Optimal Housing Taxation with Land Scarcity and Maintenance: A Mirrleesian Perspective*

Spencer Bastani† Sören Blomquist‡ Firouz Gahvari§
Luca Micheletto¶ Khayyam Tayibov‖

March 11, 2025

Abstract

We study optimal housing taxation in a Mirrleesian framework where individuals differ in both labor productivity and land ownership. Housing services are produced by combining scarce land with structures that require maintenance, which can be performed either in-house or through market purchases. We first characterize optimal allocations under information and resource constraints. We then restrict the government to the use of proportional housing taxes. Numerical simulations show that uniform taxation of land and structures is desirable only when political constraints prevent the imposition of very high land taxes. Otherwise, the optimal policy is to tax land at a much higher rate than structures, while still imposing a positive tax on structures to mitigate distortions from income taxation. A positive marginal tax on labor income incentivizes in-house over market-purchased maintenance. To prevent an inefficiently large reliance on in-house maintenance, optimal policy should generally subsidize market-purchased maintenance services.

Keywords: Optimal taxation, Housing capital, Land, Labor supply, Maintenance

JEL codes: D1, D3, HC, R2

*We are grateful for comments from David Agrawal, Aart Gerritsen, Floris Zoutman, Katarina Nordblom, Åsa Hansson, Håkan Selin and seminar participants at Linnaeus University, the UCFS workshop in Krusenberg, and the IFN workshop in Vaxholm.

†Corresponding author. IFAU and Department of Economics, Uppsala University; Research Institute of Industrial Economics (IFN); CESIfo, Germany. E-mail: spencer.bastani@ifau.uu.se.

‡Department of Economics, Uppsala University. E-mail: soren.blomquist@nek.uu.se

§Department of Economics, University of Illinois at Urbana-Champaign. Email: fgahvari@illinois.edu

¶Department of Law, University of Milan, and Dondena Centre for Research on Social Dynamics and Public Policy, Bocconi University; UCFS; CESifo. E-mail: luca.micheletto@unimi.it.

‖Department of Economics and Statistics, School of Business and Economics, Linnaeus University, Sweden. E-mail: khayyam.tayibov@lnu.se.

1 Introduction

Housing is a critical determinant of individual well-being. It serves not only as a primary consumption good for most households, but also as a major source of personal wealth, accounting for a significant share of total capital in the economy.[1] In addition, housing has unique characteristics that distinguish it from other goods. It consists of land and structures, two components with notable economic differences. The value of land is largely determined by location, and land tends to be in fixed supply. Structures, on the other hand, are durable goods that can be developed at constant marginal cost but require ongoing maintenance to prevent deterioration.

The economic importance and complexity of housing has generated a rich theoretical and empirical literature on housing taxation. For example, many studies examine the effects of preferential tax treatments for owner-occupied housing, such as the deductibility of mortgage interest and the exemption of imputed rent. These provisions influence tenure choices and may lead to overinvestment in housing.[2] Other strands of research focus on corrective (Pigouvian) taxes to address the positive externalities (e.g., social stability, human capital development) and negative externalities (e.g., sprawl, environmental degradation, status effects) associated with housing.[3] Another line of inquiry, popularized by Henry George, advocates taxing scarce land as an efficient source of revenue.[4]

Despite extensive research on housing taxation, the literature often neglects the crucial distributional issues that are essential for determining the appropriate role of housing taxes within the broader tax system. In the absence of distributional concerns, a government could hypothetically rely on a perfectly efficient lump-sum tax, making all other taxes unnecessary. This paper aims to fill this gap by investigating optimal housing taxation, focusing on taxes on land and structures, and the tax treatment of purchased maintenance services, building on the seminal optimal income tax framework of Mirrlees (1971). In this way, we study optimal housing taxation in a setting that formalizes the equity/efficiency trade-off that policymakers face when designing tax policies.

We consider an economy with heterogeneous agents who derive utility from ordinary goods and housing, and who differ in both labor productivity and land endowments. Several features of the housing market that are often overlooked in the existing literature are incorporated: (a) housing services are produced by combining land and structures, (b) land is scarce and has an endogenous price that is affected by tax policy, (c) structures require maintenance that can be provided by household effort or market purchases.

[1] As reported by Denk and Schich (2020), the median share of housing wealth in total wealth across OECD countries was 61% in 2017, ranging from 38% in Germany to 83% in Norway. See also Causa et al. (2019).

[2] See, for example, Gahvari (1985), which examines the long-run effects of eliminating these implicit subsidies on capital accumulation and welfare. Related issues include whether housing capital and business capital should face the same tax rate (Gahvari 1984a, Gahvari 1984b, Pines et al. 1985, Gervais 2002, Eerola and Määttänen 2013) and how housing capital gains should be taxed (Englund 2003).

[3] See, e.g., Rossi-Hansberg et al. (1999) and Schünemann and Trimborn (2022).

[4] See, e.g., Aura and Davidoff (2012) and Albouy (2016) for a discussion.

Our central research questions are: How should housing (structures and land) be taxed in the presence of an optimal nonlinear labor income tax? Should structures and land be taxed differently, given their different characteristics? How does the need for maintenance affect the optimal design of housing taxes, and should maintenance services be taxed or subsidized? What are the implications of an endogenous land price? Finally, how does optimal policy change if the government must rely on simpler proportional instruments rather than fully nonlinear commodity taxes? To answer these questions, we develop a theoretical framework based on optimal tax analysis under different information constraints, and then complement this framework with numerical simulations using Swedish register data.

As a conceptual starting point, it is useful to recall Atkinson and Stiglitz (1976)'s well-known result on optimal commodity taxation. In an economy with fixed producer prices where labor productivity is the only source of heterogeneity—and in the absence of externalities and internalities—the Atkinson-Stiglitz theorem implies that uniform taxation of all goods is optimal if labor supply decisions are separable from consumption decisions.[5] Our setting, however, departs from these assumptions in three important ways. First, individuals differ not only in labor productivity but also in land ownership. Second, the decision to outsource maintenance or to do it oneself creates interactions between labor supply and housing maintenance. Third, the fixed supply of land implies an endogenous land price, allowing fiscal policy to affect redistribution indirectly through changes in commodity prices.

We address two key issues that have been underexplored in the literature. First, households maintain their homes either by hiring professionals or by doing the work themselves, and high market productivity does not necessarily imply greater maintenance efficiency. Including this aspect in our model challenges the uniform taxation of goods implied by the Atkinson-Stiglitz theorem. Second, our framework separates structures from land, assuming that the latter is in fixed supply with an endogenously determined price. This setup suggests that differential commodity taxation could help mitigate distortions from income taxation by influencing commodity prices, whereas land that can be developed at constant marginal cost would not warrant differential taxation under standard separability assumptions.

The theoretical part of the paper begins by examining the optimal taxation of housing in an environment where government policy is constrained only by resource and information limitations. In this scenario, all variables except individual productivity and hours of work are observable and can therefore be taxed nonlinearly.[6] We then move to a second scenario that combines progressive nonlinear taxation of labor income with proportional taxes on other goods/services. This second scenario may arise from various political or economic constraints that preclude the adoption of a fully nonlinear tax system.

Focusing first on the fully nonlinear setting, our theoretical results indicate that the govern-

[5]See Bastani and Koehne (2024) for a detailed discussion of the Atkinson and Stiglitz (1976) result.

[6]Unlike most goods, which are consumed anonymously and taxed linearly, land and structures can be taxed nonlinearly based on individual ownership.

ment confiscates all land and redistributes it among agents as part of the optimal allocation. In practice, when land endowments are inherited rather than purchased, this implies a 100% tax on land rents. Moreover, the total labor supply of high-skilled agents is not distorted, while low-skilled agents face a downward distortion.[7] From a social perspective, it is optimal to encourage specialization in household activities: high-skilled agents outsource all maintenance work, while low-skilled agents do it themselves. This contrasts with many existing tax systems, in which high-income households bear heavy tax burdens on both income and professional maintenance services, discouraging the outsourcing of maintenance.

A key result is that, whereas it is optimal to avoid distorting the consumption of structures for high-skilled agents, low-skilled agents should be incentivized to consume more structures. The reason is that, since low-skilled agents do all the maintenance themselves, increasing their structure consumption increases their own maintenance effort. This effort increase partially offsets the downward distortion of their labor supply required for incentive compatibility (i.e., preventing high-skilled individuals from imitating low-skilled ones).

We then turn to the second scenario, where taxes on land, structures, and maintenance are all proportional. Using the framework of Edwards et al. (1994), we derive optimal linear commodity tax formulas that balance the social gains from discouraging mimicking with the efficiency costs of distorting consumption. As land endowments increase with skill, the proportional tax rate on land is pushed to its politically feasible maximum. However, because these commodity taxes are linear and land prices adjust endogenously, land rents are not fully taxed away, unlike in the fully nonlinear case. Moreover, the endogeneity of land prices appears in the optimal tax formulas for structures and maintenance, which capture the potential for indirect redistribution via changes in the price of land. These effects depend on the distribution of land across skill types and on the relative demand for land by low-skilled agents versus high-skilled agents seeking to emulate them.

In the quantitative part of the paper, we illustrate these theoretical insights with numerical simulations using Swedish register data. We adopt functional forms common in the literature and allow for a broader discrete set of agent types. In the fully nonlinear taxation scenario, the simulations confirm our theoretical findings, but provide additional insights into how distortions vary across the income distribution. In the second scenario, where the government relies on proportional taxes on structures, land, and maintenance, a key observation is that the tax rate on land always reaches its regulatory limit, but this is not always true for the tax on structures. In all of our simulations, we find that maintenance services purchased in the market should be subsidized.

There is little prior work on optimal housing taxation within a Mirrleesian framework. Among the few contributions, Cremer and Gahvari (1998) distinguish housing consumption from other goods by assuming its observability and allowing for nonlinear taxation. However,

[7] This is consistent with standard optimal income tax theory, which imposes downward distortions at the bottom to facilitate redistribution and sets the marginal tax rate to zero at the top to maximize revenue.

they do not separate structures from land or incorporate maintenance and endogenous land pricing. Similarly, Koehne (2018) emphasizes the durability of housing, but does not address the broader issues considered here.

Our focus on land links our study to Bonnet et al. (2021) and Schwerhoff et al. (2022). Bonnet et al. (2021) document a rise in wealth-to-income ratios, driven primarily by housing wealth and rising land values. Using a Ramsey framework à la Judd (1985), they argue that, despite political and practical challenges, a land tax is the most efficient and equitable form of wealth taxation. Schwerhoff et al. (2022) analyze optimal land value taxation under linear instruments and no land trade, and show that full land value taxation depends on the correlation between household land holdings and social welfare weights. Our work is also related to recent empirical studies on property taxation. Löffler and Siegloch (2021) decompose the welfare effects of marginal property tax reforms in Germany, while Määttänen and Terviö (2021) examine the welfare costs of housing transaction taxes using Finnish data.

Our paper also relates to the broader literature on optimal goods taxation, in particular research that considers interactions between consumption, labor supply, and household production. This includes studies of child care (Bastani et al., 2020; Ho and Pavoni, 2020; Moschini, 2023; Casarico et al., 2023; Koll et al., 2023), household production (Anderberg and Balestrino, 2000; Kleven et al., 2000; Gayle and Shephard, 2019), and work-related expenditures (Koehne and Sachs, 2022).[8] The fact that we study optimal goods taxation with endogenous prices means that we relate to the literature that has considered how the general equilibrium effects of tax policy interact with redistributive goals (Stiglitz 1982, Rothschild and Scheuer 2013, and Sachs et al. 2020). Most of this literature has focused on income taxation and perfect competition. Our work is perhaps more related to the literature on the use of commodity taxes in imperfectly competitive product markets (e.g, Stern 1987, Myles 1995, Cremer and Thisse 1994, Auerbach and Hines 2002), and perhaps especially those papers that have studied this problem in a Mirrleesian setting (Kushnir and Zubrickas 2019 and Jaravel and Olivi 2022).[9]

The rest of the paper is organized as follows. Section 2 describes the main features of our model. Section 3 explores the optimal taxation of housing under a fully nonlinear tax system, while section 4 focuses on the mixed tax system where the policymaker relies on a nonlinear labor income tax and proportional commodity taxes. Section 5 outlines our simulation model based on Swedish register data, and section 6 discusses the quantitative results. Section 7 discusses some of the limitations of our analysis. Finally, section 8 provides concluding remarks.

[8] It is also related to work on the role of endowment differences for optimal goods taxation (Cremer et al. 2001, Bastani et al. 2015).

[9] Our work also relates to the the emerging literature that studies the interaction between taxation, regulation, mark-ups, and distributional issues (Gürer 2022, Boar and Midrigan 2024, and Eeckhout et al. 2025).

2 The model

We consider an economic model with two types of agents: a low-skilled type (1) and a high-skilled type (2). The total population is normalized to one, and ζ^j denotes the fraction of agents of type j, where $j = 1, 2$. Agents differ in labor market productivity w^j, where $w^1 < w^2$, and derive utility from consumption of a composite good c (treated as a numéraire), housing services, and leisure. Housing services are produced by combining structures (s) and land (l). Each unit of structures requires γs units of maintenance services, which can be purchased on the market in the form of professional maintenance services, denoted by z, or performed by the agent himself/herself, denoted by h_m. In particular, we assume that

$$\gamma s = w_z z + \omega h_m, \tag{1}$$

where $\gamma > 0$, w_z is the productivity of workers hired to perform maintenance, and ω is the productivity of an agent performing maintenance at home.

For most owner-occupiers, maintenance costs are substantial, and these are usually time costs, either because the owner-occupier spends his or her own time on maintenance or because hiring professional maintenance workers is a time-consuming activity. An owner-occupier is essentially a business owner, where managerial effort or own labor is combined with capital investment to produce housing services. In reality, maintenance requires both physical goods and labor. For simplicity, we assume that only labor is required, and that the amount of labor required is strictly proportional to the amount of structures.[10] Note that a broad interpretation of maintenance is possible. Beyond direct repairs, larger houses require more household services, such as cleaning and general upkeep. Thus, our framework need not restrict maintenance to repairs, but can include a broader range of household services.

Agents derive disutility from labor supplied in the market, denoted by h, and from household maintenance h_m (since these activities reduce the amount of leisure available). This disutility is assumed to be weakly separable from consumption goods (including housing consumption). Due to the maintenance activity, there will still be interactions between an individual's time allocation and his/her consumption of goods and services, as we will explore below. The preferences of the agents are represented by the following utility function:

$$U = v(c) + g(s, l) + f(h + h_m), \tag{2}$$

where $v' > 0$, $v'' < 0$, $\partial g/\partial s > 0$, $\partial^2 g/\partial s^2 < 0$, $\partial g/\partial l > 0$, $\partial^2 g/\partial l^2 < 0$, and $f' < 0$, $f'' < 0$. In the quantitative part of the paper, we focus on the numerically convenient special case where the argument of f is $1 - h - h_m$ and where $h + h_m \leqslant 1$, where unity represents the endowment of time. Note that h and h_m are perfect substitutes in the production of disutility.

[10] The latter aspect is related to the idea that consumption requires time, which has been explored in an optimal tax context by Gahvari (2007).

This assumption is made for simplicity and is not critical to the analysis.

We consider a perfectly competitive labor market where wage equals productivity, so that w^1 and w^2 are the market wage rates of the two types. We assume that all agents have the same home-maintenance productivity ω (an assumption that will be relaxed in section 5). To simplify the analysis and make it more transparent, we assume that structures are competitively supplied at a constant unit price q_s. In contrast, there is a fixed total endowment of land, denoted by $\bar{l}$. The price of land q_l is endogenously determined by the condition that demand equals the fixed supply $\bar{l}$.[11]

Each agent of type j owns a fraction π^j of the total land $\bar{l}$. This is without loss of generality when we consider fully nonlinear taxation in section 3, because in that case the government can confiscate all land and distribute it as part of the optimal allocation. This is equivalent to a 100% land rent tax, although we recognize that in a broader dynamic perspective, land ownership does not necessarily imply pure rents, since the land was likely purchased at some point. Land ownership is important in section 4, where the government is limited to proportional land taxes, and in the simulation section 5, where we consider empirical patterns of land ownership.

Our analysis focuses on optimal tax design in the presence of asymmetric information (allowing for different assumptions about what is observable and thus taxable by the government). The optimal income tax schedule and the labor choices of individuals follow the discrete-type adaptation of the Mirrlees (1971) model developed by Stiglitz (1982), Stern (1982), and Guesnerie and Seade (1982). This setup provides an intuitive illustration of the key trade-offs that determine the optimal structure of the tax system.[12]

3 Fully nonlinear taxation

We begin by characterizing the properties of an optimal allocation under the assumption that all variables except individual productivity, market hours, and hours spent performing maintenance services at home are observable at the individual level and thus can be taxed nonlinearly. These observability assumptions imply that the tax function can potentially depend on four variables: labor income, amount of structures, land use, and resources spent on maintenance services purchased in the market. While practical concerns may limit the feasibility of a fully nonlinear tax system, its analysis provides a useful benchmark because it does not impose arbitrary constraints on the available tax instruments (see, e.g., Mirrlees 1976).

[11] An earlier version of this study (Bastani et al. 2024) included an additional rural region where land could be developed at constant marginal cost, along with exogenous zoning constraints that limit the number of structures that can be built on a given plot of land.

[12] A potential drawback is that the resulting optimal income tax formulas cannot be written in terms of empirically estimable "sufficient statistics". In this paper, however, we consider large deviations from existing tax systems, such as a comprehensive housing tax reform, that render standard sufficient-statistics approaches inapplicable. See Kleven (2021) for a discussion of when sufficient-statistics formulas are more or less useful.

Government's problem Let Y^j denote the labor income earned by agents of type j. Let α^j denote the welfare weight given to agents of type j in the government's social welfare function and assume that $\omega^1 = \omega^2 = w_z = w^1$. The government chooses allocations $\{Y^j, c^j, s^j, l^j, h_m^j, z^j\}_{j=1,2}$ to maximize:

$$\sum_{j=1}^{2} \alpha^j \zeta^j \Big[v(c^j) + g(s^j, l^j) + f\big(\tfrac{Y^j}{w^j} + h_m^j\big) \Big], \tag{3}$$

subject to the following constraints:

Incentive Compatibility Constraint (IC). High-skilled agents must be prevented from mimicking low-skilled agents:[13]

$$v(c^2) + g(s^2, l^2) + f\Big(\frac{Y^2}{w^2} + h_m^2\Big) \geqslant v(c^1) + g(s^1, l^1) + f\Big(\frac{Y^1}{w^2} + h_m^1\Big). \tag{4}$$

Resource Constraint. The total consumption of goods and services must not exceed available resources:

$$\sum_{j=1}^{2} \zeta^j \Big[Y^j - c^j - q_s s^j - w^1 z^j \Big] \geqslant 0. \tag{5}$$

Land Constraint. The total land available is fixed at $\bar{l}$:

$$\sum_{j=1}^{2} \zeta^j l^j \leqslant \bar{l}. \tag{6}$$

Maintenance Requirement. Structures must receive sufficient maintenance:

$$(h_m^j + z^j) w^1 \geqslant \gamma s^j. \quad j = 1, 2. \tag{7}$$

We let $\lambda^{2,1}$ be the Lagrange multiplier associated with the IC constraint, μ be the multiplier associated with the resource constraint, η be the multiplier associated with the land constraint, ρ^j be the multiplier associated with the maintenance requirement constraint. We assume that all constraints are binding. The incentive constraint captures the key challenge in designing the tax policy: to be able to redistribute to low-skilled individuals while preventing high-skilled agents from being tempted to behave as "mimickers" in order to qualify for the more generous tax treatment intended only for the latter.

From the first-order conditions of the government's problem, we derive the following result.

[13] We assume that the only possible binding incentive compatibility (self-selection) constraint is the one that requires high-skilled agents not to be tempted to mimic low-skilled agents. Put differently, we assume that the welfare weights α^j are such that the government seeks to redistribute from the high-skilled to the low-skilled.

Proposition 1 (Labor Supply Distortions) *Under a fully non-linear tax system:*

1. *Total labor supply* $h^j + h^j_m$*, remains undistorted for high-skilled agents* ($j = 2$) *but is downward distorted for low-skilled agents* ($j = 1$)*:*

$$1 + \frac{f'\left(\frac{Y^2}{w^2} + h^2_m\right)}{w^2 \frac{\partial v}{\partial c^2}} = 0, \tag{8}$$

$$1 + \frac{f'\left(\frac{Y^1}{w^1} + h^1_m\right)}{w^1 \frac{\partial v}{\partial c^1}} = \frac{\lambda^{2,1}}{\mu \zeta^1}\left(\frac{f'\left(\frac{Y^1}{w^2} + h^1_m\right)}{w^2} - \frac{f'\left(\frac{Y^1}{w^1} + h^1_m\right)}{w^1}\right) > 0. \tag{9}$$

2. *High-skilled agents purchase all maintenance in the market* ($z^2 = \gamma s^2/w^1$ *and* $h^2_m = 0$)*.*

3. *Low-skilled agents do all maintenance at home* ($h^1_m = \gamma s^1/w^1$ *and* $z^1 = 0$)*.*

Proof. See Appendix A.2. ■

The first result follows from standard considerations of mimicking deterrence. Since the government wants to redistribute from the high-skilled to the low-skilled, it must distort the latter's labor supply downward in order to discourage high-skilled agents from mimicking them.[14] On the other hand, since low-skilled agents have no incentive to mimic high-skilled agents, there is no reason to distort the total labor supply of the latter. The second result follows from the fact that high-skilled agents are more productive in the market than at home ($w^2 > \omega^2 = w^1$). Given the assumed functional form of labor disutility, it is first-best efficient for high-skilled individuals to fully specialize in market work and purchase all necessary maintenance services. This first-best result extends to our second-best setting due to the fact that no one has an incentive to mimic the high-skilled agents. The third result shows that it is second-best efficient for low-skilled agents to do all the maintenance themselves.

