi)e_P.$$

Several observations are straightforward. First, the intuition between both models differs. In our model, consumption and evasion are two distinct forms of consumption, while in the classic model, avoidance adds to regular consumption. Aggregate consumption is given by $c_O + e_O$ in our model, whereas it is c_P in the classic model.

Second, certain conditions must apply to both models to yield the same results. Assuming $y_P = y_O$, if aggregate consumptions coincide, $c_P = c_O + e_O$, but this implies that $e_O = e_P$ only if $e_O = e_P = 0$ or $\pi = 0$. Otherwise, if $c_P = c_O + e_O \implies e_P > e_O$; $e_P = e_O \implies c_P < c_O + e_O$. Moreover,

$c_P = c_O + e_O$ and $e_O = e_P$, only hold if $y_P \neq y_O$, implying that without knowledge of y (as we will assume later), the conclusions regarding the proportion of avoidance will differ between models.

Nevertheless, this result may not be surprising theoretically. If evasion in our model occurs through consumption within the firm, it should not be subject to taxes, which may differ from a pass-through case, where Piketty et al. model assumes it to be a case where high incomes pass labor income as capital gains, taxed at rates of 20% or even 15% in some years.

2.3. A Macro Model with Informal Consumption

Now, we extend the model developed by Heathcote and Tsujiyama (2021) by incorporating the characteristics of informal consumption. The individual's budget constraint now changes to:

$$y(\alpha, \varepsilon) = c(\alpha, \varepsilon) + e(\alpha, \varepsilon) - T(y_T(\alpha, \varepsilon)) \tag{11}$$

where y_T is the taxable income, defined by $y_T \equiv y - e$. With an HSV function, we have $c(\alpha, \varepsilon) = \lambda(y(\alpha, \varepsilon) - e(\alpha, \varepsilon))^{1-\tau}$, as discussed in the earlier subsection. The utility function of the agent is given by:

$$u(c, h, e) = \log c - \frac{h^{1+\sigma}}{1+\sigma} + \omega \log e,$$

where one can interpret the parameter ω as a composite variable encompassing various economic characteristics, including moral preferences, the extent of consumption within a firm, intricacies of tax avoidance, and regulatory measures aimed at mitigating it, among other factors. Thus, ω is a parameter that reflects both preferences for informal consumption and the technology of evasion. It can be seen as the inverse of the "Tax technology" of the government, which endeavors to ensure compliance.

We assume $\omega \in [0, \infty^+)$, where $\omega = 0$ implies no preference for informal consumption, and $\omega > 1$ indicates a household's preference for informal consumption over formal. Although the intuition for the latter is less straightforward, one might speculate that consumption within the firm can lead to enhanced status, access to exclusive goods, or expectations of future income. As noted by Clotfelter (1983), individuals can use a firm's income for "Travel & Entertainment" (T& E), which within the firm can manifest as luxuries like first-class airfare and luxury hotel accommodations. Simultaneously, these resources can be integral to various business operations, such as meetings with current or potential clients. We will revisit this discussion later.

The individual's problem is to choose $c(\alpha, \varepsilon)$, $h(\alpha, \varepsilon)$, and $e(\alpha, \varepsilon)$ to maximize:

$$\max_{\{c(\alpha,\varepsilon),h(\alpha,\varepsilon),e(\alpha,\varepsilon)\}} \int \left[\log c(\alpha, \varepsilon) - \frac{h(\alpha, \varepsilon)^{1+\sigma}}{1+\sigma} + \omega \log e(\alpha, \varepsilon) \right] dF_\varepsilon(\varepsilon), \tag{12}$$

subject to (1), (11), and (3). Due to the greater disutility associated with lower formal consumption, it is guaranteed that it will never be zero. Formally, this is established by two Inada conditions present in the model: $\lim_{c\to 0} \partial u(c, h, e)/\partial c = \infty_+$ and $\lim_{c\to\infty_+} \partial u(c, h, e)/\partial c = 0$.

Solving the first-order conditions using the HSV tax function yields the following results:

$$c(\alpha) = \left(1 + \frac{\omega}{1-\tau}\right)^{\frac{-\sigma(1-\tau)}{1+\sigma}} \lambda(1-\tau)^{\frac{1-\tau}{1+\sigma}} \left\{\mathbf{E}\left[\exp(\varepsilon)^{\frac{1+\sigma}{\sigma}}\right]\right\}^{\frac{\sigma(1-\tau)}{1+\sigma}} \exp((1-\tau)\alpha), \tag{13}$$

$$h(\varepsilon) = \left(1 + \frac{\omega}{1-\tau}\right)^{\frac{1}{1+\sigma}} (1-\tau)^{\frac{1}{1+\sigma}} \left\{\mathbf{E}\left[\exp(\varepsilon)^{\frac{1+\sigma}{\sigma}}\right]\right\}^{\frac{-1}{1+\sigma}} \exp\left(\frac{1}{\sigma}\varepsilon\right), \tag{14}$$

$$y(\alpha) = \left(1 + \frac{\omega}{1-\tau}\right)^{\frac{1}{1+\sigma}} (1-\tau)^{\frac{1}{1+\sigma}} \left\{\mathbf{E}\left[\exp(\varepsilon)^{\frac{1+\sigma}{\sigma}}\right]\right\}^{\frac{\sigma}{1+\sigma}} \exp(\alpha), \tag{15}$$

$$e(\alpha) = \omega\left(1 + \frac{\omega}{1-\tau}\right)^{\frac{-\sigma}{1+\sigma}} (1-\tau)^{\frac{-\sigma}{1+\sigma}} \left\{\mathbf{E}\left[\exp(\varepsilon)^{\frac{1+\sigma}{\sigma}}\right]\right\}^{\frac{\sigma}{1+\sigma}} \exp(\alpha). \tag{16}$$

It is worth noting that when ω equals zero, informal consumption ceases to exist, effectively returning us to the model described in Heathcote and Tsujiyama (2021) and yielding its associated outcomes.

Finally, we describe the Ramsey planners used to derive welfare optima. Given that the only parametric function being analyzed is an HSV one, the planners choose τ to maximize the Pareto-weighted utility functions. Unlike the original case, we can now define two planners: one that knows about tax avoidance and one that does not. We define the first planner as:

$$\max_{\tau} \int W(\alpha) \int u(c(\alpha, \varepsilon), h(\alpha, \varepsilon), e(\alpha, \varepsilon)) dF_\alpha(\alpha) dF_\varepsilon(\varepsilon) \tag{17}$$

$$\begin{aligned} \text{s.t.} \int\int c(\alpha, \varepsilon) + e(\alpha, \varepsilon) dF_\alpha(\alpha) dF_\varepsilon(\varepsilon) + G &= \int\int \exp(\alpha + \varepsilon) h(\alpha, \varepsilon) dF_\alpha(\alpha) dF_\varepsilon(\varepsilon) \\ &\text{Conditions (13), (14), and (16).} \end{aligned} \tag{18}$$

This planner knows about tax avoidance but cannot solve it. Therefore, its decision is to choose τ to maximize the expected utility of society, considering avoidance in the utility, and subject to the decentralized equilibrium and the condition that aggregate output will be destined for consumption, informal consumption, and the public good.

We define the second planner as:

$$\max_{\tau} \int W(\alpha) \int u(c(\alpha, \varepsilon), \tilde{h}(\alpha, \varepsilon)) dF_\alpha(\alpha) dF_\varepsilon(\varepsilon) \tag{19}$$

$$\begin{aligned} \text{s.t.} \int\int c(\alpha, \varepsilon) dF_\alpha(\alpha) dF_\varepsilon(\varepsilon) + G &= \int\int \exp(\alpha + \varepsilon) h(\alpha, \varepsilon) - e(\alpha, \varepsilon) dF_\alpha(\alpha) dF_\varepsilon(\varepsilon) \\ &\text{Conditions (13), (15), and (16).} \end{aligned} \tag{20}$$

This planner, unlike the first one, cannot observe tax avoidance. Therefore, it maximizes expected utility without informal consumption, subject to the condition that post-avoidance aggregate income is destined to pay for formal consumption and the public good. Moreover, note that the hours worked that the planner can deduce now are different from those before. To see that, remember that the planner can see only end-of-period household income, originally defined by $y(\alpha) = \int \exp(\alpha+\varepsilon)h(\alpha,\varepsilon)dF_{\varepsilon}(\varepsilon)$. Now, the planner will observe $y(\alpha) - e(\alpha)$, which will be consistent with another $h(\varepsilon)$ allocation, which we denote by $\tilde{h}(\varepsilon)$. These changes, along with the changes in avoidance, which will be seen as changes in aggregate output, will determine that this planner will overestimate the behavioral effects of the change in progressivity.

3. Homogeneous Informal Consumption

In this section, we explore the scenario where ω remains constant for all individuals. It is crucial to emphasize that this parameter is uniformly set for every household in the distribution; thus, every agent in the economy has the same preferences for informal consumption or the same technology for tax avoidance. While it is possible, under this specification, that households at the top of the distribution have more informal consumption, interpreting this as differences in the ability to avoid taxes a priori may not be accurate. This assumption is strong, and we relax it in Section 4.

3.1. Calibration

Let us begin by summarizing the calibration presented by Heathcote and Tsujiyama (2021) in their paper. Some parameters follow straightforward targets, as shown in Table 1, while others require more detailed explanation. The parameters with straightforward values include σ, set to 2 to match a Frisch elasticity ($1/\sigma$) of 0.5; τ, set to 0.181, determined using the slope of $\log(y)$ and $\log(y - T(y))$ in household surveys by ordinary least squares (OLS); and λ, set to match the household budget constraint while maintaining government purchases at 18.8 percent of GDP. A more detailed explanation of the composition of pre and post-tax income used to calculate τ is provided in Section 4.2.

Now, let us delve into the parameters that need more explanation. Firstly, the authors assume that the stochastic terms follow the distributions: $\varepsilon \sim N(-\sigma_{\varepsilon}^2/2, \sigma_{\varepsilon}^2)$ and $\alpha = \alpha_N + \alpha_E$, where $\alpha_N \sim N(\mu_{\alpha}, \sigma_{\alpha}^2)$ and $\alpha_E \sim Exp(\lambda_{\alpha})$, resulting in $\alpha \sim EMG(\mu_{\alpha}, \sigma_{\alpha}^2, \lambda_{\alpha})$. This implies that wages at the individual level follow a log-normal distribution with a Pareto tail, or a Pareto lognormal distribution. The rationale behind this departure originates from the long-standing recognition that the upper tail of income (and, as posited, wealth) distributions aligns well with a Pareto distribution, initially proposed by Pareto (1896) and commonly employed in the literature (Piketty and Saez, 2003; Piketty, 2003; Atkinson et al., 2011).

Moving on to equations 13 and 15, the equilibrium distributions for log earnings and log consumption are also EMG with:

$$Var[\log y] = \left(\frac{1+\sigma}{\sigma}\right)^2 \sigma_\varepsilon^2 + \sigma_\alpha^2 + \frac{1}{\lambda_\alpha^2}, \tag{21}$$

$$Var[\log c] = (1-\tau)^2\sigma_\alpha^2 + \frac{(1-\tau)^2}{\lambda_\alpha^2}, \tag{22}$$

a relationship maintained in our results. The author's strategy involves using an empirical distribution for log earnings to estimate the normal variance $\sigma_y^2 = \left(\frac{1+\sigma}{\sigma}\right)^2 \sigma_\varepsilon^2 + \sigma_\alpha^2$ and the tail parameter λ_α^2. The variance of log consumption is then estimated to infer σ_α^2. Residually, the variance of log earnings identifies σ_ε^2. The parameter μ_α is exclusively determined to align with the distribution on the grid as defined by the authors.

Finally, the authors consider a Pareto weight function of $W(\alpha;\theta) = \frac{\exp(-\theta\alpha)}{\int \exp(-\theta\alpha)dF_\alpha(\alpha)}$, for $\alpha \in \mathcal{A}$. Within this framework, the weight assigned to an agent with uninsurable idiosyncratic productivity α is $\exp(-\theta\alpha)$, where the parameter θ determines the planner's inclination for redistribution, with $\theta > 0$ indicating a concern for the poor that goes beyond utilitarian considerations.

Table 1. Calibration of the model of Heathcote and Tsujiyama (2021)

Parameter	Value	Target
σ	2	Frisch elasticity ($1/\sigma$) equal to 0.5.
τ	0.181	OLS using the Panel Study of Income Dynamics (PSID).
λ_α	2.2	MLE using the Survey of Consumer Finances (SCF).
σ_y^2	0.412	MLE using the Survey of Consumer Finances (SCF).
σ_α^2	0.142	Match equation 22.
σ_ε^2	0.12	Residual. Match equations 21 and 22.
λ	0.84	US government consumption G/Y of 0.188 for 2000-2006.
θ	0	Utilitarian planner.

Note: all value parameters are retained, except for λ, which is set to 0.79 to align with the budget constraints.

