Prices, Non-homotheticities, and Optimal Taxation*

Xavier Jaravel
LSE & CEPR
x.jaravel@lse.ac.uk

Alan Olivi
UCL
a.olivi@ucl.ac.uk

August 12, 2025

Abstract

We characterize theoretically and quantitatively the effects of price changes on optimal tax design in the presence of non-homothetic preferences. We find that, rather than offsetting price changes, the optimal tax system *amplifies* their redistributive effects. With a quantitative model matching observed non-homotheticities and the empirical elasticity of prices to market size in the United States, we find that due to the amplification channel, (i) the optimal tax schedule is more redistributive when accounting for non-homothetic spending patterns, (ii) observed heterogeneous inflation rates, which are lower for luxuries relative to necessities in the United States, generate a regressive tax response.

Keywords: Optimal Taxation, Returns to Scale, Non-homotheticities, Inequality.
JEL: H21, H23, O31

*For thoughtful comments, we thank Pierre Boyer, Raj Chetty, Antoine Ferey, Nathan Hendren, Camille Landais, Abdoulaye Ndiaye, Emmanuel Saez, Florian Scheuer, Stefanie Stantcheva, Aleh Tsyvinski, Matthew Weinzierl, Nicolas Werquin, Iván Werning, and participants at Aalto University, the Collège de France, MIT, and the NBER Summer Institute. For excellent research assistance, we thank Filip Klein, Gabriel Leite-Mariante, and Bas Sanders. We are grateful to Nathan Hendren for sharing some of the data used in the quantitative analysis. Researcher(s) own analyses calculated (or derived) based in part on data from Nielsen Consumer LLC and marketing databases provided through the NielsenIQ Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the NielsenIQ data are those of the researcher(s) and do not reflect the views of NielsenIQ. NielsenIQ is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

1 Introduction

How should tax policy respond when prices change? While price variation is pervasive, the optimal taxation literature has largely treated prices as fixed. Indeed, the seminal result of Diamond and Mirrlees (1971b) states that optimal tax formulas can be derived as if prices were fixed at their equilibrium level, and leaves implicit the optimal response of taxes to price changes.[1] Yet, empirically price changes are ubiquitous – and importantly, they tend to correlate with household income. Recent work shows that heterogeneous inflation rates across products consumed by low- and high-income households played an important role for purchasing power inequality in the United States (e.g., McGranahan and Paulson (2005), Kaplan and Schulhofer-Wohl (2017), Jaravel (2019), Argente and Lee (2020), Klick and Stockburger (2021), Jaravel and Lashkari (2023), Jaravel (2024)). Despite the prevalence of price changes, to date we lack the tools to characterize their potential effects on optimal taxation.

In this paper, we develop a theoretical framework to analyze the effect of prices on optimal tax design, and we quantitatively estimate their impact. We provide an explicit characterization of the impact of prices on the marginal social value of transfers, on labor supply, and on labor supply elasticities. Furthermore, we show that when the response of the tax schedule to price changes is non-trivial, a feedback loop may emerge: taxes shift demand for goods, which can induce a further change in prices through general equilibrium adjustments (e.g., returns to scale), and a new response of the tax schedule. Equilibrium prices are still a sufficient statistic in our optimal tax formulas, but finding the new equilibrium prices (in response to exogenous shocks) requires characterizing this feedback loop and, in particular, the response of the supply side of the economy.

To facilitate comparison with the prior literature, we work with a standard, static Mirrlees model: agents have preferences over multiple consumption goods and leisure, and labor is the only factor of production; preferences are weakly separable between consumption and labor, as in the Atkinson-Stiglitz benchmark. This setting allows us to capture non-homothetic spending patterns across the income distribution while focusing on a single tax instrument for redistribution, the nonlinear income tax.

The main challenge is that the channels through which prices shape redistribution are not explicit, as they appear only implicitly in the first-order conditions determining the optimal marginal tax rates.[2] Using a comparative static approach, we characterize the first-order responses of taxes to price changes in terms of observable statistics.

To isolate the role of non-homotheticities, we start by analyzing the case of linear production functions: prices are fully exogenous and do not respond to shifts in aggregate demand for goods. We identify two key channels through which prices affect the income tax schedule. We show that the impact of prices on taxes is governed by the marginal propensity to spend on the products experiencing a price change. To illustrate, consider an increase in the price of a product for which the marginal propensity to spend decreases with income, which we refer to as a "necessity product". First, a price increase on a necessity good raises the marginal price index of lower-income households relatively more than that of higher-income households, i.e. lower income households can now buy less with an additional dollar of income. Therefore, the social value of a dollar transfer from higher-income to lower-income households decreases (Channel #1). Second,

[1] Standard optimal tax formulas are first-order conditions featuring endogenous variables that depend on prices, such as the marginal utility of disposable income.

[2] Another difficulty in our case is that non-homothetic demand systems do not yield closed form solutions in general.

as the decrease in marginal purchasing power is larger at lower income levels, the price increase generates a positive income effect,[3] which is higher as income increases: a dollar transfer disincentivizes labor supply relatively more at higher income levels, and thus a price increase on a necessity good increases the efficiency cost of taxation (Channel #2). Since both the cost of taxation and the social value of transfers to higher-income increases, the marginal tax rates decreases everywhere, and redistribution to the rich increases.

Perhaps surprisingly, we thus find that, far from compensating price movements, the optimal tax system amplifies their redistributive effects: an increase in the price of necessity goods induces more redistribution at the top of the income distribution; the opposite is true when luxuries become more expensive. These channels do not operate when preferences are homothetic as all agents are equally impacted by price changes.[4] While we highlight these channels under linear production functions, they do not depend on any particular supply-side model.

Next, we consider non-linear production functions: prices become endogenous and also adjust through general equilibrium effects. With non-linear production functions, the elasticity of prices with respect to aggregate consumer demand across products becomes pivotal to evaluate the interplay between optimal taxes and prices. Using a sufficient statistics specification, we capture both the canonical Diamond-Mirrlees setting – with perfect competition and potentially decreasing returns to scale – and a wide class of free entry models allowing for increasing returns to scale through firm selection (Melitz (2003)), variable markups (Feenstra and Weinstein (2017)), and innovation (Bustos (2011)). We show theoretically that, when product prices decrease as their market expands,[5] the redistributive effects of price changes and their amplification through taxes are strengthened in general equilibrium. This amplification occurs through both substitution and income effects. For instance, when the relative price of necessities increases, consumers substitute away from them, which leads to further increases in their relative price through increasing returns. Moreover, the increase in the relative price of necessities leads to a fall in the real income of lower-income households, who consume relatively more necessities, implying a further decline in their relative demand and a further increase in their relative price. These amplification channels operate in any supply side model with elastic prices, although this effect remains implicit in the standard Diamond-Mirrlees tax formulas.[6]

Building on these theoretical insights, in the quantitative section of the paper we evaluate the optimal response of taxes to the price changes observed in the data in recent years, and we examine more generally how our benchmark specification – with non homothetic preferences and downward sloping supply curves – affects optimal redistribution policies. We first implement our comparative static approach. While it only gives the first order response of taxes to price changes, it has two advantages: we can directly use re-

[3]Denoting $\tilde{\eta}$ the income effect on labor supply, we have for an increase in the price of necessity p_l, $\partial_{p_l}\tilde{\eta} > 0$ which decreases the efficiency cost of taxation.

[4]In the main text, we consider a benchmark case where there are no income effects on labor supply (at initial prices). In Online Appendix E, we extend these results to more general labor supply functions. Our results are qualitatively similar in these more general cases.

[5]Prior work shows that product markets with larger demand tend to have higher productivity and lower prices due to several channels. For instance, higher demand increases the incentives to enter a market, to innovate, and to compete, which leads to lower marginal cost, lower markups, larger product variety, and lower consumer price indices. These channels have been analyzed in a recent empirical literature (e.g., Costinot et al. (2019), Jaravel (2019), Faber and Fally (2021)) as well as in a long-standing theoretical literature (e.g., Romer (1990a), Aghion and Howitt (1992a), Melitz (2003)).

[6]When prices increase as the market expands (in contrast with our baseline case featuring increasing returns, in line with empirical evidence), the redistributive effects of prices are muted through general equilibrium effects.

cent causal estimates of the elasticity of prices to market size to evaluate supply side responses to shifts in demand, and we can non-parametrically fit non-homothetic spending patterns. By linking the Consumer Expenditure Survey (CEX) and the Consumer Price Index (CPI) data sets, we obtain observed price changes and households' spending across 248 product categories for the period 2004 to 2015, covering the entire consumption basket of American households. Empirically, inflation was lower in product categories with higher income elasticities. We find that, in response, it is optimal to reduce redistribution and set lower marginal tax rates, with a fall in marginal tax rates of about 8 percentage points at the bottom of the income distribution (relative to the observed tax schedule).[7]

Next, we make parametric assumptions on non-homotheticities, using non-homothetic CES (nhCES) preferences as in Hanoch (1975), Matsuyama (2019), and Comin et al. (2021). We then study the quantitative importance of increasing returns to scale, non-homotheticities and price shocks for optimal tax rates and welfare across the skill distribution. By introducing parametric assumptions on preferences, these analyses are complementary with the analysis of first-order approximations, because they characterize how our new channels affect the optimum when accounting for potential non-linearities. They also allow us to characterize the quantitative importance of non-homotheticities for the optimal tax schedule.

Relative to the optimal tax schedule with homothetic preferences, we find that non-homotheticities imply more redistribution. Relative to the optimum under homothetic preferences, marginal taxes increase over the full range of the income distribution. The increase is more pronounced at the bottom of the income distribution, with an increase in marginal tax rates of about 6pp for levels of earned income below \$20,000. The increase is about 2pp at an income level of \$100,000, and then gradually decreases, reaching levels close to zero above \$300,000. Thus, the simulations show that non-homotheticities have a significant quantitative impact on optimal marginal tax rates. To document whether these tax changes and their induced price effects have meaningful distributional effects, we compute the willingness to pay of agents for the optimal tax schedule under non-homothetic preferences, relative to the optimal schedule under homothetic preferences. The equivalent variation is sizable, ranging from about 15% in the bottom decile of the income distribution to -9% in the top decile.

We show that this increase in redistribution can be explained by the change in equilibrium prices and in the marginal utility of redistribution across the skill distribution. As the relative price of the necessity bundle decreases, it is optimal to redistribute more to those with a higher marginal propensity to consume on necessities, which induces further tax changes and changes in labor supply, etc. The strength of these feedback loops depends on the parameters governing increasing returns and social preferences for redistribution, and we find them to be large in our calibration. We also document the robustness of our results to alternative parameter values.

Related literature. The main contribution of this paper is to provide a theoretical and quantitative characterization of the impact of prices on optimal tax design. We thus relate to several strands of literature. First, in prior work the effect of prices on the tax schedule has remained implicit, as standard

[7] Since empirical studies stress the importance of using granular data to properly measure inflation heterogeneity, we also estimate the impact of price changes in the subset of goods covered by the NielsenIQ scanner data. We find that the sensitivity of the tax rate to change in the prices is larger when we consider granular products rather than goods aggregated at a level comparable to the CEX. Our baseline results using the CEX-CPI data are therefore likely to underestimate the impact of price changes on optimal redistribution.

tax formulas depend on endogenous variables that depend on prices, such as the marginal utility of disposable income. Several papers have highlighted the implications of specific assumptions on consumers' preferences for tax design, including preference heterogeneity (e.g., Saez (2002), Diamond and Spinnewijn (2011)) and consumers' myopia (e.g., Allcott et al. (2019)). Instead, we show theoretically and quantitatively that prices play an important role even in the canonical setting where the utility function is separable between labor and all commodities, i.e. no indirect taxes need to be used, as in Atkinson and Stiglitz (1976).[8] We explicitly characterize the impact of prices on the tax schedule, both in partial equilibrium and general equilibrium, providing decompositions isolating the economic forces at play. Second, our results contribute to a growing strand of the optimal taxation literature that has isolated the general equilibrium effects of taxes, focusing on wages (e.g., Rothschild and Scheuer (2013), Sachs et al. (2020)); we complement these analyses by characterizing the general equilibrium impact on prices in the presence of non-homotheticities. Third, although imperfect competition is not our focus, our work relates to a growing literature on optimal taxation in the presence of imperfect competition, in which endogenous prices or wages play a role for redistribution from firm owners toward workers (e.g., Boar and Midrigan (2019), Eeckhout et al. (2021), Kushnir and Zubrickas (2020)).[9] Instead, we demonstrate the importance of non-homotheticities and show that prices play an important role even in the canonical setting with no profit or full profit taxation, as in Diamond and Mirrlees (1971b). We thus isolate a novel mechanism, the amplification of redistribution due to the interaction between price changes and non-homotheticities.

Furthermore, by studying price changes stemming from increasing returns to scale, this paper contributes to a growing literature on optimal tax design and endogenous productivity. Recent work highlights the role that taxes may have on entrepreneurial effort (e.g, Jaimovich and Rebelo (2017), Bell et al. (2018)) and draws implications for optimal taxation of top earners (e.g, Jones (2019), Bell et al. (2019)). In contrast, we study productivity effects that are induced by changes in demand, through returns to scale, and which inherently interact with the income tax schedule. We find that the impact of taxes on productivity through demand and returns to scale is quantitatively large, implying substantial adjustments to the optimal tax schedule.