Let us take a closer look at the third result, $z^1 = 0$, which is the most interesting. Suppose we have an initial allocation where $\frac{Y^1}{w^1} + h^1_m > 0$, all constraints of the government problem are satisfied, and the IC constraint (4) holds. Suppose also that, contrary to Proposition 1, $z^1 > 0$. Now consider a reform that increases h^1_m by a small amount dh^1_m, combined with a reduction of Y^1 by $dY^1 = -w^1 dh^1_m$.[15] By construction, the reform leaves the total labor supply of low-skilled agents unaffected, and thus leaves their disutility of working unchanged. This, together with the fact that the proposed reform does not vary c^1, s^1, and l^1, also implies that the reform is welfare neutral for type-1 agents. Moreover, it has no effect on the resource constraint (5) (due to the fact that $dY^1 = -w^1 dh^1_m$ and $dz^1 = -dh^1_m$). Note, however, that the reform has

[14] When high-skilled agents mimic low-skilled agents, they receive the same disposable income as low-skilled agents, but enjoy more leisure time, $Y^1/w^2 < Y^1/w^1$.

[15] Any intended variation dh^1_m can be implemented by keeping s^1 fixed and letting z^1 vary by an amount $dz^1 = -dh^1_m$.

a detrimental effect on highly skilled agents acting as mimickers.[16] This is because it forces them to increase their total labor supply by the amount:

$$dh^1_m + \frac{dY^1}{w^2} = \left(1 - \frac{w^1}{w^2}\right) dh^1_m > 0. \tag{10}$$

By weakening the incentives for high-skilled agents to behave as mimickers, the reform allows to relax the IC constraint (4). This, in turn, opens up the possibility to mitigate the downward distortion on the (total) labor supply provided by low-skilled agents, thereby achieving a higher social welfare. Since, for any given value of s^1, the reform and its effects can be replicated as long as $z^1 > 0$, it follows that a second-best optimum must necessarily be that $z^1 = 0$.[17]

Now let's look at the other properties of a social optimum.

Proposition 2 (Distortions on Structures and Land) *Under a fully non-linear tax system:*

1. *Structures* (s) *are left undistorted for high-skilled agents but encouraged for low-skilled agents:*

$$MRS^2_{s,c} \equiv \frac{\partial g/\partial s^2}{\partial v/\partial c^2} = q_s + \gamma = MRT_{s,c}, \tag{11}$$

$$MRS^1_{s,c} \equiv \frac{\partial g/\partial s^1}{\partial v/\partial c^1} = q_s + \frac{\rho^1 \gamma}{\mu \zeta^1} < MRT_{s,c}. \tag{12}$$

2. *Land* (l) *satisfies:*

$$MRS^j_{l,c} \equiv \frac{\partial g/\partial l^j}{\partial v/\partial c^j} = \frac{\eta}{\mu}, \quad j = 1, 2.$$

Proof. See Appendix A.3. ■

The first part of the proposition concerns the consumption of structures. The result for high-skilled agents shows that their marginal rate of substitution (MRS) between structures (s) and the numéraire good (c) is equal to the corresponding marginal rate of transformation (MRT). This implies that the tax system should not distort their consumption choices for structures. Since no one has an incentive to mimic the high-skilled agents, it is efficient for the government to allocate their resources based on undistorted market prices.

[16] Note that z and s are publicly observable at the individual level, which means that a high-skilled mimicker must replicate s^1 and z^1. In addition, the mimicker must replicate h^1_m because $h_m = (s\gamma - w_z z)/\omega$ and $\omega^1 = \omega^2 = w_z = w^1$.

[17] This result would also hold in a setting where $w^2 > \omega^2 > \omega^1 = w^1$, namely in a setting where high-skilled agents, while more productive in the market than at home, have higher home productivity than low-skilled agents. In such a case, even though a high-skilled mimicker would work fewer hours than a low-skilled agent both in the market and at home (as a mimicker, a high-skilled type would spend at home an amount of time given by $h^1_m \omega^1/\omega^2$), the result that $z^1 = 0$ would still hold because low-skilled agents have a comparative advantage in home production ($w^2/\omega^2 > w^1/\omega^1$). As we will clarify in section 6, the result is also driven by the assumption that low-skilled agents are equally productive when working in the market and when working at home. If they were more productive when working in the market, the mimicking-deterring effects that are achieved by raising h^1_m and lowering Y^1 should be weighed against the efficiency costs arising from the fact that the labor supply of low-skilled agents is shifted towards the activity in which they are less productive.

For low-skilled agents, the MRS between s and c is strictly lower than the corresponding MRT, implying that the tax system should *encourage* structure consumption for this group. The key intuition is that an increase in structure consumption s^1 goes hand in hand with an increase in home maintenance effort h_m^1 (given that low-skilled individuals do not purchase maintenance services in the market). As discussed in Proposition 1, for a given total amount of labor supplied by a low-skilled agent (i.e., a given value of $\frac{Y^1}{w^1} + h_m^1$), a reallocation of effort from work in the market to work at home allows mimicking deterrence effects to be achieved. But an increase in h_m^1 requires a corresponding increase in s^1. Thus, the desirability of encouraging the demand for structures by low-skilled agents can be interpreted as a byproduct of the desirability of increasing h_m^1.

The second part of the proposition concerns land consumption. The key result is that for both high- and low-skilled agents, the MRS between l and c should equal the shadow price of land, denoted by η/μ. This follows from the fact that land is in fixed supply, i.e. its value is determined by scarcity rather than production costs. Unlike structures, land does not require maintenance, and its consumption does not interact with labor supply decisions. As a result, the optimal tax system treats land as a source of economic rent that can be fully taxed away and then redistributed to agents without distorting labor supply or consumption decisions.

3.1 Implementation

Suppose that decentralization is achieved by the government owning all land and leasing it to individuals at a "rent" given by η/μ. Then, to implement the allocation that solves the government's problem, it suffices to use a nonlinear tax $T(Y, s)$. Denote partial derivatives by a subscript on $T(Y, s)$, so that $T_1 = \partial T(Y, s)/\partial Y$ and $T_2 = \partial T(Y, s)/\partial s$; also denote by an asterisk the value that a variable takes in solving the government's problem. The following proposition characterizes the properties of the implementing tax.

Proposition 3 (Implementation) *For implementation purposes, the nonlinear tax* $T(Y, s)$ *should be such that:*

i) The marginal income tax rate should be zero for high-skilled agents (i.e., $T_1(Y^{2*}, s^{2*}) = 0$*), whereas it should be positive for low-skilled agents; in particular,*

$$T_1(Y^{1*}, s^{1*}) = \frac{\lambda^{2,1}}{\mu\zeta^1} \left[\frac{f'\left(\frac{Y^{1*}}{w^2} + \frac{\gamma s^{1*}}{w^1}\right)}{w^2} - \frac{f'\left(\frac{Y^{1*}}{w^1} + \frac{\gamma s^{1*}}{w^1}\right)}{w^1} \right] > 0. \tag{13}$$

ii) Structures for high-skilled agents should be neither taxed nor subsidized at the margin (i.e., $T_2(Y^{2*}, s^{2*}) = 0$*), while low-skilled agents should face a marginal tax on structures* *equal to:*

$$T_2\left(Y^{1*}, s^{1*}\right) = \frac{\lambda^{2,1}}{\mu\zeta^1}\frac{\gamma}{w^1}\left[f'\left(\frac{Y^{1*}}{w^2} + \frac{\gamma s^{1*}}{w^1}\right) - f'\left(\frac{Y^{1*}}{w^1} + \frac{\gamma s^{1*}}{w^1}\right)\right] > 0. \tag{14}$$

iii) At the bundles $\left(Y^{2*}, s^{2*}\right)$ *and* $\left(Y^{1*}, s^{1*}\right)$ *the tax should be given by:*

$$T\left(Y^{2*}, s^{2*}\right) = Y^{2*} - c^{2*} - (q_s + \gamma)\, s^{2*} - \frac{\eta}{\mu} l^{2*}, \tag{15}$$

$$T\left(Y^{1*}, s^{1*}\right) = Y^{1*} - c^{1*} - q_s s^{1*} - \frac{\eta}{\mu} l^{1*}. \tag{16}$$

Proof. See Appendix A.4 ■

The most interesting feature of Proposition 3 is the result that low-skilled agents should face a positive marginal tax on structures. This result may seem surprising at first, given that, according to result (12) in Proposition 2, at a social optimum, the marginal rate of substitution between s and c for low-skilled agents should be lower than the corresponding marginal rate of transformation. The key to reconciling these two seemingly contradictory results lies in the observation that, for an individual providing the necessary maintenance services in-house, the demand for structures is already encouraged relative to the demand for c when income is taxed at the margin. Thus, even if an agent faces a marginal tax on structures, it is still possible that his total consumption demand is stimulated. What (14) tells us is that in the absence of a corrective marginal tax on structures, the demand for structures by low-skilled agents would be over-stimulated (given the positive marginal income tax rate they face). Therefore, in order to implement the socially optimal wedge between $MRS^1_{s,c}$ and the corresponding $MTR_{s,c}$, low-skilled agents must face a marginal tax on structures.

4 Simple tax instruments

In the previous section, we assumed that all variables except hours worked and individual wage rates could be observed at the individual level, allowing for the possibility of nonlinear taxes on all goods. In this section, we characterize the solution to the government's problem under the assumption that while labor income can be subject to a nonlinear tax, other taxes (including land taxes) are constrained to be linear (more precisely, proportional).

We maintain the assumption that the market for maintenance services is perfectly competitive and that people purchase home maintenance services from low-skilled workers. This implies that the (pre-tax) price of z is $w_z = w^1$. Structures are produced at a constant producer price q_s and the pre-tax (endogenous) unitary price of land is denoted by q_l. Good c is chosen as the untaxed numéraire of our economy; for the remaining goods/services, the consumer prices (which include taxes), are given by:

$$
\begin{aligned}
p_1 &\equiv p_z = q_1(1+\tau_z) = w^1(1+\tau_1), \\
p_2 &\equiv p_s = q_2(1+\tau_s) = q_s\,(1+\tau_2)\,, \\
p_3 &\equiv p_l = q_3(1+\tau_l) = q_l(1+\tau_3).
\end{aligned}
$$

As we will see, a key aspect of our analysis below is how the market-clearing, pre-tax, endogenous price of land (q_l) will respond to variations in tax policy.

Denote the nonlinear labor income tax by $T(Y)$, with after-tax income $B \equiv Y - T(Y)$. Agents choose between pre-tax/post-tax bundles (Y, B), with hours worked in the market being equal to Y/w. All other variables are subject to private optimality conditions. Denoting by $\pi^j \bar{l}$ the fraction of the total land $\bar{l}$ owned by each agent of type j (for $j = 1, 2$), the total ("full") disposable income, which includes the value of an individual's land endowment, for an agent of type j is denoted by D^j and given by $D^j \equiv B + q_l \pi^j \bar{l}$.

4.1 Individual's optimization problem

The individual problem can be divided into two stages. In the first stage, an agent solves, for a given (Y, B)-bundle, the problem:

$$\max_{c,s,l,z,h_m} \quad v\,(c) + g\,(s,l) + f\,(Y/w + h_m)$$

subject to

$$c + p_1 z + p_2 s + p_3 l = B + q_3 \pi^j \bar{l} = D^j, \tag{17}$$

$$(z + h_m)\, w^1 = s\gamma, \tag{18}$$

where the first constraint is the private budget constraint and the second is the required maintenance constraint.[18] Denote by $\mathbf{p}$ the vector (p_1, p_2, p_3) and by $\mathbf{x}^j(\mathbf{p}, D^j, Y)$ the vector

$$\left(x_1^j(\mathbf{p}, D^j, Y), x_2^j(\mathbf{p}, D^j, Y), x_3^j(\mathbf{p}, D^j, Y)\right) \equiv \left(z^j(\mathbf{p}, D^j, Y), s^j(\mathbf{p}, D^j, Y), l^j(\mathbf{p}, D^j, Y)\right),$$

[18] Note that the private budget constraint can be equivalently restated as $c + p_1 z + p_2 s + \tau_3 q_3 l + q_3\left(l - \pi^j \bar{l}\right) = B$. This shows that an individual pays a land tax on his total land consumption l (regardless of whether the land was bought on the market or was part of his initial endowment). In addition, an individual must pay the pre-tax price q_3 on the amount of land purchased on the market. If an agent is a net seller of land ($l - \pi^j \bar{l} < 0$), the amount $q_3\left(l - \pi^j \bar{l}\right)$ represents additional resources that supplement his/her after-tax labor income B and that can be used to finance purchases of c and s.

which provides the optimized (conditional) demand for z, s, and l by an agent of type j. Using the vector $\mathbf{x}^j(\mathbf{p}, D^j, Y)$, and considering that

$$\begin{aligned} c^j(\mathbf{p}, D^j, Y) &= D^j - \mathbf{p} * \mathbf{x}^j(\mathbf{p}, D^j, Y), \\ h^j_m(\mathbf{p}, D^j, Y) &= \left[\gamma x^j_2(\mathbf{p}, D^j, Y) - w^1 z^j(\mathbf{p}, D^j, Y)\right] / w^1, \end{aligned}$$

one can insert the optimized values for c, s, l, and h_m into the individual utility function and derive the indirect utility $V^j(\mathbf{p}, D^j, Y)$. In the second stage, taking prices as given, agents choose Y to maximize:

$$V^j\left(\mathbf{p}, Y - T(Y) + q_l \pi^j \bar{l}, Y\right). \tag{19}$$

This gives rise to the first-order condition $V^j_D\,(1 - T'(Y)) + V^j_Y = 0$, from which one can derive the following implicit characterization of the marginal income tax rate faced by an agent:

$$T'(Y) = 1 + V^j_Y / V^j_{D^j} = 1 - MRS^j_{YB} = 1 + \frac{f'\left(\frac{Y}{w^j} + h^j_m\right)}{w^j v'(c^j)}. \tag{20}$$

4.2 Government's problem

The government's problem is:

$$\max_{\{\tau_i\}_{i=1}^3, \{Y^j, B^j\}_{j=1}^2} \sum_{j=1}^{2} \alpha^j \zeta^j V^j(\cdot)$$

subject to

$$V^j\left(\mathbf{p}, D^j, Y^j\right) \geqslant V^j\left(\mathbf{p}, B^k + q_3 \pi^j \bar{l}, Y^k\right) \equiv V^{j,k} \quad \text{for } j = 1, 2 \text{ and } k \neq j, \tag{21}$$

$$\sum_{j=1}^{2} \zeta^j \left[Y^j - B^j + \sum_{i=1}^{2} \tau_i q_i x^j_i\right] + \tau_3 q_3 \sum_{j=1}^{2} \zeta^j l^j \geqslant \bar{R}, \tag{22}$$

$$\tau_3 \leqslant \bar{\tau}_3, \tag{23}$$

where $q_3 = q_l$ is the equilibrium (endogenous) price of land (which allows satisfying the market clearing condition $\sum_{j=1}^{2} \zeta^j l^j(\mathbf{p}, D^j, Y) = \bar{l}$) and $\bar{\tau}_3$ is an upper bound on the land tax to capture political feasibility considerations.[19]

Denote by $\lambda^{j,k}$ the Lagrange multiplier associated with the constraint requiring an agent of type j not to choose the income point intended for an agent of type k. Also, denote by μ and ϕ the Lagrange multipliers associated with the public budget constraint and the constraint

[19] For example, satisfying this constraint could ensure that housing tax policy is sustainable in the sense that it will continue to attract sufficient political support in the future. Scheuer and Wolitzky (2016) explores endogenous such constraints that limit the politically feasible level of capital taxation.

imposing an upper bound on the ad valorem tax τ_l, respectively. We begin by characterizing the optimal tax on land.

Proposition 4 (Land Taxation) *The optimal tax rate on land, $\tau_l = \tau_3$, should be set at the upper bound $\bar{\tau}_3$ provided that*

$$\sum_{j \neq k} \lambda^{j,k} \frac{\partial V^{j,k}}{\partial B^k} \left(\pi^j - \pi^k\right) > 0. \tag{24}$$

Proof. See Appendix B. ■

To understand Proposition 4, recall that agents have a type-specific endowment of land ($\pi^j \bar{l}$ for agents of type j), whose value depends on the endogenous price q_l, which in turn is determined by the market clearing condition $\sum_{i=1}^{2} \zeta^i l^i(\mathbf{p}, D^i, Y) = \bar{l}$. This implies that the price q_l is affected by all policy variables τ_z, τ_s, τ_l, Y^1, Y^2, B^1, and B^2. In particular, by a combined variation in τ_l and B^i (for $i = 1, 2$), the government can reallocate an individual's "full" disposable income (D) from the value of his land endowment to his after-tax labor income. In fact, note that whatever the change in q_l is induced by a variation in τ_l, we can always adjust B^i according to $dB^i = -\pi^i \bar{l} dq_l$ so that $dD^i = 0$.

Now suppose that $\lambda^{j,k} > 0$, so that an agent of type j is initially indifferent between choosing the bundle (Y^j, B^j) intended for him/her and the bundle (Y^k, B^k) intended for an agent of type k. The reduction of q_l caused by a marginal increase of τ_l would have a negative endowment effect that is more severe for an agent of type j than for an agent of type k if $\pi^j > \pi^k$. In this case, the upward adjustment $dB^k = -\pi^k \bar{l} dq_l$ required to compensate for the negative wealth effect experienced by a type k agent would not be sufficient to compensate for the corresponding wealth effect experienced by a type j agent acting as a mimicker. On the other hand, the $dB^j = -\pi^j \bar{l} dq_l$ adjustment fully offsets the negative endowment effect experienced by an individual of type j when not acting as a mimicker (i.e., when choosing the (Y^j, B^j) bundle). Thus, if $\pi^j > \pi^k$ and the $\lambda^{j,k}$ constraint is initially binding, one can design a reform that achieves mimicking deterrence by combining an increase in τ_l with a properly designed change in the nonlinear labor income tax schedule.

Finally, note that the proposed reform implies that the increase in τ_l will be fully capitalized into a lower q_l. In fact, if the market clearing condition $\sum_{i=1}^{2} \zeta^i l^i(\mathbf{p}, D^i, Y) = \bar{l}$ was satisfied in the pre-reform equilibrium, p_l must remain unchanged since $dD^i = 0$ (and $d\tau_z = d\tau_s = dY^1 = dY^2 = 0$), which means that $dq_l / d\tau_l = -q_l / (1 + \tau_l)$.

We now characterize the optimal tax rate on structures. To streamline the exposition, we define:

$$\Xi_i = \frac{\partial \left(\sum_{j=1}^{2} \zeta^j \widetilde{l}^j \right) / \partial p_i}{\partial \left(\sum_{j=1}^{2} \zeta^j \widetilde{l}^j \right) / \partial q_l}, \quad i = 1, 2,$$

where $\widetilde{l}^j$ denotes the Hicksian demand of an agent of type j. Thus, $-\Xi_i$ measures the change in q_l resulting from a compensated marginal increase in p_i.

Proposition 5 (Structures and Maintenance Taxation) *The optimal ad valorem tax on structures and maintenance, τ_s and τ_z, are implicitly characterized by the following conditions:*

$$\sum_{j=1}^{2} \zeta^j \left[\sum_{i=1}^{2} \tau_i q_i \left(\frac{\partial \widetilde{x}_i^j}{\partial p_s} - \Xi_2 \frac{\partial \widetilde{x}_i^j}{\partial q_l} \right) \right]$$
$$= \sum_{j \neq k} \frac{\lambda^{j,k}}{\mu} \frac{\partial V^{j,k}}{\partial B^k} \left[s^k - s^{j,k} - \Xi_2 \left((\pi^j - \pi^k) \bar{l} + (1 + \tau_l) (l^k - l^{j,k}) \right) \right], \quad (25)$$

$$\sum_{j=1}^{2} \zeta^j \left[\sum_{i=1}^{2} \tau_i q_i \frac{\partial \widetilde{x}_i^j}{\partial \tau_z} - \Xi_1 \sum_{i=1}^{2} \tau_i q_i \frac{\partial \widetilde{x}_i^j}{\partial q_l} \right]$$
$$= w^1 \sum_{j \neq k} \frac{\lambda^{j,k}}{\mu} \frac{\partial V^{j,k}}{\partial B^k} \left[z^k - z^{j,k} - \Xi_1 \left((\pi^j - \pi^k) \bar{l} + (1 + \tau_l) (l^k - l^{j,k}) \right) \right], \quad (26)$$

where we use a tilde symbol to denote a compensated variable.

Proof. See Appendix B.3. ■

Equations (25) and (26) closely resemble the standard optimal commodity tax formulas (see, e.g., Edwards et al. 1994), which consider the effects of a marginal change in a commodity tax accompanied by compensating adjustments in the income tax schedule designed to leave the welfare of all non-mimicking agents unaffected. At a social optimum, any net change in tax revenue due to substitution effects must be exactly offset by the welfare effects of either tightening or loosening self-selection constraints. The novelty here is the need to also consider the effect of a compensated increase in τ_s or τ_z on the endogenous price q_l, which affects both sides of equations (25) and (26).

On the left-hand side of (25), the term depending on Ξ_2 captures the revenue effects, operating through a change in q_l and the associated substitution effects on demand, of a compensated variation in τ_s. Note that these revenue effects capture only those associated with changes in the Hicksian demand for maintenance ($\tilde{x}_1$) and structures ($\tilde{x}_2$). A revenue effect on land taxes is absent because land is in fixed supply. On the right hand side of (25), Ξ_2-term derives from the fact that in order to compensate non-mimicking agents for the welfare effects of a marginal increase in τ_s, one must consider both the direct effect on p_s and the indirect effect on q_l.

The logic underlying condition (26) is similar to that underlying condition (25). The tax τ_z has both a direct effect on the price of z and an indirect effect on the price of land. The main difference is that the direct welfare effect of an increase in p_z is zero for those agents who rely only on h_m to provide the maintenance services they need. Moreover, $\Xi_1 = 0$ if all agents, when not acting as mimickers, refrain from purchasing maintenance services in the market.

The next corollary illustrates how a compensated increase in either τ_s or τ_z can act to deter mimicking, thereby reducing income tax distortions, and achieving a higher social welfare.