In this context, our goal is to configure ω to align with the observed values of tax avoidance in reality. This data is inherently challenging to directly observe. Therefore, we heavily rely on the findings of Johns and Slemrod (2010) and Guyton et al. (2021)[2], who utilize administrative data to quantify income tax avoidance. The authors employ a combination of IRS random audits data, targeted enforcement activities, and operational audits to estimate a distribution of tax evasion. This involves calculating an *income under-reporting gap* (the amount of income under-reported as a fraction of true income) and a *tax gap* (the amount of legally owed tax that is not paid, expressed as a fraction of the amount legally owed). The identified evasion takes various forms, including unreported self-employment income, overstated deductions, abuse of tax

[2]A more refined version of this paper is available in Guyton et al. (2023). Despite the alterations in the principal results, we utilize the preceding ones for our analysis.

credits, foreign intermediaries (such as foreign bank accounts), and pass-through businesses. We leverage some general data provided by the paper to calibrate our parameter ω.

Figure 4 illustrates the aggregate effects of varying values of ω. It is important to note that, while the parameter can theoretically exceed one, such values are unlikely given the aggregate results it produces. For instance, when $\omega = 1$, the aggregate tax avoidance as a percentage of GDP exceeds 50% (Panel B), the total taxes paid as a percentage of legally owed taxes is less than 40% (Panel C), and the mean ratio of informal to formal consumption approaches 2 (Panel D).

In contrast, with $\omega = 0$, the results align with the original model: a progressivity parameter beyond 0.3 (Panel A) and aggregate avoidance measured at 0 in Panels B to D. The relationship is observed to be negative for optimal progressivity and aggregate taxes paid, while it is positive for aggregate avoidance and the ratio of informal to formal consumption. According to Guyton et al. (2021), approximately 14% of aggregate income goes unreported, and 20% of taxes are not paid. An $\omega = 0.13$ closely approximates these figures, yielding values of 13.6% and 21.15%, respectively.

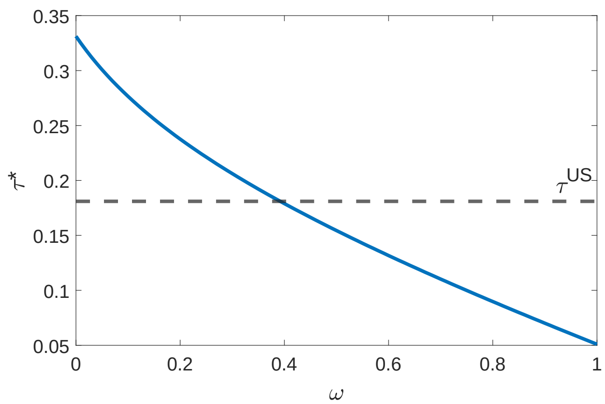
The rest of parameters remain unchanged, except for λ, which changes from 0.84 to 0.79 due to the introduction of the informal consumption decision.

3.2. Results

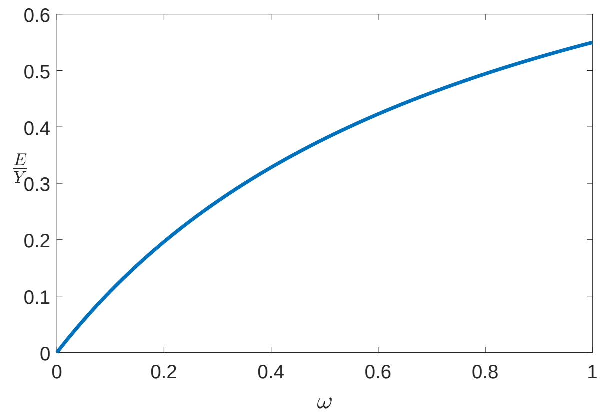
Using the latter calibration, we proceed to the quantitative analysis of the model. Figure 5 displays the simulation results, aggregating the option of tax avoidance at the household level. It is evident that households engage in tax avoidance, with the extent increasing as household income rises. This implies that households pay lower average and marginal tax rates, with a gap of roughly 10% between the rate that should be paid (which we denote as "Theoretical") and the rate that is paid (which we call "Effective") in the top income brackets. In these extreme cases, informal consumption is nearly half of formal consumption. However, contrary to expectations, there is no regressivity at the top, as indicated by empirical findings. Panel 3 sheds light on this phenomenon: due to our modeling approach, all households exhibit the same percentage of income allocated to avoidance, regardless of their varying income levels. Note that this is feasible due to the positive relation between taxes paid and income. This results in richer households consuming more informally without a corresponding increase in regressivity. This pattern is also observed in the first two panels, where the rates consistently increase.

Concerning the planner exercise, it is noteworthy that in the original study by Heathcote and Tsujiyama (2021), the optimal progressivity value was 0.331—considerably high, especially when compared to the empirical findings reported by Holter et al. (2019). According to their data, Denmark had the highest value at 0.258, while Japan had the minimum at 0.101. Even considering the possibility of downward bias in this data (as evidenced by those authors calculating a parameter value of 0.137 for the U.S.), the progressivity parameter calculated by Heathcote remains notably high.

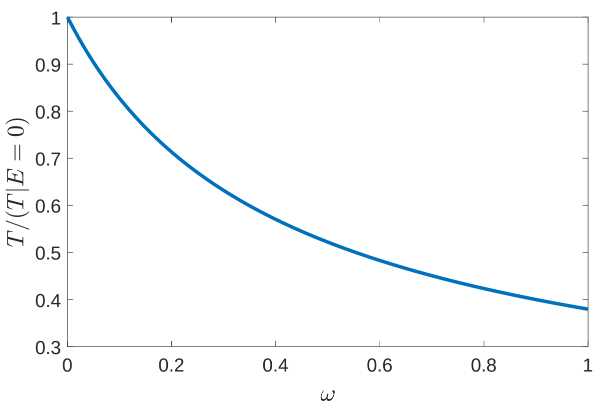
Figure 4. Aggregate fiscal consequences of different values of ω



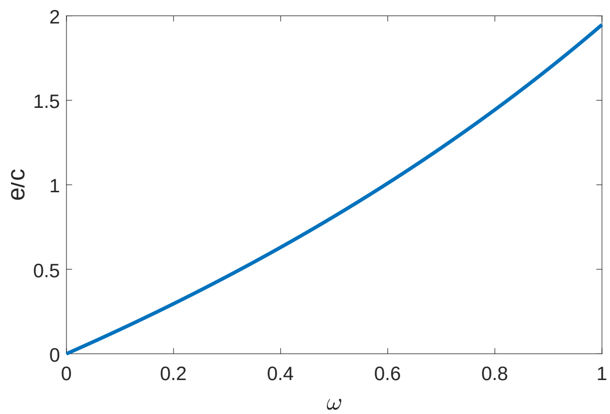
A. Optimal progressivity.



B. Income under-reporting gap.



C. (Inverse of) Tax gap.



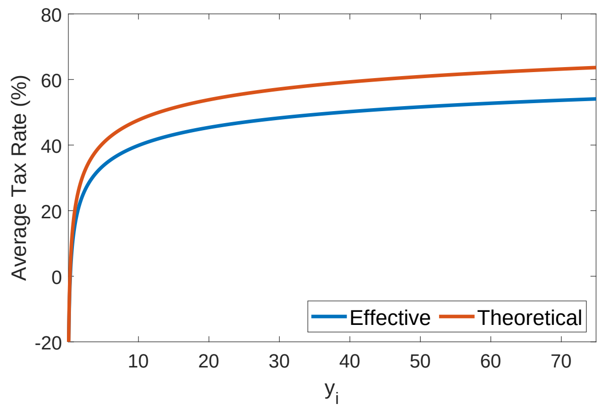
D. Average informal over formal consumption.

Note: In all panels except Panel A, the y-axis is in proportions, where a value of 1 represents 100%. Here, τ^* denotes the optimal progressivity rate set by the planner, E denotes aggregate informal consumption, Y denotes aggregate GDP, T denotes aggregate taxes paid, and e and c denote average informal and formal consumption, respectively.

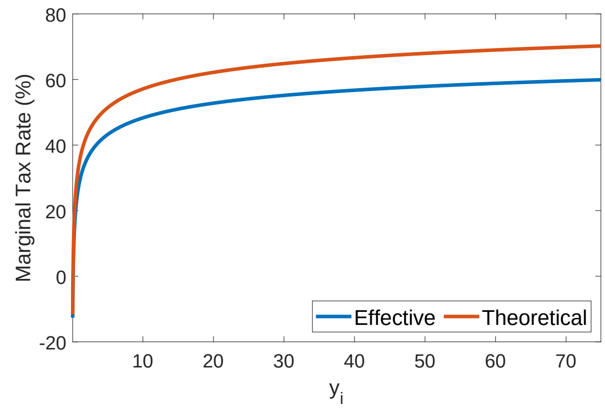
Following the authors, we conduct a similar planner analysis, as shown in Table 2. In the table, HSV^{US} denotes the baseline HSV approximation to the current U.S. tax and transfer system, and HSV represents the optimal Ramsey planner decision in their analysis. Additionally, $HSV^{US}_{e>0}$ is our approximation to the current U.S. tax and transfer system, which includes tax avoidance. We also introduce HSV^1 and HSV^2 as the planner's optimum decisions, where the first planner is aware of the existence of tax avoidance but cannot address it, and the second planner is unaware of any avoidance, as discussed earlier.

The initial observation reveals disparities between our baseline approximation and that of Heathcote and Tsujiyama (2021), particularly evident in scenarios involving tax avoidance. Notably, there is a systematic reduction in the fiscal burden, characterized by a lower average income-weighted marginal tax and a diminished volume of transfers. Despite the observed

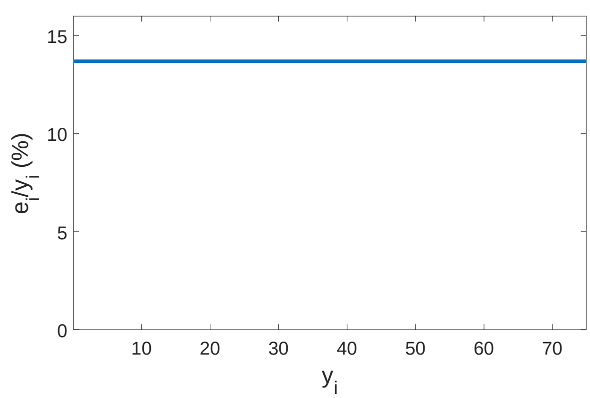
Figure 5. Simulation results when ω is homogeneous



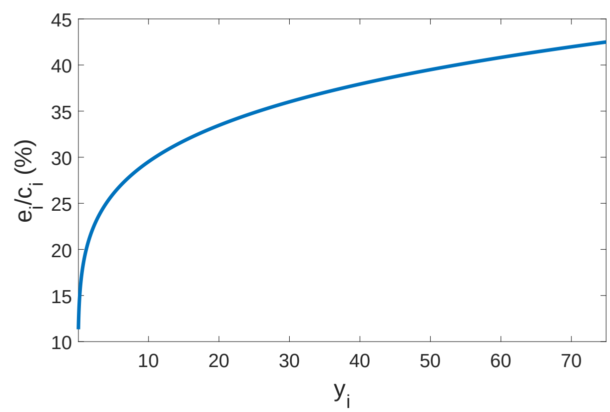
A. Average Tax Rate.



B. Marginal Tax Rate.



C. Informal consumption over income.



D. Informal consumption over formal consumption.

Note: we denote the rate that should be paid as "Theoretical" and the rate that is paid as "Effective." Here, y_i represents household income, c_i denotes consumption, and e_i stands for informal consumption.

differences, the numerical variances are not substantive.

The outcomes of the planner exercises carry particular significance. Our planners consistently generate lower values of τ, indicating a reluctance to escalate progressivity to the extent observed in the authors' analysis. This hesitation can be attributed to the diminished effectiveness of tax increases on the affluent, primarily due to prevalent tax avoidance within this group. It is crucial to note that high-income households experience a change in the relative prices between informal and formal consumption, rendering formal consumption more expensive once the tax policy shifts towards greater progressivity.

The calculated value of τ approximates 0.264 for the first planner and approximately 0.2 for the second. The former closely aligns with Denmark's estimate in Holter et al. (2019), while the latter corresponds to the UK estimate in the same study. Naturally, the first planner advocates for heightened progressivity, leveraging their ability to discern genuine movements in aggre-

Table 2. Ramsey Optimal Taxation when ω is homogeneous

System	Parameters		Outcomes					
	λ	τ	$\bar{T}'(\%)$	$Tr(\$)$	$\frac{Tr}{Y}(\%)$	$\frac{Tr+G}{Y}(\%)$	$WG(\%)$	$\Delta Y(\%)$
Heathcote and Tsujiyama (2021)								
HSV^{US}	0.84	0.181	33.5	1,753	2.3	21.1	-	-
HSV	0.82	0.331	46.6	4,632	6.4	26.5	1.65	-6.53
Own								
$HSV^{US}_{e>0}$	0.79	0.181	31.0	1,335	1.7	20.5	-	-
HSV^1	0.77	0.264	30.6	2,519	3.4	22.7	0.60	-2.99
HSV^2	0.79	0.200	30.9	1,590	2.1	21.0	0.24	-0.67

Note: $\bar{T}'$ is the average income-weighted marginal tax rate in percent. Tr is transfers defined as consumption minus income for the lowest earning household in 2007 dollars. $\frac{Tr}{Y}$ is transfers as a percentage of average income. $\frac{Tr+G}{Y}$ is total government spending, measured as transfers plus government purchases, as a percentage of average income. WG is the welfare gain of moving from the current tax system T to the optimal one $\hat{T}$, defined as the percentage increase in consumption for all agents under policy T that leaves the planner indifferent between T and $\hat{T}$. ΔY is the associated percentage change in aggregate output.