[8] Naito (1999) considers a model with two types of labor inputs (i.e., workers are not fully substitutable in the firm production functions) and shows that if the two types of labor cannot be taxed at different rates (a deviation from the Diamond and Mirrlees (1971b) benchmark), then subsidizing the good of the sector that uses relatively more low-type workers can be optimal, as it increases their wages. While Naito (1999)'s channel requires segmented labor markets and multiple factors or production but can operate with homothetic utility, our analysis can have a single factor of production (with differences in efficiency units across agents) and requires non-homothetic utility. In this sense, our analysis is conceptually distinct from the contribution of Naito (1999).

[9] These papers highlight the importance of rents that accrue to firm owners, which can be redistributed through taxation of income, endogenous price changes, and commodity taxes. We instead characterize different channels, which continue to apply in settings with no rents, i.e. with full profit taxation or zero profit. In particular, Kushnir and Zubrickas (2020) study optimal taxation with endogenous prices, decreasing returns to scale, positive firm profits, and homothetic utility. While their Appendix A.3 examines the case of non-homothetic preferences, the impacts of non-homotheticities and prices remain implicit in their tax formulas through endogenous variables that depend on prices, such as the marginal utility of disposable income. Our analysis thus complements the work of Kushnir and Zubrickas (2020) as our results do not depend on the taxation of profits and provide a full characterization of the role of non-homotheticities. The intuition for their main result is that, when profits are not fully taxed, the social planner uses the price level as an additional redistributing tool: a decrease in the price level benefits low-productivity agents as they can afford more consumption, but hurts high-productivity agents through a decrease in firm profits. We characterize a different channel: our price effects operate through non-homotheticities and changes in the marginal utility of income at different income levels. As entry is free, firms make no profit on average in our model. We do not need to keep track of the distribution of profits across households as we assume that households hold a fully diversified portfolio of firms, with zero profits on average. We can thus cleanly separate our analysis from complementary prior work focusing on the distribution of firm profits.

Outline. The remainder of the paper is organized as follows. Section 2 presents the model. Section 3 derives the optimal income and commodity taxation formula in terms of sufficient statistics. Section 4 uses the comparative static approach to characterize the sensitivity of optimal tax rates to price shocks. The quantitative analysis is carried out in Section 5. Supplemental results and proofs are reported in the Online Appendix.[10]

2 Model

To streamline our analysis, we present in the main text the simplest possible model allowing us to illustrate the mechanisms described in the introduction.The general model is relegated to Appendix E.

We consider a two-sector economy. Sectors are indexed by $k = l, h$. There is a mass 1 of households with different productivity types θ distributed according to $\pi(\theta)$. To facilitate reading, Online Appendix Table A1 provides a list of the variables and notation we use.

Households. Households' preferences over goods c_h, c_l and hours worked z/θ, where z denotes pre-tax income, are given by:

$$u(c_l, c_h) - \frac{1}{1+\frac{1}{\varepsilon}} \left(\frac{z}{\theta}\right)^{1+\frac{1}{\varepsilon}},$$

with u concave, increasing and $\mathcal{C}^3$, and $\varepsilon \leq 1$. Preferences are of the Atkinson-Stiglitz type so that consumption choices only depend on consumer prices and post tax income z^*. This specification allows us to capture non-homothetic spending patterns across the income distribution, and thus the unequal effects of price changes, while focusing on a single tax instrument for redistribution, the nonlinear income tax.[11]

We denote by V the indirect utility of the agent, by v the indirect sub-utility out consumption (i.e., the maximum of $u(c_l, c_h)$ at fixed post tax income z^*), and by v_{z^*} the marginal utility of income. V depends on the agent type θ, on consumer prices and on the tax schedule. Aggregate demand for k across all households is denoted by C_k.

Firms. We adopt a supply-side formulation that nests both perfect competition and monopolistic competition with free entry. The competitive case, following the canonical Diamond-Mirrlees framework, remains central in public finance. In that case, prices are pinned down by technology and demand. However, incorporating monopolistic competition is crucial for capturing how prices respond to demand shifts. Empirical evidence (e.g., Costinot et al. (2019), Jaravel (2019), Faber and Fally (2021)) shows that prices tend to fall as market size grows—a result driven by increased entry and declining markups. Our

[10] Appendix A presents the proofs of all theoretical results in the main text; Appendix B describes additional quantitative results; additional figures and tables are reported in Appendix D; finally, Appendix E presents additional theoretical results.

[11] With more general preferences, it would be possible to use the consumer prices of certain goods to better discriminate between different taxpayers (e.g., Saez (2002), Ferey et al. (2023)). However, we focus on characterizing how unequal price changes for consumption baskets along the income distribution affect the desirability of redistribution policies. In the interest of providing a streamlined analysis, it is sufficient to generate heterogeneous baskets of consumption through non-homothetic Atkinson-Stiglitz preferences rather than idiosyncratic preferences. Indeed, the heterogeneous welfare impacts of price changes only depend on the heterogeneity in households' expenditure shares, whether they stem from idiosyncratic preferences or households' income levels.

framework is meant to accommodate both forces: technological determination of prices and endogenous markups that respond to demand.[12]

In each sector, good k is produced either competitively or monopolistically using labor as the sole input. Under these assumptions, we can summarize the production process through a cost function $\chi_k(C_k, \xi_k)$ (capturing the total labor cost needed to produce C_k units of good k), and a pricing function $p_k = \phi_k(C_k, \xi_k)$ where ξ is a cost shifter. When the good is produced competitively, the pricing function is simply the marginal cost of production: $\phi_k(C_k, \xi_k) = \partial_{C_k}\chi_k(C_k, \xi_k)$. When the good is produced monopolistically, we further assume that firms can freely enter market k, by paying a fixed labor cost, in which case total cost is equal to total revenue: $\chi_k(C_k, \xi_k) = C_k\phi_k(C_k, \xi_k)$. In Appendix E.1, we provide microfoundations for this specification of the supply side.

An important statistic for our analysis is the elasticity of the price p_k to market size, $-C_k/p_k\, \partial_{C_k}\phi_k$. In the main text, we impose that this elasticity is constant and equal to α in all sectors. This elasticity will be crucial: tax changes shift households' incomes and thus aggregate demand for goods, which affects prices.

To illustrate our supply side model, we provide some simple parametric microfoundations as examples. First, consider a competitive case with a representative firm in sector k producing its good at cost $\chi_k(C_k, \xi_k) = \xi_k C_k^{1-\alpha}/(1-\alpha)$. The pricing function is then $\phi_k(C_k, \xi_k) = \xi_k C_k^{-\alpha}$ and the elasticity of the price p_k to market size is α. In particular, $\alpha = 0$ corresponds to the standard case where production functions are linear and prices are exogenous, given by ξ_k.

For the monopolistic case, consider the Melitz-Chaney model (Melitz (2003), Chaney (2008)). Producers of differentiated varieties of product k can freely enter market k by paying a fixed labor cost $\xi_{e,k}$. Upon entering, they draw their productivity type $\gamma(i)$ from a Pareto distribution $1 - \Psi_k(\gamma) = \gamma^{-\gamma_k}$. To start production, firms have to pay a second fixed cost, $\xi_{p,k}$. The variable labor cost of producing $c_{k,i}$ units of variety i is $c_{k,i}/\gamma(i)$. Competitive retailers then aggregate the varieties according to $C_k = \left(\int_{i\in\mathcal{I}_k} c_{k,i}^{1-\alpha} di\right)^{\frac{1}{1-\alpha}}$, with $0 < \alpha < 1$, where $\mathcal{I}_k$ is the set of producing firms. We then obtain the price of good k is $p_k = C_k^{-\alpha}\varphi_k(\boldsymbol{\xi}_k)$[13] and the elasticity of the price p_k to market size is α.

[12] A longstanding literature emphasizes that larger markets stimulate entry, which can increase product variety and lower prices, through both reduced marginal costs and decreased markups. The notion that market size endogenously drives productivity improvements originates with Linder (1961) and Schmookler (1966), and was subsequently formalized in foundational models by Dixit and Stiglitz (1977), Krugman (1979), Shleifer (1986), Romer (1990a), Aghion and Howitt (1992a), Acemoglu (2002), and Melitz (2003). Empirically, several studies confirm that both entry and total factor productivity (TFP) respond to changes in market size. Acemoglu and Linn (2004) show that pharmaceutical innovation and the entry of new drugs in the U.S. are shaped by market size. Weiss and Boppart (2013), using national accounts data, find that TFP growth is higher in income-elastic sectors. In Chinese manufacturing, Beerli et al. (2020) estimate that a 1% increase in market size raises TFP by 0.46%. In the context of local housing markets, Diamond (2016) and Couture et al. (2020) document that amenities adjust endogenously to shifts in local demand. A smaller but growing literature studies the impact of market size on prices. Bartelme et al. (2019), exploiting trade shocks, estimate scale elasticities across two-digit U.S. manufacturing sectors, averaging 0.13 with a range from 0.07 to 0.25. Jaravel (2019), using a shift-share IV design in consumer packaged goods, finds that a 1% increase in demand reduces prices by 0.42%. Structural estimates suggest that over half of this effect reflects falling markups rather than declining marginal costs. Similarly, Faber and Fally (2021), using Nielsen scanner data, estimate a structural model showing that quality production exhibits increasing returns to scale. Together, these findings underscore the importance of accounting for increasing returns when analyzing optimal taxation, particularly from a long-run perspective that incorporates entry dynamics. In contrast, benchmark optimal tax models—such as Mirrlees (1971) and Saez (2001)—typically assume constant returns to scale, which may be more relevant in short-run settings.

[13] With $\varphi_k(\boldsymbol{\xi}_k) = \frac{1}{\alpha}\xi_{p,k}\left(\frac{\alpha(1+\gamma_k)-1}{1-\alpha}\frac{\xi_{e,k}}{\xi_{p,k}}\right)^{-\frac{1}{\gamma_k}\alpha}\left(\frac{\alpha}{1-\alpha}\frac{1}{\xi_{p,k}}\right)^{\frac{1}{1-\alpha}}$.

Planner's problem. The social planner has access to a full set of commodity taxes and to a non-linear income tax and a full tax on profits (in the monopolistic case, profits are 0). As our agents have Atkinson-Stiglitz preferences, the role of commodity taxation is limited but we include commodity taxes for completeness.

The planner maximizes the following social welfare function,

$$\int_{\underline{\theta}}^{\bar{\theta}} G(V(\theta), \theta)\pi(\theta)d\theta,$$

with G concave and increasing in its first argument. The planner sets consumer prices $\{q_h, q_l\}$ and the income tax $T(z)$ subject to three constraints. First, the producer prices p_h, p_l are given by the functions ϕ_k, with $p_k = \phi_k(C_k, \boldsymbol{\xi}_k)$, where C_k denotes the aggregate demand for k. Second, households optimally choose consumption and labor supply under q_h, q_l and $T(z)$. We denote by $f(z)$ the resulting distribution of income, with $dz/d\theta f(z) = \pi(\theta)$, and by z^* disposable income, with $z^* = z - T(z)$. Finally, the government's budget constraint is given by $\sum_{k=h,l}(q_k - p_k)C_k + \mathbb{E}_z(T(z)) + \sum_{k=h,l}\left(p_k C_k - \chi_k\left(C_k, \xi_k\right)\right) \geq 0$.

With this formulation of the planner problem, we encompass both the standard Diamond-Mirrlees framework where firms are competitive and profits are fully taxed and the monopolistic case with free entry where firms earn zero profits on average ($\chi_k\left(C_k, \xi_k\right) = C_k \phi_k\left(C_k, \xi_k\right)$).

Missing Tax. Our benchmark specification does not have, in general, an efficient supply side. Therefore, the solution of the government problem will be *constrained efficient*. With more tax instruments, the social planner could regulate firms and improve the allocation. For example, the planner could directly choose the number of firms in each market to minimize the total cost of production, which includes the variable cost of production and the entry cost.

Note however that the planner can regulate supply in a revenue neutral fashion: for a given industrial policy τ that depends on aggregate quantities, there is a new reduced-form pricing function $p_k = \phi_k^\tau(C_k, \boldsymbol{\xi}_k)$ which depends on the regulatory regime. The solution of the planner problem characterizes the optimal choice of consumption and income taxes for a given industrial policy, which may or may not be optimal.[14] In that sense, industrial policies and redistribution are separable: for a given regulatory rule of the supply side, we take the induced pricing function (and the market size elasticity) as given and derive the optimal redistributive policy.[15] Our results will therefore be valid whether or not industrial policies are optimal, or missing altogether.

Notation. We use standard notation throughout the paper. $\zeta \equiv \varepsilon/\left(1 - \varepsilon \partial_{\ln(z)} \ln\left(v_{z^*}\right)\right)$ is the compensated labor supply elasticity and $\tilde{\zeta}$ is the compensated labor supply elasticity corrected for non-linearities in the budget constraint: $\tilde{\zeta} = \zeta/(1 + z\zeta T''/(1 - T'))$. Similarly, $\eta = \zeta \partial_{\ln(z)} \log\left(v_{z^*}\right)$ is the income effect with a linear budget constraint and $\tilde{\eta}$ the corrected income effect.[16] Regarding spending patterns, $e_k = q_k c_k(z^*, q)$ denotes the agent's expenditure on k, $s_k(z^*, q) = e_k/z^*$the share of k consumption, $\partial_{z^*} e_k$

[14] Suppose that under the industrial policy τ the pricing function is $\phi_k^\tau(C_k, \boldsymbol{\xi}_k)$, so the cost is $C_k \phi_k^\tau(C_k, \boldsymbol{\xi}_k)$, and the fiscal cost is $C_k \psi_k^\tau(C_k, \boldsymbol{\xi}_k)$ for an arbitrary function ψ_k^τ. By imposing a sales tax $t_k(C_k, \boldsymbol{\xi}_k) = \psi_k^\tau(C_k, \boldsymbol{\xi}_k)$ the industrial policy is budget neutral and the income taxation problem is equivalent with a new pricing function $\tilde{\phi}_k^\tau(C_k, \boldsymbol{\xi}_k) = \phi_k^\tau(C_k, \boldsymbol{\xi}_k) + t_k(C_k, \boldsymbol{\xi}_k)$.