Corollary 1 *Suppose that high-skilled agents own more land, $\pi^2 > \pi^1$, and that they are net land sellers, $l^2 - \pi^2\bar{l} < 0$. Suppose further that structures and land are Hicksian complements, and that low-skilled agents provide in-house for the maintenance services that they need whereas high-skilled agents purchase these services in the market. Then, starting from an initial situation where $\tau_s = \tau_z = \tau_l = 0$, introducing a small tax on structures or maintenance (while adjusting the nonlinear labor income tax schedule to offset the welfare effects for all non-mimicking agents) allows increasing social welfare.*

Proof. See Appendix B.4. ■

The reason for the result regarding τ_s in Corollary 1 is twofold. First, neither a low-skilled agent nor a high-skilled mimicker buys any maintenance in the market, $z^1 = z^{2,1} = 0$, which implies that they both rely exclusively on home maintenance. But this also implies that, for any given amount of s, the marginal effective price of structures is lower for a high-skilled mimicker than for a low-skilled agent: whereas the unitary consumer price of structures is $p_s = (1 + \tau_s)\, q_s$ for both, the (effort) cost of the additional maintenance services required by a marginal increase in s is larger for a low-skilled agent, since $-f'\left(\frac{Y^1}{w^1} + h_m^1\right) = -f'\left(\frac{Y^1}{w^1} + \frac{\gamma s}{w^1}\right) > -f'\left(\frac{Y^1}{w^2} + \frac{\gamma s}{w^1}\right) = -f'\left(\frac{Y^1}{w^1} + h_m^{2,1}\right)$. It then follows that $s^1 - s^{2,1} < 0$. A tax on structures therefore hurts a mimicker more than a true low-skilled agent. Second, the increase in τ_s exerts a general equilibrium effect on q_l that lowers its value (due to the assumption that s and l are Hicksian complements), and this endogenous price effect is more detrimental to a high-skilled mimicker than to a true low-skilled agent.

The reason for the result regarding τ_z in Corollary 1 is entirely due to the general equilibrium effect on q_l, since both a low-skilled agent and a high-skilled mimicker refrain from purchasing maintenance services in the market. Although $z^1 = z^{2,1} = 0$, an increase in τ_z has a depressive effect on q_l due to the fact that $z^2 > 0$. However, the importance of this result should not be overestimated. Corollary 1 considers the effects of varying τ_z starting from a situation where $\tau_s = \tau_z = \tau_l = 0$. When all of these policy instruments are optimized together, the desired general equilibrium effects on q_l are likely to be achieved more effectively by relying on τ_s and τ_l. Most importantly, Corollary 1 is based on a two-type setting where low-skilled agents are equally productive in the market and at home. As we will see in section 5, where we consider a richer setting with more types, even though the intended beneficiaries of the redistributive policy have a comparative advantage in home production, the optimal τ_z is likely to be negative. The intuition for this result stems from the observation that a positive marginal income tax rate already incentivizes in-house maintenance. This is good, for mimicking-deterrent reasons, if, under the binding IC constraints, the mimicked agents have a comparative advantage over the mimickers in working at home. However, if the mimicked agents are also more productive in the market than at home, inducing them to choose in-house maintenance entails an efficiency cost. If this cost is large (because the mimicked agents are significantly more productive in the market than at home), a negative τ_z (a subsidy on maintenance services purchased in the market) serves to ensure that the efficient maintenance mode is chosen.

Let's now consider the optimality conditions for marginal income taxation.

Proposition 6 *Let Ω be the derivative of the Lagrangian of the government problem with respect to the land price q_l. The optimal marginal income tax rates are given by:*

$$T'\left(Y^1\right) = \frac{\lambda^{2,1}}{\mu\zeta^1}\frac{\partial V^{2,1}}{\partial B^1}\left(MRS^1_{Y,B} - MRS^{2,1}_{Y,B}\right) - \sum_{i=1}^{2}\tau_i q_i\left(\frac{\partial x^1_i}{\partial Y^1} + MRS^1_{Y,B}\frac{\partial x^1_i}{\partial B^1}\right) - \frac{\Omega}{\mu\zeta^1}\left(\frac{\partial q_l}{\partial Y^1} + MRS^1_{Y,B}\frac{\partial q_l}{\partial B^1}\right), \tag{27}$$

$$T'\left(Y^2\right) = \frac{\lambda^{1,2}}{\mu\zeta^2}\frac{\partial V^{1,2}}{\partial B^2}\left(MRS^2_{Y,B} - MRS^{1,2}_{Y,B}\right) - \sum_{i=1}^{2}\tau_i q_i\left(\frac{\partial x^2_i}{\partial Y^2} + MRS^2_{Y,B}\frac{\partial x^2_i}{\partial B^2}\right) - \frac{\Omega}{\mu\zeta^2}\left(\frac{\partial q_l}{\partial Y^2} + MRS^2_{Y,B}\frac{\partial q_l}{\partial B^2}\right). \tag{28}$$

Proof. See Appendix B. ■

Equations (27) and (28) characterize the optimal marginal tax rates for the two groups of agents. The main difference between these formulas and those characterizing a standard Mirrleesian framework is the presence of the term depending on Ω, which reflects the social welfare impact of land price changes. Since eqs. (27)–(28) share a common structure, we limit our discussion to (27). For interpretation, it is useful to rewrite (27) as:

$$T'\left(Y^1\right) + \sum_{i=1}^{2}\tau_i q_i\left(\frac{\partial x^1_i}{\partial Y^1} + MRS^1_{Y,B}\frac{\partial x^1_i}{\partial B^1}\right) = \frac{\lambda^{2,1}}{\mu\zeta^1}\frac{\partial V^{2,1}}{\partial B^1}\left(MRS^1_{Y,B} - MRS^{2,1}_{Y,B}\right) - \frac{\Omega}{\mu\zeta^1}\left(\frac{\partial q_l}{\partial Y^1} + MRS^1_{Y,B}\frac{\partial q_l}{\partial B^1}\right). \tag{29}$$

Consider the left-hand side of (29). It represents the change in total tax revenue if low-skilled agents marginally increase their labor income along an indifference curve. Given an initial bundle (Y^1, B^1), the term $MRS^1_{Y,B} \equiv -\frac{\partial V^1}{\partial Y^1}/\frac{\partial V^1}{\partial B^1}$ represents the minimum amount of additional disposable income required to induce low-skilled agents to earn an additional dollar of income Y^1. Thus, if low-skilled agents were induced to marginally increase their labor income Y^1, the change in income tax revenue would be given by $1 - MRS^1_{Y,B} = T'\left(Y^1\right)$, while the change in revenue from other tax bases would be given by the second term in brackets.[20] In other words, the left-hand side of (29) can be interpreted as the marginal effective tax rate for type-1 agents.

[20]Note that these revenue effects include only those associated with changes in the demand for maintenance (x_1) and structures (x_2). A revenue effect on land taxation is absent because land is in fixed supply; for a given τ_l, the government's revenue from land taxation changes only due to variations in q_l.

At an optimum, there are two reasons why this marginal effective tax rate should be different from zero. The first has to do with the difference between $\mathrm{MRS}^1_{Y,B}$ and $\mathrm{MRS}^{2,1}_{Y,B}$. If Y^1 is slightly increased while B^1 is adjusted by $dB^1 = \mathrm{MRS}^1_{Y,B}$, the resulting change in the utility of a highly skilled mimicker is given by $dV^{2,1} = \frac{\partial V^{2,1}}{\partial B^1}\left(\mathrm{MRS}^1_{Y,B} - \mathrm{MRS}^{2,1}_{Y,B}\right)$. If $\mathrm{MRS}^1_{Y,B} > \mathrm{MRS}^{2,1}_{Y,B}$, then $V^{2,1}$ increases, tightening the corresponding incentive constraint (i.e., the $\lambda^{2,1}$ constraint becomes more binding) and creating an incentive to distort the labor supply of type 1 agents downward. The second reason why low-skilled agents face a non-zero marginal effective tax rate is related to general equilibrium effects on q_l. These affect the incentives to distort the labor supply of low-skilled agents via the second term on the right hand side of (29).

As we show in the appendix, the sign of Ω is generally ambiguous and depends on the interaction between the self-selection terms and the revenue terms. For the sake of interpretation, suppose that $\Omega < 0$, i.e., that a marginal reduction in q_l would be socially valuable. Further suppose that $\frac{\partial q_l}{\partial Y^1} + \mathrm{MRS}^1_{Y,B}\frac{\partial q_l}{\partial B^1} > 0$, i.e., that q_l increases when low-skilled agents are induced to marginally increase their pre-tax labor income Y^1. Under such assumptions, it would follow that, according to the second term on the right-hand side of (29), the general equilibrium effects on q_l provide a separate motive to distort downward the labor supply of low-skilled agents.

5 Quantitative model

While theoretical analysis provides insight into the underlying mechanisms, it does not determine the optimal structure of income and commodity tax rates. To fill this gap, we turn to illustrative numerical simulations using Swedish population register data.

5.1 Generalized government problem in the fully nonlinear tax system

We begin by generalizing the government's problem under nonlinear commodity taxation, as outlined in section 3. As before, let ζ^j denote the fraction of agents with productivity j in the population and α^j denote the corresponding welfare weight. In the numerical simulations, the government's problem with nonlinear commodity taxes is formulated as follows:

$$\max_{\{Y^j, c^j, s^j, l^j, h^j_m, z^j\}} \sum_{j=1}^{N} \alpha^j \zeta^j \left[v(c^j) + g(s^j, l^j) + f\left(\frac{Y^j}{w^j} + h^j_m\right)\right] \tag{30}$$

subject to

$$v(c^j) + g(s^j, l^j) + f\left(\frac{Y^j}{w^j} + h^j_m\right) \geqslant$$
$$v(c^{j'}) + g(s^{j'}, l^{j'}) + f\left(\frac{Y^{j'}}{w^j} + \frac{\omega^{j'}}{\omega^j} h^{j'}_m\right), \quad j, j' \in \{1, \ldots, N\} \tag{31}$$
$$\sum_{j=1}^{N} \zeta^j \left(Y^j - c^j - q_s s^j - w^{j_m} z^j\right) - \sum_{j=1}^{N} \zeta^j q_l l^j \geqslant 0, \tag{32}$$
$$\sum_{j=1}^{N} \zeta^j l^j \leqslant \bar{l}, \tag{33}$$
$$\left(h^j_m \omega^j + z^j w^{j_m}\right) \geqslant \gamma s^j, \quad j = 1, \ldots, N, \tag{34}$$
$$\sum_{j=1}^{j_m} \zeta^j Y^j \geqslant \sum_{j=1}^{N} \zeta^j (\gamma s^j - h^j_m \omega^j). \tag{35}$$

The structure of the above problem is identical to that in section 3, with a few exceptions. First, the notation is richer to account for the presence of N instead of 2 types. Second, the set of incentive constraints is expanded due to the larger number of types. Third, we add the constraint (35) which ensures that the total output of low-productivity workers is at least as large as the total market value of market-purchased maintenance services.[21]

5.2 Generalized government problem in the mixed tax system

Similar to what we did for the nonlinear case, we also generalize the case with simple (linear) tax instruments that we considered in section 4. In the numerical simulations, the government's problem with linear commodity taxes takes the following form:

$$\max_{\{Y^j, B^j\}_{j=1}^N, \boldsymbol{\tau}} \sum_{j=1}^{N} \zeta^j \alpha^j V^j \tag{36}$$

[21]The latter implies, for example, that if a positive amount of maintenance services are purchased in the market, the government cannot set $Y^j = 0$ for $j = 1, \ldots, j_m$, because then there would be no workers available to perform those market-purchased maintenance services. This constraint is added for realism in the numerical part.

subject to

$$V^j\left(\boldsymbol{\tau}, B^j, Y^j\right) \geqslant V^j\left(\boldsymbol{\tau}, B^{j'}, Y^{j'}\right), \quad j, j' \in \{1, \dots, N\}; \tag{37}$$

$$\sum_{j=1}^{N} \zeta^j \left[Y^j - B^j + \tau_z w^{j_m} z^j + \tau_s s^j\right] + \tau_l q_l \bar{l} \geqslant \bar{R}, \tag{38}$$

$$\sum_{j=1}^{j_m} \zeta^j Y^j \geqslant \sum_{j=1}^{N} \zeta^j (\gamma s^j - h_m^j \omega^j), \tag{39}$$

$$\tau_j \leqslant \bar{\tau}, j = s, l \tag{40}$$

where q_l is the equilibrium price of land solving $\sum_{j=1}^{N} \zeta^j l^j = \bar{l}$ and the variables s^j, l^j, z^j refer to the individual (optimized) demand functions. Equation (37) contains the incentive constraints, while (38) represents the public resource constraint with $\bar{R}$ as the exogenous revenue requirement. Equation (39) is identical to (35). Finally, (40) are the upper bounds imposed on the tax rates on structures and land, where we initially set $\bar{\tau} = 0.1$ but gradually consider higher values in the numerical simulations.

5.3 Wage distributions

Wage distributions are calibrated using market hourly wage rates for people living in Sweden in 2016, as shown in figure 1.[22] We consider $N = 10$ different types of productivity, denoted by $w^j, j = 1, \dots, N$. We let w^1 represent the 10th wage percentile, w^2 the 20th percentile, and so on, except for the top wage w^{10}, which represents the 95th percentile of the wage distribution.

Wage rank (j)	1	2	3	4	5	6	7	8	9	10
Hourly wage rate (w^j)	141.59	155.63	165.69	175.63	186.78	200.00	216.88	242.50	292.92	356.25

Table 1: Hourly market wage rates in Sweden in 2016.

5.4 Maintenance productivity

The maintenance productivity of a worker of productivity type j is denoted by ω^j. We assume that firms providing professional maintenance services use labor as their only input and employ workers whose rank in the productivity distribution is j_m. This implies that the producer price of maintenance services is equal to w^{j_m}. For agents $j \leqslant j_m$, we assume that they are equally productive whether they work in the market or perform maintenance at home, namely $\omega^j = w^j \leqslant w^{j_m}$. For agents with $j > j_m$, we assume that their higher productivity in the market does not translate into higher productivity in household production, namely $\omega^j = w^{j_m}$. We set

[22]We use a register-based dataset covering monthly full-time equivalent wages for all employees in the public sector and about half of all employees in the private sector. The dataset is maintained by the Swedish National Mediation Office, which collects data annually from all private companies with more than 500 employees and a stratified representative sample of about 8500 smaller companies. We divide these monthly wages by the usual monthly working hours to obtain hourly wages.

$j_m = 3$, corresponding to the 3rd decile of the wage distribution.

5.5 Land endowments

The proportion of land owned by individuals of productivity type j is denoted by π^j. Based on the available data, it is difficult to assess whether the land owned by an individual is an endowment, in the sense of representing inherited wealth, or whether it was purchased on the market. For simplicity, we use the Swedish property tax register and calculate for each wage group the total taxable market value of land ownership within that wage group and divide by the total taxable market value of land in the property tax register.[23] The resulting land shares are shown in table 2.

Wage rank (j)	1	2	3	4	5	6	7	8	9	10
Land ownership share (π^j)	0.020	0.041	0.052	0.060	0.071	0.086	0.104	0.129	0.163	0.274

Table 2: Land ownership shares in Sweden in 2016.

5.6 Functional forms and parameterization

Inspired by e.g., Muth (1975), housing services are produced according to a CES function:

$$H(s,l) = (\kappa_0 s^{\kappa_1} + (1-\kappa_0) l^{\kappa_1})^{\frac{1}{\kappa_1}}, \tag{41}$$

where κ_1 reflects how substitutable land and structures are in the production of housing services ($\kappa^1 = (\sigma - 1)/\sigma$ where σ is the elasticity of substitution), and κ_0 captures the relative importance of structures and land. Following e.g., Bastani et al. (2020), the components of the utility functions are specified as

$$v(c) = \frac{c^{1-\gamma_0}}{1-\gamma_0}, \quad g(s,l) = \chi_0 \frac{H(s,l)^{1-\gamma_1}}{1-\gamma_1}, \quad f(h+h_m) = \chi_1 \frac{(1-h-h_m)^{1-\gamma_2}}{1-\gamma_2}. \tag{42}$$

The parameters γ_0, γ_1, and γ_2 control the curvature of the utility derived from the consumption of ordinary goods, housing services, and leisure, respectively. The relative importance of the difference components of the utility functions is controlled by χ_0 and χ_1 (the utility weight on the consumption of ordinary goods is normalized to one).

Table 3 presents the parameters used in our baseline specification. The functional forms and parameter values were chosen to ensure that the private decision variables are reasonably scaled, making the components of the utility function comparable in magnitude.

[23] The Swedish property tax was abolished in 2008 and replaced by a capped municipal property tax (*fastighetsavgift*). Despite this change, the Swedish Tax Agency continues to collect data on taxable property values (*taxeringsvärden*), which are divided into separate assessments for buildings and land.

Table 3: Parameters in baseline calibration

Parameter	Description	Value
γ_0	Curvature of consumption	1.01 (log-utility)
γ_1	Curvature of the housing function	1.01 (log-utility)
γ_2	Curvature of leisure	1.01 (log-utility)
χ_0	Weight on housing services in utility function	1
χ_1	Weight on leisure in utility function	1
κ^0	Weight on structures in housing production	0.25
κ^1	Substitution parameter in housing production	-1
$\bar{l}$	Total supply of land	1
R	Exogenous revenue requirement	$0.1 \approx 10\%$ of GDP
α^j	Welfare weights	$1/j, j = 1, \ldots, N.$

As shown in Table 3, the three components of the utility function—consumption, housing, and leisure—are all approximately logarithmic, reflected by an exponent of 1.01. Each component has the same relative weight, since $\chi_0 = \chi_1 = 1$. The share parameter κ^0, which controls the role of structures in housing production, is set to 0.25. This implies that land plays a relatively important role in the production of housing services. The elasticity of substitution between land and structures, κ^1, which is set to -1. This choice implies some degree of complementarity between the two inputs, yielding a substitution elasticity of $\sigma = 0.5$.[24] The choice of $\theta = 1$ imposes the constraint that structures cannot exceed land.[25] Finally, the revenue requirement is set to $R = 0.01$, which is about 10% of GDP in the benchmark economy.

All optimization problems are solved using the state-of-the-art optimization package KNITRO. When commodity taxes are constrained to be linear, the government's problem takes the form of a bi-level optimization program, where the optimal policy must be determined while taking into account the optimal behavior of individual agents. In this case, we solve the government's problem by incorporating the first-order conditions of individual optimization as nonlinear complementarity constraints, allowing for both interior and corner solutions. To compute welfare gains, we follow Bastani et al. (2013) and consider an equivalent variation-type welfare gain measure, where we calculate the amount of additional revenue that must be injected into the government's budget constraint in the pre-reform (zero commodity tax) economy in order to reach the social welfare level of the post-reform economy. We then divide this figure by total output in the pre-reform economy to obtain a relative measure (as a percentage of GDP).

[24] This value is consistent with five of the 12 studies reviewed in McDonald (1981), which report estimates of σ in the range of 0.4-0.6 (see also Jackson et al. 1984).

[25] One interpretation is that a two-story house can be built, but it cannot occupy more than 50% of a given lot.

6 Quantitative results

6.1 Baseline results for the fully nonlinear and mixed tax system

We begin by presenting results for our baseline case, where housing taxes are capped at 10% in the mixed tax system. The top panel of Figure 1 shows the optimal values of leisure, home maintenance, and work hours across the type distribution. With respect to the choice of maintenance mode, low-productivity agents prefer home maintenance, while high-productivity agents rely entirely on market-purchased maintenance. Most individuals devote about 40–60 percent of their time to market work, while home maintenance accounts for between 0 and 10 percent.

The bottom panel of figure 1 shows the structure of optimal marginal income tax rates across the type distribution. While marginal tax rates are relatively flat, they are somewhat higher at lower income levels to support transfers to low-income individuals while discouraging mimicking by higher-income individuals. Conversely, they are lower for high-income individuals to maximize tax revenue. In the mixed tax system, the top marginal tax rate is not zero, a result expected due to the presence of linear commodity taxes. Note, however, that the top marginal tax rate is typically negative in standard models analyzing mixed tax systems; the reason is that the top marginal tax rate internalizes the positive revenue effects (from commodity taxation) that arise when agents work more, thereby earning higher after-tax incomes and increasing the demand for taxed commodities. In our setting, the fact that the top marginal tax rate is positive is due to the endogeneity of q_l. By raising the top marginal tax rate (which causes type 10 agents to earn less), the government can induce a reduction in the equilibrium value of q_l. The key difference between the mixed tax system and the fully nonlinear case is that the labor distortions reflected in the marginal tax rates are consistently lower when the government has access to nonlinear commodity taxation.

Looking at the average tax rates, we find that income is redistributed from high to low productivity agents, with the redistribution being more pronounced under nonlinear commodity taxation. In the mixed tax regime, the optimal tax rates on structures and land reach their upper bounds, i.e. $\tau_s = \tau_l = 0.1$, while the linear tax on maintenance is negative, $\tau_z = -0.209$, and the endogenous price of land is $q_l = 0.222$ (in a setting where $\tau_z = \tau_s = \tau_l = 0$, we have that $q_l = 0.247$). The reason why the tax τ_l should be set to its upper bound is that we are considering a setting where the land endowment is larger for higher ability agents, which implies that at all binding IC constraints, the land endowment of a mimicker is larger than that of the agent being mimicked. Then, according to the result in Proposition 4, τ_l should be pushed as high as possible.

There are two reasons for a positive tax rate on structure. First, at all binding IC-constraints, the demand for structures by the agent acting as a mimicker exceeds that of the agent being mimicked, implying that the tax hurts the former more than the latter.[26] Second, the tax on

[26]In our setup, where higher-ability agents also have a larger land endowment, the relevant IC-constraints are the downward constraints. At these IC-constraints, a mimicker's demand for structures exceeds that of the agent

structures exerts downward pressure on the endogenous price q_l due to the fact that, in our setup, good s and l are Hicksian complements (the parameter κ_l in (41) takes the value -1); Given our assumption that the land endowment is larger for higher ability agents, a reduction in q_l tends to be more detrimental to a mimicker than to a mimicked agent (at each of the relevant IC constraints, namely the downward ones). An important aspect to consider is that, due to the assumption that structures and land are Hicksian complements, the policymaker can exert downward pressure on q_l both by increasing τ_l and by increasing τ_s. While τ_l can be regarded as a more efficient instrument to influence q_l, the importance of the role played by τ_s depends crucially on where the exogenous upper bound on τ_l is set. In our baseline scenario, this upper bound is set at 10%, which is a rather low value. The fact that the government is severely constrained in exploiting the general equilibrium effects that could be achieved through τ_l reinforces the importance of relying on τ_s to achieve them. This helps explain why $\tau_s = \tau_l$, with both set at the common upper bound. As we will see in Section 6.3, as this upper bound is gradually raised, one eventually reaches a point where the two tax rates move in opposite directions (with τ_l increasing and τ_s decreasing).