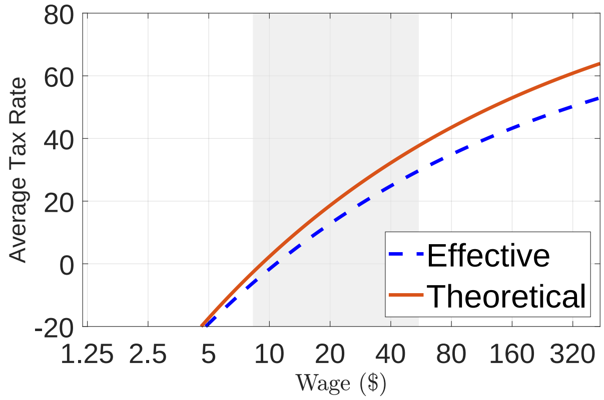
gate income. Conversely, the second planner, lacking insight into authentic income dynamics, interprets tax evasion as an aggregate income reduction. In the same vein, under both optimal schedules, the reduction in aggregate product is less than in the original optimum.

Concluding the analysis, it is pertinent to note that the fiscal burden, quantified as the average income-weighted marginal tax rate in percentage terms, is not significantly increased under the planner optimum; instead, a reduction is observed. This outcome may be attributed to two effects: a progressivity effect, as over 50% of households witness a decline in taxes under this system (to be discussed later), and an avoidance effect, given that those experiencing increased taxes are primarily those engaged in heightened tax avoidance.

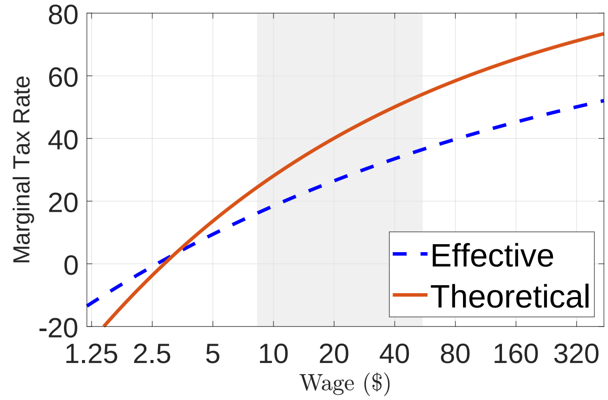
Figure 6 depicts the average and marginal taxes that decentralize the constrained efficient allocation, plotted against observed hourly wages. The optimal average rate initiates at a positive level around the 5th percentile, remaining below 40% under the optimum that an omniscient planner would impose at the 95th percentile. Naturally, under the myopic planner, average rates are higher for poorer households. Optimal marginal rates exceed 50% for the 95th percentile and reach up to 73% for hourly wages surpassing $320, closely aligning with figures found in the literature (see, for example, Diamond and Saez, 2011). "Effective" rates are consistently lower under every schedule, consistent with the preceding analysis. Notably, for the less affluent households, the negative rates are closer to zero than expected if everyone pays, which is intuitive given the lower revenue collection than the ideal scenario, resulting in fewer transfers being feasible.

One question that may arise is whether the planner raises progressivity at the cost of the lower or middle classes paying more, resulting in nominal progressivity increase but regressive effects. This could occur if behavioral effects in the top income brackets are significant enough to

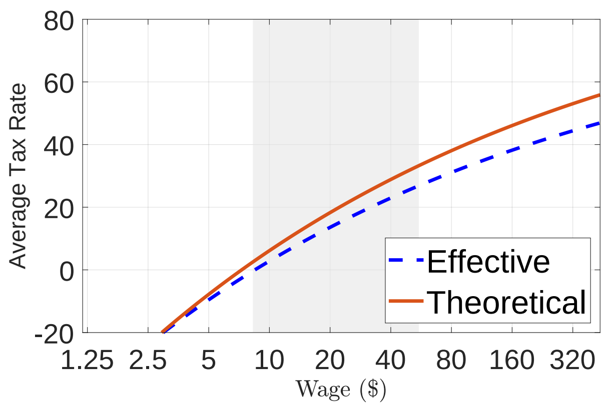
Figure 6. Ramsey Optimal Taxation when ω is homogeneous



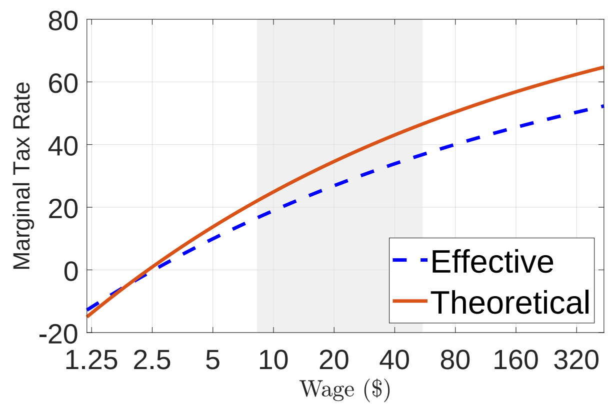
A. Average Tax Rate (%). First Planner.



B. Marginal Tax Rate (%). First Planner.



C. Average Tax Rate (%). Second Planner.



D. Marginal Tax Rate (%). Second Planner.

Note: The x axis for each plot shows the household average hourly wage, $\bar{w}\exp(\alpha)$. The area between the 5th and 95th percentiles is shaded gray.

cause higher payments from those in the middle. However, given the monotonically increasing functions observed in Figure 5, this scenario does not seem likely. We will revisit this discussion when analyzing the heterogeneous case.

Table 3 offers insights into the proportion of total taxes attributable to each percentile, calculated using Effective Average Tax Rates, inclusive of avoidance. Moving from the baseline case to the optimum set by the planner (either the myopic or the omniscient) decreases taxes for percentiles 0 to 80 and increases thereafter. This increase becomes progressively larger across percentiles, but the top 0.1% experiences a slightly smaller increase compared to those preceding them (given that $\frac{8.43-7.34}{7.34} > \frac{6.31-5.53}{5.53}$). Nevertheless, high-income households bear the additional burden, as evident in the table. Even if the middle classes share the rise in taxes with high-income households, the progressivity is still maintained in the total structure of taxes

Table 3. Total taxes paid by percentile (% of the total)

	Taxes paid by the group (%)		
Percentile	$HSV_{e>0}^{US}$	HSV^1	HSV^2
0-10	0.18	-1.16	-0.13
10-20	1.29	0.02	0.99
20-30	2.22	1.03	1.93
30-40	3.19	2.12	2.93
40-50	4.28	3.36	4.05
50-60	5.64	4.90	5.45
60-70	7.48	7.01	7.36
70-80	10.19	10.14	10.16
80-90	15.51	16.29	15.68
90-95	12.59	13.76	12.86
95-99	19.40	21.87	20.00
99-99.5	5.17	5.91	5.35
99.5-99.9	7.34	8.43	7.62
99.9-100	5.53	6.31	5.74

Note: $HSV_{e>0}^{US}$ represents our approximation of the current U.S. tax and transfer system, accounting for tax avoidance. HSV^1 and HSV^2 denote the regimes under the planner's optimal decisions, where the first planner is omniscient, and the second planner is myopic. The columns sum to 100.

4. Heterogeneous Informal Consumption

We now explore the scenario where there can be a heterogeneous parameter of utility for formal consumption, denoted as $\omega = \omega(\alpha)$ for all $\alpha \in \mathcal{A}$. This assumption is intuitive when interpreting the parameter as a technology of evasion, suggesting differences between households. Specifically, households with a higher idiosyncratic uninsurable shock α might possess a more effective technology of evasion: they may possess deeper knowledge of legal intricacies, adeptness in managing firm expenditures, and a higher capacity to engage in tax avoidance, among other factors. Furthermore, they might place a different value on informal consumption. As previously discussed, individuals can use their income within the firm to purchase luxury goods, such as T&E, and this might be more prevalent among individuals in the highest ranks of firm administration. For illustrative purposes, one might posit that individuals with lower skills receive only fringe benefits if they engage in tax avoidance, compared to T&E, which might be enjoyed by individuals with high skills. The underlying mechanism may vary, but the basic idea is that individuals could have different preferences for evasion.

Importantly, this does not imply a monotonically increasing function. Individuals in the ultra-top might value tax avoidance less because they are more exposed to public scrutiny, or there may be a higher likelihood of, for instance, a random audit being conducted on them. They might even choose not to avoid taxes due to concerns about reputation. We will revisit this point later when discussing the calibration of ω.

4.1. Calibration

Unlike the original calibration by Heathcote and Tsujiyama (2021), we encounter different issues that prompt a change in our approach. Firstly, observe that using equation 13 no longer yields $Var[\ln c(\alpha)] = Var[(1-\tau)\alpha]$, so we no longer have equation 22. In fact, we now have:

$$Var[\ln y(\alpha, \varepsilon)] = Var\left[\frac{1}{1+\sigma}\ln\left(1+\frac{\omega(\alpha)}{1-\tau}\right)+(1-\tau)\alpha\right]+\left(\frac{1+\sigma}{\sigma}\right)\sigma_\varepsilon^2, \tag{23}$$

$$Var[\ln c(\alpha)] = Var\left[\frac{-\sigma(1-\tau)}{1+\sigma}\ln\left(1+\frac{\omega(\alpha)}{1-\tau}\right)+(1-\tau)\alpha\right]. \tag{24}$$

The approach of allowing a heterogeneous parameter introduces several challenges and compromises certain desirable properties. Firstly, the relationship between α and income is no longer guaranteed, as the assurance of the income being *EMG* is compromised. Secondly, the connections between the variances of both log consumption and log income are severed. Additionally, determining the variance of ω and the covariance between ω and α becomes necessary, introducing dependencies on the parameter μ_α—which was previously determined residually—and the parameters governing the ω function. Thirdly, in conjunction with the aforementioned challenges, a specific function for ω must be assumed, adding more parameters to be estimated. Given these complexities, we are faced with a high number of variables for only two equations.

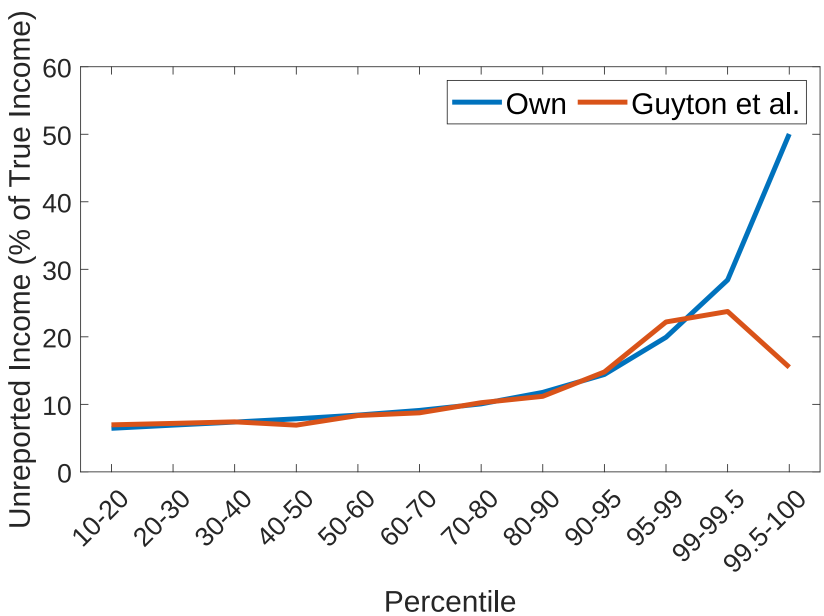
Our strategy involves assuming a functional form for $\omega(\alpha; \delta)$, now dependent on a vector of parameters δ. We then choose these parameters to match some of the original moments of the distribution used by Heathcote and Tsujiyama. We propose a four-step calibration process with this aim.

Let us denote by x^0 a parameter x that Heathcote and Tsujiyama derived in their calibration; for example, σ_ε^0 represents the calibrated insurable shock variance, and y^0 represents the income distribution they found. Next, we follow the calibration procedure outlined below. The objective is to match some of the original moments of the distribution used by the authors and an empirical distribution of tax avoidance. Considering the latter, we use the distribution found by Guyton et al. (2021), which can be seen in Figure 7. This distribution shows a relatively constant unreported income of 7% for poorer households and subsequently increases to more than 20% until the 99th percentile, where it starts to decline. Then, we proceed as follows.