[15] In the quantitative analysis, we use the estimated market size elasticity in the United States between 2004 and 2015, which depends implicitly on the regulatory regime in that period.

[16] See for instance Appendix A1 of Scheuer and Werning (2016).

the marginal propensity to spend on k, E_k and $\bar{s}_k$ are the aggregate spending and the aggregate spending share on k, and $\partial_{z^*}E_k$ the average marginal propensity to spend on k. Finally, S is the matrix of cross price derivatives of the aggregate Hicksian demand function, with $S_{jk} = \mathbb{E}(\partial_{q_k}c_j + \partial_{z^*}c_jc_k)$, $\mathcal{S}$ the matrix of price elasticities $\mathcal{S}_{jk} = q_k/C_jS_{jk}$, and $\sigma = -\mathcal{S}_{hh} + \mathcal{S}_{lh} = -\mathcal{S}_{ll} + \mathcal{S}_{hl}$ the elasticity of substitution.

3 Optimal Taxation: First-Order Approach

In this section, we characterize the optimal commodity and income taxes. While we provide heuristic derivations in this section, Online Appendix A reports the formal proofs.

Commodity Tax. Consider a small change in the consumer price of k, dq_k, compensated with an income tax change $dT(z) = -c_k(z^*, q)dq_k$. As explained in Saez (2002), this compensation keeps the welfare and labor supply of all agents constant. Therefore, the impact on government revenue is:

$$\underbrace{dq_kC_k}_{\text{Mechanical effect}} + \underbrace{\mathbb{E}(dT(z))}_{\text{Cost of the compensation}} + \underbrace{\sum_{j=h,l}(q_j - \partial_{C_j}\chi_j(C_j, \xi_j))dC_j}_{\text{Households' behavioral response}}.$$

The increase in q_k first mechanically raises revenues from the tax on k by dq_kC_k. Households are compensated for the consumer price increase through the income tax, so revenue from the income tax decreases: $\mathbb{E}(dT) = -C_kdq_k$. The mechanical effect and the cost of the compensation exactly offset each other: $dq_kC_k + \mathbb{E}(dT(z)) = 0$. Since dq_k is compensated, aggregate consumption reacts through a substitution effect ($dC_k/C_k = \mathcal{S}_{k,j}(dq_j/q_j)$) and the impact on government revenue of the households' change in consumption is $\sum_{j=h,l}(q_j - \partial_{C_j}\chi_j(C_j, \xi_j))dC_j = \sum_{j=h,l}(q_j - \partial_{C_j}\chi_j(C_j, \xi_j))C_j\mathcal{S}_{j,k}(dq_k/q_k)$. Given our supply side specification, we have $\partial_{C_j}\chi_j(C_j, \xi_j) = (1 - t_w)p_j$ with $t_w = \alpha$ in the monopolistic case () and $t_w = 0$ in the competitive case.[17] This change in marginal cost captures the response of the supply side to the shift in demand, and the adjustment of producer prices. At the optimal consumption prices, revenue should remain unchanged, so we obtain:

$$\sum_{j=h,l}(q_j - (1 - t_w)p_j)C_j\mathcal{S}_{j,k}\frac{dq_k}{q_k} = 0.$$

Since this must hold for both dq_l and dq_h, we have $q_k = \beta p_k$ at the optimum, with β an arbitrary scaling constant. Without loss of generality, we choose the scaling so that on average commodity taxes raise no revenue, which implies that we get $q_k = p_k$.[18] We therefore obtain a version of the standard Atkinson-Stiglitz result that commodity taxes are not used at the optimum. The derivation also makes clear that when pricing inefficiencies, $\phi_k(C_k, \xi_k) - \partial_{C_k}\chi_k(C_k, \xi_k)$, varies across sectors, then commodity tax should

[17] In the monopolistic case, we have $\partial_{C_k}\chi_k(C_k, \xi_k) = \phi_k(C_k, \xi_k) + C_k\partial_{C_k}\phi_k(C_k, \xi_k) = (1 - \alpha)p_k$, in the competitive case, we have $\partial_{C_k}\chi_k(C_k, \xi_k) = \phi_k(C_k, \xi_k) = p_k$.

[18] Conceptually, the average commodity tax should be zero. If it was instead positive (or negative), consumer prices would be on average higher than producer prices, which is an implicit income tax (or an implicit income subsidy). If the revenue from the tax was positive (or negative), the government would have to rebate the tax revenue optimally (or raise funds to finance the subsidy). Thus, there would be a non trivial interaction between redistribution and corrective commodity taxation. Imposing zero revenue from commodity taxes on average cleanly separates the redistributive and corrective motives.

be used to correct relative inefficiencies across sectors (i.e., heterogeneous α_k's).[19] This is not the case in our benchmark specification where pricing inefficiencies are the same in both sectors.

Income Tax. Next, consider the perturbation of Saez (2001), a small change of marginal tax $d\tau$ in a neighborhood dz of z and a change in tax $dzd\tau$ above. This has four effects: a mechanical change in revenue, a welfare effect, a fiscal externality due to labor supply responses, and a fiscal externality due to shifts in aggregate consumption and producer price adjustments.

Mechanical and Welfare Effects. Households above z pay an additional $dzd\tau$ in taxes. Their welfare loss is $v_{z^*}dzd\tau$, valued $G'v_{z^*}dzd\tau/\lambda$ by the planner, where λ is the Lagrange multiplier on the government budget constraint. The total effect on social welfare is:

$$\mathbb{E}_{z'>z}\left(1 - G'v_{z^*}/\lambda\right) dzd\tau.$$

Labor Supply Effects. The change in tax rate at z generates a compensated wage effect on labor supply, while the change in the tax burden above z creates an income effect. The change in government revenue is:

$$-f(z)\frac{T'}{1-T'}z\tilde{\zeta}dzd\tau - \mathbb{E}_{z'>z}\left(\frac{T'}{1-T'}\tilde{\eta}\right)dzd\tau.$$

Price and Demand Effects. The change in the tax schedule affects households' disposable income both mechanically and through labor supply responses. This leads to a change in aggregate demand for goods, through substitution and income effects, and in producer prices and costs. The total impact on government revenue, through the receipts of the commodity and profit taxes, is given by:

$$\begin{aligned}\sum_{k=h,l}(q_k - \partial_{C_k}\chi_k(C_k,\xi_k))dC_k &= \sum_{k=h,l}(p_k - (1-t_w)p_k)\,dC_k \\ &= t_w\sum_{k=h,l}p_k dC_k \\ &= -t_w\left(f(z)z\tilde{\zeta} + \mathbb{E}_{z'>z}(1+\tilde{\eta})\right)dzd\tau.\end{aligned}$$

In the derivation above, the first line uses $\partial_{C_k}\chi_k(C_k,\xi_k) = (1-t_w)\,p_k$and $q_k = p_k$. The third line uses the fact that, from the budget constraint of households, $\sum_{k=h,l}p_k dc_k$ is equal to the change in disposable income. To interpret this effect, note that when $t_w > 0$, prices are above marginal costs and demand for goods is inefficiently low. An increase in the income tax decreases labor income and further depresses demand, accentuating the initial inefficiency.

Summing up all of the effects derived above gives the first order conditions for the optimal tax rate. We denote by g the Pareto weights $g = G'v_{z^*}/((1-t_w)\lambda)$, where the $1-t_w$ normalization is such that $\mathbb{E}(g) = 1$ when there are no income effects. Collecting our earlier result on optimal commodity taxes, we then obtain our first Proposition:

[19] We show that commodity taxes are used to correct pricing inefficiencies in the proof of Proposition E1 in Appendix E, which covers the general case.

Proposition 1. *Commodity taxes are not used at the optimum. The optimal non-linear income tax schedule is characterized by:*

$$\frac{T'}{1-T'} = -t_w + \frac{1-t_w}{z\tilde{\zeta}f(z)}\left\{\mathbb{E}_{z'>z}\left(1-g\right) - \frac{1}{1-t_w}\mathbb{E}_{z'>z}\left(\left(t_w + \frac{T'}{1-T'}\right)\tilde{\eta}\right)\right\}, \tag{1}$$

where $t_w = \alpha$ in the monopolistic case and $t_w = 0$ in the competitive case. With $\alpha = 0$, we obtain the standard optimal tax formula in both cases.

Proof: See Appendix A.1. Proposition E1 in Appendix E provides a generalization of this result, with general household preferences and in a multi-sector economy with heterogeneous returns to scale, with potential spillovers across sectors.

Proposition 1 first shows that, as in the standard Atkinson-Stiglitz framework, commodity taxes are not needed. This is not the case when α varies across sectors: commodity taxes then have a corrective role but are not used for redistribution.

Turning to the optimal income tax schedule, Proposition 1 suggests that when the average market size elasticity is positive and firms are not competitive ($t_w = \alpha > 0$), labor supply is subsidized and optimal tax rates are reduced.[20] Intuitively, there is an externality from working: more labor supply increases aggregate income, i.e. market size, and leads to a fall in prices through returns to scale. If the endogenous quantities (the Pareto weights, the income distribution, and the labor supply elasticities) remain constant as α varies, then the formula tells us that the tax rate with $\alpha > 0$ is such that $1 - T' = (1 - T')_{\alpha=0}/(1-\alpha)$. In that case, the planner implements a uniform wage subsidy $1/(1-\alpha)$ on top of the standard non linear tax, and it appears that there is no interaction between the corrective tax (the wage subsidy) and redistributive motives. This interpretation is however naive, as prices and all endogenous quantities are likely to vary as α changes: the interaction between corrective tax and redistributive motives is hidden in the formula.

These observations highlight an important limitation of the standard optimal tax formula, which leaves the effects of prices completely implicit and therefore provides little insight about how prices affect optimal redistribution. In the next section, we provide a characterization of the role of prices for optimal taxes.

4 Understanding the Impact of Prices and Non-homotheticities

In this section, we use a comparative statics approach to understand the mechanisms through which optimal tax rates respond to prices in the presence of non-homotheticities.

[20] Note that this correction could be implemented through consumer prices using uniform commodity taxes, dropping the requirement that commodity taxes should be budget neutral. Indeed, a homogeneous reduction in prices is equivalent to a wage subsidy, so the optimal income tax could be given by the formula of Proposition 1 with $\alpha = 0$ if all consumer prices were multiplied by $1 - \alpha$.

4.1 Assumptions

To streamline the analysis, we make the following assumptions. First, while sectors h and l were arbitrary in the previous section, we now impose that l is a "necessity" good (therefore h is a "luxury" good) to highlight the importance of non-homotheticities.[21] Second, we make an assumption on the distribution of skills.

Assumption A1. *At initial prices, l is a necessity good: $\partial_{z^*} e_l$ is decreasing in post-tax income and $\partial_{z^*} E_l \leq \bar{s}_l$, where $\bar{s}_l$ and $\partial_{z^*} E_l$ are the aggregate spending share and average marginal propensity to spend on l.*

Assumption A2. $\underline{\theta}\pi(\underline{\theta}) = 0$ and $\left(1 + \theta\pi'(\theta)/\pi(\theta)\right)\epsilon/(1+\epsilon) \leq 1$ *for all θ.*

These assumptions allow us to derive clean theoretical results in the following section but are not substantive restrictions. They do not affect the tax formulas of Proposition 2: $A1$ is used to sign the tax response and $A2$ to characterize the monotonicity of the welfare response. We relax $A1$ in our quantitative analysis in Section 5 to use observed spending patterns for all products and the income distribution. The empirical income distribution in the United States satisfies $A2$,as discussed in Section 5.

Next we normalize the income effect of labor supply to 0 at initial prices:

Assumption A3. *There are no income effects at initial prices, i.e. $v_{z^*} = 1 \, \forall z^*$.*

This assumption is common in the optimal taxation literature and provides a useful benchmark to facilitate comparisons between our results and prior work; it is relaxed Appendix E.[22] To clarify its role, note that household preferences can be written as $\Psi\left(u\left(c_l, c_h\right)\right) - \left(1+\epsilon^{-1}\right)^{-1}\left(z/\theta\right)^{1+\epsilon^{-1}}$. Here, the function Ψ parametrizes the income effect of labor supply without affecting the household's consumption demand functions $c_k\left(z^*, p\right)$.[23] This normalization simplifies the labor supply side of the model, allowing us to focus on how non-homotheticities in consumption preferences influence optimal redistribution, compared to prior work where the benchmark formulas similarly do not feature income effects (e.g., Diamond (1998), Saez (2001)).[24]

Finally, we make an assumption on social preferences:

Assumption A4. *The social welfare function is linear, i.e. $G(v(\theta), \theta) = \lambda_\theta v(\theta)$.*

Note that λ_θ can be used arbitrarily to parametrize the level of redistribution at initial prices. Assumption $A.4$ is important to obtain closed form solutions to our comparative statics exercise in Proposition 2.