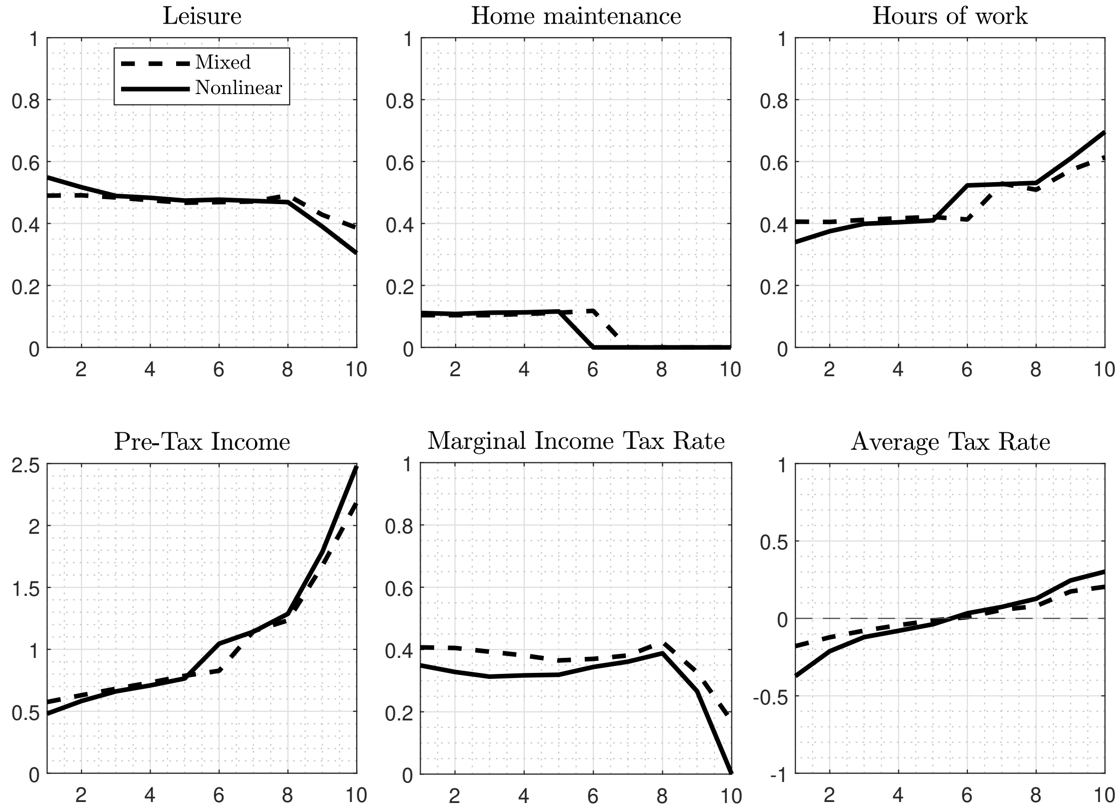
The primary role of τ_z, and the main reason why it is negative (i.e., there is a subsidy for market-purchased maintenance services), is that without such a subsidy, in-house maintenance would be over-incentivized. If $\tau_z = 0$, an agent will prefer in-house maintenance to market-based maintenance if $w(1 - T') < \omega$, where w is the wage rate, T' is the marginal tax rate, and ω is the agent's productivity in performing in-house maintenance. Thus, positive marginal tax rates may imply that an agent chooses to perform in-house maintenance even though it would be better from an efficiency perspective if he didn't (since $w > \omega$). Instead, if $\tau_z \neq 0$, an agent prefers in-house maintenance to market-based maintenance if $w(1 - T') < \omega(1 + \tau_z)$. Thus, by setting $\tau_z < 0$, the government can restore the incentives that lead to efficient maintenance mode choice, at least for some groups of agents.

Compared to a setting in which the only policy instrument is a nonlinear labor income tax, supplementing it with linear commodity taxes (capped at 10%) yields a welfare gain of 2.08% of GDP. Most of these welfare gains come from the positive taxes on structures and land. Allowing positive taxes on structures and land while setting the maintenance tax to zero still yields a welfare gain of 1.80% of GDP, implying that about 86% of the welfare gains come from the taxation of structures and land.[27] Note, however, that while the qualitative features of the optimal tax structure are fairly robust to varying parameters, the exact magnitude of the welfare gains is not.

being mimicked for one or both of the following reasons: i) the effective marginal price of structures (which includes the maintenance costs required by structures) is lower for a mimicker (since he spends less time working in the market, the effort cost of doing maintenance work at home is lower for a mimicker than for the agent being mimicked); ii) total disposable income (which includes the value of an individual's land endowment) is higher for a mimicker than for the agent being mimicked (they have the same after-tax labor income, but the mimicker has a larger land endowment).

[27] Relaxing the upper bound on the land tax yields a welfare gain of 16.84%. In contrast, the fully nonlinear commodity tax system yields a welfare gain of 17.25%.

Figure 1: Results for the fully nonlinear and mixed tax systems



Note: On the x-axis, agent types are reported. The solid lines correspond to the fully nonlinear tax case, while the dashed lines correspond to the mixed tax regime (where only the income tax is nonlinear). The marginal income tax rates ($MITR$) are defined as $MITR = 1 - MRS_{Y,c} = 1 + f'/(wv')$; the average tax rates (ATR) are computed as $\frac{Y^j - c^j - q_s s^j - q_z z^j}{Y^j}$ in the fully nonlinear case and as $\frac{Y^j - B^j + \tau_z q_z z^j + \tau_s q_s s^j}{Y^j}$ in the mixed tax system.

6.2 Further results for the fully nonlinear tax system

In the fully nonlinear case, building on the material in section 3, we compute the implicit wedges between $MRS_{s,c}$ and $MRT_{s,c}$ at the constrained efficient allocation. The values of this wedge for the different types of agents are shown in figure 2, along with the implementing values for the marginal taxes on structures and maintenance.

The first thing to notice (leftmost panel) is that this wedge is zero for most agents. In particular, the wedge is non-zero only for agents of type 3, 4, and 5. This may seem puzzling at first, especially when compared to the result we obtained in the first part of Proposition 2, where we showed that mimicking deterrence considerations justify encouraging the demand for structures by low-skilled agents. There, we interpreted this result as a by-product of the fact that, since low-skilled agents have a comparative advantage in home production (vis-à-vis a high-skilled agent), it is desirable to encourage their efforts in this dimension in order to make it less tempting for high-skilled agents to behave as mimickers.

In the 10-type model we consider here, $w^i = \omega^i$ for $i = 1, 2, 3$, while for types $j = 4, \ldots, 10$ higher market productivity does not lead to higher household productivity. Thus, the comparative advantage argument for encouraging effort at home (to deter imitation by more skilled agents) holds only for agents of type 3 and above. Consequently, for the first two groups of agents, there is no motive to create a wedge between $MRS_{s,c}$ and $MRT_{s,c}$. For agents of type 10, there is no reason to create a wedge because no one is tempted to imitate them. The wedge is also zero for agents of type $j \in \{6, 7, 8, 9\}$. In this case, the reason is that although encouraging their effort along the household margin would reduce the incentives for type $j + 1$ agents to behave as mimickers (because type j agents have a comparative advantage in household production relative to type $j + 1$ agents), this benefit would be more than offset by the efficiency costs of shifting type j agents' effort from the sector where they are more productive (market work) to the sector where they are less productive (household work).

The middle and right panels show how maintenance and structure taxes are combined to implement the required wedge. Types 1–2 are the only ones who choose a mixture of the two available maintenance modes. Since they both face a positive marginal tax rate, a subsidy on z is required to induce them to combine h_m with z.[28] The other groups of agents facing a subsidy on market-purchased maintenance are represented by types 6, 7, and 8. These agents rely only on z. However, given the rate at which their income is taxed at the margin, they would be induced to choose h_m in the absence of a subsidy for market-purchased maintenance.[29] For types 3, 4, and 5, maintenance services are provided exclusively at home, and this choice of maintenance mode is supported by the marginal income tax rate they face. Finally, types 9 and 10 rely only on z. Since these two groups are much more productive in the market than at home, there is no need to support this choice with a subsidy on z.

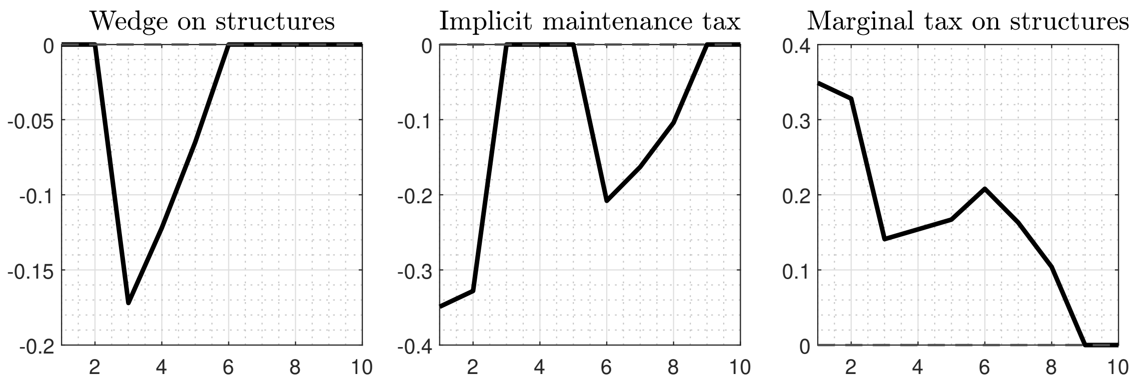
Given that types 1 and 2 benefit from a subsidy on z, and given that such a subsidy would tend to make $MRS^i_{s,c}$ lower than $MRT_{s,c}$, one must let them face a marginal tax on structures in order to have $MRS^i_{s,c} = MRT_{s,c}$. The same is true for types 6, 7, and 8. Finally, for types $j \in \{3, 4, 5\}$ we have that even though $MRS^j_{s,c} < MRT_{s,c}$, they all face a positive marginal tax on structures. This is logically possible since, as we pointed out in our discussion of Proposition 3, for someone who provides the necessary maintenance services in-house, the demand for structures is already encouraged relative to the demand for c when income is taxed at the

[28]Consider a general tax $T\left(Y, s, w^{j_m} z\right)$ with T_i as the i-th partial derivative. For a (w, ω)-agent, the individual FOC for Y is $w(1-T_1)v' = -f'$, and the FOC for s is $\partial g/\partial s = (q_s + T_2)\, v' + \min\{(1 + T_3)\, \gamma v', -\gamma f'/\omega\}$, where the $\min\{\}$ captures that the maintenance cost of raising s can be met either by buying services in the market (utility cost $(1 + T_3)\, \gamma v'$) or by raising h_m (utility cost $-\gamma f'/\omega$). Using the FOC for Y, the FOC for s can be rewritten as $\partial g/\partial s = (q_s + T_2)\, v' + \gamma v' \min\{1 + T_3, w(1 - T_1)/\omega\}$. For an agent to be indifferent between z and h_m, it must be true that $(1 + T_3)\, \omega = (1 - T_1)w$. For agents of types 1 and 2, we have that $w = \omega$, which implies that $T_3 < 0$ (and $T_3 = -T_1$) to make them indifferent between the two maintenance modes.

[29]Consider a general tax $T\left(Y, s, w^{j_m} z\right)$ and let T^i_1 denote the MITR of type i. For $i \in \{6, 7, 8\}$, the implicit marginal tax on maintenance services purchased in the market is $T^i_3 = \frac{(1-T^i_1)w^i}{\omega^i} - 1 < 0$. Any marginal subsidy on $w^{j_m} z \geqslant \left|\frac{(1-T^i_1)w^i}{\omega^i} - 1\right|$ induces agents of type i to prefer market maintenance to household maintenance.

margin.[30]

Figure 2: Further results for the fully nonlinear tax system



Note: On the x-axis, agent types are reported. The definition of the wedge on structures, implicit maintenance tax, and marginal tax on structures are provided in the main text.

6.3 Further results for the mixed tax system

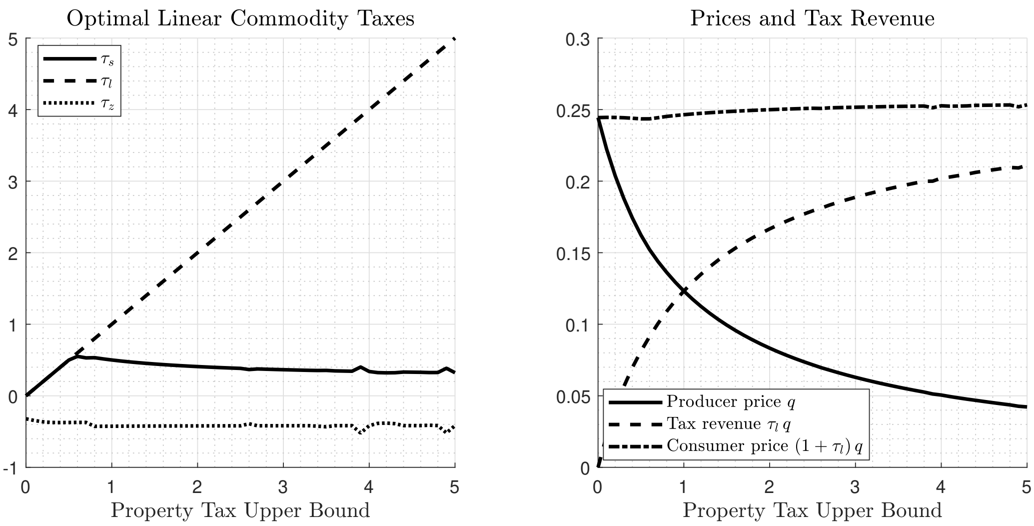
In our base case, taxes on land and structures were capped at 10%. We now examine the effect of gradually varying this cap from 0% to 500%. The results are shown in figure 3. A key observation is that market-purchased maintenance services are subsidized in all cases. Moreover, while the tax on land is always at the upper bound, the upper bound on the tax rate on structures eventually becomes non-binding, and as the tax on land reaches very high values, the tax on structures tends to zero. This happens for two reasons.

The first is that τ_s becomes less important as an instrument to achieve the intended general equilibrium effects on q_l; a more generous upper bound on τ_l provides enough room to achieve these effects directly through land taxation. The second reason is that as τ_l is allowed to reach higher values, interpersonal wealth differences (arising from individual differences in land endowments) become less pronounced, as a higher τ_l is capitalized into a lower q_l. This tends to reduce the difference in demand for structures between a low-skilled agent and a higher-skilled agent acting as a mimicker, since part of this demand gap is due to differences in total disposable income, which includes the value of an individual's land endowment. As a result, the importance of relying on τ_s as a mimicking deterrent is mitigated.

[30]Consider a general tax $T\left(Y, s, w^{j_m} z\right)$. From the definition of the wedge between s and c, i.e. $wedge_{s,c} = MRS_{s,c} - MRT_{s,c} = \frac{\partial g/\partial s}{v'} - q_s - \gamma$, and the individual FOC for s, the marginal tax on structures for agents of type i is $T_2^i = wedge^i_{s,c} + \gamma - \gamma \min\left\{1 + T_3^i, w^i(1 - T_1^i)/\omega^i\right\}$. This implies that, for $i = 1, 2$, $T_2^i = wedge^i_{s,c} + \gamma T_1^i$; for $i = 3, 4, 5$, $T_2^i = wedge^i_{s,c} + \gamma\left[1 - w^i(1 - T_1^i)/\omega^i\right]$; for $i = 6, 7, 8$, $T_2^i = wedge^i_{s,c} - \gamma T_3^i$; for $i = 9, 10$, $T_2^i = wedge^i_{s,c}$.

Finally, notice from the right panel of Figure 3 that land taxes are substantially capitalized into land prices, as reflected in the declining curve for q. While raising the housing tax cap does not significantly change the tax-inclusive price of land (although the pre-tax price falls sharply), it does have a significant impact on government revenues. However, this effect on tax revenues gradually flattens out as the housing tax cap is raised. Thus, capitalization effects limit the potential of housing taxation as a revenue-raising instrument.

Figure 3: Results of varying the housing tax cap in the mixed tax system



Note: On the x-axis we vary the upper bound, $\bar{\tau}$, which applies to both the tax rate on structures and the tax rate on land. This boundary is varied from 0 to 5 in steps of 0.1.

7 Discussion

Housing markets and housing taxation are among the most important and complex areas of economics. Housing plays a dual role as both a consumption good and a store of wealth, and its taxation must balance equity, efficiency, and market distortions. Our analysis contributes to this discussion by examining the redistributive and efficiency implications of housing taxation within a Mirrleesian framework. While our model captures important features of the housing market, it relies on several simplifying assumptions.

We abstract from the investment role of housing and focus only on its consumption function, ignoring saving decisions and interactions with capital taxation (such as the taxation of capital gains related to housing transactions). We also do not consider migration responses to housing taxation, instead assuming that households do not move abroad in response to tax changes. In reality, migration decisions could impose a natural upper bound on housing taxation. Moreover, in some parts of our analysis, we impose an ad hoc upper bound on housing taxation without explicitly modeling the political constraints that could lead to such limits (Scheuer and Wolitzky

2016, Bastani and Waldenström 2021).

Our framework also abstracts from several institutional and spatial considerations. We do not consider the efficient sorting of individuals into communities based on their willingness to pay for public services (Tiebout 1956). Similarly, we do not consider optimal zoning policies that shape housing quality sorting (Hamilton 1975). While we focus on owner-occupied housing, we do not explicitly consider rental housing, which may have important implications for fiscal neutrality and housing market dynamics (Englund 2003). In addition, we do not consider commuting costs or urban sprawl.

A common justification for housing taxes is the benefit principle, which argues that housing taxes serve as a mechanism to finance local public goods and services such as police, sanitation, and parks. In this view, housing taxes are not primarily a redistributive tool, but rather a means of ensuring that those who benefit from local amenities contribute to their provision. This is an important perspective that we hope to address in future work. Our current approach differs in that we focus on housing taxation within an optimal tax framework, balancing equity and efficiency objectives by targeting land rents and considering interactions with labor income taxation.

8 Concluding remarks

Housing is a critical consumption good for most households, yet there is a lack of studies that examine its optimal taxation within a Mirrleesian framework that accounts for the unique characteristics of housing. In this paper, we aim to address this gap by investigating whether and how the government should tax housing to balance equity and efficiency objectives. We construct a model that incorporates the production of housing using both structures and land, as well as the maintenance of structures, which can be performed either at home or purchased in the market. A key feature of our model is the scarcity of land.

Our main findings suggest that housing should generally be taxed, even when a progressive labor income tax is available. The main rationale for taxing housing is to capture land rents, which are a source of inequality and can be efficiently taxed. In contrast, structures should be taxed only when land taxation is constrained. Our results also show that housing taxes affect redistribution indirectly by influencing land prices. In addition, we show that professional maintenance services should be subsidized to ensure an efficient allocation of time between market work and home maintenance and to mitigate the distortions created by income taxation.

Our analysis is based on a stylized model and illustrative simulations designed to capture key mechanisms in a tractable manner. While our framework provides important first steps in understanding the optimal taxation of housing, future research could build on this foundation by developing a richer model that incorporates additional complexities, such as dynamic considerations, endogenous mobility responses, and the interaction between housing taxation and broader wealth accumulation. Moreover, while our calibration is illustrative, a more detailed empirical approach, drawing on richer microdata and incorporating further institutional details, could refine the quantitative implications of our results. Despite these simplifications, our study provides novel insights into the equity and efficiency trade-offs in housing taxation and lays the groundwork for further theoretical and empirical research in this area.

References

Albouy, D. (2016). What are cities worth? land rents, local productivity, and the total value of amenities. *The Review of Economics and Statistics*, 98(3):477–487.

Anderberg, D. and Balestrino, A. (2000). Household production and the design of the tax structure. *International Tax and Public Finance*, 7(4):563–584.

Atkinson, A. B. and Stiglitz, J. E. (1976). The design of tax structure: Direct versus indirect taxation. *Journal of Public Economics*, 6(1-2):55–75.

Auerbach, A. J. and Hines, J. R. (2002). Taxation and economic efficiency. In Auerbach, A. J. and Feldstein, M., editors, *Handbook of Public Economics*, volume 3, pages 1347–1421. Elsevier, Amsterdam.

Aura, S. and Davidoff, T. (2012). An analysis of constrained property taxes in a simple optimal tax model. *CESifo Economic Studies*, 58(3):525–543.

Bastani, S., Blomquist, S., and Micheletto, L. (2013). The welfare gains of age-related optimal income taxation. *International Economic Review*, 54(4):1219–1249.

Bastani, S., Blomquist, S., and Micheletto, L. (2020). Child care subsidies, quality, and optimal income taxation. *American Economic Journal: Economic Policy*, 12(4):1–37.

Bastani, S., Blomquist, S., Micheletto, L., and Tayibov, K. (2024). Taxation of housing in a Mirrleesian context when urban land is scarce. In *Essays on Optimal Income Taxation*. Linnéuniversitetet. Doctoral dissertation by Khayyam Tayibov.

Bastani, S., Blomquist, S., and Pirttilä, J. (2015). How should commodities be taxed? A counter-argument to the recommendation in the Mirrlees Review. *Oxford Economic Papers*, 67(2):455–478.

Bastani, S. and Koehne, S. (2024). How should consumption be taxed? *FinanzArchiv: European Journal of Public Finance*, 80(3):259–302.

Bastani, S. and Waldenström, D. (2021). Wealth and inheritance taxation: Theory and evidence from the Nordic countries. In *Tax by Design for the Netherlands*. Oxford University Press.

Boar, C. and Midrigan, V. (2024). Markups and inequality. *The Review of Economic Studies*. rdae103.

Bonnet, O., Chapelle, G., Trannoy, A., and Wasmer, E. (2021). Land is back, it should be taxed, it can be taxed. *European Economic Review*, 134:103696.

Casarico, A., Del Rey, E., and Silva, J. I. (2023). Child care costs, household liquidity constraints, and gender inequality. *Journal of Population Economics*, 36(3):1461–1487.

Causa, O., Woloszko, N., and Leite, D. (2019). Housing, wealth accumulation and wealth distribution: Evidence and stylized facts. OECD Economics Department Working Papers 1588, OECD Publishing.

Cremer, H. and Gahvari, F. (1998). On optimal taxation of housing. *Journal of Urban Economics*, 43(3):315–335.

Cremer, H., Pestieau, P., and Rochet, J.-C. (2001). Direct versus indirect taxation: The design of the tax structure revisited. *International Economic Review*, 42(3):781–799.