1. Starting with the assumption that both shocks are distributed in the same way as in the original calibration (i.e., $\alpha = \alpha_0$ and $\varepsilon = \varepsilon_0$), assume a functional form, such as $\omega(\alpha) = \delta_0 + \delta_1 \exp(\alpha)$. Utilize a grid to simulate the model with evasion for various values of δ_0 and δ_1, selecting the combination that minimizes the residuals of the empirical distribution of tax avoidance.

Two related points are worth noting. First, the exponential functional form is a strong assumption, implying that households in the top have a tax avoidance that is significantly greater than what is observed empirically. The underlying idea is that wealthy individuals evade more than can be observed even in high-quality data. While we maintain this assumption throughout the paper, it's important to acknowledge that alternative functional forms, such as polynomial, logistic or mixture distributions, could also match the empirical distribution. Second, given our assumption that richer households avoid more than what is empirically observed, we focus on matching the distribution up to the 99th percentile. The results are shown in Figure 7.

Figure 7. Unreported income in our model and corrected random audit data



Note: this figure displays the distribution of under-reported income in the 2006-2013 National Research Program (NRP) data, adjusted using detection-controlled estimation (DCE) as derived by Guyton et al. (2021). It also includes the simulated results of our model.

2. Let y represent the distribution of observable incomes, and e be the distribution of tax evasion, both as vectors of the same length. Assume $y - e = y^0$. This implies that the observed empirical data of income, whether in the Heathcote and Tsujiyama model or a corrected household survey, is the data of income after evasion—an assumption we acknowledge.

   After this assumption, we note that using equations 15 and 16, we can obtain y^0 depending only on α and a set of parameters.

$$
\begin{aligned}
y^0 &= y - e \\
y^0 &= \left(1 + \frac{\omega(\alpha)}{1-\tau}\right)^{\frac{1}{1+\sigma}} (1-\tau)^{\frac{1}{1+\sigma}} \left\{\mathbf{E}\left[\exp(\varepsilon)^{\frac{1+\sigma}{\sigma}}\right]\right\}^{\frac{\sigma}{1+\sigma}} \exp(\alpha) \\
&\qquad - \omega(\alpha)\left(1 + \frac{\omega(\alpha)}{1-\tau}\right)^{\frac{-\sigma}{1+\sigma}} (1-\tau)^{\frac{-\sigma}{1+\sigma}} \left\{\mathbf{E}\left[\exp(\varepsilon)^{\frac{1+\sigma}{\sigma}}\right]\right\}^{\frac{\sigma}{1+\sigma}} \exp(\alpha) \\
y^0 &= \left\{\mathbf{E}\left[\exp(\varepsilon)^{\frac{1+\sigma}{\sigma}}\right]\right\}^{\frac{\sigma}{1+\sigma}} \exp(\alpha) \left[\left(1 + \frac{\delta_0 + \delta_1 \exp(\alpha)}{1-\tau}\right)^{\frac{1}{1+\sigma}} (1-\tau)^{\frac{1}{1+\sigma}} \right. \\
&\qquad \left. - (\delta_0 + \delta_1 \exp(\alpha))\left(1 + \frac{\delta_0 + \delta_1 \exp(\alpha)}{1-\tau}\right)^{\frac{-\sigma}{1+\sigma}} (1-\tau)^{\frac{-\sigma}{1+\sigma}}\right]
\end{aligned}
\tag{25}
$$

In particular, if we set $\mathbf{E}\left[\exp(\varepsilon)^{\frac{1+\sigma}{\sigma}}\right]$, which is a function of σ_ε, then we can fully determine a vector of α that exactly corresponds to the vector of y^0.

Presume $\sigma_\varepsilon = \sigma_\varepsilon^0$, which implies that the distribution of the insurable shock remains the same. Then, we are allowing the extraction of a unique α for each income using numerical methods and household conditions. We will have a vector of α, which will have a certain distribution.

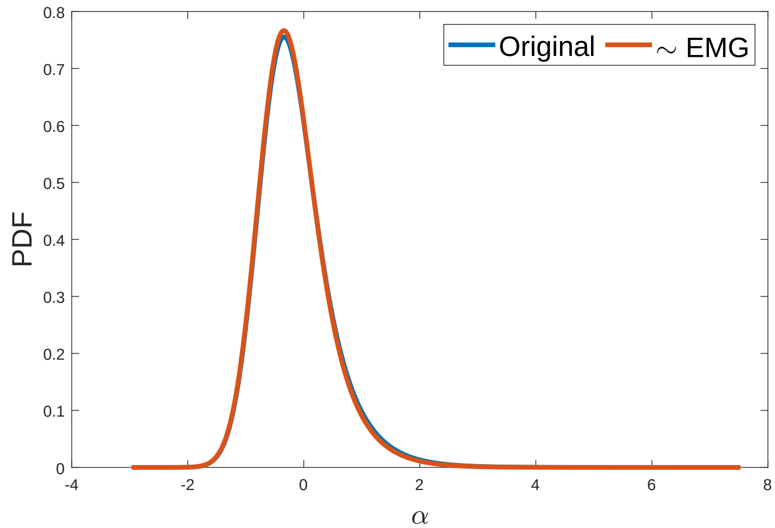
3. Assume $\alpha \sim EMG(\mu_\alpha, \sigma_\alpha, \lambda_\alpha)$. This means that the distribution of α is of the same family as the original model. We then estimate these parameters through Maximum Likelihood Estimation (MLE) using numerical methods. The results are shown in Figure 8. Although the fit is not perfect, the *EMG* distribution approximates very closely the original distribution.

4. Iterate with new grids until converging on values for both the distribution of α and δ_0 and δ_1.

The remaining parameters remain unchanged, except for those requiring adjustment to satisfy specific constraints, such as λ, as discussed earlier. These results yield $\delta_0 = 0.032$, $\delta_1 = 0.052$, $\mu_\alpha = -0.6446$, $\sigma_\alpha^2 = 0.3787$, $\lambda_\alpha = 2.1091$. Notably, this calibration produces results that differ from Guyton et al. (2021) data on aggregate avoidance. Specifically, our aggregate underreported income equals 14.25%, and the percentage of taxes not paid equals 25.34%, both higher than the original estimates.

The results of our proposed calibration can be seen in figure 9. As evident, the income distribution after avoidance closely aligns with the original distribution by Heathcote and Tsujiyama (2021), with only a slight extension in the Pareto tail. Note that this figure is truncated for practical reasons, as the original distribution extends to approximately $y_i = 72$. On the other hand, the ω parameter reaches values as high as 100 for individuals at the top of the skills distribution, which is considerably elevated. Once again, this is plausible given the considerations discussed when proposing this parameter.

Figure 8. Calibrated distribution of uninsurable shocks



Note: this figure displays the probability density function of the uninsurable shocks. "Original" denotes the numerical distribution obtained by Equation 25, while "EMG" signifies the maximum likelihood estimation (MLE) approximation.

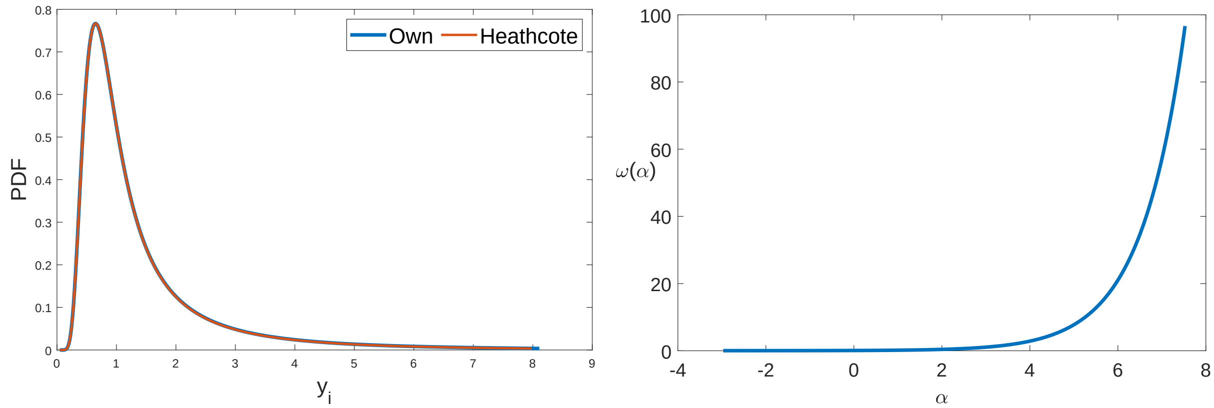
A notable result is that our real income distribution extends as high as 9 thousand, indicating that, accounting for avoidance, the actual income of the highest-ranked household is more than 100 times the income reported to authorities. Consequently, ultra-wealthy households seem to consume a significant portion of their income within the firm, leading to a remarkably low proportion of taxes paid. While we do not delve into the significance of this result, it makes clear that there is a proportion of aggregate income substantially underconsidered by the actual tax distribution, assuming our parametric avoidance distribution is true.

Finally, note that now the elasticity of income related to the progressivity parameter follows the relationship:

$$\xi_{y,1-\tau} = \frac{\partial y(\alpha)}{\partial(1-\tau)}\frac{1-\tau}{y(\alpha)} = \frac{1}{1+\sigma}\left(\frac{1-\tau}{1-\tau+\omega(\alpha)}\right), \tag{26}$$

This is interesting in two ways. First, now the elasticity itself is not policy invariant, so changes in tax policy (i.e., a change in τ) will alter it. Note that $\frac{\partial \xi_{y,1-\tau}}{\partial \tau} < 0$, indicating that changes in policy from lower progressivity rates will have more behavioral effects than when the rates are high, probably because the adjustment is greater beginning with low taxes. Second, $\frac{\partial \xi_{y,1-\tau}}{\partial \omega} < 0$, implying that people with higher tech or preference for avoidance will have a minor response, probably due to their capacity for evasion.

Figure 9. Results of the proposed calibration



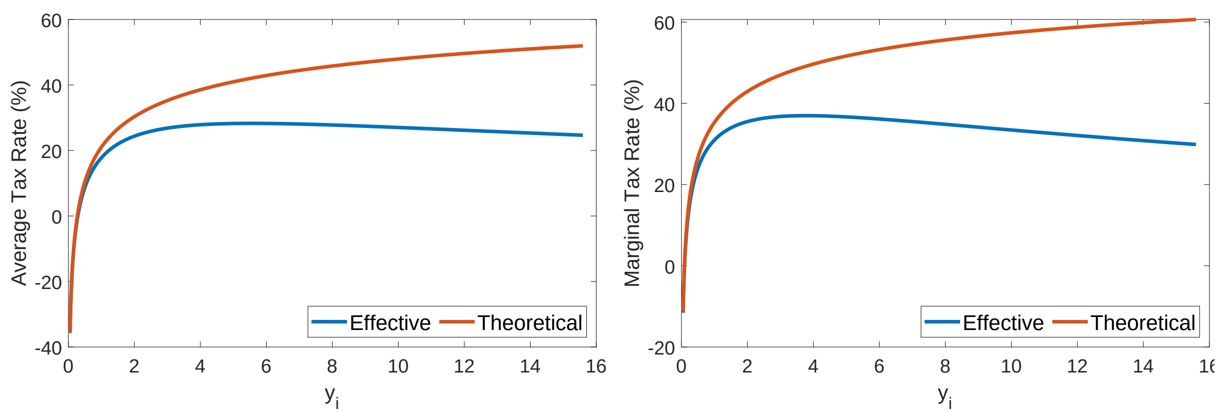
A. Income distribution.

B. Avoidance parameter.

Note: Panel A depicts the probability density function against the income (y) distribution of Heathcote and Tsujiyama (2021) and the post-avoidance income in our calibration ($y - e$). This figure is truncated for practical reasons.

4.2. Results

Figure 10. Simulation results when ω is a function of α



A. Average Tax Rate.

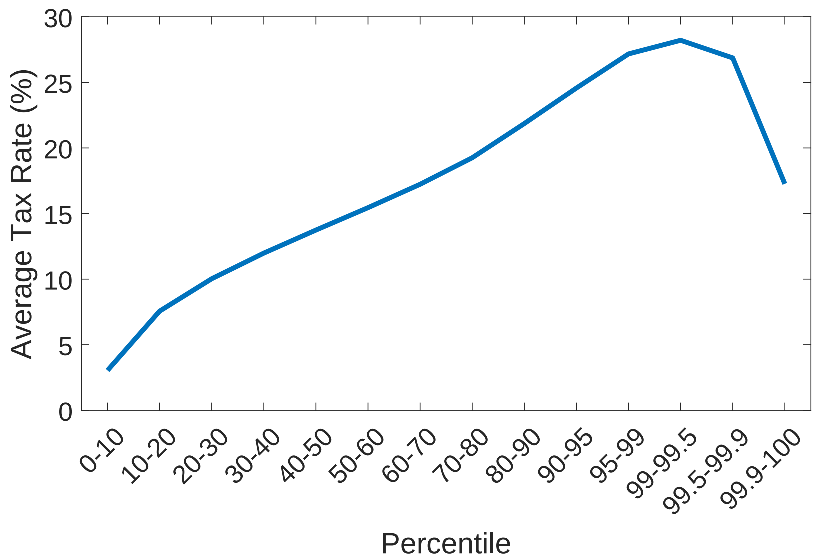
B. Marginal Tax Rate.