[21] The comparative statics approach we use in this section allows us to provide an explicit characterization of the impact of non-homotheticities and market size effects on optimal taxation, in terms of observable statistics. With non-homotheticities, it is not possible to obtain an explicit solution of the integral equation characterizing the optimal tax schedule in partial or general equilibrium. First, non-homothetic demand systems do not yield closed form expressions for both demand functions and marginal utility of income – with some exceptions e.g. Stone-Geary preferences, which are too limited to capture the impact of price changes observed in the data. Second, in general, the Pareto weights will depend on agents' disposable income and thus on on the income tax $T(z)$. With homothetic utility, quasilinear preferences in consumption, and a linear social welfare function, a closed form solution can be obtained in some cases (e.g., Eeckhout et al. (2021)).

[22] While the traditional empirical literature suggests income effects are small (e.g., Imbens et al. (2001), Cesarini et al. (2017)), more recent evidence is more mixed (Giupponi (2024), Golosov et al. (2024), Vivalt et al. (2024)).

[23] Details on how to choose Ψ are provided in Section 5: see footnote 43.

[24] In Appendix E, we provide comparative statics formula with non-zero income effects. We also provide in Proposition E3 a qualitative characterization of the response of redistributive policies to price changes that generalizes the results of Proposition 3. Finally, we provide closed-form formulas for the top tax rates and show that, for reasonable calibrations, income effects can exacerbate the regressive impact of the price changes described in this section.

We provide comparative statics formulas with a general G in the appendix, a qualitative characterization in Proposition 3, and a quantitative analysis in Section 5.

In the remainder of this section, we mostly analyze how the optimal tax schedule and household welfare respond to changes in the relative price of the necessity good. To isolate redistribution effects from aggregate efficiency concerns, we consider a compensated price change: the price of the necessity rises while that of the luxury falls, keeping the average price level constant. Formally, we consider an increase $d\ln\bar{p}_l$, such that $d\ln p_l = \bar{s}_h d\ln\bar{p}_l$ and $d\ln p_h = -\bar{s}_l d\ln\bar{p}_l$.[25]

4.2 Response to a Price Change with Linear Production Functions: Channels #1 and #2

To first isolate the role of non-homotheticities, we first focus on the response of the optimal income tax schedule to a price change when prices are fully exogenous. This is the case when $\alpha = 0$, which imposes $p_k = \phi_k(\xi_k)$ and $\chi_k(C_k, \xi_k) = \phi_k(\xi_k) C_k$ with both competitive and monopolistic firms (linear production functions). We show that the redistributive effects of prices are amplified through the tax system via two channels: a change in the value of redistribution (Channel #1) and a change in the efficiency cost of taxation (Channel #2).

Response with a Linear Social Welfare Function

We first consider the case of a linear social function which allows us to derive closed form solutions. Prices affect both the value and the cost of redistribution around θ (i.e. around the $F(z(\theta))$ percentile of the income distribution). While these effects are only implicit in Proposition 1, we now make them explicit for a marginal change in the price of good k, p_k. Recall that in Proposition 1 the optimal tax rate is defined implicitly by

$$\frac{T'}{1-T'} = \frac{1}{z\tilde{\zeta} f(z)} \left\{ \underbrace{\mathbb{E}_{z'>z}(1-g)}_{\text{Welfare effects}} - \underbrace{\mathbb{E}_{z'>z}\left(\frac{T'}{1-T'}\tilde{\eta}\right)}_{\text{Behavioral effects}} \right\}. \tag{2}$$

Under $A3$, we have $z\tilde{\zeta} f(z) = \epsilon/(\epsilon+1)\,\theta\pi(\theta)$, independent from prices and taxes, so we only need to derive the change in the Welfare and Behavioral Effects terms, the two channels of the adjustment of the tax system. We now provide heuristic derivations these channels, presenting the proofs in Appendix A.2.

Welfare Effects (Channel #1). Under $A3-A4$, the derivative of Pareto weights with respect to prices p_k satisfies $p_k \partial_{p_k} g = -g \partial_{z^*} e_k$, with $\partial_{z^*} e_k$ the marginal propensity to spend on good k. Indeed, at the

[25] Any price increase can be viewed as the sum of (i) a uniform price increase and (ii) a change in relative prices. The relative price effect is central to our analysis while the uniform price change simply scales real wages and has limited interaction with consumption heterogeneity. One of the main insights of the section is that the planner favors high income households when the relative price of necessity increases, i.e. the welfare of low income households falls. The advantage of considering a relative price change is that, as shown in Proposition 3, the planner could, in principle, redistribute resources to keep the welfare of low-income households unchanged. If we analyzed an uncompensated increase in the price of necessities, the overall productivity of the economy would decline, reducing the capacity of the planner to redistribute. In that case, a reduction in the utility of low-income households would be less surprising, as full compensation would not be feasible. We report the comparative statics for uniform price increases in Appendix A.

initial prices the social value of a dollar transfer to an agent with income z is given by the Pareto weight g. With this additional dollar, the agent spends $\partial_{z^*}e_k$ on good k. When the price of k increases, the purchasing power of the agent is therefore reduced at the margin by $\partial_{z^*}e_k$. Since an agent at z can buy less with an additional dollar, the value of a dollar transfer is reduced. This channel can be thought of as a "terms of trade effect for redistribution."

The change in the optimal income tax for type θ is determined by the change in individual purchasing power for type θ *relative to* the average change in purchasing power, $\mathbb{E}\left(g\partial_{z^*}e_k\right)$. The planner decreases the tax rate at θ – to redistribute more to agent with income $z(\theta') > z(\theta)$ – if a dollar transfer above θ buys relatively more welfare than below θ, that is if the (marginal) purchasing power after the price change decreases relatively less above than below. For the necessity good l, the marginal propensity to spend decreases with income, so the tax rate is lowered everywhere in response to an increase in p_l.

Therefore, the adjustment in tax rates $\frac{p_k d}{dp_k}\left\{\frac{T'}{1-T'}\right\}$ through Channel #1, the change in the value of redistribution at θ, is given by:

$$\frac{1}{z\tilde{\zeta}f(z(\theta))}\mathbb{E}_{z>z(\theta)}\Big(g\left(\partial_{z^*}e_k - \mathbb{E}\left(g\partial_{z^*}e_k\right)\right)\Big).$$

By contrast, when preferences are homothetic, there is no effect on the tax rate since the change in purchasing power is uniform along the income distribution.

Behavioral Effects (Channel #2). Under $A3-A4$, the derivative of the income effect with respect to consumer prices p_k satisfies $p_k\partial_{q_k}\tilde{\eta} = -z\tilde{\zeta}((1-T')\partial_{z^*z^*}e_k$. Under $A1$, the marginal propensity to spend on good l decreases, $\partial_{z^*z^*}e_k > 0$. An increase in the price of l causes a fall in the households' (marginal) purchasing power, which is smaller at higher income levels. Therefore a dollar transfer to an agent makes work more valuable – since they now have a higher real wage at the margin – and stimulates labor supply through an income effect.

Consequently, an increase in the price of the necessity good l increases the cost of taxation at θ: raising the tax rate at θ lowers the income of all agents with $\theta' > \theta$, and reduces their labor supply. Through this mechanism, the tax rate should be reduced at θ. More formally, as $\partial_{z^*}e_l$ is decreasing, an increase in p_l convexifies the indirect utility of consumption and therefore increases the income effect.

As above, what ultimately determines the change in the optimal income tax is the increase in the cost of taxation at θ relative to the average change in the cost of taxation across the distribution. Thus, the adjustment in optimal tax rates due to Channel #2, the change in the cost of redistribution at θ, is given by:

$$\frac{1}{z\tilde{\zeta}f(z(\theta))}\mathbb{E}_{z>z(\theta)}\left(T'z\tilde{\zeta}\partial_{z^*z^*}e_k - g\mathbb{E}\left(T'z\tilde{\zeta}\partial_{z^*z^*}e_k\right)\right).$$

As the first channel, this channel does not operate when preferences are homothetic, since $\partial_{z^*z^*}e_k = 0$.

Taking stock. Summing and rearranging the two channels above,[26] we obtain a Proposition characterizing the response of the tax schedule to consumer price changes. This Proposition also shows formally that the adjustment of the tax system amplifies the redistributive effects of price changes.

[26]We use the optimality of the schedule – in equation 2– to express the Pareto weight in function of the tax schedule.

Proposition 2. *Under A3 − A4, the response of the optimal tax rate at θ to an increase in the price of k when $\alpha = 0$ is:*

$$\frac{p_k d}{dp_k}\left\{\frac{T'}{1-T'}\right\} = \frac{1}{z\tilde{\zeta} f(z(\theta))}\mathbb{E}_{z>z(\theta)}\left(\partial_{z^*} e_k - \partial_{z^*} E_k\right) - \frac{T'}{1-T'}\left(\partial_{z^*} e_k - \partial_{z^*} E_k\right). \tag{3}$$

With homothetic preferences ($\partial_{z^*} e_h = s_h$), we have $d_{p_k} T' = 0$. With non-homothetic preferences, under A1 the response of the optimal tax schedule to changes in the price of the necessity ($k = l$) and luxury ($k = h$) goods satisfies :

$$\frac{p_l d}{dp_l}\left\{\frac{T'}{1-T'}\right\} < 0 \quad \text{and} \quad \frac{p_h d}{dp_h}\left\{\frac{T'}{1-T'}\right\} = -\frac{p_l d}{dp_l}\left\{\frac{T'}{1-T'}\right\} > 0 \quad \forall\theta.$$

Proof: See Appendix A.2.2. Proposition E2 in Appendix E provides a generalization of this result, with general household preferences and in a multi-sector economy with potential spillovers across sectors.

The advantage of Proposition 2 is twofold. First, it allows us to quantify the effect of prices on the tax rate as a function of *observable* quantities. For example, we do not need to explicitly specify Pareto weights to evaluate the impact of prices.[27]

Second, we can unequivocally sign the impact of prices on taxes. When the marginal propensity to spend on good k decreases (i.e, k is a "necessity" good), the tax rate decreases everywhere in response to an increase in p_k. The tax burden decreases at the top of the distribution and increases at the bottom: the planner redistributes to higher-income households. This result might seem surprising, because the optimal tax schedule amplifies the redistributive effects of price changes instead of offsetting them, but our analysis explains why: when k is a necessity good, the social value of a dollar transfer decreases less for higher-income than lower-income households (Channel #1), and the income effects increase more at the top (Channel #2).

It is then simple to characterize the welfare consequences of price changes. By itself, an increase in the relative price of necessities reduces the utility of households at the bottom of the income distribution, as necessities are a larger part of their budget. As the planner responds to the price increase by decreasing tax rates, the transfer to low-income households is reduced, which lowers their utility further. Through the same channels, high-income households strictly benefit from the price increase and the welfare gains are increasing in income. Note that it would be feasible to compensate all households[28] in a budget neutral fashion, but this is not optimal because of Channel #1 and #2. These observations are formalized in the

[27]While social preferences do not appear in our formulas, they still play an implicit role. Note that the derivative of the tax rate $d_{p_k} T'$ is of order $(1 - T')^2$. The stronger the preference for redistribution, the higher the (initial) tax rate, and the lower the sensitivity of the tax rate to changes in prices. To illustrate, assume that the marginal propensity to spend on k is constant above an arbitrary θ_0. Denoting $\bar{g}(\theta) = \mathbb{E}(g \mid \theta' > \theta)$ the average Pareto weight for households with ability larger than θ, we have for $\theta \geq \theta_0$:

$$\frac{p_k d}{dp_k}\left\{\frac{T'}{1-T'}\right\} = \frac{\bar{g}(\theta)}{1-\bar{g}(\theta)}\frac{T'}{1-T'}\left(\partial_{z^*} e_k - \partial_{z^*} E_k\right).$$

For a luxury good ($k = h$), an increase in p_k reduces the value of a transfer to high income households and tax rates are set higher at the top of the distribution. However, if $\bar{g}(\theta)$ is small then the increase in tax rates is small. Intuitively, if the planner does not value the welfare of higher ability households, price changes have no effects on top tax rates as long as they do not change the cost of taxation through labor supply. The derivation can be found in Appendix A.2.2 on page A13.

[28]By "compensating", we mean keeping their utility equal to their pre-price change level.

corollary below.

Corollary 1. *For an increase in the relative price of necessities $dln\bar{p}_l$, with $dlnp_l = \bar{s}_h dln\bar{p}_l$ and $dlnp_h = -\bar{s}_l dln\bar{p}_l$, the compensating scheme $dT(z(\theta)) = -(s_l - \bar{s}_l) z^*(\theta) dln\bar{p}_l$ is feasible but only optimal when preferences are homothetic. With non-homothetic preferences, under $A1 - A4$ we have $dV(\underline{\theta})/d\bar{p}_l < 0$ and $dV(\bar{\theta})/d\bar{p}_l > 0$; $dV(\theta)/d\bar{p}_l$ is increasing in θ and $\mathbb{E}(gdV(\theta)/d\bar{p}_l) = 0$.*
Proof: See Appendix A.2.2.

Rawlsian social preferences. The formulas of Proposition 2 can be adapted when social preferences are Rawlsian. In that case, we have:

$$\frac{p_k d}{dp_k}\left\{\frac{T'}{1-T'}\right\} = \frac{T'}{1-T'}\left(\mathbb{E}\left(\partial_{z^*} e_k \mid z' > z(\theta)\right) - \partial_{z^*} e_k\right)$$

Even in the extreme case were the social planner only values the welfare of the poorest agent, an increase in the price of necessities leads to more redistribution towards higher income households. This is entirely due to the impact of the price change on labor supply. An increase in the price of necessities increases the income effect on labor supply ($\partial_{q_l}\tilde{\eta} > 0$) and decreases the income tax.