Cremer, H. and Thisse, J.-F. (1994). Commodity taxation in a differentiated oligopoly. *International Economic Review*, 35(3):613–633.

Denk, O. and Schich, S. (2020). Housing, wealth accumulation and wealth distribution: Evidence and stylized facts. Technical report, OECD Publishing.

Edwards, J., Keen, M., and Tuomala, M. (1994). Income tax, commodity taxes and public good provision: A brief guide. *FinanzArchiv / Public Finance Analysis*, 4(51):472–487.

Eeckhout, J., Fu, C., Li, W., and Weng, X. (2025). Optimal taxation and market power. *American Economic Review*. Forthcoming.

Eerola, E. and Määttänen, N. (2013). The optimal tax treatment of housing capital in the neoclassical growth model. *Journal of Public Economic Theory*, 15(6):912–938.

Englund, P. (2003). Taxing residential housing capital. *Urban Studies*, 40(5-6):937–952.

Gahvari, F. (1984a). Incidence and efficiency aspects of differential taxation of residential and industrial capital in a growing economy. *Journal of Public Economics*, 25(2):211–233.

Gahvari, F. (1984b). The optimal taxation of housing. *Public Finance = Finances publiques*, 39(2):213–225.

Gahvari, F. (1985). Taxation of housing, capital accumulation, and welfare: A study in dynamic tax reform. *Public Finance Quarterly*, 13(2):132–160.

Gahvari, F. (2007). On optimal commodity taxes when consumption is time consuming. *Journal of Public Economic Theory*, 9(1):1–27.

Gayle, G.-L. and Shephard, A. (2019). Optimal taxation, marriage, home production, and family labor supply. *Econometrica*, 87(1):291–326.

Gervais, M. (2002). Housing taxation and capital accumulation. *Journal of Monetary Economics*, 49(7):1461–1489.

Guesnerie, R. and Seade, J. (1982). Nonlinear pricing in a finite economy. *Journal of Public Economics*, 17(2):157–179.

Gürer, E. (2022). Rising markups and optimal redistributive taxation. *International Tax and Public Finance*, 29(5):1227–1259.

Hamilton, B. W. (1975). Zoning and property taxation in a system of local governments. *Urban Studies*, 12(2):205–211.

Ho, C. and Pavoni, N. (2020). Efficient child care subsidies. *American Economic Review*, 110(1):162–99.

Jackson, J. R., Johnson, R. C., and Kaserman, D. L. (1984). The measurement of land prices and the elasticity of substitution in housing production. *Journal of Urban Economics*, 16(1):1–12.

Jaravel, X. and Olivi, A. (2022). Prices, non-homotheticities, and optimal taxation: The amplification channel of redistribution. Working Paper.

Judd, K. L. (1985). Redistributive taxation in a simple perfect foresight model. *Journal of Public Economics*, 28(1):59–83.

Kleven, H. J. (2021). Sufficient statistics revisited. *Annual Review of Economics*, 13(1):515–538.

Kleven, H. J., Richter, W. F., and Sørensen, P. B. (2000). Optimal taxation with household production. *Oxford Economic Papers*, 52(3):584–594.

Koehne, S. (2018). On the taxation of durable goods. *International Economic Review*, 59(2):825–857.

Koehne, S. and Sachs, D. (2022). Pareto-improving reforms of tax deductions. *European Economic Review*, 148:104214.

Koll, D., Sachs, D., Stürmer-Heiber, F., and Turon, H. (2023). Equity and efficiency of childcare subsidies: A dynamic structural approach. Working Paper.

Kushnir, A. I. and Zubrickas, R. (2019). Optimal income taxation with endogenous prices. *SSRN Electronic Journal*.

Löffler, M. and Siegloch, S. (2021). Welfare effects of property taxation. ZEW Discussion Paper 21-026, ZEW - Leibniz Centre for European Economic Research, Mannheim.

McDonald, J. F. (1981). Capital-land substitution in urban housing: A survey of empirical estimates. *Journal of Urban Economics*, 9(2):190–211.

Mirrlees, J. A. (1971). An exploration in the theory of optimum income taxation. *Review of Economic Studies*, 38(2):175–208.

Mirrlees, J. A. (1976). Optimal tax theory : A synthesis. *Journal of Public Economics*, 6(4):327–358.

Moschini, E. G. (2023). Childcare subsidies and child skill accumulation in one- and two-parent families. *American Economic Journal: Macroeconomics*, 15(1):475–516.

Muth, R. F. (1975). Numerical solution of urban residential land-use models. *Journal of Urban Economics*, 2(4):307–332.

Myles, G. D. (1995). *Public Economics*. Cambridge University Press, Cambridge, UK.

Määttänen, N. and Terviö, M. (2021). Welfare effects of housing transaction taxes: A quantitative analysis with an assignment model. *The Economic Journal*, 132(644):1566–1599.

Pines, D., Sadka, E., and Sheshinski, E. (1985). The normative and positive aspects of the taxation of imputed rent on owner-occupied housing. *Journal of Public Economics*, 27(1):1–23.

Rossi-Hansberg, E., Sarte, P.-D. G., and Owens III, R. E. (1999). Economics of housing externalities. *International Economic Review*, 40(1):1–25.

Rothschild, C. and Scheuer, F. (2013). Redistributive taxation in the roy model*. *The Quarterly Journal of Economics*, 128(2):623–668.

Sachs, D., Tsyvinski, A., and Werquin, N. (2020). Nonlinear tax incidence and optimal taxation in general equilibrium. *Econometrica*, 88(2):469–493.

Scheuer, F. and Wolitzky, A. (2016). Capital taxation under political constraints. *American Economic Review*, 106(8):2304–28.

Schünemann, J. and Trimborn, T. (2022). Boosting taxes for boasting about houses? status concerns in the housing market. *Journal of Economic Behavior & Organization*, 205:120–143.

Schwerhoff, G., Edenhofer, O., and Fleurbaey, M. (2022). Equity and efficiency effects of land value taxation. *IMF Working Papers*, 2022(263):1–29.

Stern, N. (1987). The effects of taxation, price control and government contracts in oligopoly and monopolistic competition. *Journal of Public Economics*, 32(2):133–158.

Stern, N. H. (1982). Optimum taxation with errors in administration. *Journal of Public Economics*, 17:181–211.

Stiglitz, J. E. (1982). Self-selection and pareto efficient taxation. *Journal of Public Economics*, 17(2):213–240.

Tiebout, C. M. (1956). A pure theory of local expenditures. *Journal of Political Economy*, 64(5):416–424.

A Proof of Propositions in Section 3

A.1 First order conditions for the government's problem

The results in section 3 follow algebra based on the first-order conditions to the governments problem. Denote respectively by $\lambda^{2,1}$, μ, η, and ρ^j (for $j = 1, 2$) the Lagrange multipliers associated with the constraints (4), (5), (6) and (7). The Lagrangian is:

$$\begin{aligned}
\mathcal{L} = & \sum_{j=1}^{2} \alpha^j \zeta^j \left[v\left(c^j\right) + g\left(s^j, l^j\right) + f\left(\frac{Y^j}{w^j} + h_m^j\right) \right] \\
& + \lambda^{2,1} \left[v\left(c^2\right) + g\left(s^2, l^2\right) + f\left(\frac{Y^2}{w^2} + h_m^2\right) - v\left(c^1\right) - g\left(s^1, l^1\right) - f\left(\frac{Y^1}{w^2} + h_m^1\right) \right] \\
& + \mu \left[\sum_{j=1}^{2} \zeta^j \left(Y^j - c^j - q_s s^j - w^1 z^j \right) \right] + \eta \left[\bar{l} - \sum_{j=1}^{2} \zeta^1 l^j \right] \\
& + \sum_{j=1}^{2} \rho^j \left[\left(h_m^j + z^j \right) w^1 - \gamma s^j \right].
\end{aligned}$$

First order conditions with respect to the variables pertaining to type 1 agents

$$\frac{\partial \mathcal{L}}{\partial c^1} = (\alpha^1 \zeta^1 - \lambda^{2,1}) \frac{\partial v}{\partial c^1} - \mu \zeta^1 = 0, \tag{A1}$$

$$\frac{\partial \mathcal{L}}{\partial s^1} = (\alpha^1 \zeta^1 - \lambda^{2,1}) \frac{\partial g(s^1, l^1)}{\partial s^1} - \mu q_s \zeta^1 - \rho^1 \gamma = 0, \tag{A2}$$

$$\frac{\partial \mathcal{L}}{\partial l^1} = (\alpha^1 \zeta^1 - \lambda^{2,1}) \frac{\partial g(s^1, l^1)}{\partial l^1} - \eta \zeta^1 = 0, \tag{A3}$$

$$\frac{\partial \mathcal{L}}{\partial Y^1} = \frac{\alpha^1 \zeta^1}{w^1} f' \left(\frac{Y^1}{w^1} + h_m^1 \right) - \frac{\lambda^{2,1}}{w^2} f' \left(\frac{Y^1}{w^2} + h_m^1 \right) + \mu \zeta^1 = 0, \tag{A4}$$

$$\frac{\partial \mathcal{L}}{\partial h_m^1} = \alpha^1 \zeta^1 f' \left(\frac{Y^1}{w^1} + h_m^1 \right) - \lambda^{2,1} f' \left(\frac{Y^1}{w^2} + h_m^1 \right) + \rho^1 \omega^1 \leqslant 0 \quad \text{with } h_m^1 \frac{\partial \mathcal{L}}{\partial h_m^1} = 0, \tag{A5}$$

$$\frac{\partial \mathcal{L}}{\partial z^1} = -\mu w^1 \zeta^1 + \rho^1 w^1 \leqslant 0 \quad \text{with } z^1 \frac{\partial \mathcal{L}}{\partial z^1} = 0. \tag{A6}$$

First order conditions with respect to the variables pertaining to type 2 agents

$$\frac{\partial \mathcal{L}}{\partial c^2} = (\alpha^2 \zeta^2 + \lambda^{2,1}) \frac{\partial v}{\partial c^2} - \mu \zeta^2 = 0, \tag{A7}$$

$$\frac{\partial \mathcal{L}}{\partial s^2} = (\alpha^2 \zeta^2 + \lambda^{2,1}) \frac{\partial g(s^2, l^2)}{\partial s^2} - \mu q_s \zeta^2 - \rho^2 \gamma = 0, \tag{A8}$$

$$\frac{\partial \mathcal{L}}{\partial l^2} = (\alpha^2 \zeta^2 + \lambda^{2,1}) \frac{\partial g(s^2, l^2)}{\partial l^2} - \eta \zeta^2 = 0, \tag{A9}$$

$$\frac{\partial \mathcal{L}}{\partial Y^2} = \frac{\alpha^2 \zeta^2 + \lambda^{2,1}}{w^2} f' \left(\frac{Y^2}{w^2} + h_m^2 \right) + \mu \zeta^2 = 0, \tag{A10}$$

$$\frac{\partial \mathcal{L}}{\partial h_m^2} = \left(\alpha^2\zeta^2 + \lambda^{2,1}\right) f'\left(\frac{Y^2}{w^2} + h_m^2\right) + \rho^2 w^1 \leqslant 0 \quad \text{with } h_m^2 \frac{\partial \mathcal{L}}{\partial h_m^2} = 0, \tag{A11}$$

$$\frac{\partial \mathcal{L}}{\partial z^2} = -\mu w^1 \zeta^2 + \rho^2 w^1 \leqslant 0 \quad \text{with } z^2 \frac{\partial \mathcal{L}}{\partial z^2} = 0. \tag{A12}$$

A.2 Proof of Proposition 1

Part i) Combining (A10) and (A7) we get:

$$\frac{1}{w^2} \frac{f'\left(\frac{Y^2}{w^2} + h_m^2\right)}{\frac{\partial v}{\partial c^2}} = -1.$$

Combining (A4) and (A1), we get:

$$\frac{f'\left(\frac{Y^1}{w^1} + h_m^1\right)}{w^1 \frac{\partial v}{\partial c^1}} \left[\lambda^{2,1} \frac{\partial v}{\partial c^1} + \mu\zeta^1\right] = \frac{\lambda^{2,1}}{w^2} f'\left(\frac{Y^1}{w^2} + h_m^1\right) - \mu\zeta^1,$$

and therefore

$$1 + \frac{f'\left(\frac{Y^1}{w^1} + h_m^1\right)}{w^1 \frac{\partial v}{\partial c^1}} = \frac{\lambda^{2,1}}{\mu\zeta^1} \left[\frac{f'\left(\frac{Y^1}{w^2} + h_m^1\right)}{w^2} - \frac{f'\left(\frac{Y^1}{w^1} + h_m^1\right)}{w^1}\right].$$

Part ii) From the first order condition (A10) we have that:

$$(\alpha^2\zeta^2 + \lambda^{2,1}) f'\left(\frac{Y^2}{w^2} + h_m^2\right) = -\mu\zeta^2 w^2. \tag{A13}$$

Using (A13) to substitute in (A11) gives:

$$-\mu\zeta^2 w^2 + \rho^2 w^1 \leqslant 0. \tag{A14}$$

Considering (A14) jointly with (A12), and noticing that one of the two must hold as an equality, we can conclude that:

$$\rho^2/\mu\zeta^2 = 1. \tag{A15}$$

Thus, high-skilled agents choose $h_m^2 = 0$.

Part iii) From (A4) we have that:

$$\alpha^1\zeta^1 f'\left(\frac{Y^1}{w^1} + h_m^1\right) = \lambda^{2,1} \frac{w^1}{w^2} f'\left(\frac{Y^1}{w^2} + h_m^1\right) - \mu w^1 \zeta^1. \tag{A16}$$

Using (A16) to substitute in (A5) gives:

$$\lambda^{2,1}\frac{w^1}{w^2}f'\left(\frac{Y^1}{w^2}+\frac{\omega^1}{\omega^2}h_m^1\right)-\lambda^{2,1}f'\left(\frac{Y^1}{w^2}+h_m^1\right)-\mu w^1\zeta^1+\rho^1\omega^1\leqslant 0,$$

or, collecting terms,

$$\lambda^{2,1}\frac{w^1-w^2}{w^2}f'\left(\frac{Y^1}{w^2}+h_m^1\right)-\mu w^1\zeta^1+\rho^1\omega^1\leqslant 0. \tag{A17}$$

Considering (A17) jointly with (A6), and noticing that one of the two must hold as an equality, we can conclude that $\rho^1/\mu=\frac{\lambda^{2,1}}{\mu}\frac{w^2-w^1}{w^1w^2}f'\left(\frac{Y^1}{w^2}+h_m^1\right)+\zeta^1$, implying that $z^1=0$.

A.3 Proof of Proposition 2

From the first order conditions to the government's problem presented in section A.1, we get:

$$\frac{\frac{\partial g(s^1,l^1)}{\partial s^1}}{\frac{\partial v}{\partial c^1}}=q_s+\frac{\rho^1\gamma}{\mu\zeta^1} \tag{A18}$$

$$=q_s+\frac{\lambda^{2,1}\gamma}{\mu\zeta^1}\frac{w^2-w^1}{w^1w^2}f'\left(\frac{Y^1}{w^2}+h_m^1\right)+\gamma, \tag{A19}$$

$$\frac{\frac{\partial g(s^1,l^1)}{\partial l^1}}{\frac{\partial v}{\partial c^1}}=\frac{\eta}{\mu}, \tag{A20}$$

$$\frac{\frac{\partial g(s^2,l^2)}{\partial s^2}}{\frac{\partial v}{\partial c^2}}=q_s+\frac{\rho^2\gamma}{\mu\zeta^2}=q_s+\gamma, \tag{A21}$$

$$\frac{\frac{\partial g(s^2,l^2)}{\partial l^2}}{\frac{\partial v}{\partial c^2}}=\frac{\eta}{\mu}, \tag{A22}$$

where the last equality in (A21) follows from (A15).

A.4 Proof of Proposition 3

Consider the optimization problem solved by an individual under a nonlinear tax $T(y,s)$ and given a rental price η/μ for land:

$$\max\ v\left(Y-T(Y,s)-w_zz-q_ss-\frac{\eta}{\mu}l\right)+g(s,l)+f\left(\frac{Y}{w}+h_m\right) \tag{A23}$$

subject to the constraint $\gamma s=w_zz+\omega h_m$. The first order condition for Y gives:

$$w\left(1-T_1(Y,s)\right)v'=-f' \quad\Longrightarrow\quad T_1(Y,s)=1+f'/(wv'). \tag{A24}$$

Taking into account that we have assumed $w_z = w^1$, the first order condition for s is

$$\frac{\partial g}{\partial s} = \left(q_s + T_2(Y, s)\right) v' + \min\left\{\gamma v', -\gamma \frac{f'}{\omega}\right\}, \tag{A25}$$

where the min{} operator captures that the maintenance cost associated with a marginal increase in s can be met either by purchasing z (entailing a utility cost given by $\gamma v'$) or by increasing h_m (entailing a utility cost given by $-\gamma f'/\omega$). Exploiting (A24), eq. (A25) can be rewritten as:

$$\frac{\partial g}{\partial s} = \left(q_s + T_2(Y, s)\right) v' + \gamma v' \min\left\{1, \frac{w\left(1 - T_1(Y, s)\right)}{\omega}\right\}. \tag{A26}$$

According to eq. (A26), a high-skilled agent (for whom $w = w^2 > \omega = w^1$) will refrain from doing in-house maintenance if $T_1(Y^{2*}, s^{2*}) = 0$. On the other hand, a low-skilled agent (for whom $w = \omega = w^1$) will refrain from buying z if $T_1(Y^{1*}, s^{1*}) > 0$. Finally, the first order condition for l is given by

$$\frac{\partial g}{\partial l} = \frac{\eta}{\mu} v'. \tag{A27}$$

Let $T_1(Y^{2*}, s^{2*}) = T_2(Y^{2*}, s^{2*}) = 0$; a high-skilled agent will then choose to purchase maintenance services in the market, and the allocation intended for him by the planner will be consistent with his private first order conditions

$$1 + \frac{f'\left(\frac{Y^{2*}}{w^2}\right)}{w^2 v'} = 0, \tag{A28}$$

$$\frac{\partial g\left(s^{2*}, l^{2*}\right)}{\partial s} / v' = q_s + \gamma, \tag{A29}$$

$$\frac{\partial g\left(s^{2*}, l^{2*}\right)}{\partial s} / v' = \frac{\eta}{\mu}. \tag{A30}$$

Let $T_1(Y^{1*}, s^{1*})$ be given by the positive rate (13); a low-skilled agent will then refrain from buying maintenance services in the market. Given that, under a positive marginal income tax rate, the first order condition for s of a low skilled agent is given by

$$\frac{\partial g}{\partial s} = \left[q_s + T_2(Y, s) + \gamma\left(1 - T_1(Y, s)\right)\right] v', \tag{A31}$$

it follows that, letting

$$T_2(Y, s) = \frac{\rho^1 \gamma}{\mu \zeta^1} - \gamma(1 - T_1(Y, s)), \tag{A32}$$

a low-skilled agent makes choices that are consistent with the following condition:

$$\frac{\partial g}{\partial s} / v' = q_s + \frac{\rho^1 \gamma}{\mu \zeta^1}. \tag{A33}$$

Therefore, letting $T_1(Y^{1*}, s^{1*})$ be given by (13), and $T_2(Y^{1*}, s^{1*})$ be given by $T_2(Y^{1*}, s^{1*}) = \frac{\rho^1\gamma}{\mu\zeta^1} - \gamma(1 - T_1(Y^{1*}, s^{1*}))$, implies that the allocation intended by the planner for the low-skilled agents is consistent with their private first order conditions. Eq. (14) is then obtained by substituting into $T_2(Y^{1*}, s^{1*}) = \frac{\rho^1\gamma}{\mu\zeta^1} - \gamma(1 - T_1(Y^{1*}, s^{1*}))$ the value for $T_1(Y^{1*}, s^{1*})$ provided by (13), and taking into account that, as shown in Appendix A.1,

$$\frac{\rho^1\gamma}{\mu\zeta^1} = \left[1 - \frac{\lambda^{2,1}}{\mu\zeta^1}\frac{w^1 - w^2}{w^1 w^2} f'\left(\frac{Y^{1*}}{w^2} + h_m^{1*}\right)\right]\gamma. \tag{A34}$$

Finally, from (15)-(16) we have that

$$\sum_{j=1}^{2} \zeta^j T\left(Y^{j*}, s^{j*}\right) = \zeta^1\left[Y^{1*} - c^{1*} - q_s s^{1*} - \frac{\eta}{\mu} l^{1*}\right] + \zeta^2\left[Y^{2*} - c^{2*} - \left(q_s + \gamma\right) s^{2*} - \frac{\eta}{\mu} l^{2*}\right],$$

from which it follows that

$$\frac{\eta}{\mu}\bar{l} + \sum_{j=1}^{2} \zeta^j T\left(Y^{j*}, s^{j*}\right) = -\gamma\zeta^2 s^{2*} + \sum_{j=1}^{2} \zeta^j\left[Y^{j*} - c^{j*} - q_s s^{j*}\right]. \tag{A35}$$

Given that high-skilled agents (at (Y^{2*}, s^{2*})) purchase maintenance services in the market and low-skilled agents (at (Y^{1*}, s^{1*})) do maintenance in-house, $z^2 = \gamma s^{2*}/w^1$ and $z^1 = 0$. The right hand side of (A35) is then equal to zero given that the allocation that solves the government's problem satisfies the resource constraint (5). Thus, according to (A35), the public budget constraint is balanced and the relationship between $T\left(Y^{1*}, s^{1*}\right)$ and $T\left(Y^{2*}, s^{2*}\right)$ is given by $T\left(Y^{1*}, s^{1*}\right) = -\left[\zeta^2 T\left(Y^{2*}, s^{2*}\right) + \frac{\eta}{\mu}\bar{l}\right]/\zeta^1$.

Two remarks are in order. As is well known, in optimal income tax models with a discrete number of types, the implementing tax function is typically non-differentiable at the relevant income levels. This is also true in our model where the tax schedule $T(Y, s)$ should be kinked at (Y^{1*}, s^{1*}), to ensure implementability of the socially optimal allocation. More precisely, the value provided by the right-hand sides of (13) and (14) should be more properly interpreted as the left-hand derivatives of the implementing tax function, i.e. $T_1^-(Y^{1*}, s^{1*})$ and $T_2^-(Y^{1*}, s^{1*})$.