Note: see notes to figure 5. This figure is truncated for practical reasons and displays the rates between the 0th and 99.9th percentiles.

Using the latter calibration, we proceed to the quantitative analysis of the model. Figure 10 presents the simulation results, allowing ω to depend on α. Now, it is evident that the rates increase up to a certain point, where they begin to decline due to the exponential growth in the technology or preferences for evasion. The figure is truncated at the 99.9th percentile, revealing a very long tail of individuals in the top 0.1%, where this rate rapidly slows.

More importantly, significant results are displayed in Figure 11, which is the focal point

Figure 11. Average tax rates by income group (% of income)



Note: the figure depicts the average tax rate by income groups in our simulation. Transfers are included in the model.

of this study. The figure shows the average tax rate per percentile, increasing until the 99th percentile, where it begins to decrease. Here, it is observed that the model effectively replicates the regressivity at the very top, a phenomenon seen empirically and documented by Saez and Zucman (2019) for the United States, Piketty (2020) for France and Milligan (2022) for Canada, and one that the model with a homogeneous ω parameter fails to capture.

A few noteworthy points arise from this figure. Firstly, the "HSV" parametrization, as explained by Heathcote et al. (2017), considers that pregovernment gross household income includes various components such as labor earnings, self-employment income, private transfers, plus income from interest, dividends, and rents. Taxable income is then calculated as gross income minus deductions. Postgovernment income equals pregovernment income minus taxes plus transfers, where taxes include federal and state income taxes as well as the total Federal Insurance Contributions Act (FICA) taxes, and transfers include public cash transfers (welfare receipts, unemployment benefits, workers' compensation, and veterans' pensions).

Regarding the comparison with Figure 1, a consideration is needed. The HSV parametrization includes only personal taxes, so Panel A should be discarded. Also, gains from dividends are included, so Panel D should be discarded as well. However, the remaining comparison is not as straightforward: the HSV formulation includes payroll taxes, but also transfers, with the latter not being included in any panel. Given this, we propose that our results might be compared with Panel C, under the assumption that payroll taxes and transfers roughly cancel each other. Naturally, this is the figure that is closest to our results, but it is worth noting that our regressivity starts earlier, around the top 0.1%, while in their results it begins around the top 0.01%, likely due to our specific calibration.

Table 4. Ramsey Optimal Taxation when ω is heterogeneous

System	Parameters		Outcomes					
	λ	τ	$\bar{T}'$(%)	Tr($)	$\frac{Tr}{Y}$(%)	$\frac{Tr+G}{Y}$(%)	WG(%)	ΔY(%)
Heathcote and Tsujiyama (2021)								
HSV^{US}	0.84	0.181	33.5	1,753	2.3	21.1	-	-
HSV	0.82	0.331	46.6	4,632	6.4	26.5	1.65	-6.53
Own								
$HSV^{US}_{e>0}$	0.79	0.181	30.9	1,307	1.69	20.49	-	-
HSV^1	0.78	0.227	30.7	1,948	2.56	21.67	0.164	-1.623
HSV^2	0.79	0.181	30.9	1,310	1.70	20.50	0.001	-0.007

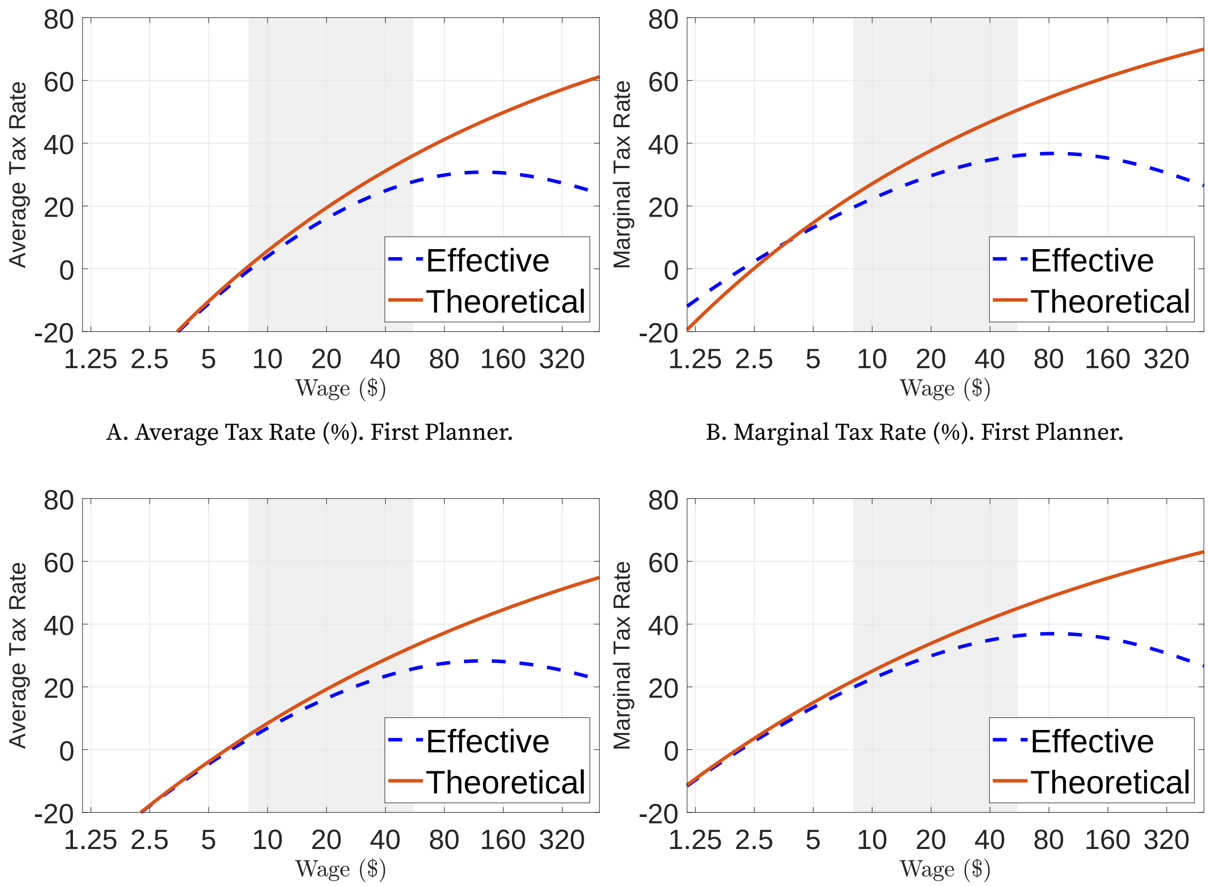
Note: See the notes to table 2.

Now we turn to the analysis of the planner's optimal decisions. As shown in Table 4, we use the same notation as before. The results in direction are similar to the case with a homogeneous ω parameter, but two things are worth noting. First, the progressivity proposed by the omniscient planner is now less than before. In particular, $\tau^* = 0.227$, which is a progressivity rate that is present empirically in countries like Sweden or Ireland, again following Holter et al. (2019). Second, and more importantly, the myopic planner basically doesn't propose a progressivity rise at all. This is consistent with what we can see in reality, where tax rates have been maintained in the last years or even reduced since 1980, given that the possible behavioral effects are magnified.

Figure 12 depicts the average and marginal taxes that decentralize the constrained efficient allocation, plotted against observed hourly wages. The optimal average rate starts to be positive again around the 5th percentile and remains below 40% in the 95th percentile. Optimal marginal rates are below 50% in the 95th percentile but increase to more than 60% for larger hourly wages. Naturally, we observe that effective rates decrease at a certain point, starting after the 95th percentile, indicating that evasion is much more prevalent in richer households. The result of minor transfers is maintained.

Finally, it's worth re-evaluating the analysis of whether progressivity is generated at the cost of the middle classes. This is more probable now, given that the ultra-rich have more options for avoidance. Table 5 shows the same analysis that we did before, but now with the heterogeneous model. As we can see, the results don't change significantly. The bottom of the distribution is still facing lower taxes under the optimal regime, but now the increase begins one decile before, at the 70th-80th percentiles. The change is lower in proportion for the top 0.1% than for the group exactly before, but there are still more taxes paid by the ultra-rich than before.

Figure 12. Ramsey Optimal Taxation when ω is heterogeneous



A. Average Tax Rate (%). First Planner.

B. Marginal Tax Rate (%). First Planner.

C. Average Tax Rate (%). Second Planner.

D. Marginal Tax Rate (%). Second Planner.

Note: The x axis for each plot shows the household average hourly wage, $\bar{w}\exp(\alpha)$. The area between the 5th and 95th percentiles is shaded gray.

5. Other models and discussion

This section aims to examine the impact on our results when considering alternative specifications. Both of these alternatives entail specifications that render unnecessary the assumption of a specific parametric function for tax avoidance. Instead, tax avoidance is determined by parameters representing the dynamics between utility and disutility associated with it.

5.1. CES utility function

Let us consider the model detailed in Appendix A.1, where we change the utility of consumption and avoidance from a Cobb-Douglas to a CES utility. That is, instead of having $u(c, e) = \log c + \omega \log e$, we now have $u(c, e) = \log C(c, e)$ where $C(c, e)$ is an aggregate consumption term commonly used in International Finance literature of the form:

Table 5. Total taxes paid by percentile (% of the total)

Percentile	Taxes paid by the group (%)		
	$HSV^{US}_{e>0}$	HSV^1	HSV^2
0-10	0.53	-0.18	0.52
10-20	1.78	1.12	1.77
20-30	2.79	2.19	2.78
30-40	3.85	3.32	3.84
40-50	5.02	4.58	5.01
50-60	6.46	6.13	6.46
60-70	8.34	8.18	8.34
70-80	11.17	11.25	11.17
80-90	16.46	16.99	16.46
90-95	12.81	13.49	12.81
95-99	18.28	19.52	18.29
99-99.5	4.33	4.66	4.34
99.5-99.9	5.35	5.75	5.36
99.9-100	2.83	3.01	2.83

Note: see notes to table 3.

$$C(c,e) = \left[ac(\alpha,\varepsilon)^{1-\frac{1}{\xi}} + (1-a)e(\alpha,\varepsilon)^{1-\frac{1}{\xi}}\right]^{\frac{1}{1-\frac{1}{\xi}}}.$$

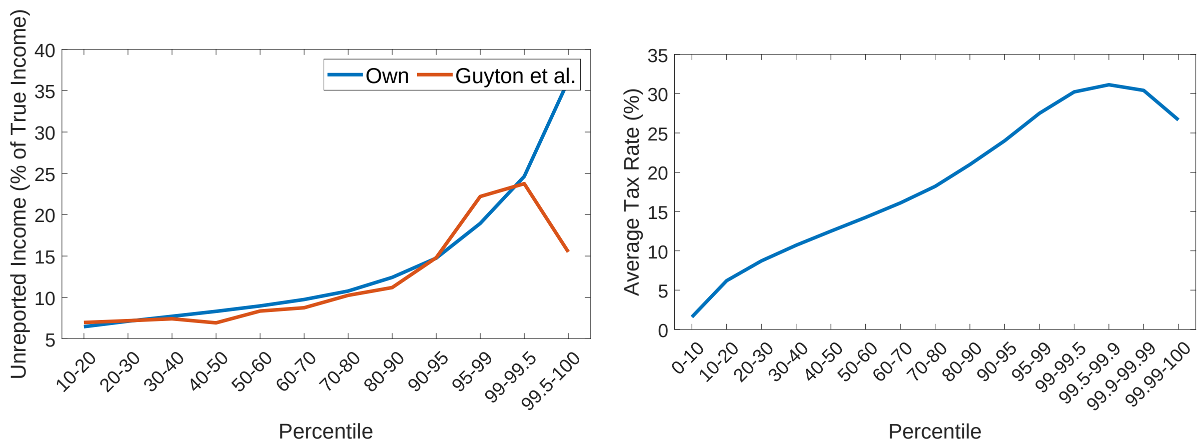
Here, a is a share parameter and ξ the elasticity of substitution between goods.

The resolution of this model closely resembles that of the original model: we begin by solving the first-order conditions and then proceed to integrate towards the distribution of ε to determine the optimum. However, unlike the previous model, this optimum does not possess an analytical form, necessitating numerical solutions. Then, the calibration is quite similar to the heterogeneous model: we start with an initial guess for a and ξ, we assume $y-e$ is equal to the original distribution of income, from which we can extract a vector of α for every income level. We then assume α follows an EMG distribution, and we estimate those parameters using Maximum Likelihood. We repeat all the steps until convergence for both the income distribution and the parameters a and ξ, resulting in a distribution of avoidance that is consistent with what is seen empirically, except at the top of the distribution.