Response with a Non-Linear Social Welfare Function

We now consider the case of a non-linear social welfare function, relaxing assumption $A4$. While the concavity of the welfare function does not affect the initial level of redistribution,[29] it impacts the redistributive effects of price changes. With a concave social welfare function, an increase in the tax burden or in prices has a direct "income effect" on Pareto weights: reducing the (real) disposable income of an agent directly increases the social value of a transfer to this agent. There is therefore an incentive for the social planner to compensate lower-income households when they face higher prices.

As the analysis with a concave welfare function is complex,[30] we present first a very simple model to convey the intuition. Consider a finite type version of our model: the poor (p), middle-class (m) and rich (r) households have types $0 = \theta_p < \theta_m < \theta_r$, so that only the middle income agent faces a positive marginal tax rate.[31] Households have the same preferences as in the continuous type version (satisfying $A1$-$A3$). More specifically, we assume a household of type i has a marginal propensity to spend on the necessity product, $\partial_{z^*} e_{l,i}$, satisfying $\partial_{z^*} e_{l,p} > \partial_{z^*} e_{l,m} = \partial_{z^*} E_l > \partial_{z^*} e_{l,r}$. In words, the marginal propensity to spend on the necessity good is highest for the low-income households, followed by the middle class – whose propensity to spend on this good is assumed to be equal to the aggregate propensity – and the rich. As before, with a linear social welfare function ($G''(V) = 0$), the change in tax rate in response to an increase in the relative price of necessities $dT'_m/d\bar{p}_l$ satisfies:

$$\frac{1}{1-(\theta_m/\theta_h)^{1+\frac{1}{\epsilon}}}\pi_m \frac{d}{dln\bar{p}_l}\left\{\frac{T'_m}{1-T'_m}\right\} = \left(\partial_{z^*} e_{l,h} - \partial_{z^*} E_l\right)\pi_h < 0, \tag{4}$$

[29] As long as $G'(V(\theta)) \propto \lambda_\theta$,the initial tax rate is the same.

[30] With a concave social welfare function, the system determining optimal taxes becomes an integro-differential equation with non-constant coefficients, which does not have a closed form solution.

[31] The low income agent does not work and there is no distortion at the top. We introduce a three-agent model rather than a two-agent model to avoid introducing a new version of assumption $A2$.

which is essentially our formula 3evaluated at $\partial_{z^*} e_l = \partial_{z^*} E_l$. The change in welfare of household i, $dV_i/d\bar{p}_l$, satisfies $dV_p/d\bar{p}_l = dV_m/d\bar{p}_l < 0 < dV_r/d\bar{p}_l$.

With a concave social function G satisfying $G''(V) < 0$, we denote the change in welfare of household i as $dV_i^G/d\bar{p}_l$. The change in welfare of household i satisfies $dV_i^G/d\bar{p}_l = dV_i/d\bar{p}_l/(1+\mathcal{G})$ with $\mathcal{G} > 0$.[32] The parameter $\mathcal{G}$ captures the income effect on Pareto weights: the price change and the tax reform of equation 4 reduce the welfare of poor and middle income households, which raises their Pareto weights. As a result, we have $dV_p/d\bar{p}_l < dV_p^G/d\bar{p}_l < 0$: a concave social welfare function mutes the incentive to redistribute to the high income household. However, the planner still does not offset the loss to poorer households—even though such compensation is implementable – and the price change remains regressive for any concave G. [33]

We now show that this remains true with continuous types. In the appendix, we also provide the comparative statics formulas that we use to quantitatively evaluate the impact of price changes with a concave welfare function.

Proposition 3. *Asssume that $-G''(V,\theta)/G'(V,\theta)$ is positive and non increasing (e.g., a CARA or CRRA function) and that $A1 - A3$ are satisfied. For an increase in the relative price of necessities, the compensating scheme $dT(z(\theta)) = -(s_l - \bar{s}_l) z^* dln\bar{p}_l$ is feasible but only optimal when preferences are homothetic.*

With non-homothetic preferences, the change in welfare of agent θ, $dV^G/d\bar{p}_l(\theta)$, satisfies $dV/d\bar{p}_l(\underline{\theta}) < dV^G/d\bar{p}_l(\underline{\theta}) < 0$, $dV^G/d\bar{p}_l(\theta) - dV^G/d\bar{p}_l(\underline{\theta}) < dV/d\bar{p}_l(\theta) - dV/d\bar{p}_l(\underline{\theta})$ and $\mathbb{E}\left(gdV^G(\theta)/d\bar{p}_l\right) = 0$, where $dV/d\bar{p}_l(\theta)$ is the welfare impact of price change with a linear social welfare function satisfying $\lambda_\theta \propto G'(V(\theta),\theta)$.

If in addition $\bar{\theta} = \infty$, the distribution of type is bounded by a Pareto distribution, $\theta\pi'(\theta)/\pi(\theta) \leq -1-\gamma$ for θ large enough, and $G(V,\theta)$ is either CARA or CRRA, then we have $dV^G/d\bar{p}_l(\theta) \backsim dV/d\bar{p}_l(\theta)$ at infinity.

Proof: See Appendix A.2.2. Proposition E3 in Appendix E provides a generalization of this result, with general household preferences.

As in Proposition 2, it would be feasible to compensate all households in a budget neutral fashion for relative price changes but the planner still optimally decreases the welfare of lower income households. Intuitively, fully compensating agents for a price change leaves their disposable income unchanged. This neutralizes the "income effect" of prices on social welfare weights, but leaves unchanged the valuation effects of prices and their effect on labor supply derived in Proposition 2. For an increase in the relative price of necessities, fully compensating agents cannot be optimal: there is an incentive to redistribute to higher income through Channel#1 and #2. Note however that the concavity of the welfare function mutes the impacts of Channel#1 and #2: lower income households lose less than with a linear function ($dV/d\bar{p}_l(\underline{\theta}) < dV^G/d\bar{p}_l(\underline{\theta}) < 0$) and utility increases at a lower rate across types ($dV^G/d\bar{p}_l(\theta) - dV^G/d\bar{p}_l(\underline{\theta}) < dV/d\bar{p}_l(\theta) - dV/d\bar{p}_l(\underline{\theta})$).

[32] The derivations for the simple example can be found in Appendix A.2.2on page A.19.

[33] The change in tax rate with a concave G, $dT'^G_m/d\bar{p}_l$, satisfies $dT'^G_m/d\bar{p}_l = \left(dT'_m/d\bar{p}_l - \mathcal{G}\left(1-T'_m\right)\left(\partial_{z^*}E_l - \bar{s}_l\right)\right)/\left(1+\mathcal{G}\right)$. The first term is the change in taxes with a linear G while the second is the compensating scheme that would fully offset the effect of price changes.

Finally, the last item of Proposition 3 shows that the concavity of the social welfare function does not matter for the change in welfare of high income households. Intuitively, if the Pareto weights are small at the top of the income distribution, the planner does not directly value a change in utility of high income households. The only determinant of taxes are the one described by Channel#1 and #2 and, as a direct corollary, the income tax rate at the top is left unchanged by the concavity of the welfare function.

4.3 Response to a Price Change with Non-Linear Production Functions

We now turn to the case of non-linear production functions. Non-linearity introduces interaction between demand and supply. As prices change, so does demand for the two goods, which generates a supply side response, further changes in demand, and so on. In this section, we investigate this feedback loop. We show that when $\alpha > 0$, the response of taxes to an increase in the price of necessity is amplified and leads to further redistribution towards higher-income households.

To formally define an exogenous price change with a non-linear production function, we normalize the supply shifter ξ_k such that an increase in ξ_k corresponds to an increase in the price of k, p_k. We consider a cost shifter $p_k^* = 1/\partial_{\xi_k}\phi_k$, which implies $\partial_{p_k^*}\phi_k = 1$ and $\partial_{p_k^*}\chi_k = (1-\alpha+t_w)^{-1} C_k$.[34] As before, we define an increase in the relative price of the necessity $d\ln\bar{p}_l$, such that $d\ln p_l^* = \bar{s}_h d\ln\bar{p}_l$ and $d\ln p_h^* = -\bar{s}_l d\ln\bar{p}_l$.

In addition, we introduce τ_l which captures the impact of non-homothecities on the sensitivity of aggregate demand to prices:

$$\tau_l(z) \equiv (1-t_w)(1-T')\left(\frac{1}{z\tilde{\zeta}f(z)}\mathbb{E}_{z'>z}\left(\partial_{z^*}e_l - \partial_{z^*}E_l\right) + \partial_{z^*}e_l - \partial_{z^*}E_l\right) < 0.$$

To understand the role of τ_l, consider the impact of an increase in the relative price of l when $\alpha = 0$ (no supply side adjustments). As shown in the previous section, optimal tax rates decrease in response to the price increase. Furthermore, we show in Lemma A3 of Appendix A.2.3 that the change in relative demand for necessity satisfies:[35]

$$\hat{C}_l - \hat{C}_h = \underbrace{-\sigma}_{\text{Substitution}} \underbrace{-\frac{\zeta}{\bar{s}_h\bar{s}_l}\mathbb{E}\left(\frac{z}{E}\left(\tau_l + \partial_{z^*}E_l - \bar{s}_l\right)^2\right)}_{\text{Income}}.$$

When preferences are homothetic, the income effect is zero and demand for l decreases only through a substitution channel. When preferences are nonhomothetic, relative demand for necessities further decreases through an income effect. Indeed, an increase in the price of necessities relative to luxuries has a negative income effect on lower income households, as necessities constitute a larger portion of their consumption basket.[36] Since tax rates decrease, the income of poorer households is further lowered. Lower income households have a higher propensity to spend on necessities, so the aggregate share of necessities decreases through income effects: income is reallocated away from necessities.

[34] With monopolistic competition ($\tau_w = \alpha$), this is obvious since $\chi_k = C_k\phi_k$ so $\partial_{p_k^*}\chi_k = C_k$. With competitive firms ($\tau_w = 0$), we can rewrite the pricing function as $\phi_k(\xi_k, C_k) = \tilde{\phi}_k(\xi_k)C_k^{-\alpha} = \partial_{C_k}\chi_k(\xi_k, C_k)$ so $\chi_k(\xi_k, C_k) = \phi_k(\xi_k, C_k)C_k/(1-\alpha) + \underline{\chi_k}$, where the potential fixed cost $\underline{\chi_k}$is assumed to be independent from ξ_k.

[35] In this expression, the change in demand is due to the combination of the change in prices and the change in taxes. With the change in prices alone, we would have $\hat{C}_l - \hat{C}_h = -\sigma - \frac{\zeta}{\bar{s}_h\bar{s}_l}\mathbb{E}\left(\frac{z}{E}\left(\partial_{z^*}e_l - \bar{s}_l\right)\left(\tau_l + \partial_{z^*}E_l - \bar{s}_l\right)\right)$.

[36] The income effect on consumption due to heterogeneity in spending is captured in the first term in τ_l. The second term captures the change in real wages when the price of l increases.

With these definitions, we can now characterize the response of the tax rate to price changes with non-linear production functions. We first consider the partial equilibrium response, when prices do not endogenously respond, so that $dp_k/dp_k^* = \partial p_k/\partial p_k^* = 1$. We then derive the general equilibrium response, when prices adjust to their new equilibrium level.

Proposition 4. *Under A3 – A4, the partial equilibrium response of the income tax to a change in the relative price of necessities is:*

$$\frac{\partial}{\partial ln\bar{p}_l}\left\{\frac{T'}{1-T'}\right\} = \frac{1-t_w}{z\tilde{\zeta}f(z(\theta))}\mathbb{E}_{z>z(\theta)}\left(\partial_{z^*}e_l - \partial_{z^*}E_l\right) - \left(\frac{T'}{1-T'}+t_w\right)\left(\partial_{z^*}e_l - \partial_{z^*}E_l\right).$$

Under A1, $\partial_{\bar{p}_l}T'$ is negative for all θ.

In general equilibrium, the response of the income tax to a change in the relative price of necessities is:

$$\underbrace{\frac{dT'}{d\bar{p}_l}}_{\text{GE response}} = \left(1-\alpha\left(\sigma+\Omega\right)\right)^{-1}\underbrace{\frac{\partial T'}{\partial\bar{p}_l}}_{\text{PE response}},$$

with $\Omega = \frac{1}{1-t_w}\frac{\zeta}{\bar{s}_h\bar{s}_l}\left(\mathbb{E}_z\left((\tau_l+\partial_{z^*}E_l-\bar{s}_l)^2\right)+\frac{\alpha\zeta}{1-t_w-\alpha\zeta}\left(\mathbb{E}_z(\tau_l+\partial_{z^*}E_l-\bar{s}_l)\right)^2\right) > 0$. When $\alpha > 0$, $\frac{dT'}{d\bar{p}_l} < \frac{\partial T'}{\partial\bar{p}_l} < 0$; when $\alpha < 0$, $\frac{\partial T'}{\partial\bar{p}_l} < \frac{dT'}{d\bar{p}_l} < 0$.

Proof: See Appendix A.2.3. Proposition E4 in Appendix E.3.2 provides a generalization of this result, with general household preferences and in a multisector economy with potential spillovers across sectors.