Given that the IC-constraint (4) guarantees that a high-skilled agent is weakly better off by at the bundle intended for him than at the bundle intended for a low-skilled agent, implementability requires that the best deviating strategy for a high-skilled agent is indeed represented by the choice of the bundle $\left(Y^{1*}, s^{1*}\right)$. For a high-skilled agent, the first order condition (A24) is satisfied at $\left(Y^{1*}, s^{1*}\right)$ provided that $T_1(Y^{1*}, s^{1*}) = 1 + \frac{f'\left(\frac{Y^{1*}}{w^2} + \frac{\gamma s^{1*}}{w^1}\right)}{w^2 v'(c^{1*})}$, which is larger than $1 + \frac{f'\left(\frac{Y^{1*}}{w^1} + \frac{\gamma s^{1*}}{w^1}\right)}{w^1 v'(c^{1*})}$ (which is the value required for a low-skilled agent). Thus, $T_1^+(Y^{1*}, s^{1*})$

should exceed the right-hand side value of eq. (13) by an amount given by

$$\frac{f'\left(\frac{Y^{1*}}{w^2}+\frac{\gamma s^{1*}}{w^1}\right)}{w^2 v'\left(c^{1*}\right)}-\frac{f'\left(\frac{Y^{1*}}{w^1}+\frac{\gamma s^{1*}}{w^1}\right)}{w^1 v'\left(c^{1*}\right)}>0.$$

Now consider the marginal tax on structure provided by (14). It represents a value that is too low to ensure that a high-skilled agent finds optimal to choose $s=s^{1*}$ when earning Y^{1*}. To understand this point, consider the first order condition, with respect to s, for an agent of type i earning Y^{1*} and relying on do-it-yourself activity to provide the required maintenance services. This is given by

$$\frac{\partial g}{\partial s}=\left(q_s+T_2\left(Y^{1*},s\right)\right)v'-\frac{\gamma}{w^1}f'\left(\frac{Y^{1*}}{w^i}+\frac{\gamma s}{w^1}\right).$$

Suppose now that $T_2\left(Y^{1*},s\right)$ is such that, for $w^i=w^1$, the first order condition above implies that the amount s^{1*} is chosen. The same $T_2\left(Y^{1*},s\right)$ would imply that, for $w^i=w^2>w^1$, the amount s^{1*} would be suboptimally low. Thus, unless the right-hand derivative $T_2^+(Y^{1*},s^{1*})$ is sufficiently larger than the rate provided by (14), a high-skilled mimicker would have an incentive to choose an amount of structure larger than s^{1*}. To ensure that, when earning Y^{1*}, the best course of action for a high-skilled is to buy an amount of structures s^{1*}, $T_2^+(Y^{1*},s^{1*})$ should exceed the right-hand side value of eq. (14) by an amount given by

$$\frac{\gamma}{w^1}\left[f'\left(\frac{Y^{1*}}{w^2}+\frac{\gamma s^{1*}}{w^1}\right)-f'\left(\frac{Y^{1*}}{w^1}+\frac{\gamma s^{1*}}{w^1}\right)\right].$$

B Proof of Propositions in Section 4

B.1 Comparative static results

Before presenting the first order conditions of the government's problem, we begin by providing the comparative statics results for the price of urban land. The before tax price of urban land, q_l, is given by the condition

$$\sum_{j=1,2}\zeta^j l^j\left(p_z,p_s,p_l,D^j,Y^j\right)=\bar{l} \tag{B1}$$

If we disregard the effect of a change in τ_l on the disposable income D, the comparative statics are simple. The consumer price p_l will not change and we will have $dq_l/d\tau_l=-q_l/(1+\tau_l)$. However, taking the effect on disposable income into account we get:

$$\frac{dq_l}{d\tau_l}=-\frac{q_l\sum_{j=1,2}\zeta^j\partial l^j/\partial p_l}{\sum_{j=1,2}\left[(1+\tau_l)\,\partial l^j/\partial p_l+\pi^j\bar{l}\partial l^j/\partial D^j\right]\zeta^j} \tag{B2}$$

and

$$\frac{dp_l}{d\tau_l} = \frac{q_l \sum_{j=1,2} \zeta^j \pi^j \bar{l} \partial l^j / \partial D^j}{\sum_{j=1,2} \left[(1+\tau_l) \partial l^j / \partial p_l + \pi^j \bar{l} \partial l^j / \partial D^j \right] \zeta^j}. \tag{B3}$$

Other comparative static results that will be useful at a later stage include:

$$\frac{dq_l}{dB^2} = -\frac{\zeta^2 \partial l^2 / \partial D^2}{\sum_{j=1,2} \left[(1+\tau_l) \partial l^j / \partial p_l + \pi^j \bar{l} \partial l^j / \partial D^j \right] \zeta^j} \tag{B4}$$

$$\begin{aligned} \frac{dq_l}{dY^2} &= -\frac{\zeta^2 \partial l^2 / \partial Y^2}{\sum_{j=1,2} \left[(1+\tau_l) \partial l^j / \partial p_l + \pi^j \bar{l} \partial l^j / \partial D^j \right] \zeta^j} \\ &= -\frac{1}{w^2} \frac{\zeta^2 \partial l^2 / \partial h^2}{\sum_{j=1,2} \left[(1+\tau_l) \partial l^j / \partial p_l + \pi^j \bar{l} \partial l^j / \partial D^j \right] \zeta^j} \end{aligned} \tag{B5}$$

$$\frac{dq_l}{dB^1} = -\frac{\zeta^1 \partial l^1 / \partial D^1}{\sum_{j=1,2} \left[(1+\tau_l) \partial l^j / \partial p_l + \pi^j \bar{l} \partial l^j / \partial D^j \right] \zeta^j} \tag{B6}$$

$$\begin{aligned} \frac{dq_l}{dY^1} &= -\frac{\zeta^1 \partial l^1 / \partial Y^1}{\sum_{j=1,2} \left[(1+\tau_l) \partial l^j / \partial p_l + \pi^j \bar{l} \partial l^j / \partial D^j \right] \zeta^j} \\ &= -\frac{1}{w^1} \frac{\zeta^1 \partial l^1 / \partial h^1}{\sum_{j=1,2} \left[(1+\tau_l) \partial l^j / \partial p_l + \pi^j \bar{l} \partial l^j / \partial D^j \right] \zeta^j} \end{aligned} \tag{B7}$$

$$\frac{dq_l}{d\tau_s} = -\frac{\sum_{j=1,2} \zeta^j \partial l^j / \partial p_s}{\sum_{j=1,2} \left[(1+\tau_l) \partial l^j / \partial p_l + \pi^j \bar{l} \partial l^j / \partial D^j \right] \zeta^j} \tag{B8}$$

$$\frac{dq_l}{d\tau_z} = -\frac{\sum_{j=1,2} \zeta^j \partial l^j / \partial p_z}{\sum_{j=1,2} \left[(1+\tau_l) \partial l^j / \partial p_l + \pi^j \bar{l} \partial l^j / \partial D^j \right] \zeta^j} \tag{B9}$$

Given that the equilibrium value for q_l satisfies eq. (B1), we have that $\sum_{j=1,2} \tau_l q_l \zeta^j l^j = \tau_l q_l \bar{l}$. We take this into account when studying the government's problem below. Defining $D^{j,k}$ as $D^{j,k} = B^k + q_l \pi^j \bar{l}$, the Lagrangian of the government's problem can be written as

$$\begin{aligned} \Lambda &= \sum_{j=1}^{2} \alpha^j \zeta^j V^j(\mathbf{p}, D^j, Y^j) + \sum_{j=1}^{2} \left\{ \sum_{k \neq j} \lambda^{j,k} \left[V^j(\mathbf{p}, D^j, Y^j) - V^j(\mathbf{p}, D^{j,k}, Y^k) \right] \right\} \\ &\quad + \mu \left\{ \tau_3 q_3 \bar{l}_u + \sum_{j=1}^{2} \zeta^i \left(Y^i - B^i + \sum_{i=1}^{2} \tau_i q_i x_i^j \right) - \bar{R} \right\} + \phi \left[\bar{\tau}_l - \tau_3 \right]. \end{aligned}$$

Define $\Omega \equiv \frac{\partial \Lambda}{\partial q_l} = \frac{\partial \Lambda}{\partial q_3}$ as follows (where π^i is the proportion of land that is owned by each agent of type i, with $i = 1, 2$):

$$\begin{aligned}\Omega \quad \equiv \quad & \frac{\partial \Lambda}{\partial q_l} = \sum_{j=1}^{2} \left(\alpha^j \zeta^j + \sum_{k \neq j} \lambda^{j,k} \right) \left[\frac{\partial V^j}{\partial D^j} \pi^j \bar{l} + (1 + \tau_l) \frac{\partial V^j}{\partial p_l} \right] \\ & - \sum_{j=1}^{2} \left[\sum_{k \neq j} \lambda^{j,k} \left(\frac{\partial V^{j,k}}{\partial D^{j,k}} \pi^j \bar{l} + (1 + \tau_l) \frac{\partial V^{j,k}}{\partial p_l} \right) \right] \\ & + \mu \sum_{j=1}^{2} \left\{ \zeta^j \sum_{i=1}^{2} \tau_i q_i \left[\frac{\partial x_i^j}{\partial D^j} \pi^j \bar{l} + (1 + \tau_l) \frac{\partial x_i^j}{\partial p_l} \right] \right\} + \mu \tau_l \bar{l}.\end{aligned} \tag{B10}$$

The first order conditions of the government's problem, for the policy variables Y^1, B^1, Y^2, B^2, are respectively given by:

$$\left(\alpha^1 \zeta^1 + \lambda^{1,2} \right) \frac{\partial V^1}{\partial Y^1} = \lambda^{2,1} \frac{\partial V^{2,1}}{\partial Y^1} - \mu \zeta^1 \left[1 + \sum_{i=1}^{2} \tau_i q_i \frac{\partial x_i^1}{\partial Y^1} \right] - \Omega \frac{\partial q_l}{\partial Y^1} \tag{B11}$$

$$\left(\alpha^1 \zeta^1 + \lambda^{1,2} \right) \frac{\partial V^1}{\partial B^1} = \lambda^{2,1} \frac{\partial V^{2,1}}{\partial B^1} + \mu \zeta^1 \left[1 - \sum_{i=1}^{2} \tau_i q_i \frac{\partial x_i^1}{\partial B^1} \right] - \Omega \frac{\partial q_l}{\partial B^1} \tag{B12}$$

$$\left(\alpha^2 \zeta^2 + \lambda^{2,1} \right) \frac{\partial V^2}{\partial Y^2} = \lambda^{1,2} \frac{\partial V^{1,2}}{\partial Y^2} - \mu \zeta^2 \left[1 + \sum_{i=1}^{2} \tau_i q_i \frac{\partial x_i^2}{\partial Y^2} \right] - \Omega \frac{\partial q_l}{\partial Y^2} \tag{B13}$$

$$\left(\alpha^2 \zeta^2 + \lambda^{2,1} \right) \frac{\partial V^2}{\partial B^2} = \lambda^{1,2} \frac{\partial V^{1,2}}{\partial B^2} + \mu \zeta^2 \left[1 - \sum_{i=1}^{2} \tau_i q_i \frac{\partial x_i^2}{\partial B^2} \right] - \Omega \frac{\partial q_l}{\partial B^2} \tag{B14}$$

The first order condition with respect to τ_l is given by

$$\begin{aligned}& \sum_{j=1}^{2} \left(\alpha^j \zeta^j + \sum_{k \neq j} \lambda^{j,k} \right) \frac{\partial V^j}{\partial \tau_l} - \sum_{j=1}^{2} \left(\sum_{k \neq j} \lambda^{j,k} \frac{\partial V^{j,k}}{\partial \tau_l} \right) \\ & + \mu q_l \bar{l} + \mu \sum_{j=1}^{2} \left(\zeta^j \sum_{i=1}^{2} \tau_i q_i \frac{\partial x_i^j}{\partial \tau_l} \right) + \Omega \frac{\partial q_l}{\partial \tau_l} - \phi = 0,\end{aligned}$$

whereas the first order condition for $\tau_2 = \tau_s$ is given by

$$\begin{aligned}& \sum_{j=1}^{2} \left(\alpha^j \zeta^j + \sum_{k \neq j} \lambda^{j,k} \right) \frac{\partial V^j}{\partial \tau_2} - \sum_{j=1}^{2} \left(\sum_{k \neq j} \lambda^{j,k} \frac{\partial V^{j,k}}{\partial \tau_2} \right) \\ & + \mu \sum_{j=1}^{2} \left[\zeta^j \left(\sum_{i=1}^{2} \tau_i q_i \frac{\partial x_i^j}{\partial \tau_2} + q_2 x_2^j \right) \right] + \Omega \frac{\partial q_l}{\partial \tau_2} = 0,\end{aligned}$$

and the first order condition with respect to $\tau_1 = \tau_z$ is given by:

$$\sum_{j=1}^{2}\left(\alpha^j\zeta^j+\sum_{k\neq j}\lambda^{j,k}\right)\frac{\partial V^j}{\partial\tau_1}-\sum_{j=1}^{2}\left(\sum_{k\neq j}\lambda^{j,k}\frac{\partial V^{j,k}}{\partial\tau_1}\right)$$
$$+\mu\sum_{j=1}^{2}\left[\zeta^j\left(\tau_2 q_2\frac{\partial x_2^j}{\partial\tau_1}+\tau_1 w^1\frac{\partial z^j}{\partial\tau_1}+w^1 z^j\right)\right]+\Omega\frac{\partial q_l}{\partial\tau_1}=0.$$

B.2 Proof of Proposition 4

Applying Roy's identity, we can rewrite the first order condition with respect to $\tau_1 = \tau_l$ as:

$$-\sum_{j=1}^{2}\left(\alpha^j\zeta^j+\sum_{k\neq j}\lambda^{j,k}\right)q_l l^j\frac{\partial V^j}{\partial D^j}+\sum_{j=1}^{2}\left(\sum_{k\neq j}\lambda^{j,k}q_l l^{j,k}\frac{\partial V^{j,k}}{\partial D^{j,k}}\right)$$
$$+\mu\left[q_l\bar{l}+\sum_{j=1}^{2}\left(\zeta^j\sum_{i=1}^{2}\tau_i q_i\frac{\partial x_i^j}{\partial\tau_l}\right)\right]+\Omega\frac{\partial q_l}{\partial\tau_l}-\phi=0.$$

Multiply the first order condition with respect to B^1 by $q_l l^1$ and the first order condition with respect to B^2 by $q_l l^2$; adding up the resulting equations with the first order condition with respect to τ_l, we get:

$$\sum_{j=1}^{2}\left(\sum_{k\neq j}\lambda^{j,k}q_l l^{j,k}\frac{\partial V^{j,k}}{\partial D^{j,k}}\right)-\lambda^{2,1}q_l l^1\frac{\partial V^{2,1}}{\partial B^1}-\lambda^{1,2}q_l l^2\frac{\partial V^{1,2}}{\partial B^2}$$
$$+\mu\left\{\sum_{j=1}^{2}\left(\zeta^j\sum_{i=1}^{2}\tau_i q_i\frac{\partial x_i^j}{\partial\tau_l}\right)+q_l\sum_{j=1}^{2}\left[\zeta^j l^j\left(\sum_{i=1}^{2}\tau_i q_i\frac{\partial x_i^j}{\partial B^j}\right)\right]\right\}$$
$$+\Omega q_l l^1\frac{\partial q_l}{\partial B^1}+\Omega q_l l^2\frac{\partial q_l}{\partial B^2}+\Omega\frac{\partial q_l}{\partial\tau_l}-\phi=0.$$

The equation above can be rewritten as

$$q_l\sum_{k=1}^{2}\left[\sum_{j\neq k}\lambda^{j,k}\left(l^{j,k}-l^k\right)\frac{\partial V^{j,k}}{\partial B^k}\right]$$
$$+\mu\left\{\sum_{j=1}^{2}\left(\zeta^j\sum_{i=1}^{2}\tau_i q_i\frac{\partial x_i^j}{\partial\tau_l}\right)+q_l\sum_{j=1}^{2}\left[\zeta^j l^j\left(\sum_{i=1}^{2}\tau_i q_i\frac{\partial x_i^j}{\partial B^j}\right)\right]\right\}$$
$$+\Omega q_l l^1\frac{\partial q_l}{\partial B^1}+\Omega q_l l^2\frac{\partial q_l}{\partial B^2}+\Omega\frac{\partial q_l}{\partial\tau_l}-\phi=0,$$

or equivalently, denoting by a tilde a compensated variable (for instance, $\zeta^j q_l l^j \frac{\partial z^j}{\partial B^j} + \zeta^j \frac{\partial z^j}{\partial \tau_l} = \zeta^j \frac{\partial \tilde{z}^j}{\partial \tau_l}$):

$$
\begin{aligned}
& q_l \sum_{k=1}^{2} \left[\sum_{j \neq k} \lambda^{j,k} \left(l^{j,k} - l^k \right) \frac{\partial V^{j,k}}{\partial B^k} \right] \\
& + \mu \left[\sum_{j=1}^{2} \left(\zeta^j \sum_{i=1}^{2} \tau_i q_i \frac{\partial \tilde{x}_i^j}{\partial \tau_l} \right) \right] + \left(\frac{\partial q_l}{\partial \tau_l} + q_l \sum_{j=1}^{2} l^j \frac{\partial q_l}{\partial B^j} \right) \Omega - \phi = 0,
\end{aligned}
$$

where we have $\frac{\partial \tilde{z}^k}{\partial \tau_l} = \frac{\partial \tilde{z}^k}{\partial p_l} q_l$ and $\frac{\partial \tilde{x}_i^k}{\partial \tau_l} = \frac{\partial \tilde{x}_i^k}{\partial p_l} q_l$. The equation above can also be rewritten as follows:

$$
\begin{aligned}
& \left(\frac{\partial q_l}{\partial \tau_l} + q_l \sum_{j=1}^{2} l^j \frac{\partial q_l}{\partial B^j} \right) \Omega - \phi + \mu \left[\sum_{j=1}^{2} \left(\zeta^j \sum_{i=1}^{2} \tau_i q_i \frac{\partial \tilde{x}_i^j}{\partial \tau_l} \right) \right] \\
= \; & -q_l \sum_{j=1}^{2} \left[\sum_{k \neq j} \lambda^{j,k} \frac{\partial V^{j,k}}{\partial B^k} \left(l^{j,k} - l^k \right) \right].
\end{aligned}
$$

Consider now term Ω. Applying a Slutsky-type decomposition, we can rewrite it as:

$$
\begin{aligned}
\Omega \; \equiv \; & \frac{\partial \Lambda}{\partial q_l} = \sum_{j=1}^{2} \left(\alpha^j \zeta^j + \sum_{k \neq j} \lambda^{j,k} \right) \left[\pi^j \bar{l} - (1+\tau_l) l^j \right] \frac{\partial V^j}{\partial D^j} \\
& - \sum_{j=1}^{2} \left[\sum_{k \neq j} \lambda^{j,k} \left(\pi^j \bar{l} - (1+\tau_l) l^{j,k} \right) \frac{\partial V^{j,k}}{\partial D^{j,k}} \right] + \mu \tau_l \bar{l} \\
& + \mu \sum_{j=1}^{2} \left\{ \zeta^j \sum_{i=1}^{2} \tau_i q_i \left[(1+\tau_l) \frac{\partial \tilde{x}_i^j}{\partial p_l} - \left((1+\tau_l) l^j - \pi^j \bar{l} \right) \frac{\partial x_i^j}{\partial D^j} \right] \right\}.
\end{aligned}
$$

Rewriting the first order conditions with respect to B^1 and B^2 respectively as

$$
\begin{aligned}
& -\Omega \frac{\partial q_l}{\partial B^1} \left[\pi^1 \bar{l} - (1+\tau_l) l^1 \right] \\
= \; & \left(\alpha^1 \zeta^1 + \lambda^{1,2} \right) \frac{\partial V^1}{\partial B^1} \left[\pi^1 \bar{l} - (1+\tau_l) l^1 \right] - \left[\pi^1 \bar{l} - (1+\tau_l) l^1 \right] \lambda^{2,1} \frac{\partial V^{2,1}}{\partial B^1} \\
& - \mu \zeta^1 \left[1 - \sum_{i=1}^{2} \tau_i q_i \frac{\partial x_i^1}{\partial B^1} \right] \left[\pi^1 \bar{l} - (1+\tau_l) l^1 \right],
\end{aligned}
$$

$$
\begin{aligned}
& -\Omega \frac{\partial q_l}{\partial B^2}\left[\pi^2 \bar{l}-(1+\tau_l) l^2\right] \\
= \; & \left(\alpha^2 \zeta^2+\lambda^{2,1}\right) \frac{\partial V^2}{\partial B^2}\left[\pi^2 \bar{l}-(1+\tau_l) l^2\right]-\left[\pi^2 \bar{l}-(1+\tau_l) l^2\right] \lambda^{1,2} \frac{\partial V^{1,2}}{\partial B^2} \\
& -\mu \zeta^2\left[1-\sum_{i=1}^2 \tau_i q_i \frac{\partial x_i^2}{\partial B^2}\right]\left[\pi^2 \bar{l}-(1+\tau_l) l^2\right].
\end{aligned}
$$

we can re-express Ω as:

$$
\begin{aligned}
\Omega = \; & -\Omega \sum_{j=1}^2 \frac{\partial q_l}{\partial B^j}\left[\pi^j \bar{l}-(1+\tau_l) l^j\right]+ \\
& +\left[\pi^1 \bar{l}-(1+\tau_l) l^1\right] \lambda^{2,1} \frac{\partial V^{2,1}}{\partial B^1}+\left[\pi^2 \bar{l}-(1+\tau_l) l^2\right] \lambda^{1,2} \frac{\partial V^{1,2}}{\partial B^2}+\mu \tau_l \bar{l} \\
& +\mu \zeta^1\left[1-\sum_{i=1}^2 \tau_i q_i \frac{\partial x_i^1}{\partial B^1}\right]\left[\pi^1 \bar{l}-(1+\tau_l) l^1\right]+\mu \zeta^2\left[1-\sum_{i=1}^2 \tau_i q_i \frac{\partial x_i^2}{\partial B^2}\right]\left[\pi^2 \bar{l}-(1+\tau_l) l^2\right] \\
& +\mu \sum_{j=1}^2\left\{\zeta^j \sum_{i=1}^2 \tau_i q_i\left[(1+\tau_l) \frac{\partial \widetilde{x}_i^j}{\partial p_l}-\left((1+\tau_l) l^j-\pi^j \bar{l}\right) \frac{\partial x_i^j}{\partial D^j}\right]\right\} \\
& -\sum_{j=1}^2\left[\sum_{k \neq j} \lambda^{j,k}\left(\pi^j \bar{l}-(1+\tau_l) l^{j,k}\right) \frac{\partial V^{j,k}}{\partial D^{j,k}}\right],
\end{aligned}
$$

or equivalently, simplifying terms:

$$
\begin{aligned}
\Omega = \; & -\Omega \sum_{j=1}^2 \frac{\partial q_l}{\partial B^j}\left[\pi^j \bar{l}-(1+\tau_l) l^j\right] \\
& +\sum_{j=1}^2\left\{\sum_{k \neq j} \lambda^{j,k} \frac{\partial V^{j,k}}{\partial B^k}\left[\left(\pi^k \bar{l}-(1+\tau_l) l^k\right)-\left(\pi^j \bar{l}-(1+\tau_l) l^{j,k}\right)\right]\right\} \\
& +\mu \sum_{j=1}^2\left\{\zeta^j \sum_{i=1}^2 \tau_i q_i\left[(1+\tau_l) \frac{\partial \widetilde{x}_i^j}{\partial p_l}\right]\right\}.
\end{aligned}
$$