The results of the calibration are $a = 0.7$ and $\xi = 4.5$. The latter is intuitive, showing that the goods consumed after paying taxes or inside the firm are more substitutable than with a Cobb-Douglas function. This can be based on the notion that workers or managers inside the firm are faced with restrictions on the goods they can consume: workers may have fringe benefits that usually come in the form of a bundle, limiting their ability to choose specific brands, or managers may use business travels to engage in entertainment activities in the particular area of travel (even, as mentioned by Clotfelter, if it is for an important business meeting). However, they may not use these funds anywhere.

Given all that, the results of our estimation are summarized in Figure 13, where we see that this model can explain the regressive Average Tax Rate at the top, via the channel of households substituting from formal to informal consumption once they have more income, as we explained in earlier versions of our model. This simulation gives us the result of an *income under-reporting gap* of 13.58% and a *tax gap* of 23.67%, the latter being slightly higher than what was originally found but lesser than that of the heterogeneous model.

Figure 13. Taxes evaded as percentage of taxes owed and Average Tax Rates, by pre-tax income groups



A. Distribution of under-reported income. B. Average tax rates by income group (% of income)

Now let us delve into the results of the planner optimization, which remain similar to what we developed earlier. When we compute its results, our original conclusion changes: as illustrated in Table 6, the planner with perfect knowledge chooses $\tau = 0.1568$, while the planner with imperfect knowledge chooses $\tau = 0.181$, practically without changing it. In that sense, our second planner is conservative not in the sense that it allows for a reduction in progressivity, but in the sense that it maintains the status quo.

Table 6. Ramsey Optimal Taxation when utility function is CES

System	Parameters		Outcomes				
	λ	τ	Tr($)	$\frac{Tr}{Y}$(%)	$\frac{Tr+G}{Y}$(%)	WG(%)	ΔY(%)
$HSV^{US}_{e>0}$	0.80	0.181	1338	1.73	20.53	-	-
HSV^1	0.81	0.1568	1017	1.33	20.36	0.19	-1.18
HSV^2	0.8	0.181	1338	1.73	20.53	-0.00	-0.00

Note: See the notes to table 2. The results for the second planner are equivalent to the baseline.

This puzzle requires further explanation. For intuition, first note that it is possible to express the worked hours as a function of income: given that $y(\alpha) = \int \exp(\alpha + \varepsilon) h(\alpha, \varepsilon) dF_\varepsilon(\varepsilon)$, then, as noted by Heathcote and Tsujiyama (2021), $h(\varepsilon) = \Omega(\sigma_\varepsilon)\frac{y(\alpha)}{exp(\alpha)}$, where $\Omega(\sigma_\varepsilon)$ is a function of the distribution of ε with the only important parameter of that distribution being σ_ε by construction.

Considering the aforementioned, let us examine the decision confronting a planner tasked with selecting between two tax rates: $\tau_0 = 0.181$ and $\tau_1 = 0.1568$.

Given that the model accommodates heterogeneous agents, let us consider an example of the optimal decision, which is determined by evaluating the utility associated with the mean of the distributions of relevant variables. Although this may not directly correspond to assessing the average utility of all agents (which may not vary in the same direction), it can still offer illustrative insights. Herein, y_0, c_0, and e_0 represent the average income, consumption, and evasion of a household under τ_0, while y_1, c_1, and e_1 denote the corresponding values under τ_1. Given that reduced progressivity diminishes the incentives for tax evasion, it is plausible to assume that $c_1 > c_0$ and $e_1 < e_0$. Although it is conceivable that $y_1 < y_0$, it generally follows that $y_1 - e_1 > y_0 - e_0$. Indeed, these relationships are observed in our simulation. Consequently, the decision-making process of the planner is intricately linked to Table 7.

Table 7. Choice of the planner under CES utility

	Planner 1	Planner 2
$\tau_0 = 0.181$	$\log C(c_0, e_0) - \frac{1}{1+\sigma}\left(\Omega(\sigma_\varepsilon)\frac{y_0}{\exp(\alpha)}\right)^{1+\sigma}$	$\log c_0 - \frac{1}{1+\sigma}\left(\Omega(\sigma_\varepsilon)\frac{y_0-e_0}{\exp(\alpha)}\right)^{1+\sigma}$
$\tau_1 = 0.15$	$\log C(c_1, e_1) - \frac{1}{1+\sigma}\left(\Omega(\sigma_\varepsilon)\frac{y_1}{\exp(\alpha)}\right)^{1+\sigma}$	$\log c_1 - \frac{1}{1+\sigma}\left(\Omega(\sigma_\varepsilon)\frac{y_1-e_1}{\exp(\alpha)}\right)^{1+\sigma}$

For the first planner, the decision is straightforward: the utility of consumption increases and the disutility of work diminishes, as c grows and y declines for the average household following the reduction in progressivity. However, for the second planner, the outcome differs. Even though they may consider an increase in the utility of consumption (which, given their lack of observation of e, may not be identical), they also perceive an elevation in the disutility of work, given that $y_1 - e_1 > y_0 - e_0$. In this context, the second planner exhibits a conservative approach, driven by the same mechanism observed in previous models: an overestimation of behavioral effects, stemming from a perception that working hours are greater than they actually are in reality. This result should hold when comparing $\int u(c_j, h_j, e_j)dF_\alpha(\alpha)dF_\varepsilon(\varepsilon)$ for the first planner with $j = 1, 2$ and $\int u(c_j, \tilde{h}_j, e_j)dF\alpha(\alpha)dF_\varepsilon(\varepsilon)$ for the second planner. The critical trade-off arises when considering the effects of $\int \frac{1}{1+\sigma}\left(\Omega(\sigma_\varepsilon)\frac{y_j-e_j}{\exp(\alpha)}\right)^{1+\sigma}dF_\alpha(\alpha)dF_\varepsilon(\varepsilon)$.

Two final observations are noteworthy. Firstly, as showed by the analysis (where $y_1 < y_0$) the elasticity of income with respect to the progressivity parameter may differ between this model and the heterogeneous one. Although we cannot ascertain this with certainty, given the absence of an analytical expression for $\xi_{y,1-\tau}$ in this model, Table 6 indicates that a reduction in τ leads to a decrease in aggregate income, which contrasts with the outcome observed in the heterogeneous model where aggregate income decreases as τ increases. In this regard, unlike the heterogeneous model where $\xi_{y,1-\tau} > 0 \ \forall \tau \in (0,1)$, in this model, there likely exists a more concave relationship between the two.

Table 8. Total taxes paid by percentile (% of the total)

Percentile	Taxes paid by the group (%)		
	$HSV^{US}_{e>0}$	HSV^1	HSV^2
0-10	0.26	0.59	0.26
10-20	1.41	1.69	1.41
20-30	2.35	2.59	2.35
30-40	3.35	3.53	3.34
40-50	4.48	4.60	4.47
50-60	5.87	5.92	5.86
60-70	7.72	7.67	7.71
70-80	10.48	10.30	10.47
80-90	15.81	15.39	15.80
90-95	12.72	12.33	12.72
95-99	19.25	18.74	19.26
99-99.5	4.94	4.88	4.95
99.5-99.9	6.72	6.78	6.74
99.9-100	4.64	4.97	4.66

Note: see notes to table 3.

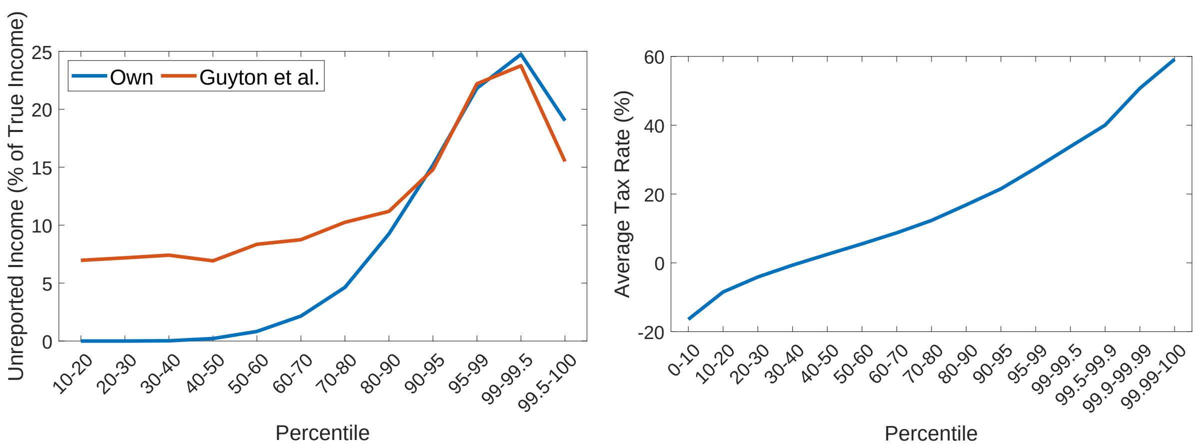
Secondly and finally, as demonstrated in Table 8, the decrease in progressivity within this model does not necessarily translate to an increase in the share of tax revenue for poorer households and a reduction for richer ones in a straightforward manner. Rather, there are heterogeneous dynamics at play, which are quite intriguing: poorer households witness an increase in their share, middle-class households (ranging from decile 6 to the 99th percentile) experience a reduction, and the top 1% observe an increase in their share. This may be attributed to the fact that the reduction in firm consumption is more pronounced for the top percentile compared to other effects. Hence, the decrease in τ does not uniformly signify a decrease in progressivity across the entire income distribution, yielding heterogeneous outcomes.

5.2. Classical model

Finally, it is noteworthy that a classical model can be formulated based on the micro models developed in Section 2.2. Section A.2 offers comprehensive insights into this formulation. In this framework, avoidance enhances total consumption, albeit at a twofold cost: a minor tax payment and a disutility increasing and convex in avoidance. However, it is important to emphasize that this model, akin to the one featuring CES utility, lacks an analytical solution. Moreover, it is not feasible to establish an analytical relationship between e and c at present, leading to a system of two nonlinear equations. Consequently, solving the model becomes more time-consuming.

Nevertheless, this model can still be calibrated using a similar approach to previous models, albeit with the challenge of calibrating three parameters (π, ν, and ω) along with a vector of α. This implies that the calibration process becomes notably slower and more intricate, rendering it challenging to execute efficiently with this iterative and grid-based procedure. Our best approximation is provided by $\nu = 0.28$, and $\omega = 0.25$ if we set $\pi = 15\%$, yielding the results depicted in Figure 14.

Figure 14. Taxes evaded as percentage of taxes owed and Average Tax Rates, by pre-tax income groups



A. Distribution of under-reported income. B. Average tax rates by income group (% of income)

Two notable observations arise from these preliminary results. Firstly, our initial estimate for π is considerably high, based on Piketty et al. (2014). This can be attributed to the fact that Piketty et al. developed a model primarily focused on the top of the income distribution. They posit that individuals in this bracket may shelter income by categorizing labor income as capital gains, which typically incur lower tax rates (recently between 15%-20% in the US). Consequently, the assumption implies that the rate of avoidance among lower percentiles would be relatively lower, as they are subject to taxation at rates generally lower than those applied to capital gains. This discrepancy is reflected in the figure, where the modeled distribution of under-reported income does not align well for those below the 80th percentile.

Secondly, and perhaps more intriguingly, this model is capable of reproducing the observed reduction in avoidance at the upper end of the income distribution. However, this comes at the expense of not capturing the regressivity present at the top end of the average tax rate, which is a direct consequence of the model's assumptions. Nonetheless, these conclusions are merely preliminary, and a thorough resolution of these issues is warranted. There may still be a combination of parameters that effectively reconcile both the distribution of under-reported income and regressivity at the top. However, achieving this balance may come at the expense of rendering the parameter π less easily interpretable.

6. Conclusion

We revisited the classic question of the optimal structure of the income tax schedule in an economy calibrated to match the earning distribution in the United States, incorporating private insurance and the possibility of tax avoidance, manifesting as informal consumption within a firm or preceding the payment of taxes. We emphasize two key findings from our analysis.

Firstly, the introduction of tax avoidance mechanisms allows us to account for the observed regressivity at the top of the income distribution. In cases where the possibilities of informal consumption are heterogeneous—increasing in the insurable shock, at least up to a certain point or including an elasticity of substitution greater than 1—wealthier households tend to engage in more avoidance, thereby reducing the amount of personal income taxes they pay. This observation sheds light on the paradoxical situation where, despite income at the top of the distribution being primarily derived from human rather than financial capital, as argued by Piketty and Saez (2003), Piketty (2014), and Smith et al. (2019), high-income individuals contribute minimally in labor income taxes, often due to the utilization of tax avoidance mechanisms.

Secondly, the introduction of tax avoidance mechanisms significantly alters the optimal income tax progressivity, as demonstrated in our numerical exercises. An omniscient planner, aware of the existence of avoidance but powerless to eliminate it, allows for an increase in progressivity, albeit substantially less than initially postulated by Heathcote and Tsujiyama (2021) in a model where informal consumption utility is increasing in skills. Alternatively, when we allow for a mechanism of greater elasticity of substitution, the planner may even prefer a reduction of progressivity. Moreover, a myopic planner, ignorant of avoidance and interpreting all avoidance as a reduction in aggregate production, thereby amplifying the behavioral effects of the change in tax policy, proposes practically no change in the optimal progressivity parameter.