The main insight of Proposition 4 is that the response of the supply side amplifies the response of the tax rate to supply shifters when $\alpha > 0$ and mutes it when $\alpha < 0$. With $\alpha > 0$, the amplification is driven by the equilibrium response of the relative price of l, given by:

$$\frac{dln\left(p_l/p_h\right)}{dln\bar{p}_l} = -\frac{1}{1-\underbrace{\alpha\sigma}_{\substack{\text{Amplification}\\\text{through subsititution effects}}}\underbrace{-\alpha\Omega}_{\substack{\text{Amplification}\\\text{through income effects}}}}.$$

The increase in the relative price of l is amplified through income and substitution effects in general equilibrium. First, as the relative price of l increases, agents substitute the necessity good for the luxury good. The market for h expands, so the price of h further decreases when there are returns to scale ($\alpha > 0$), which creates more substitution. This is the only channel of amplification when preferences are homothetic, since the shares of h and l remain constant as income shifts.

When preferences are nonhomothetic, the share of l further decreases through income effects. This amplification is denoted by Ω and operates through two channels, driven respectively by changes in relative prices and in the average price index.

The first term in Ω, $\mathbb{E}_z\left((\tau_l+\partial_{z^*}E_l-\bar{s}_l)^2\right)$, corresponds to the reduction in the share of l in response to an increase in the relative price of l. Lower income households are more affected, and, as they consume more of good l at the margin, the direct impact of the price increase is to reduce the aggregate expenditure share of l. In addition, tax rates decrease everywhere, as it is more valuable to redistribute to higher income households. This optimal adjustment of the schedule, given by $-\partial_{\bar{p}_l} T'$, amplifies the reallocation of income towards the luxury good. As the aggregate expenditure share of l[37] decreases, the relative price of l increases through returns to scale ($\alpha > 0$), and agents further reallocate their income towards the necessity bundle.

The second term in Ω, $\frac{\alpha\zeta}{1-t_w-\alpha\zeta}\left(\mathbb{E}_z(\tau_l + \partial_{z^*} E_l - \bar{s}_l)\right)^2$, captures a further reduction in the share of l stemming from an endogenous fall in the average price index. When the relative price of h decreases, it can be shown that real wages increase on average across agents, hence labor supply increases.[38] Higher labor supply raises nominal incomes and demand for both the necessity and the luxury goods, which leads to a fall in both prices. Thus, on average households' real incomes grow and they reallocate their expenditures towards the luxury good. The relative price of the luxury good therefore decreases through returns to scale ($\alpha > 0$) and the planner responds by lowering tax rates. Lower taxes further stimulate aggregate labor supply and increase real incomes, generating a further fall in the price index, and, as households become richer on average, more reallocation towards luxuries, inducing a further fall in the price of the luxury good through returns to scale, and so on.[39]

Redistribution towards higher income households is therefore amplified, through general equilibrium effects, when the relative price of l increases. The amplification is stronger when the price elasticity α or the elasticity of substitution σ are larger. Moreover, the amplification is stronger when non-homotheticities are more pronounced, as they accentuate reallocation towards necessities through income effects in Ω.[40]

5 Quantitative Analysis

In this section, we examine the quantitative importance of our theoretical results about increasing returns, non-homotheticities and price shocks for the optimal tax schedule. We first present the setting and main specifications (Section 5.1). We then implement our comparative static approach, studying a general first-order approximation (Section 5.2). Finally, we make additional parametric assumptions on non-homotheticities to study the optimal tax schedule and the feedback loops between redistribution and endogenous prices (Section 5.3).

5.1 Setting

Starting from the general model with multiple goods in Appendix E.1, we consider a standard additively separable specification (e.g., Saez (2001)):

$$U(z^*, z, \boldsymbol{p}, \theta) = v\left(z^*, \boldsymbol{p}\right) - \psi\left(\frac{z}{\theta}\right),$$

[37]The coefficient $\bar{s}_h \bar{s}_l$in the formula captures the decrease in the share of l relative to h as $d\bar{s}_l/\bar{s}_l - d\bar{s}_h/\bar{s}_h = d\bar{s}_l/(\bar{s}_h \bar{s}_l)$.

[38]Keeping taxes fixed, an increase in the relative price of l increases income by $\zeta \mathbb{E}_z(\bar{s}_l) - \partial_{z^*} e_l)$. Indeed, the change in real wage for household θ is $\hat{w}_t^*(\theta) = -\partial_{z^*} e_l(\theta)\, \hat{p}_l^* - \partial_{z^*} e_h(\theta)\, \hat{p}_h^* = -\left(\partial_{z^*} e_l(\theta) - \bar{s}_l\right) d\log \bar{p}_l$, so $\mathbb{E}\left(\hat{w}_t^*(\theta)\right) = -\left(\partial_{z^*} E_l - \bar{s}_l\right) d\log \bar{p}_l > 0$. Thus, the change in real wages is positive under our Assumption A2, $\partial_{z^*} E_l \le \bar{s}_l$. In addition, the decrease in tax rates further stimulates labor supply. The total gain in aggregate income is given by $-\zeta \mathbb{E}_z(\partial_{z^*} E_l - \bar{s}_l + \tau_l)$.

[39]We show in our working paper (Jaravel and Olivi (2024)) that the general equilibrium response can be decomposed into five channels, which are summarized and discussed in Appendix Figure A1.

[40]More precisely, comparing two economies A and B where $\partial_{z^*} e_h^A - \partial_{z^*} e_h^B$ is increasing and A and B are otherwise identical, then we have $\Omega^A \ge \Omega^B$ and the amplification through income effects is stronger.

where $\psi\left(\frac{z}{\theta}\right) \equiv \frac{1}{1+\frac{1}{\varepsilon_z}}\left(\frac{z}{\theta}\right)^{1+\frac{1}{\varepsilon_z}}$ is the isoelastic utility cost of earning z given ability θ, and $v\left(z^*, \mathbf{p}\right)$ is the indirect utility function given prices and disposable income. Following the nonparametric evidence of Kleven and Schultz (2014), we set $\varepsilon_z = 0.214$; for robustness, we consider $\varepsilon_z = 0.33$ as in the meta-analysis of Chetty (2012). We calibrate the skill distribution $f(\theta)$ nonparametrically to match the income distribution at the observed tax schedule, using data from Hendren (2020).

Returns to scale. As before, returns to scale are governed by the parameter α. There is an emerging empirical consensus that increasing demand leads to higher productivity and lower prices (in the long run), and recent papers provide causal estimates for α. Using a shift-share instrument with NielsenIQ data in the U.S., Jaravel (2019) finds that when demand increases by one percent, consumer prices for continued products fall by 0.42 percent. When accounting for changes in product variety, the consumer price index falls by 0.62 percentage points. Leveraging data on durable good industries in the Chinese manufacturing sector and an IV design based on potential market size, Beerli et al. (2020) estimate that increasing market size by one percent leads to a TFP increase of 0.46 percent. Using trade shocks as instruments, Bartelme et al. (2019) estimate sector-level economies of scale and find statistically significant scale elasticities in every 2-digit manufacturing sector, with an average of 0.13.[41] Given this range of estimates, we set $\alpha = 0.30$ in our baseline specification and study sensitivity.[42] For our comparative statics exercise below, we also analyze the case $\alpha = 0$ (linear production functions) which is the benchmark of the public finance literature and allows us to isolate the impact of non-homotheticities in consumption preferences.

For the comparative static analysis in Section 5.1, α can be viewed as the "local" returns to scale. When studying the optimal tax schedule in Section 5.3, we specify the global relationship between the price p_i of the good produced in sector i and equilibrium quantities in that sector, setting $p_i = \gamma_i Q_i^{-\alpha} \quad \forall i \in \mathcal{I}$. We use the observed equilibrium to calibrate the set of parameters γ_i, as discussed in Online Appendix D.1.2.

Preferences. We set the indirect utility function $v\left(z^*, \mathbf{p}\right)$ to be either homothetic or non-homothetic in the analysis below to isolate the quantitative impact of non-homotheticities on the optimal schedule.

A non-homothetic utility function introduces curvature in the agent's indirect utility from consumption, which affects the social marginal utility of disposable income. Therefore, we normalize the curvature of utility at fixed prices, so that we mechanically reach the same optimum with homothetic and non-homothetic utility under constant returns to scale.[43] This approach ensures that the comparison between

[41] Other papers provide empirical evidence for returns to scale in different settings. Acemoglu and Linn (2004) provide empirical evidence that market size influences entry of new drugs and U.S. pharmaceutical innovation. Weiss and Boppart (2013) show that TFP growth is higher in more income-elastic sectors, using national accounts data covering the entire U.S. economy. Analyzing NielsenIQ scanner data across local markets, Handbury (2019) finds that the products and prices offered in markets are correlated with local income-specific tastes. Focusing on housing and local amenities, Diamond (2016) and Couture et al. (2020) find that amenities adjust endogenously to an increase in local demand and lower the price index.

[42] The closest empirical evidence to discipline our model is provided by Jaravel (2019), who looks directly at consumer prices rather than TFP. For continued products, the estimate for α varies between 0.23 and 0.458, depending on the set of controls, and $\alpha = 0.30$ cannot be rejected in any of the specifications. The estimates are larger when product entry is accounted for, hovering between 0.38 and 0.67 depending on the specification.

[43] We work with a "deflated indirect utility function" $\widetilde{v}(z^*, \boldsymbol{p})$, defined such that

$$\widetilde{v}(z^*, \boldsymbol{p}) = v^{-1}(v(z^*, \mathbf{p}), \mathbf{p}_{CRS}),$$

where $\mathbf{p}_{CRS}$ are the prices prevailing under constant returns (which are normalized to one in the simulations, without loss of the homothetic and non-homothetic specifications captures the channel of interest, namely differences stemming from endogenous prices and their impact on the marginal utility of disposable income across the skill distribution, rather than assumptions about curvature *per se*. Our results are thus comparable to the benchmark models of Mirrlees (1971) and Saez (2001), with no additional curvature and no additional income effects absent returns to scale, despite the introduction of non-homothetic utility.

For the comparative static analysis, we directly use the formulas derived in Section 4. In partial equilibrium, we only need to know the local marginal propensities to consume across goods for agents across the income distribution, $\partial_{z^*} e_i$. We measure marginal propensities to consume non-parametrically from expenditure shares across 248 product categories. This data set covers the full consumption basket of American households, linking the CPI price data set to the consumption patterns of the Consumer Expenditure Survey (CEX) to the Consumer Price Index (CPI), following Jaravel (2019).

As shown in Proposition 4, the demand elasticity of substitution σ between products plays an important role for the feedback loops in general equilibrium. We take estimates from the literature as bounds for the elasticity of substitution between our product categories. Based on estimates of the elasticity of substitution between goods and services, two broad categories of consumption which are likely to be less substitutable than our 248 categories, we set $\sigma = 0.6$ as a lower bound (see Comin et al. (2021) and Cravino and Sotelo (2019)). Given estimates on the substitutability between products within the same detailed product category, we take $\sigma = 2$ as our upper bound (e.g., Broda and Weinstein (2006), Broda and Weinstein (2010), DellaVigna and Gentzkow (2019), and Handbury (2019)).

To study the optimal tax schedule beyond the comparative static approach, we need parametric assumptions on the utility function. As further described in Subsection 5.3.2, we use non-homothetic CES preferences as in Hanoch (1975), Matsuyama (2019), and Comin et al. (2021).

Social preferences for redistribution. For the comparative static approach in Section 5.2, the formulas derived in Section 4 show that social preferences for redistribution can be recovered from the initial tax schedule. Taking the observed tax schedule as optimal obviates the need for specifying the social welfare function explicitly. For the analysis of the optimal tax schedule in Section 5.3, the planner's social welfare function, $G\left(U(\theta, \mathbf{p})\right)$, is assumed to be CRRA, with a relative risk aversion coefficient of one in our baseline specification and 0.5 for sensitivity.

5.2 Comparative Statics

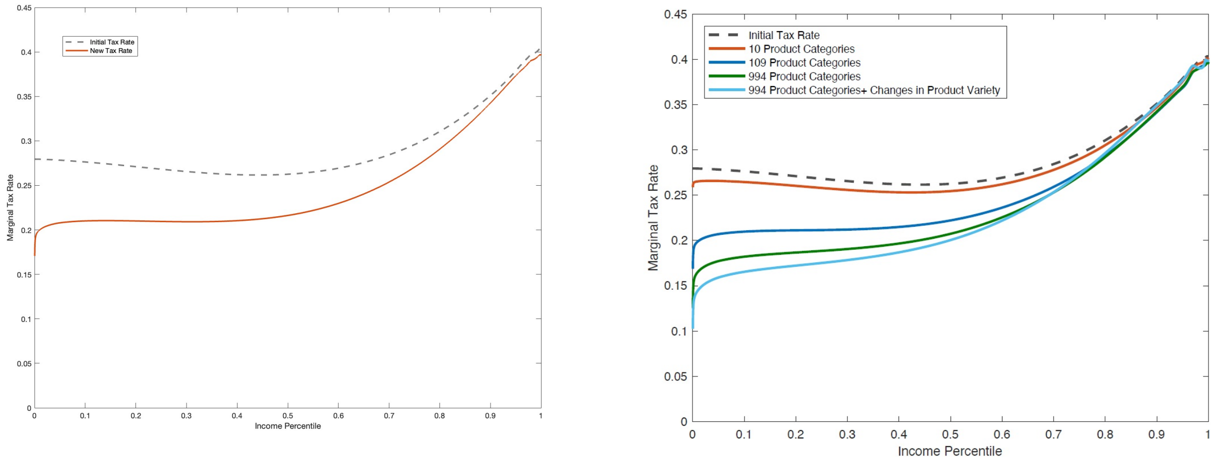
Using the comparative static approach introduced in Section 4, we now examine the quantitative response of the tax schedule to exogenous price shocks.