Thus, we have

$$
\begin{aligned}
& \Omega\left\{1+\sum_{j=1}^2 \frac{\partial q_l}{\partial B^j}\left[\pi^j \bar{l}-(1+\tau_l) l^j\right]\right\} \\
= \; & \sum_{j=1}^2\left\{\sum_{k \neq j} \lambda^{j,k} \frac{\partial V^{j,k}}{\partial B^k}\left[\left(\pi^k \bar{l}-(1+\tau_l) l^k\right)-\left(\pi^j \bar{l}-(1+\tau_l) l^{j,k}\right)\right]\right\} \\
& +\mu \sum_{j=1}^2\left\{\zeta^j \sum_{i=1}^2 \tau_i q_i\left[(1+\tau_l) \frac{\partial \widetilde{x}_i^j}{\partial p_l}\right]\right\},
\end{aligned}
$$

or equivalently, exploiting the fact that $\frac{\partial \widetilde{x}_i^j}{\partial q_l} = \frac{\partial \widetilde{x}_i^j}{\partial p_l}(1+\tau_l)$ and $\frac{\partial \widetilde{z}^j}{\partial q_l} = \frac{\partial \widetilde{z}^j}{\partial p_l}(1+\tau_l)$:

$$
\begin{aligned}
& \Omega\left\{1+\sum_{j=1}^{2}\frac{\partial q_l}{\partial B^j}\left[\pi^j\bar{l}-(1+\tau_l)l^j\right]\right\} \\
= & \sum_{j=1}^{2}\left\{\sum_{k\neq j}\lambda^{j,k}\frac{\partial V^{j,k}}{\partial B^k}\left[\left(\pi^k\bar{l}-(1+\tau_l)l^k\right)-\left(\pi^j\bar{l}-(1+\tau_l)l^{j,k}\right)\right]\right\} \\
& +\mu\sum_{j=1}^{2}\left(\zeta^j\sum_{i=1}^{2}\tau_i q_i\frac{\partial \widetilde{x}_i^j}{\partial q_l}\right).
\end{aligned}
$$

Define Ψ as $\Psi \equiv \left\{1+\sum_{j=1}^{2}\frac{\partial q_l}{\partial B^j}\left[\pi^j\bar{l}-(1+\tau_l)l^j\right]\right\}$. We have

$$
\begin{aligned}
\Omega = & \sum_{j=1}^{2}\left\{\sum_{k\neq j}\lambda^{j,k}\frac{\partial V^{j,k}}{\partial B^k}\left[\left(\pi^k\bar{l}-(1+\tau_l)l^k\right)-\left(\pi^j\bar{l}-(1+\tau_l)l^{j,k}\right)\right]\right\}/\Psi \\
& +\frac{\mu}{\Psi}\sum_{j=1}^{2}\left(\zeta^j\sum_{i=1}^{2}\tau_i q_i\frac{\partial \widetilde{x}_i^j}{\partial q_l}\right), \qquad \text{(B15)}
\end{aligned}
$$

where:

$$
\begin{aligned}
\Psi = & 1+\sum_{j=1}^{2}\frac{\partial q_l}{\partial B^j}\left[\pi^j\bar{l}-(1+\tau_l)l^j\right] \\
= & 1-\frac{\left\{\left[\pi^1\bar{l}-(1+\tau_l)l_u^1\right]\zeta^1\partial l^1/\partial B^1\right\}+\left\{\left[\pi^2\bar{l}-(1+\tau_l)l^2\right]\zeta^2\partial l^2/\partial B^2\right\}}{\sum_{j=1}^{2}\left[(1+\tau_l)\,\partial l^j/\partial p_l+\pi^j\bar{l}\partial l^j/\partial B^j\right]\zeta^j} \\
= & \frac{\sum_{j=1}^{2}\left[(1+\tau_l)\,\partial l^j/\partial p_l+\pi^j\bar{l}\partial l^j/\partial B^j\right]\zeta^j-\sum_{j=1}^{2}\left[\pi^j\bar{l}-(1+\tau_l)l^j\right]\zeta^j\partial l^j/\partial B^j}{\sum_{j=1}^{2}\left[(1+\tau_l)\,\partial l^j/\partial p_l+\pi^j\bar{l}\partial l^j/\partial B^j\right]\zeta^j} \\
= & \frac{(1+\tau_l)\sum_{j=1}^{2}\left(\partial l^j/\partial p_l+l^j\partial l^j/\partial B^j\right)\zeta^j}{\sum_{j=1}^{2}\left[(1+\tau_l)\,\partial l^j/\partial p_l+\pi^j\bar{l}\partial l^j/\partial B^j\right]\zeta^j} \\
= & \frac{(1+\tau_l)\sum_{j=1}^{2}\zeta^j\partial\widetilde{l}^j/\partial p_l}{\sum_{j=1}^{2}\left[(1+\tau_l)\,\partial l^j/\partial p_l+\pi^j\bar{l}\partial l^j/\partial B^j\right]\zeta^j} \\
= & \frac{\sum_{j=1}^{2}\zeta^j\partial\widetilde{l}^j/\partial q_l}{\sum_{j=1}^{2}\left[(1+\tau_l)\,\partial l^j/\partial p_l+\pi^j\bar{l}\partial l^j/\partial B^j\right]\zeta^j}.
\end{aligned}
$$

Going back to the optimality condition for τ_l, i.e.:

$$
\begin{aligned}
& \left(\frac{\partial q_l}{\partial \tau_l}+q_l\sum_{j=1}^{2}l^j\frac{\partial q_l}{\partial B^j}\right)\Omega-\phi+\mu\left[\sum_{j=1}^{2}\left(\zeta^j\sum_{i=1}^{2}\tau_i q_i\frac{\partial \widetilde{x}_i^j}{\partial \tau_l}\right)\right] \\
= & -q_l\sum_{j=1}^{2}\left[\sum_{k\neq j}\lambda^{j,k}\frac{\partial V^{j,k}}{\partial B^k}\left(l^{j,k}-l^k\right)\right],
\end{aligned}
$$

we can substitute the expressions for $\frac{\partial q_l}{\partial B^1}$, $\frac{\partial q_l}{\partial B^2}$ and $\frac{\partial q_l}{\partial \tau_l}$ (provided by (B6), (B4) and (B2)) and obtain

$$\frac{-\Omega q_l\left(l^1\zeta^1\partial l^1/\partial B^1+l^2\zeta^2\partial l^2/\partial B^2\right)}{\sum_{j=1}^2\left[(1+\tau_l)\,\partial l^j/\partial p_l+\pi^j\bar{l}\partial l^j/\partial B^j\right]\zeta^j}$$
$$-\Omega\frac{q_l\sum_{j=1}^2\zeta^j\partial l^j/\partial p_l}{\sum_{j=1}^2\left[(1+\tau_l)\,\partial l^j/\partial p_l+\pi^j\bar{l}\partial l^j/\partial B^j\right]\zeta^j}+\mu\left[\sum_{j=1}^2\left(\zeta^j\sum_{i=1}^2\tau_i q_i\frac{\partial\tilde{x}_i^j}{\partial\tau_l}\right)\right]-\phi$$
$$=-q_l\sum_{j=1}^2\left[\sum_{k\neq j}\lambda^{j,k}\frac{\partial V^{j,k}}{\partial B^k}\left(l^{j,k}-l^k\right)\right],$$

or, equivalently:

$$\Omega\left[-\frac{q_l\sum_{j=1}^2\zeta^j\partial\tilde{l}^j/\partial p_l}{\sum_{j=1}^2\left[(1+\tau_l)\,\partial l^j/\partial p_l+\pi^j\bar{l}\partial l^j/\partial B^j\right]\zeta^j}\right]+\mu\left[\sum_{j=1}^2\left(\zeta^j\sum_{i=1}^2\tau_i q_i\frac{\partial\tilde{x}_i^j}{\partial\tau_l}\right)\right]-\phi$$
$$=-q_l\sum_{j=1}^2\left[\sum_{k\neq j}\lambda^{j,k}\frac{\partial V^{j,k}}{\partial B^k}\left(l^{j,k}-l^k\right)\right],$$

or, equivalently:

$$-\Psi\Omega\frac{q_l}{1+\tau_l}+\mu\left[\sum_{j=1}^2\left(\zeta^j\sum_{i=1}^2\tau_i q_i\frac{\partial\tilde{x}_i^j}{\partial\tau_l}\right)\right]-\phi=-q_l\sum_{j=1}^2\left[\sum_{k\neq j}\lambda^{j,k}\frac{\partial V^{j,k}}{\partial B^k}\left(l^{j,k}-l^k\right)\right],$$

or equivalently:

$$-\Psi\Omega\frac{q_l}{1+\tau_l}+\mu q_l\left[\sum_{j=1}^2\left(\zeta^j\sum_{i=1}^2\tau_i q_i\frac{\partial\tilde{x}_i^j}{\partial p_l}\right)\right]-\phi=-q_l\sum_{j=1}^2\left[\sum_{k\neq j}\lambda^{j,k}\frac{\partial V^{j,k}}{\partial B^k}\left(l^{j,k}-l^k\right)\right].$$

Finally, substituting the expression for Ω that we have derived above, we obtain:

$$-\frac{q_l}{1+\tau_l}\sum_{j=1}^2\left\{\sum_{k\neq j}\lambda^{j,k}\frac{\partial V^{j,k}}{\partial B^k}\left[\left(\pi^k\bar{l}-(1+\tau_l)l^k\right)-\left(\pi^j\bar{l}-(1+\tau_l)l^{j,k}\right)\right]\right\}$$
$$-\mu\frac{q_l}{1+\tau_l}\sum_{j=1}^2\left(\zeta^j\sum_{i=1}^2\tau_i q_i\frac{\partial\tilde{x}_i^j}{\partial q_l}\right)+\mu q_l\left[\sum_{j=1}^2\left(\zeta^j\sum_{i=1}^2\tau_i q_i\frac{\partial\tilde{x}_i^j}{\partial p_l}\right)\right]-\phi$$
$$=-q_l\sum_{j=1}^2\left[\sum_{k\neq j}\lambda^{j,k}\frac{\partial V^{j,k}}{\partial B^k}\left(l^{j,k}-l^k\right)\right].$$

Simplifying and rearranging terms gives:

$$\frac{q_l}{1+\tau_l}\sum_{j=1}^2\left[\sum_{k\neq j}\lambda^{j,k}\frac{\partial V^{j,k}}{\partial B^k}\left(\pi^j-\pi^k\right)\right]\bar{l}-\phi=0.$$

B.3 Proof of Proposition 5

Optimal commodity tax formula for τ_s Applying Roy's identity we can rewrite the first order condition for $\tau_2 = \tau_s$ as follows:

$$-\sum_{j=1}^{2}\left(\alpha^j\zeta^j + \sum_{k\neq j}\lambda^{j,k}\right) q_2 s^j \frac{\partial V^j}{\partial B^j} + \sum_{j=1}^{2}\left(\sum_{k\neq j}\lambda^{j,k}\frac{\partial V^{j,k}}{\partial B^k}s^{j,k}\right) q_2$$
$$+\mu\sum_{j=1}^{2}\left[\zeta^j\left(\sum_{i=1}^{2}\tau_i q_i \frac{\partial x_i^j}{\partial \tau_2} + q_2 x_2^j\right)\right] + \Omega\frac{\partial q_l}{\partial \tau_2} = 0.$$

Using the Slutsky decomposition this can be written as

$$-\sum_{j=1}^{2}\left(\alpha^j\zeta^j + \sum_{k\neq j}\lambda^{j,k}\right) q_2 s^j \frac{\partial V^j}{\partial B^j} + \sum_{j=1}^{2}\left(\sum_{k\neq j}\lambda^{j,k}\frac{\partial V^{j,k}}{\partial B^k}s^{j,k}\right) q_2$$
$$+\mu q_2\sum_{j=1}^{2}\left[\zeta^j\left(\sum_{i=1}^{2}\tau_i q_i \left(\frac{\partial \widetilde{x}_i^j}{\partial p_2} - x_2^j\frac{\partial x_i^j}{\partial B^j}\right) + x_2^j\right)\right] + \Omega\frac{\partial q_l}{\partial \tau_2} = 0.$$

Using the first order conditions for B^1 and B^2 we can rewrite the previous equation as follows:

$$-q_2 s^1\lambda^{2,1}\frac{\partial V^{2,1}}{\partial B^1} - \mu\zeta^1\left[1 - \sum_{i=1}^{2}\tau_i q_i\frac{\partial x_i^1}{\partial B^1}\right] q_2 s^1 + \Omega\frac{\partial q_l}{\partial B^1}q_2 s^1$$
$$-q_2 s^2\lambda^{1,2}\frac{\partial V^{1,2}}{\partial B^2} - \mu\zeta^2\left[1 - \sum_{i=1}^{2}\tau_i q_i\frac{\partial x_i^2}{\partial B^2}\right] q_2 s^2 + \Omega\frac{\partial q_l}{\partial B^2}q_2 s^2$$
$$+\sum_{j=1}^{2}\left(\sum_{k\neq j}\lambda^{j,k}\frac{\partial V^{j,k}}{\partial B^k}s^{j,k}\right) q_2$$
$$+\mu q_2\sum_{j=1}^{2}\left[\zeta^j\left(\sum_{i=1}^{2}\tau_i q_i \left(\frac{\partial \widetilde{x}_i^j}{\partial p_2} - x_2^j\frac{\partial x_i^j}{\partial B^j}\right) + x_2^j\right)\right] + \Omega\frac{\partial q_l}{\partial \tau_2} = 0.$$

Simplifying terms we can rewrite the equation above as

$$\sum_{j=1}^{2}\left[\zeta^j\left(\sum_{i=1}^{2}\tau_i q_i\frac{\partial \widetilde{x}_i^j}{\partial p_2}\right)\right] = \sum_{j=1}^{2}\left[\sum_{k\neq j}\frac{\lambda^{j,k}}{\mu}\frac{\partial V^{j,k}}{\partial B^k}\left(s^k - s^{j,k}\right)\right] - \frac{\Omega}{\mu}\left(\frac{1}{q_2}\frac{\partial q_l}{\partial \tau_2} + \frac{\partial q_l}{\partial B^1}s^1 + \frac{\partial q_l}{\partial B^2}s^2\right). \tag{B16}$$

Since we have that

$$\begin{aligned}
\frac{\Omega}{\mu} &= \sum_{j=1}^{2}\left\{\sum_{k\neq j}\lambda^{j,k}\frac{\partial V^{j,k}}{\partial B^k}\left[\pi^k\bar{l}-(1+\tau_l)l^k\right]\right\}\frac{\sum_{j=1}^{2}\left[(1+\tau_l)\,\partial l^j/\partial p_l+\pi^j\bar{l}\partial l^j/\partial B^j\right]\zeta^j}{\mu\sum_{j=1}^{2}\zeta^j\partial\widetilde{l}^j/\partial q_l} \\
&\quad -\sum_{j=1}^{2}\left\{\sum_{k\neq j}\lambda^{j,k}\frac{\partial V^{j,k}}{\partial B^k}\left[\pi^j\bar{l}-(1+\tau_l)l^{j,k}\right]\right\}\frac{\sum_{j=1}^{2}\left[(1+\tau_l)\,\partial l^j/\partial p_l+\pi^j\bar{l}\partial l^j/\partial B^j\right]\zeta^j}{\mu\sum_{j=1}^{2}\zeta^j\partial\widetilde{l}^j/\partial q_l} \\
&\quad +\frac{\sum_{j=1}^{2}\left[(1+\tau_l)\,\partial l^j/\partial p_l+\pi^j\bar{l}\partial l^j/\partial B^j\right]\zeta^j}{\sum_{j=1}^{2}\zeta^j\partial\widetilde{l}^j/\partial q_l}\sum_{j=1}^{2}\left(\zeta^j\sum_{i=1}^{2}\tau_i q_i\frac{\partial\widetilde{x}_i^j}{\partial q_l}\right),
\end{aligned}$$

and

$$\begin{aligned}
\frac{1}{q_2}\frac{\partial q_l}{\partial\tau_2}+\frac{\partial q_l}{\partial B^1}s^1+\frac{\partial q_l}{\partial B^2}s^2 &= -\frac{\sum_{j=3}^{4}\zeta^j\partial l_u^j/\partial p_{su}}{\sum_{j=1}^{2}\left[(1+\tau_l)\,\partial l^j/\partial p_l+\pi^j\bar{l}\partial l^j/\partial B^j\right]\zeta^j} \\
&\quad -\frac{s^1\zeta^1\partial l^1/\partial B^1}{\sum_{j=1}^{2}\left[(1+\tau_l)\,\partial l^j/\partial p_l+\pi^j\bar{l}\partial l^j/\partial B^j\right]\zeta^j} \\
&\quad -\frac{s^2\zeta^2\partial l^2/\partial B^2}{\sum_{j=1}^{2}\left[(1+\tau_l)\,\partial l^j/\partial p_l+\pi^j\bar{l}\partial l^j/\partial B^j\right]\zeta^j} \\
&= -\frac{\sum_{j=1}^{2}\zeta^j\partial\widetilde{l}^j/\partial p_s}{\sum_{j=1}^{2}\left[(1+\tau_l)\,\partial l^j/\partial p_l+\pi^j\bar{l}\partial l^j/\partial B^j\right]\zeta^j},
\end{aligned}$$

we also have that

$$\begin{aligned}
&-\frac{\Omega}{\mu}\left(\frac{1}{q_2}\frac{\partial q_l}{\partial\tau_2}+\frac{\partial q_l}{\partial B^1}s^1+\frac{\partial q_l}{\partial B^2}s^2\right) \\
&= \frac{\sum_{j=1}^{2}\zeta^j\partial\widetilde{l}^j/\partial p_s}{\mu\sum_{j=1}^{2}\zeta^j\partial\widetilde{l}^j/\partial q_l}\sum_{j=1}^{2}\left\{\sum_{k\neq j}\lambda^{j,k}\frac{\partial V^{j,k}}{\partial B^k}\left[\left(\pi^k\bar{l}-(1+\tau_l)l^k\right)-\left(\pi^j\bar{l}-(1+\tau_l)l^{j,k}\right)\right]\right\} \\
&\quad +\frac{\sum_{j=1}^{2}\zeta^j\partial\widetilde{l}^j/\partial p_s}{\sum_{j=1}^{2}\zeta^j\partial\widetilde{l}^j/\partial q_l}\sum_{j=1}^{2}\left(\zeta^j\sum_{i=1}^{2}\tau_i q_i\frac{\partial\widetilde{x}_i^j}{\partial q_l}\right).
\end{aligned}$$

Defining Ξ_s as $\Xi_s\equiv\frac{\sum_{j=1}^{2}\zeta^j\partial\widetilde{l}^j/\partial p_s}{\sum_{j=1}^{2}\zeta^j\partial\widetilde{l}^j/\partial q_l}$ and substituting for $-\frac{\Omega}{\mu}\left(\frac{1}{q_2}\frac{\partial q_l}{\partial\tau_2}+\frac{\partial q_l}{\partial B^1}s^1+\frac{\partial q_l}{\partial B^2}s^2\right)$ in (B16) the RHS of the equation above gives:

$$\begin{aligned}
&\sum_{j=1}^{2}\left[\zeta^j\left(\sum_{i=1}^{2}\tau_i q_i\frac{\partial\widetilde{x}_i^j}{\partial p_2}\right)\right]-\Xi_s\sum_{j=1}^{2}\left(\zeta^j\sum_{i=1}^{2}\tau_i q_i\frac{\partial\widetilde{x}_i^j}{\partial q_l}\right) \\
&= \sum_{j=1}^{2}\sum_{k\neq j}\frac{\lambda^{j,k}}{\mu}\frac{\partial V^{j,k}}{\partial B^k}\left\{\left(s^k-s^{j,k}\right)+\Xi_s\left[\left(\pi^k\bar{l}-(1+\tau_l)l^k\right)-\left(\pi^j\bar{l}-(1+\tau_l)l^{j,k}\right)\right]\right\}
\end{aligned}$$

i.e.,

$$\sum_{j=1}^{2}\zeta^j\left[\sum_{i=1}^{2}\tau_i q_i\left(\frac{\partial\widetilde{x}_i^j}{\partial p_2}-\Xi_s\frac{\partial\widetilde{x}_i^j}{\partial q_l}\right)\right]$$

$$= \sum_{j=1}^{2}\left\{\sum_{k\neq j}\frac{\lambda^{j,k}}{\mu}\frac{\partial V^{j,k}}{\partial B^k}\left[s^k-s^{j,k}+\Xi_s\left(\left(\pi^k\bar{l}-(1+\tau_l)l^k\right)-\left(\pi^j\bar{l}-(1+\tau_l)l^{j,k}\right)\right)\right]\right\}.$$