Insights into policy considerations are not straightforward. Firstly, despite the optimal progressivity parameter being lower than initially estimated, there remains a welfare gain in increasing progressivity, leading to higher taxes for high incomes, potentially surpassing marginal rates of 60%, as suggested by a heterogeneous model. Conversely, a model incorporating a mechanism of greater elasticity of substitution recommends a slightly lower progressivity. Secondly, even with the augmented progressivity resulting in higher contributions from the affluent in the first model, significant tax avoidance persists. This dynamic leads to a portion of the increased burden being shouldered by the middle classes. Consequently, alternative policies, such as the removal of tax credits associated with firm expenses, may be advocated to mitigate top regressivity.

We acknowledge several limitations in our work. Firstly, our heavy reliance on parametric assumptions for the distribution of tax avoidance, which increases exponentially over the income distribution, differs from empirical findings. Secondly, our macroeconomic models abstract from mechanisms that often lead to lower optimal progressivity rates, such as risky human capital accumulation.

Addressing these limitations offers avenues for future investigation. Enriching our model environment can be achieved through various means. First, conducting robustness checks by incorporating different parametric assumptions. Second, utilizing more extensive data on the distribution of tax avoidance to eliminate the need for parametric assumptions. Finally, integrating human capital accumulation into the model. While these adjustments may alter the results, the current conclusion suggests that more progressive taxes can be welfare-enhancing, but it depends on the assumptions of our models.

References

Allingham, M. G. and Sandmo, A. (1972). Income tax evasion: A theoretical analysis. *Journal of public economics*, 1(3-4):323–338.

Andreoni, J., Erard, B., and Feinstein, J. (1998). Tax compliance. *Journal of economic literature*, 36(2):818–860.

Atkinson, A. B., Piketty, T., and Saez, E. (2011). Top incomes in the long run of history. *Journal of economic literature*, 49(1):3–71.

Benabou, R. (2000). Unequal societies: Income distribution and the social contract. *American Economic Review*, 91(1):96–129.

Chetty, R. and Saez, E. (2010). Dividend and corporate taxation in an agency model of the firm. *American Economic Journal: Economic Policy*, 2(3):1–31.

Clotfelter, C. T. (1983). Tax-induced distortions and the business-pleasure borderline: the case of travel and entertainment. *The American Economic Review*, 73(5):1053–1065.

Diamond, P. and Saez, E. (2011). The case for a progressive tax: From basic research to policy recommendation. *Journal of Economic Perspectives*, 25(4):165–190.

Diamond, P. A. (1998). Optimal income taxation: an example with a u-shaped pattern of optimal marginal tax rates. *American Economic Review*, pages 83–95.

Farhi, E. and Werning, I. (2013). Insurance and taxation over the life cycle. *Review of Economic Studies*, 80(2):596–635.

Feldstein, M. (1999). Tax avoidance and the deadweight loss of the income tax. *Review of Economics and Statistics*, 81(4):674–680.

Feldstein, M. S. (1969). The effects of taxation on risk taking. *Journal of Political Economy*, 77(5):755–764.

Guyton, J., Langetieg, P., Reck, D., Risch, M., and Zucman, G. (2021). Tax evasion at the top of the income distribution: Theory and evidence. Technical report, National Bureau of Economic Research.

Guyton, J., Langetieg, P., Reck, D., Risch, M., and Zucman, G. (2023). Tax evasion at the top of the income distribution: Theory and evidence. Technical report, National Bureau of Economic Research.

Heathcote, J., Storesletten, K., and Violante, G. L. (2017). Optimal tax progressivity: An analytical framework. *The Quarterly Journal of Economics*, 132(4):1693–1754.

Heathcote, J. and Tsujiyama, H. (2021). Optimal income taxation: Mirrlees meets ramsey. *Journal of Political Economy*, 129(11):3141–3184.

Holter, H. A., Krueger, D., and Stepanchuk, S. (2019). How do tax progressivity and household heterogeneity affect laffer curves? *Quantitative Economics*, 10(4):1317–1356.

Johns, A. and Slemrod, J. (2010). The distribution of income tax noncompliance. *National Tax Journal*, 63(3):397–418.

Kindermann, F. and Krueger, D. (2022). High marginal tax rates on the top 1 percent? lessons from a life-cycle model with idiosyncratic income risk. *American Economic Journal: Macroeconomics*, 14(2):319–366.

Leite, D. (2024). The firm as tax shelter: Micro evidence and aggregate implications of consumption through the firm.

Liu, C. and Yermack, D. (2012). Where are the shareholders' mansions? ceos' home purchases, stock sales, and subsequent company performance. In *Corporate governance: Recent developments and new trends*, pages 3–28. Springer.

Malmendier, U. and Tate, G. (2009). Superstar ceos. *The Quarterly Journal of Economics*, 124(4):1593–1638.

Mayshar, J. (1991). Taxation with costly administration. *The Scandinavian Journal of Economics*, pages 75–88.

Milligan, K. (2022). How progressive is the canadian personal income tax? a buffett curve analysis. *Canadian Public Policy*, 48(2):211–224.

Mirrlees, J. A. (1971). An exploration in the theory of optimum income taxation. *The review of economic studies*, 38(2):175–208.

Pareto, V. (1896). Cours d'economie politique, genève, droz, novena edición por gh bousquet et g. *Bucino (1964) misma editorial*.

Persson, M. (1983). The distribution of abilities and the progressive income tax. *Journal of Public Economics*, 22(1):73–88.

Piketty, T. (2003). Income inequality in france, 1901–1998. *Journal of political economy*, 111(5):1004–1042.

Piketty, T. (2014). *Capital in the twenty-first century*. Harvard University Press.

Piketty, T. (2020). *Capital and ideology*. Harvard University Press.

Piketty, T. and Saez, E. (2003). Income inequality in the united states, 1913–1998. *The Quarterly journal of economics*, 118(1):1–41.

Piketty, T. and Saez, E. (2013). Optimal labor income taxation. In *Handbook of public economics*, volume 5, pages 391–474. Elsevier.

Piketty, T., Saez, E., and Stantcheva, S. (2014). Optimal taxation of top labor incomes: A tale of three elasticities. *American economic journal: economic policy*, 6(1):230–271.

Ramsey, F. P. (1927). A contribution to the theory of taxation. *The economic journal*, 37(145):47–61.

Saez, E. (2001). Using elasticities to derive optimal income tax rates. *The review of economic studies*, 68(1):205–229.

Saez, E. and Zucman, G. (2019). *The triumph of injustice: How the rich dodge taxes and how to make them pay*. WW Norton & Company.

Smith, A. (1776). *An inquiry into the nature and causes of the wealth of nations*. London: printed for W. Strahan; and T. Cadell, 1776.

Smith, M., Yagan, D., Zidar, O., and Zwick, E. (2019). Capitalists in the twenty-first century. *The Quarterly Journal of Economics*, 134(4):1675–1745.

Stantcheva, S. (2017). Optimal taxation and human capital policies over the life cycle. *Journal of Political Economy*, 125(6):1931–1990.

Trabandt, M. and Uhlig, H. (2011). The laffer curve revisited. *Journal of Monetary Economics*, 58(4):305–327.

Yermack, D. (2006). Flights of fancy: Corporate jets, ceo perquisites, and inferior shareholder returns. *Journal of financial economics*, 80(1):211–242.

Zucman, G. (2019). Global wealth inequality. *Annual Review of Economics*, 11:109–138.

Appendix A. Alternative Models

A.1. CES utility function

The objective of this model is to eschew the utilization of a heterogeneous function $\omega(\alpha)$, thereby ensuring that evasion as a percentage of income increases due to a substitution elasticity different from 1, thereby providing a more parsimonious specification. We draw upon models commonly found in the International Finance literature, such as an aggregated consumption that conforms to a CES utility function.

Suppose the household utility of consumption has a CES form for the aggregator between normal and firm consumption. This implies that the household solves the following optimization problem:

$$\max_{\{c(\alpha,\varepsilon),h(\alpha,\varepsilon),e(\alpha,\varepsilon)\}} \int \left[\log C(c,e) - \frac{h(\alpha,\varepsilon)^{1+\sigma}}{1+\sigma} \right] dF_{\varepsilon}(\varepsilon) \tag{A1}$$

$$s.t.\ C(c,e) = \left[ac(\alpha,\varepsilon)^{1-\frac{1}{\xi}} + (1-a)e(\alpha,\varepsilon)^{1-\frac{1}{\xi}} \right]^{\frac{1}{1-\frac{1}{\xi}}} \tag{A2}$$

$$\int \mathbf{B}(\alpha,\varepsilon)\mathbf{Q}(\varepsilon)d\varepsilon = 0 \tag{A3}$$

$$y(\alpha,\varepsilon) = c(\alpha,\varepsilon) + T(y(\alpha,\varepsilon) - e(\alpha,\varepsilon)) + e(\alpha,\varepsilon) \tag{A4}$$

$$y(\alpha,\varepsilon) = \exp(\alpha+\varepsilon)h(\alpha,\varepsilon) + B(\alpha,\varepsilon). \tag{A5}$$

Where $T(y) = y - \lambda y^{1-\tau}$. We solve the first-order conditions and then integrate for the distribution of ε, yielding the following results:

$$e = \left[\frac{1-a}{a} \left(\frac{1}{1-\tau}\right) \left(\frac{1}{\lambda}\right)^{\frac{1}{1-\tau}} \right]^{\xi} c^{\frac{1+\tau(\xi-1)}{1-\tau}} \tag{A6}$$

(A7)

$$\left(\frac{c}{\lambda}\right)^{\frac{1}{1-\tau}} + e = \left[c^{-\frac{1+\tau(\xi-1)}{\xi(1-\tau)}} \frac{a}{ac^{1-1/\xi} + (1-a)e^{1-1/\xi}} \left(\frac{1}{1-\tau} \left(\frac{1}{\lambda}\right)^{\frac{1}{1-\tau}} \right)^{-1} \right]^{\frac{1}{\sigma}} \exp(\alpha)^{\frac{1+\sigma}{\sigma}} \mathbf{E}\left[\exp(\varepsilon)^{\frac{1+\sigma}{\sigma}} \right].$$

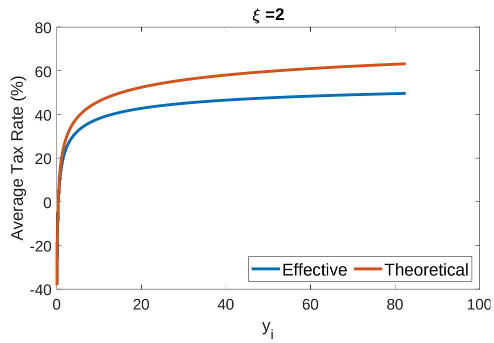
In this formulation, there exists an analytical expression for e as a function of c, but there is no analytical form of c as a function of parameters. Consequently, the model must be solved using numerical methods.

We calibrate the model in a similar manner to the previous heterogeneous model. However, the inclusion of avoidance that is changing in income implies that the original distribution and variances will not be preserved. This calibration includes:

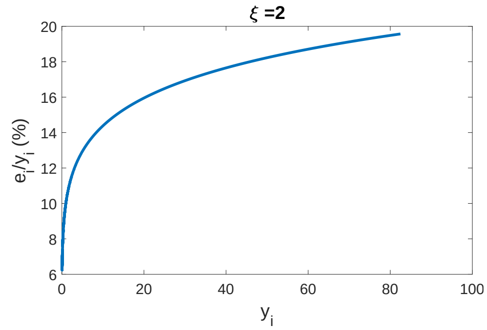
- Objective: choose a and ξ that minimize the distance for most percentiles of the evasion distribution. This will affect the income distribution, so we also need to determine a distribution of α that is consistent with the original model.
- We assume a calibration similar to that of the heterogeneous model:
  - We assume an initial guess for a and ξ. With this, assuming $y - e = y^0$, we can extract a unique α for each income. For this, since we do not have an analytical solution, we need to assume c and e for the first iteration, which will be equal to those that are consistent with the initial guess.
  - We assume that $\alpha \sim \text{EMG}(\mu_\alpha, \sigma_\alpha, \lambda_\alpha)$ and calculate these parameters by maximum likelihood estimation (MLE) using numerical methods.
  - We solve the model and iterate with new grids until we converge on values for both the distribution of α and a and ξ.

Figure A1 illustrates how the calibration changes when we set $a = 0.8$ and allow ξ to vary both above and below 1. As depicted, an elasticity of substitution between goods greater than 1 is required to yield an avoidance distribution that increases with income. Figure A2 demonstrates how the calibration changes when we set $\xi = 5$ and vary a. A lower a implies greater avoidance, as the share of avoidance utility in the CES aggregator increases.