Starting from the observed tax schedule, we implement the formulas from Section 4 characterizing the response to a price change. We obtain observed price shocks for the period 2004 to 2015 over 248 product categories covering the full consumption basket of American households, linking the CPI price data set to the consumption patterns of the CEX. Empirically, inflation is lower in product categories with higher income elasticities: how large is the impact on the optimal tax schedule? To isolate the role of non-homotheticities, we first assume linear production functions ($\alpha = 0$, as in Section 4.2). We then

generality). We have $\widetilde{v}(z^*, \mathbf{p}_{CRS}) = z^*$,which is identical to the homothetic specification. Online Appendix D.2.2 discusses the properties of the deflated non-homothetic indirect utility function.

Figure 1 Optimal Tax Schedule and Observed Price Shocks, Linear Production Functions

A. CEX data B. NielsenIQ scanner data



Notes: both panels of this figure focus on the partial equilibrium response of the optimal tax schedule to price shocks, as in Proposition 2. The initial tax schedule is taken from Hendren (2020). In Panel A, the CEX-CPI data set is used to compute inflation rates from 2004 to 2015 across the income distribution. Panel B uses the NielsenIQ scanner data at different levels of aggregation to measure differences in inflation rates across the income distribution between 2004 and 2015. The adjustment for product variety uses a CES price index.

explore how general equilibrium effects alter this response under increasing returns to scale ($\alpha > 0$, as in Section 4.3).

Main results, linear production functions. Panel A of Figure 1 shows the changes in the optimal tax schedule in response to observed price shocks from 2004 to 2015, using the full set of 248 product categories from the CEX. Cumulative inflation during this period ranges from 29.4% of the distribution at the bottom of the income distribution to 24.5% at the top (Online Appendix Figure A2.A). As in Proposition 2, we focus on characterizing the response of the optimal tax schedule to changes in relative prices, keeping the average price level constant: relative prices rise by 3.3% at the bottom of the income distribution and fall by 1.6% at the top (Online Appendix Figure A2.B).[44]

[44]The literature has documented a long-run trend of inflation inequality in the United States, with a rate of divergence in relative price indices across the income distribution similar to the specific period we study here. While in our data the annual inflation difference is 37 basis points between the bottom and the top of the income distribution between 2004 and 2015, Jaravel (2024), who use prices indices that follow exactly the same procedure as the official CPI, obtains an annual inflation gap of 41 basis points between 2002 and 2024. Jaravel and Lashkari (2023) find that, on average over the 1955–2019 period, the annual inflation rate was about 35 basis points lower for the top relative to the bottom of the income distribution. While the inflation inequality trends are sustained over the long run, the patterns differ in certain periods. Garner et al. (1996) found there was no meaningful inflation inequality between 1984 and 1994. Jaravel (2024) shows that the patterns of inflation inequality were also different during the Covid-19 pandemic: between May 2020 and May 2022, the cumulative inflation rates were inverse U-shaped, increasing from 13% at the bottom of the income distribution to 14.7% for the middle class, and falling back to 13.5% at the top of the income distribution. While long-term inflation inequality in countries other than the United States has not been studied systematically, a growing literature studies inflation inequality around the world in response to exchange rate shocks, which heterogeneous findings depending on the context. Cravino and Levchenko (2017) find that the 1994 Mexican devaluation fueled inflation inequality: two years after the devaluation, the cost of living for the bottom income decile rose over 50% more than the cost of living for the top income decile. In contrast, Breinlich et al. (2022) find no evidence of inflation inequality when studying the depreciation of the pound sterling caused by the Brexit referendum: the price indices of income groups across the distribution were similarly affected by the shock. Finally, Auer et al. (2024) find that the 2015 Swiss Franc appreciation disproportionately benefited low-income households, whose price indices fell relatively more than those of high-income households.

We compute the tax schedule response using the formula in Proposition 2, which remains valid in an economy with n sectors, as shown in Appendix A.2.4. Under linear production, the tax response depends only on marginal propensities to spend and captures Channels #1 and #2 discussed in Section 4.2. This allows us to isolate the impact of non-homothetic preferences. The response is substantial: marginal tax rates fall by about 10pp at the bottom of the income distribution; the marginal tax rates gradually converge back to the observed tax schedule at the top. Indeed, because inflation is lower in the product categories for which higher-skill agents have a higher marginal propensity to consume, it is optimal for the planner to redistribute toward them, which can be done most efficiently by reducing marginal tax rates at the bottom of the income distribution.[45] This result shows that inflation inequality generates a sizable regressive response of the tax schedule in partial equilibrium.

To understand the magnitude of the tax response at the bottom, note that for an arbitrary price change, the formula of Proposition 2 can be re-expressed as:[46]

$$dT'(\underline{\theta}) = -\underbrace{\frac{g(\underline{\theta})}{g(\underline{\theta})-1}T'\left(1-T'\right)}_{\text{Preferences for redistribution}}\underbrace{\left(dlnp\left(\underline{\theta}\right)-dln\bar{p}\right)}_{\text{Marginal Price Index}},$$

where the marginal price index for household $\underline{\theta}$ is defined as $dlnp\left(\underline{\theta}\right) = \sum_{k=1}^{n} \partial_{z^*} e_k\left(\underline{\theta}\right) dlnp_k$, and the average marginal price index is $dln\bar{p} = \sum_{k=1}^{n} \partial_{z^*} E_k dlnp_k$.

Using this formula, we can plug in values to understand what drives the sizable tax response at the bottom of the income distribution. First, we can use observed tax rates to back out preferences for redistribution. In the data (Hendren (2020)), the pre-shock marginal tax rate range from 27% at the bottom to 40% at the top. These rates are relatively low compared to the theoretical optimal tax rates obtained in Saez (2001) using social welfare function with a CRRA coefficient of 1 (ranging from 80% at the bottom to 60% at the top). The observed tax rates thus imply limited redistribution to low-income households, and hence a relatively low Pareto weight at the bottom, which we estimate to be $g(\underline{\theta}) = 1.1$.[47] This value indicates that the planner values the bottom-income agent only 10% more than the average. Substituting into the formula, the redistribution term is approximately equal to 2.2.

Second, to gauge the role of non-homothetic preferences, we approximate the change in marginal price index using the change in relative price index observed in the CEX.[48] Given that prices increase by 3.3% more at the bottom of the income distribution, relative to average, the formula implies a fall of the bottom tax rate of 7.2 percentage points.[49] This sizable response, of the same order of magnitude as the one reported in Figure 1, is primarily driven by the low social preference for redistribution – i.e., $g(\underline{\theta})$ is close to one. Increasing $g(\underline{\theta})$ to 1.3 would more than halve the tax rate response to price changes.

[45] This mechanism is standard: high marginal tax rates at the bottom are paid by all agents earning higher levels of income, without distorting their marginal incentives to work, and all revenue is rebated to the lower-income households through the intercept of the tax schedule.

[46] See Appendix A.2.2 for the derivation.

[47] Hendren (2020) reports that Pareto weights at the bottom of the distribution are between 1.1 and 1.2 depending on the elasticity ε. Note that the Pareto weight can be recovered from the optimal tax formula of Proposition 1, setting $\alpha = 0$, $\frac{d}{dz}\left\{z\tilde{\zeta}T'/\left(1-T'\right)f(z)\right\} = -\left(1-g\left(z\right)\right)f\left(z\right)$.

[48] The relative price index is measured in terms of budget shares rather than marginal budget shares. Figure A2.B reports changes in prices indices across the income distribution, relative to average.

[49] Plugging into the formula above, the calculation is: $dT'(\underline{\theta}) = -\frac{1.1}{1.1-1}\cdot 0.27\cdot(1-0.27)\cdot 3.3 = -7.2$.

At the top of the income distribution, the response of the tax rate is given by the same formula at $\bar{\theta}$:

$$dT'(\bar{\theta}) = - \underbrace{\frac{g(\bar{\theta})}{g(\bar{\theta}) - 1} T' \left(1 - T'\right)}_{\text{Preferences for redistribution}} \underbrace{\left(dlnp\left(\bar{\theta}\right) - dln\bar{p}\right)}_{\text{Marginal Price Index}}.$$

Here, the tax rate is $T' = 40\%$, the Pareto weight is lower, with $g(\bar{\theta}) = 0.65$, and relative prices fall by 1.6% (Figure A2.B). As a result, the sensitivity of the tax rate is one order of magnitude lower: the tax rate decreases by 0.7 percentage points. If $g(\bar{\theta})$ was closer to one (e.g., $g(\bar{\theta}) = 0.9$), the sensitivity would be similar to that observed at the bottom.

In terms of welfare, we find in Appendix Figure A4 that households at the bottom are more affected by the response of taxes to price shocks than by the price shocks themselves. Specifically, they would be willing to sacrifice 35% of their income to avoid the tax change, compared to just 2%[50] to avoid the price change itself. This highlights that the regressive impact of the optimal tax adjustment can be an order of magnitude larger than the direct effect of inflation inequality.[51]

Response to inflation inequality in scanner data. Using NielsenIQ scanner data, panel B of Figure 1 shows that it is important to measure changes in prices at a detailed level to draw the implications of price changes for the optimal tax schedule. We repeat the previous exercise using the NielsenIQ Homescan Consumer Panel data set instead of the CEX-CPI data. The NielsenIQ data set covers only fast-moving consumer goods (about 15% of total expenditure and 40% of expenditures on goods), but it has the advantage of being available at a much higher level of granularity than the product categories from the CEX-CPI linked data set and allows for the measurement of changes in product variety. We use the NielsenIQ data to illustrate the role of aggregation bias and product variety for the optimal tax policy response to inflation inequality. Price changes and changes in product variety are measured from 2004 to 2015 at different levels of aggregation.

We first focus on price indices computed at different levels of aggregation, using products available across consecutive years, and that do not account for changes in product variety. With the 994 most detailed product categories, called "product modules", we find that it is optimal for marginal tax rates to fall by about 12 percentage points at the bottom of the income distribution in response to the price changes observed between 2004 and 2015. With 109 larger product categories, called "product groups", the differences in inflation across the income distribution are attenuated: the fall in the tax schedule is only about 7.5 percentage points at the bottom of the distribution. With the ten broad "departments", measured inflation inequality is much smaller and the fall in tax rates is under 2 percentage points. These differences would be amplified further in general equilibrium, accounting for endogenous price changes and changes in demand.

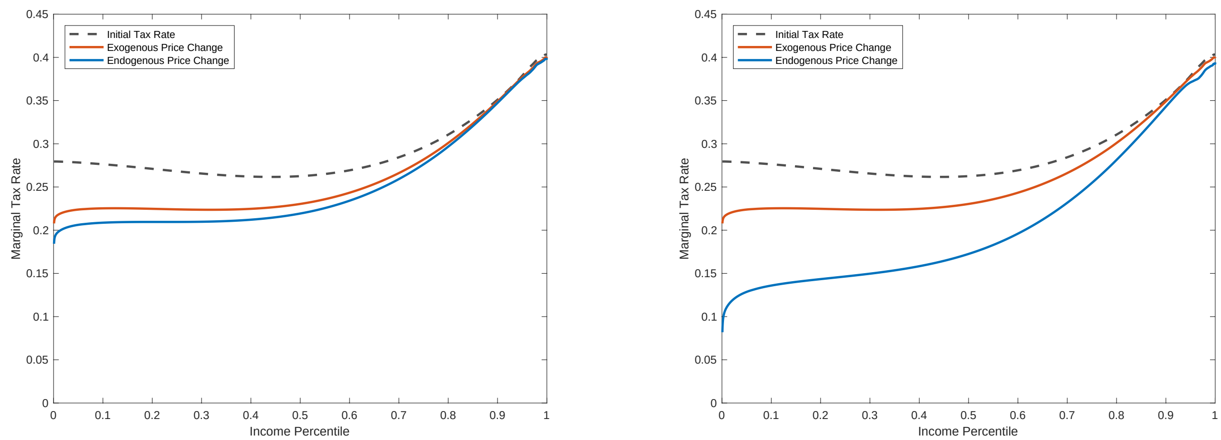
Furthermore, we introduce a correction for changes in product variety, using a CES price index as

[50] As the relative price of necessities fall, we find in our calculation that labor supply increases on average. This implies that, at the initial tax rates, more revenue is collected by the government and this additional revenue is rebated to households. As a result the cost of the price change for households at the bottom of the distribution (2%) is lower than the 3.3% observed in the CEX (Figure A2.B).

[51] The willingness to pay to avoid the tax reform is large at the bottom of the distribution. The reason is that the intercept of the tax schedule and therefore the income of households at the bottom of the distribution is small. Thus, a lower average tax rate leads to a large fall in income for households at the bottom.