Optimal commodity tax formula for τ_z Applying Roy's identity, we can rewrite the first order condition for $\tau_1=\tau_z$ as follows:

$$-\sum_{j=1}^{2}\left(\alpha^j\zeta^j+\sum_{k\neq j}\lambda^{j,k}\right)w^1z^j\frac{\partial V^j}{\partial B^j}+\sum_{j=1}^{2}\left(\sum_{k\neq j}\lambda^{j,k}\frac{\partial V^{j,k}}{\partial B^k}w^1z^{j,k}\right)$$
$$+\mu\sum_{j=1}^{2}\left[\zeta^j\left(\tau_2q_2\frac{\partial x_2^j}{\partial\tau_z}+\tau_zw^1\frac{\partial z^j}{\partial\tau_z}+w^1z^j\right)\right]+\Omega\frac{\partial q_l}{\partial\tau_z}=0.$$

Using the Slutsky decomposition this can be written as

$$-\sum_{j=1}^{2}\left(\alpha^j\zeta^j+\sum_{k\neq j}\lambda^{j,k}\right)w^1z^j\frac{\partial V^j}{\partial B^j}+\sum_{j=1}^{2}\left(\sum_{k\neq j}\lambda^{j,k}\frac{\partial V^{j,k}}{\partial B^k}w^1z^{j,k}\right)$$
$$+\mu\sum_{j=1}^{2}\left[\zeta^j\left(\tau_2q_2\left(\frac{\partial\widetilde{x}_2^j}{\partial p_1}-z^j\frac{\partial x_i^j}{\partial B^j}\right)w^1+\tau_z\left(w^1\right)^2\left(\frac{\partial\widetilde{z}^j}{\partial p_1}-z^j\frac{\partial z^j}{\partial B^j}\right)+w^1z^j\right)\right]+\Omega\frac{\partial q_l}{\partial\tau_z}=0.$$

Using the first order conditions for B^1 and B^2, we can rewrite the previous equation as follows:

$$-w^1z^1\lambda^{2,1}\frac{\partial V^{2,1}}{\partial B^1}-\mu\zeta^1\left[1-\sum_{i=1}^{2}\tau_iq_i\frac{\partial x_i^1}{\partial B^1}\right]w^1z^1+\Omega\frac{\partial q_l}{\partial B^1}w^1z^1$$
$$-w^1z^2\lambda^{1,2}\frac{\partial V^{1,2}}{\partial B^2}-\mu\zeta^2\left[1-\sum_{i=1}^{2}\tau_iq_i\frac{\partial x_i^2}{\partial B^2}\right]w^1z^2+\Omega\frac{\partial q_l}{\partial B^2}w^1z^2$$
$$+\sum_{j=1}^{2}\left(\sum_{k\neq j}\lambda^{j,k}\frac{\partial V^{j,k}}{\partial B^k}w^1z^{j,k}\right)$$
$$+\mu\sum_{j=1}^{2}\left[\zeta^j\left(\tau_2q_2\left(\frac{\partial\widetilde{x}_2^j}{\partial p_1}-z^j\frac{\partial x_i^j}{\partial B^j}\right)w^1+\tau_z\left(w^1\right)^2\left(\frac{\partial\widetilde{z}^j}{\partial p_1}-z^j\frac{\partial z^j}{\partial B^j}\right)+w^1z^j\right)\right]+\Omega\frac{\partial q_l}{\partial\tau_z}$$
$$= 0.$$

Simplifying terms we can rewrite the equation above as

$$\begin{aligned}
&\sum_{j=1}^{2}\left(\sum_{k\neq j}\lambda^{j,k}\frac{\partial V^{j,k}}{\partial B^{k}}w^{1}z^{j,k}\right)-w^{1}z^{1}\lambda^{2,1}\frac{\partial V^{2,1}}{\partial B^{1}}-w^{1}z^{2}\lambda^{1,2}\frac{\partial V^{1,2}}{\partial B^{2}}\\
&+\mu\sum_{j=1}^{2}\left[\zeta^{j}\left(\tau_{2}q_{1}\frac{\partial\widetilde{x}_{2}^{j}}{\partial p_{1}}w^{1}+\tau_{z}\left(w^{1}\right)^{2}\frac{\partial\widetilde{z}^{j}}{\partial p_{1}}\right)\right]\\
=\;&-\Omega\left(\frac{\partial q_{l}}{\partial\tau_{z}}+\frac{\partial q_{l}}{\partial B^{1}}w^{1}z^{1}+\frac{\partial q_{l}}{\partial B^{2}}w^{1}z^{2}\right),
\end{aligned}$$

or equivalently:

$$\begin{aligned}
&\sum_{j=1}^{2}\left[\zeta^{j}\left(\tau_{2}q_{2}\frac{\partial\widetilde{x}_{2}^{j}}{\partial p_{1}}w^{1}+\tau_{z}\left(w^{1}\right)^{2}\frac{\partial\widetilde{z}^{j}}{\partial p_{1}}\right)\right]\\
=\;&w^{1}\left[\frac{\lambda^{1,2}}{\mu}\frac{\partial V^{1,2}}{\partial B^{2}}\left(z^{2}-z^{1,2}\right)+\frac{\lambda^{2,1}}{\mu}\frac{\partial V^{2,1}}{\partial B^{1}}\left(z^{1}-z^{2,1}\right)\right]\\
&-\frac{\Omega}{\mu}\left(\frac{\partial q_{l}}{\partial\tau_{z}}+\frac{\partial q_{l}}{\partial B^{1}}w^{1}z^{1}+\frac{\partial q_{l}}{\partial B^{2}}w^{1}z^{2}\right).
\end{aligned}\tag{B17}$$

Since we have that

$$\begin{aligned}
\frac{\partial q_{l}}{\partial\tau_{z}}+\frac{\partial q_{l}}{\partial B^{1}}w^{1}z^{1}+\frac{\partial q_{l}}{\partial B^{2}}w^{1}z^{2}\;=\;&-\frac{w^{1}\sum_{j=1}^{2}\zeta^{j}\partial l^{j}/\partial p_{z}}{\sum_{j=1}^{2}\left[\left(1+\tau_{l}\right)\partial l^{j}/\partial p_{l}+\pi^{j}\bar{l}\partial l^{j}/\partial B^{j}\right]\zeta^{j}}\\
&-\frac{w^{1}z^{1}\zeta^{1}\partial l^{1}/\partial B^{1}}{\sum_{j=1}^{2}\left[\left(1+\tau_{l}\right)\partial l^{j}/\partial p_{l}+\pi^{j}\bar{l}\partial l^{j}/\partial B^{j}\right]\zeta^{j}}\\
&-\frac{w^{1}z^{2}\zeta^{2}\partial l^{2}/\partial B^{2}}{\sum_{j=1}^{2}\left[\left(1+\tau_{l}\right)\partial l^{j}/\partial p_{l}+\pi^{j}\bar{l}\partial l^{j}/\partial B^{j}\right]\zeta^{j}}\\
=\;&-\frac{w^{1}\sum_{j=1}^{2}\zeta^{j}\partial\widetilde{l}^{j}/\partial p_{z}}{\sum_{j=1}^{2}\left[\left(1+\tau_{l}\right)\partial l^{j}/\partial p_{l}+\pi^{j}\bar{l}\partial l^{j}/\partial B^{j}\right]\zeta^{j}},
\end{aligned}$$

we also have that

$$\begin{aligned}
&-\frac{\Omega}{\mu}\left(\frac{\partial q_{l}}{\partial\tau_{z}}+\frac{\partial q_{l}}{\partial B^{1}}w^{1}z^{1}+\frac{\partial q_{l}}{\partial B^{2}}w^{1}z^{2}\right)\\
=\;&\frac{w^{1}\sum_{j=1}^{2}\zeta^{j}\partial\widetilde{l}^{j}/\partial p_{z}}{\mu\sum_{j=1}^{2}\zeta^{j}\partial\widetilde{l}^{j}/\partial q_{l}}\sum_{j=1}^{2}\left\{\sum_{k\neq j}\lambda^{j,k}\frac{\partial V^{j,k}}{\partial B^{k}}\left[\left(\pi^{k}\bar{l}-(1+\tau_{l})l^{k}\right)-\left(\pi^{j}\bar{l}-(1+\tau_{l})l^{j,k}\right)\right]\right\}\\
&+\frac{w^{1}\sum_{j=1}^{2}\zeta^{j}\partial\widetilde{l}^{j}/\partial p_{z}}{\sum_{j=1}^{2}\zeta^{j}\partial\widetilde{l}^{j}/\partial q_{l}}\sum_{j=1}^{2}\left(\zeta^{j}\sum_{i=1}^{2}\tau_{i}q_{i}\frac{\partial\widetilde{x}_{i}^{j}}{\partial q_{l}}\right).
\end{aligned}$$

Defining Ξ_z as $\Xi_z \equiv \frac{\sum_{j=1}^{2}\zeta^j \partial\widetilde{l}^j/\partial p_z}{\sum_{j=1}^{2}\zeta^j \partial\widetilde{l}^j/\partial q_l}$ and substituting for $-\frac{\Omega}{\mu}\left(\frac{\partial q_l}{\partial \tau_z}+\frac{\partial q_l}{\partial B^1}w^1 z^1+\frac{\partial q_l}{\partial B^2}w^1 z^2\right)$ in (B17) the RHS of the equation above gives:

$$\begin{aligned}
& w^1\sum_{j=1}^{2}\left[\zeta^j\left(\tau_2 q_2\frac{\partial\widetilde{x}_2^j}{\partial p_1}+\tau_z w^1\frac{\partial\widetilde{z}^j}{\partial p_1}-\Xi_z\sum_{i=1}^{2}\tau_i q_i\frac{\partial\widetilde{x}_i^j}{\partial q_l}\right)\right] \\
= \; & w^1\left[\frac{\lambda^{1,2}}{\mu}\frac{\partial V^{1,2}}{\partial B^2}\left(z^2-z^{1,2}\right)+\frac{\lambda^{2,1}}{\mu}\frac{\partial V^{2,1}}{\partial B^1}\left(z^1-z^{2,1}\right)\right] \\
& +w^1\Xi_z\sum_{j=1}^{2}\left\{\sum_{k\neq j}\frac{\lambda^{j,k}}{\mu}\frac{\partial V^{j,k}}{\partial B^k}\left[\left(\pi^k-\pi^j\right)\bar{l}+\left(1+\tau_l\right)\left(l^{j,k}-l^k\right)\right]\right\},
\end{aligned}$$

i.e.,

$$\begin{aligned}
& \sum_{j=1}^{2}\left[\zeta^j\left(\tau_2 q_2\frac{\partial\widetilde{x}_2^j}{\partial \tau_z}+\tau_z w^1\frac{\partial\widetilde{z}^j}{\partial \tau_z}-\Xi_z\sum_{i=1}^{2}\tau_i q_i\frac{\partial\widetilde{x}_i^j}{\partial q_l}\right)\right] \\
= \; & w^1\left[\frac{\lambda^{1,2}}{\mu}\frac{\partial V^{1,2}}{\partial B^2}\left(z^2-z^{1,2}\right)+\frac{\lambda^{2,1}}{\mu}\frac{\partial V^{2,1}}{\partial B^1}\left(z^1-z^{2,1}\right)\right] \\
& +w^1\Xi_z\sum_{j=1}^{2}\left\{\sum_{k\neq j}\frac{\lambda^{j,k}}{\mu}\frac{\partial V^{j,k}}{\partial B^k}\left[\left(\pi^k-\pi^j\right)\bar{l}+\left(1+\tau_l\right)\left(l^{j,k}-l^k\right)\right]\right\},
\end{aligned}$$

i.e.,

$$\begin{aligned}
& \sum_{j=1}^{2}\left[\zeta^j\left(\tau_2 q_2\frac{\partial\widetilde{x}_2^j}{\partial \tau_z}+\tau_z w^1\frac{\partial\widetilde{z}^j}{\partial \tau_z}-\Xi_z\sum_{i=1}^{2}\tau_i q_i\frac{\partial\widetilde{x}_i^j}{\partial q_l}\right)\right] \\
= \; & w^1\sum_{j,k\in\{1,2\},j\neq k}\frac{\lambda^{j,k}}{\mu}\frac{\partial V^{j,k}}{\partial B^k}\left(z^k-z^{j,k}\right) \\
& +w^1\Xi_z\sum_{j=1}^{2}\left\{\sum_{k\neq j}\frac{\lambda^{j,k}}{\mu}\frac{\partial V^{j,k}}{\partial B^k}\left[\left(\pi^k-\pi^j\right)\bar{l}+\left(1+\tau_l\right)\left(l^{j,k}-l^k\right)\right]\right\},
\end{aligned}$$

i.e.,

$$\begin{aligned}
& \sum_{j=1}^{2}\zeta^j\left(\sum_{i=1}^{2}\tau_i q_i\frac{\partial\widetilde{x}_i^j}{\partial \tau_z}-\Xi_z\sum_{i=1}^{2}\tau_i q_i\frac{\partial\widetilde{x}_i^j}{\partial q_l}\right) \\
= \; & w^1\sum_{j,k\in\{1,2\},j\neq k}\frac{\lambda^{j,k}}{\mu}\frac{\partial V^{j,k}}{\partial B^k}\left(z^k-z^{j,k}\right) \\
& +w^1\Xi_z\sum_{j=1}^{2}\left\{\sum_{k\neq j}\frac{\lambda^{j,k}}{\mu}\frac{\partial V^{j,k}}{\partial B^k}\left[\left(\pi^k-\pi^j\right)\bar{l}+\left(1+\tau_l\right)\left(l^{j,k}-l^k\right)\right]\right\}.
\end{aligned}$$

B.4 Proof of Corollary 1

Welfare effects of τ_s Consider a compensated increase in τ_s. We must consider both the direct effect on p_s and the indirect effect on q_l. Since $\partial p_s/\partial \tau_s = q_s$, disposable income B^j should be varied by $dB^j = q_s s^j d\tau_s, j = 1,2$ to compensate non-mimicking agents for the welfare effects of a marginal change in τ_s. Instead, to compensate non-mimicking agents for the welfare effects of changing q_l by dq_l, B^j should be changed by

$$dB^j = \left[(1+\tau_l)\, l^j - \pi^j \bar{l}\right] dq_l, j = 1,2, \tag{B18}$$

where the first term in square brackets in (B18) captures the compensation required for the change in the tax-inclusive land price p_l, and the second term captures the compensation required for the change in the value of an agent's land endowment.

Given that initially $\tau_s = \tau_z = 0$, the revenue effect (given by the left hand side of (25)) of a compensated increase in τ_s is zero. With high-skilled holding a larger land endowment ($\pi^2 > \pi^1$), the relevant IC-constraint (at least for any inequality-averse planner) is the one requiring high-skilled agents not to be tempted to mimic low-skilled agents. For a high-skilled mimicker, the welfare effect of the reform is given by

$$dV^{2,1} = \left\{s^1 - s^{2,1} - \Xi_2 \left[\left(\pi^2 - \pi^1\right) \bar{l} + l^1 - l^{2,1}\right]\right\} q_s. \tag{B19}$$

Assume that $z^1 = 0$.[31] It then follows that $z^{2,1} = 0$; this is because $f'\left(\frac{Y^1}{w^1} + h_m\right) < f'\left(\frac{Y^1}{w^2} + h_m\right)$ for any given h_m (mimickers work less in the market and therefore the effort cost of raising h_m is for them smaller). This also means that, for any given amount of s, the marginal effective price of structures is lower for a high-skilled mimicker than for a low-skilled agent: whereas the unitary market price of structures is $p_s = (1+\tau_s)\, q_s$ for both, the cost of the additional maintenance services required by a marginal increase in s is larger for a low-skilled agent. This difference in the marginal effective price of structures, together with the fact that the total disposable income of a high-skilled mimicker is larger than that of a low-skilled agent (due to the assumption that $\pi^2 > \pi^1$, $D^{2,1} \equiv B^1 + q_l \pi^2 \bar{l} > D^1 \equiv B^1 + q_l \pi^1 \bar{l}$), implies that $s^{2,1} > s^1$.

Since equilibrium in the land market requires that $l^1 - \pi^1 \bar{l} = -\left(l^2 - \pi^2 \bar{l}\right) \zeta^2/\zeta^1$, we can

[31] This is necessarily the case if $\tau_z = 0$ and the marginal income tax rate is positive at Y^1. In fact, from the maintenance constraint $\gamma s = w_z z + \omega h_m$, we have that $dz/ds = \gamma/w_z$ if the additional maintenance services required by a marginal increase in structure are purchased on the market, and $dh_m/ds = \gamma/\omega$ if they are performed in-house. Therefore, the individual first order condition for s is given by $\frac{\partial g}{\partial s} = (1+\tau_s)\, q_s v' + \gamma \min\left\{\frac{(1+\tau_z) w_z}{w_z} v', -\frac{f'}{\omega}\right\}$. Exploiting (20), this first order condition can be rewritten as $\frac{\partial g}{\partial s} = (1+\tau_s)\, q_s v' + \gamma \min\left\{(1+\tau_z)\, v', \frac{w(1-T')}{\omega} v'\right\}$. Thus, in-house maintenance is preferred when $\frac{w(1-T'(Y))}{\omega(1+\tau_z)} < 1$.

rewrite the bracketed expression in (B19) as

$$\begin{aligned}\left(\pi^2-\pi^1\right)\bar{l}+l^1-l^{2,1} &= -\left(l^2-\pi^2\bar{l}\right)\frac{\zeta^2}{\zeta^1}-\left(l^{2,1}-\pi^2\bar{l}\right)\\ &= -\left(l^2-\pi^2\bar{l}\right)\left(\frac{\zeta^2}{\zeta^1}+1\right)+\left(l^2-\pi^2\bar{l}\right)-\left(l^{2,1}-\pi^2\bar{l}\right)\\ &= -\frac{l^2-\pi^2\bar{l}}{\zeta^1}+l^2-l^{2,1}.\end{aligned}$$

Using the assumptions $\pi^2 > \pi^1$ and $l^2 - \pi^2\bar{l} < 0$, and since $l^2 > l^{2,1}$ (given that the after-tax income of a highly skilled agent is lower when he behaves as a mimicker), it follows that $\left(\pi^2-\pi^1\right)\bar{l}+l^1-l^{2,1} > 0$. This, together with the fact that $\Xi_2 > 0$ (due to the assumption that structures and land are Hicksian complements) and $s^1 < s^{2,1}$, implies that from (B19) we get that $dV^{2,1} < 0$. By relaxing the binding IC-constraint, the reform allows achieving a higher social welfare.

Welfare effects of τ_z In the case of a compensated increase in τ_z, the revenue effect given by the left-hand side of (26), is zero, while the effect on the utility of a high-skilled mimicker is given by

$$dV^{2,1} = \left\{z^1 - z^{2,1} - \Xi_1\left[\left(\pi^2-\pi^1\right)\bar{l}+l^1-l^{2,1}\right]\right\}w^1. \tag{B20}$$

As noted above, the assumption that $z^1 = 0$ implies that $z^{2,1} = 0$. For high-skilled agents, given the assumption that $z^2 > 0$, the increase in τ_z is equivalent to an increase in the effective price of structures.[32] This is because, for someone purchasing maintenance services in the market, the marginal effective price of s is $p_s + \gamma\left(1+\tau_z\right)w^1/w^1 = p_s + \gamma\left(1+\tau_z\right)$, given that $dz/ds = \gamma/w^1$. Thus, given the assumption that structures and land are Hicksian complements, $\Xi_1 > 0$. This, together with the fact that $\left(\pi^2-\pi^1\right)\bar{l}+l^1-l^{2,1} > 0$ (see the discussion above), implies that $dV^{2,1} < 0$.

B.5 Proof of Proposition 6

Based on the implicit characterization $T'\left(Y^j\right) = 1 - MRS^j_{YB} = 1 + \frac{\partial V^j}{\partial Y^j}/\frac{\partial V^j}{\partial B^j}$, and combining the first order conditions (B11) and (B12), we obtain

$$\begin{aligned}&\frac{\frac{\partial V^1}{\partial Y^1}}{\frac{\partial V^1}{\partial B^1}}\left\{\lambda^{2,1}\frac{\partial V^{2,1}}{\partial B^1}+\mu\zeta^1\left[1-\sum_{i=1}^{2}\tau_i q_i\frac{\partial x_i^1}{\partial B^1}\right]-\Omega\frac{\partial q_l}{\partial B^1}\right\}\\ =\;& \lambda^{2,1}\frac{\partial V^{2,1}}{\partial Y^1}-\mu\zeta^1\left[1+\sum_{i=1}^{2}\tau_i q_i\frac{\partial x_i^1}{\partial Y^1}\right]-\Omega\frac{\partial q_l}{\partial Y^1},\end{aligned}$$

[32] In general, we have that $z^2 > 0$ when the inequality $\left[1-T'\left(Y^2\right)\right]w^2 > \left(1+\tau_z\right)w^1$ is satisfied.

i.e.,

$$1+\frac{\frac{\partial V^1}{\partial Y^1}}{\frac{\partial V^1}{\partial B^1}} = \frac{\lambda^{2,1}}{\mu\zeta^1}\frac{\partial V^{2,1}}{\partial B^1}\left(\frac{\frac{\partial V^{2,1}}{\partial Y^1}}{\frac{\partial V^{2,1}}{\partial B^1}}-\frac{\frac{\partial V^1}{\partial Y^1}}{\frac{\partial V^1}{\partial B^1}}\right)-\sum_{i=1}^{2}\tau_i q_i\left(\frac{\partial x_i^1}{\partial Y^1}-\frac{\frac{\partial V^1}{\partial Y^1}}{\frac{\partial V^1}{\partial B^1}}\frac{\partial x_i^1}{\partial B^1}\right)$$
$$-\frac{\Omega}{\mu\zeta^1}\left(\frac{\partial q_l}{\partial Y^1}-\frac{\frac{\partial V^1}{\partial Y^1}}{\frac{\partial V^1}{\partial B^1}}\frac{\partial q_l}{\partial B^1}\right),$$

and therefore

$$T'\left(Y^1\right) = \frac{\lambda^{2,1}}{\mu\zeta^1}\frac{\partial V^{2,1}}{\partial B^1}\left(MRS^1_{Y,B}-MRS^{2,1}_{Y,B}\right)$$
$$-\sum_{i=1}^{2}\tau_i q_i\left(\frac{\partial x_i^1}{\partial Y^1}+MRS^1_{Y,B}\frac{\partial x_i^1}{\partial B^1}\right)-\frac{\Omega}{\mu\zeta^1}\left(\frac{\partial q_l}{\partial Y^1}+MRS^1_{Y,B}\frac{\partial q_l}{\partial B^1}\right).$$

Combining (B13) and (B14) and proceeding in a similar way, we obtain (28).