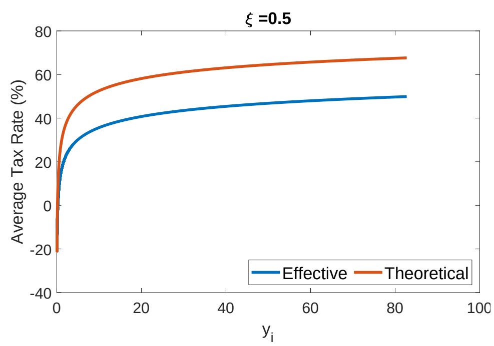
Figure A1. Average Tax rates and Taxes evaded as percentage of taxes owed, by pre-tax income groups



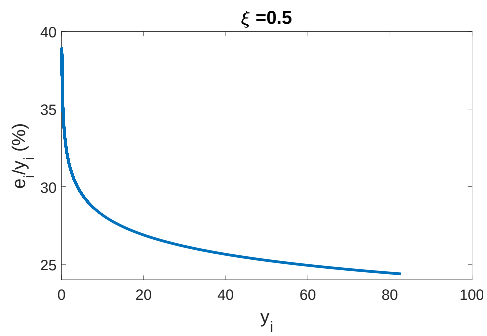
A. All taxes



B. Only personal taxes



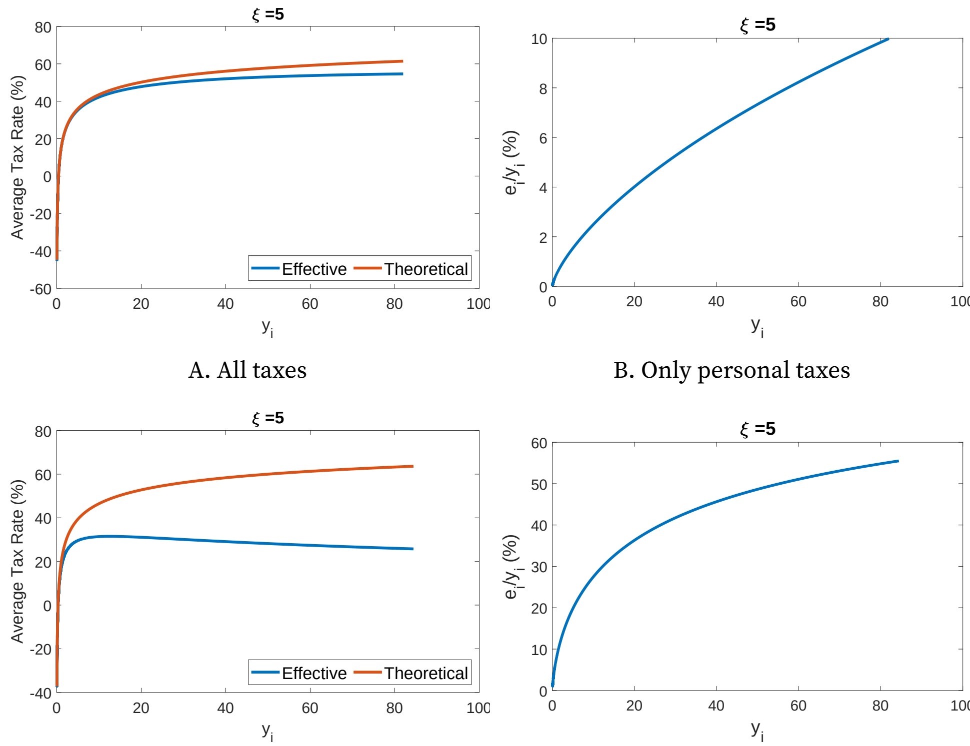
C. Personal taxes without payroll taxes



D. Only Income taxes

Note: a = 0.8

Figure A2. Taxes evaded as percentage of taxes owed, by pre-tax income groups



A. All taxes

B. Only personal taxes

C. Personal taxes without payroll taxes

D. Only Income taxes

Note: a = 0.8 in the first row and a = 0.7 in the second.

A.2. Classical model

The objective of this model is to align it more closely with current literature, which considers evasion not as separate consumption but as sheltered income. We draw upon the fundamentals of the model proposed by Piketty et al. (2014), but assume functions for the disutility of both labor and avoidance. With this assumption, the household solves the following optimization problem:

$$\max_{\{c(\alpha,\varepsilon),h(\alpha,\varepsilon),e(\alpha,\varepsilon)\}} \int \left[c(\alpha,\varepsilon) - \frac{h(\alpha,\varepsilon)^{1+\sigma}}{1+\sigma} - \nu \frac{e(\alpha,\varepsilon)^{1+\omega}}{1+\omega} \right] dF_{\varepsilon}(\varepsilon) \tag{A8}$$

$$s.t. \int \mathbf{B}(\alpha,\varepsilon)\mathbf{Q}(\varepsilon)d\varepsilon = 0 \tag{A9}$$

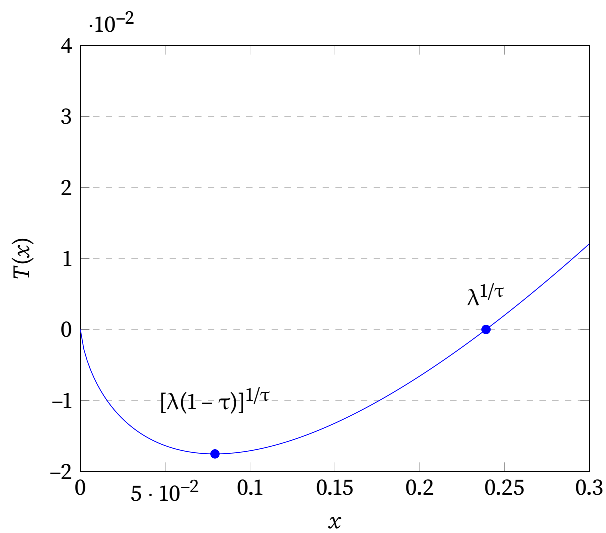
$$c(\alpha,\varepsilon) = y(\alpha,\varepsilon) - T(y(\alpha,\varepsilon) - e(\alpha,\varepsilon)) - \pi e(\alpha,\varepsilon) \tag{A10}$$

$$y(\alpha,\varepsilon) = \exp(\alpha+\varepsilon)h(\alpha,\varepsilon) + B(\alpha,\varepsilon) \tag{A11}$$

$$e(\alpha,\varepsilon) \geq 0 \tag{A12}$$

The last restriction may be necessary due to the following situation. Figure A3 illustrates $T(x) = x - \lambda x^{1-\tau}$, where x represents pretax income. It is shown that for small values of x, the function has a decreasing segment. Now, let $x = y - e$. If $y - e < [\lambda(1-\tau)]^{1/\tau}$, then $e < 0$ (which is equivalent to declaring more income than the real one) increases the transfers. Its effect on consumption can be negative, since $c = \lambda(y-e)^{1-\tau} + (1-\pi)e$ (the first term increases transfers and the second one eliminates them), but there may be a positive effect on utility since $-\frac{e^{1+\omega}}{1+\omega} > 0$. This dilemma is addressed by supposing the last restriction, for simplicity.

Figure A3. Tax function.



Given that, the resolution of the model involves solving the first-order conditions and then integrating the distribution of ε. The results are then as follows:

$$\text{(A13)} \qquad ve^{\omega} = 1 - \pi - (1-\tau)\lambda\left(\frac{c-(1-\pi)e}{\lambda}\right)^{\frac{-\tau}{1-\tau}}$$

$$\text{(A14)} \qquad \left(\frac{c-(1-\pi)e}{\lambda}\right)^{\frac{1}{1-\tau}} + e = \left[(1-\tau)\lambda\left(\frac{c-(1-\pi)e}{\lambda}\right)^{\frac{-\tau}{1-\tau}}\right]^{\frac{1}{\sigma}} \exp(\alpha)^{\frac{1+\sigma}{\sigma}}\,\mathbf{E}\left[\exp(\varepsilon)^{\frac{1+\sigma}{\sigma}}\right].$$

Here, the model lacks an analytical solution, both for e and for c. Consequently, we must solve both equations numerically, resulting in a nonlinear system of two equations.

Appendix B. Other minor ideas

B.1. Elasticities

Piketty et al. (2014) analyze their model based on two fundamental elasticities (three if we include bargaining). Can we adopt a similar approach in this case?

- We already know that $\xi_{y,1-\tau} = \frac{\partial y(\alpha)}{\partial(1-\tau)}\frac{1-\tau}{y(\alpha)} = \frac{1}{1+\sigma}\left(\frac{1-\tau}{1-\tau+\omega(\alpha)}\right)$.
- Similarly, we can determine $\xi_{e,1-\tau} = \frac{\partial e(\alpha)}{\partial(1-\tau)}\frac{1-\tau}{e(\alpha)} = \frac{-\sigma}{1+\sigma}\left(\frac{1-\tau}{1-\tau+\omega(\alpha)}\right)$.

 Let $y_T \equiv y - e$. Then,

$$\xi_{y_T,1-\tau} = \frac{\partial y_T(\alpha)}{\partial(1-\tau)}\frac{1-\tau}{y_T(\alpha)} = \Gamma^{-1}\left(\xi_{y,1-\tau}\frac{y(\alpha)}{y(\alpha)-e(\alpha)} - \xi_{e,1-\tau}\frac{e(\alpha)}{y(\alpha)-e(\alpha)}\right) \tag{A15}$$

with $\Gamma = \left\{\mathbf{E}\left[\exp(\varepsilon)^{\frac{1+\sigma}{\sigma}}\right]\right\}^{\frac{\sigma}{1+\sigma}}\exp(\alpha)$.

B.2. Macrodynamics

In their 2017 paper, Heathcote et al. proposed their model in the following form: an individual lives with probability δ, chooses a skill level s_i at $t = 0$, and enters the labor market, where they face productivity z_i, and solves:

$$\max_{c_{it},h_{it}} U_i = -v_i(s_i) + (1-\beta\delta)\mathrm{E}_0\sum_{t=0}^{\infty}(\beta\delta)^t\left[\log c_{it} - \frac{\exp[(1+\sigma)\varphi_i]}{1+\sigma}h_{it}^{1+\sigma} + \chi\log G\right] \tag{A16}$$

$$\text{s.t.} \int Q(\varepsilon)B(\varepsilon)d\varepsilon = 0 \tag{A17}$$

$$c_{it} = \lambda y_{it}^{1-\tau} + B(\varepsilon_{it}) \tag{A18}$$

$$y_{it} = p(s_i)\exp(\alpha_{it} + \varepsilon_{it}) \tag{A19}$$

Given $\log(z_{it} = \alpha_{it} + \varepsilon_{it})$, where α_{it} is an AR(1) with i.i.d. innovation distributed normally and ε an i.i.d. shock distributed normally. φ_i is also distributed normally. The disutility of investment in skills is given by $v(s_i) = \frac{\kappa_i^{-1/\psi}}{1+1/\psi}(s_i)^{1+1/\psi}$, where $\kappa \sim Exp(\eta)$.

However, when solving this model, they do it in the following way:

$$\max_{c(\varepsilon),h(\varepsilon)} \int_E \left\{\log c(\varepsilon) - \frac{\exp[(1+\sigma)\varphi]}{1+\sigma}h(\varepsilon)^{1+\sigma} + \chi\log G\right\} dF_\varepsilon \tag{A20}$$

$$\text{s.t.} \int_E c(\varepsilon)dF_\varepsilon = \lambda\int_E \exp[(1-\tau)(p(s)+\alpha+\varepsilon)]h(\varepsilon)^{1-\tau}dF_\varepsilon \tag{A21}$$

Where s_i, φ_i are determined by the rest of the variables and parameters of the model or by assumptions of distributions. The authors define it as an *island-specific planner problem*.

Subsequently, in their 2021 model, Heathcote and Tsujiyama simplified the above to what we discussed in Section 2.1, which is also equivalent to:

$$\max_{c(\varepsilon),h(\varepsilon)} \int_E \left\{ \log c(\varepsilon) - \frac{h(\varepsilon)^{1+\sigma}}{1+\sigma} \right\} dF_\varepsilon \tag{A22}$$

$$\text{s.t. } y(\alpha) = \int \exp(\alpha + \varepsilon) h(\alpha, \varepsilon) dF_\varepsilon(\varepsilon) \tag{A23}$$

$$\int c(\alpha, \varepsilon) dF_\varepsilon(\varepsilon) = y(\alpha) - T(y(\alpha)) \tag{A24}$$

where $T(y)$ is the HSV function as discussed earlier. Apart from being simplified, the model has minor changes, for example, now the insurance claims B purchased are constituent of pretax income, and before were after-tax income. Given all of these relationships, this means that our heterogeneous model should be similar to estimating:

$$\max_{c_{it},h_{it},e_{it}} U_i = \mathrm{E}_0 \sum_{t=0}^{\infty} \beta^t \left[\log c_{it} - \frac{h_{it}^{1+\sigma}}{1+\sigma} + \omega_i \log e_{it} \right] \tag{A25}$$

subject to the corresponding constraints.