Figure 2 Optimal Tax Schedule and Observed Price Shocks, Nonlinear Production Functions

A. $\sigma = 0.6$ B. $\sigma = 2$



Notes: the IRS parameter is set to $\alpha = 0.3$ and the labor supply elasticity to $\varepsilon = 0.21$. The CEX-CPI data set is used in both panels to measure the price shocks, obtained by computing inflation rates from 2004 to 2015 across the income distribution. The initial tax schedule is taken from Hendren (2020). The "exogenous price change" results, depicted in red, are obtained by applying Proposition 2. The "endogenous price change" results, depicted in blue, follow from Proposition 4 and account for all channels in Figure A1.

in Feenstra (1994) and Broda and Weinstein (2010). In the data, product variety expands faster in product categories purchased by high-income households, which further reduces the price index faced by high-skill agents. Consequently, the fall in optimal marginal tax rates is amplified. At the bottom of the distribution, marginal tax rates fall by an additional 2.5 percentage points.[52]

Main results, non-linear production functions. We now turn to the case of non linear production functions. We set $\alpha = 0.30$ and consider the monopolistic case. Figure 2 presents the results. We report the changes in the tax schedule in response to price shocks, depending on the value of σ and contrasting the responses in partial and general equilibrium.

Using the first formula of Proposition 4, we compute the partial equilibrium response. This response – shown in red and labeled "exogenous price change" in the figure – is independent of σ and only captures Channels #1 and #2. The results are reported in red in the figure, with the label "exogenous price change". The pattern of the tax response is qualitatively similar to the case with linear production ($\alpha = 0$, shown on Figure 1), but quantitatively muted: marginal tax rates at the bottom fall by around 6pp, compared to 10pp in Figure 1. This muted response arises from the interaction between returns to scale and the strength of redistribution preferences. Introducing $\alpha > 0$ implies that stronger redistribution preferences are required to rationalize the same observed tax schedule. To understand why, we have to come back to Proposition 1: with $\alpha > 0$, a wage subsidy is necessary to incentivize labor supply, exploiting increasing returns to scale to reduce prices. This shifts the effective retention rate to $(1 - \alpha)(1 - T')$, which is lower than in the linear case. As a result, even at the same observed tax rate, the planner is implicitly favoring low-income households more under non-linear production. Empirically, we find that this stronger redistribution motive is reflected in a higher Pareto weight at the bottom: $g(\underline{\theta}) = 1.2 > 1.1$,

[52]Online Appendix Figure A3 reports inflation inequality patterns in the NielsenIQ data, with and without changes in product variety.

which dampens the impact of price shocks on the optimal tax rate relative to the linear case. Thus, under increasing returns, inflation inequality still generates regressive tax responses, but their magnitude is partially offset by a stronger underlying redistributive motive.

Moreover, using the second part of Proposition 4, we find that the response of the tax schedule is amplified in general equilibrium. The results are reported in blue in the figure, with the label "endogenous price change". With $\sigma = 0.6$ in Panel A, the planner reduces marginal tax rates by an additional two percentage points at the bottom of the income distribution. With $\sigma = 2$ in Panel B, the amplification is much larger and the optimal marginal tax rate is reduced to only 10% at the bottom of the income distribution. To understand the magnitude of the amplification, recall from Proposition 4 that the general equilibrium response is the partial equilibrium response scaled by $(1 - \alpha(\sigma + \Omega))^{-1}$. In our estimation, Ω is small compared to σ, so this scaling term is well approximated by $(1 - \alpha\sigma)^{-1}$ and is equal to 1.22 with $\sigma = 0.6$ and 2.5 with $\sigma = 2$, i.e the amplification ranges from about 20% to 150% depend on the elasticity of substitution. Intuitively, in general equilibrium consumers reallocate their expenditures toward the goods that become relatively cheaper, which amplifies the price changes through increasing returns and further reduces the relative price of products with a high income elasticity. These endogenous price changes create an additional motive for the social planner to redistribute toward higher-skill agents, which leads to further price changes, and so on. These results show that the general equilibrium response of prices plays a quantitatively important role for optimal tax policy.[53]

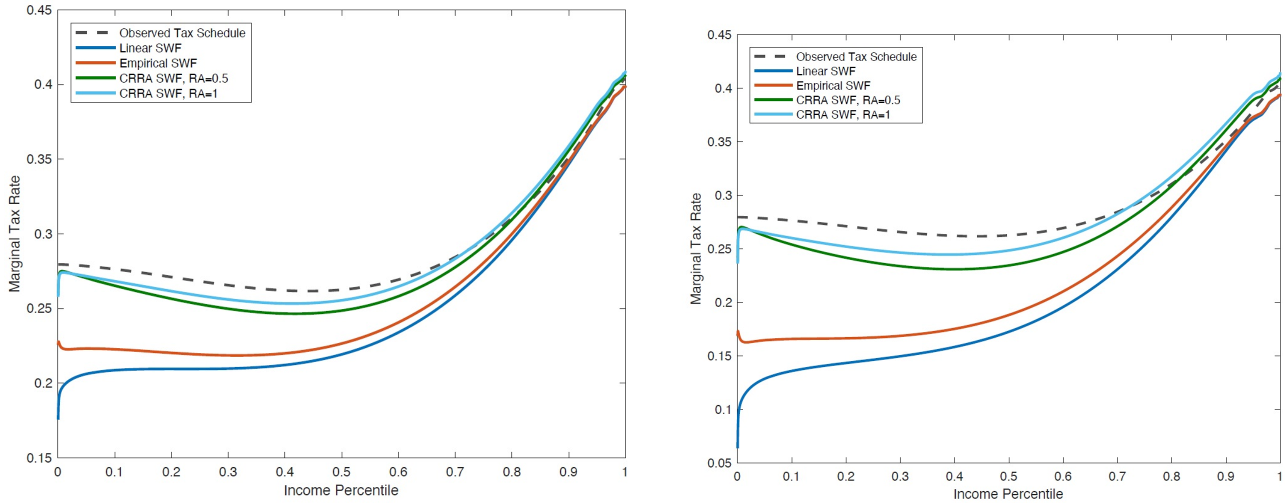
The role of non-linear social preferences. Next, Figure 3 quantifies the role of the curvature of the social welfare function for the response of the tax schedule, illustrating the theoretical insights from Proposition 3. While Figure 2 gives the results with linear social welfare weights, we now introduce curvature by taking the inverse optimum weights at the observed tax schedule as our empirical social welfare function. Specifically, with a linear social welfare function $G(V, \theta) = \lambda(\theta)V$, $\lambda(\theta)$ is chosen such that the observed schedule is optimal. With the empirical social welfare function, $G(V, \theta) = G(V)$, G is directly chosen such that the observed schedule is optimal and is not allowed to depend on θ. Figure 3 reports the results with a linear social welfare function (in blue) and the empirical social welfare function (in red). The degree of optimal redistribution toward the rich is slightly muted with the empirical social welfare function, because the social value of redistributing income toward high-skill agents falls endogenously as they get more transfers. With $\sigma = 0.6$, marginal tax rates increase by about 2.5pp at the bottom of the income distribution with the empirical nonlinear social welfare function, compared with the baseline case with linear Pareto weights. With $\sigma = 2$ in Panel B, optimal marginal tax rates are about 7pp larger than with linear Pareto weights, but the fall in optimal marginal tax rates induced by price shocks remains sizable.

To further investigate the role of the curvature of the social welfare function, we specify the social welfare function as $G(V, \theta) = \lambda(\theta)\frac{V^{1-\gamma}}{1-\gamma}$, with γ the CRRA coefficient and setting $\lambda(\theta)$ such that the

[53] In the Diamond-Mirrlees case, the partial equilibrium response are the same as in Figure 1. The amplification of the response through general equilibrium effects is the same as in the monopolistic case presented here: the partial equilibrium response is scaled by 1.22 with $\sigma = 0.6$ and 2.5 with $\sigma = 2$. A limitation of the results with endogenous price changes presented in Figure 2 is that we use a single elasticity of substitution across all goods. The existing literature offers no widely accepted estimates of heterogeneous elasticities of substitution across detailed product categories: developing such estimates would be a fruitful avenue in future work. This limitation does not affect the results we obtain in partial equilibrium or when production functions are linear.

Figure 3 The Role of the Curvature of the Social Welfare Function

A. $\sigma = 0.6$ B. $\sigma = 2$



Notes: the IRS parameter is set to $\alpha = 0.3$ and the labor supply elasticity to $\varepsilon = 0.21$. The CEX-CPI data set is used in both panels to measure the price shocks, obtained by computing inflation rates from 2004 to 2015 across the income distribution. The initial tax schedule is taken from Hendren (2020). Each panel reports the optimal tax schedule under different social welfare function: the linear social welfare function from Figure 2, the empirical nonlinear social welfare function described in the main text, and social welfare functions with a coefficient of relative risk aversion (CRRA) of 0.5 or 1. The dark blue lines in the figure are the same as in Figure 2.

observed tax schedule is optimal.[54] With stronger curvature, e.g. a social welfare function with a CRRA coefficient of 0.5 or 1, redistribution toward the rich falls further. The result that marginal tax rates fall in response to the price shocks is attenuated but not overturned: the marginal tax rates remain about one to five percentage points below the observed tax schedule in the first six deciles, and gradually converge to the observed schedule at higher percentiles.

Overall, these results show that non-linearities in social preferences for redistribution may play a significant role for the optimal response of the tax schedule. With the empirical non-linear social welfare function, the fall in taxes remains substantial for both values of σ.[55] In unreported analyses, we find that the planner continues to increase (rather than decrease) marginal tax rates in response to lower inflation for high-income households even when the curvature of the social welfare function is much higher, e.g. with a CRRA coefficient of 10.

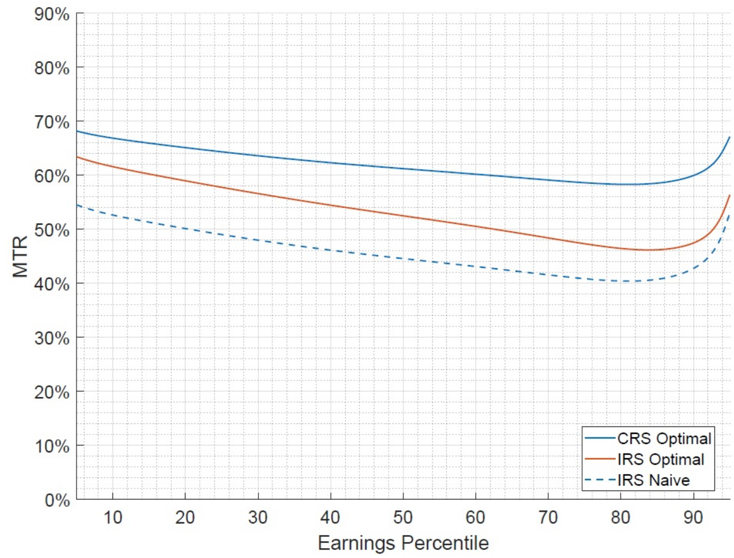
5.3 Optimal Tax Schedule

We now analyze the quantitative importance of increasing returns to scale and non-homotheticities for optimal tax rates and welfare across the skill distribution. We first document the impact of increasing returns to scale in a homothetic model, then isolate the impact of non-homotheticities. Finally, we study the response of the tax schedule to exogenous price shocks. By introducing parametric assumptions on preferences, these analyses are complementary with the first-order approximations of Section 5.2, because they characterize how our new channels affect the optimum when accounting for potential non-linearities. Online Appendix D.4 provides a complete discussion of the solution algorithm.

[54] With $\gamma = 1$, we set $G(V, \theta) = \lambda(\theta)\log(V)$.

[55] Online Appendix Figure A5 shows that similar results apply with alternative values of the labor supply elasticity.

Figure 4 Returns to Scale and the Optimal Tax Schedule



Notes: This figure plots optimal marginal tax rates under constant returns to scale (CRS, $\alpha = 0$) and increasing returns to scale (IRS, $\alpha = 0.3$). The social welfare function is logarithmic (CRRA=1) and the elasticity of labor supply is $\varepsilon = 0.21$. With increasing returns, the "naive" correction uses the formula $1 - T'_{NAIVE}(\theta) = \frac{1}{1-\alpha}\left(1 - T'_{CRS}(\theta)\right)$. The optimal tax schedule solves the full optimization problem, accounting for endogenous changes in the value of redistribution across the income distribution.

5.3.1 The Interaction between Returns to Scale and Redistributive Motives

We first investigate the impact of returns to scale on the optimal tax schedule under homothetic utility, i.e. with $v\left(z^*, \mathbf{p}\right) \equiv \frac{z^*}{p}$. We consider a setting with a single sector, such that α can be interpreted as "aggregate" returns to scale. With aggregate returns α, the "naive" interpretation of Proposition 1 is that, relative to the CRS tax schedule, the planner should uniformly subsidize nominal wages $1 - T'$ at a constant rate $1/(1-\alpha)$ throughout the distribution.

The solid blue line in Figure 4 shows the baseline optimal tax schedule under CRS and a logarithmic social welfare function. The optimal marginal tax rates start around 68% at the bottom of the income distribution, fall gradually to 58% at the 80th percentile, and then increase toward 68% at the top. The dashed blue line depicts the tax schedule with the naive correction for increasing returns to scale, with $\alpha = 0.30$, whereby the net-of-tax wage is increased by 43% everywhere. This result already conveys that it is important to take into account returns to scale for optimal tax design: the effect on optimal tax rates is large.

The solid red line shows the optimal tax schedule under returns to scale with a logarithmic social welfare function. The fall in marginal tax rates is smaller than with the naive correction. This result shows that the curvature of the social welfare function plays a quantitatively important role in determining the correction for increasing returns to scale, i.e. there is an important interaction with redistributive motives. It is optimal for the cost of the work subsidy to be predominantly paid by high-skill agents, hence marginal tax rates do not fall as much as with the naive correction. In Appendix Figure A6, we show that the interaction remains quantitatively large with other parameter values for the labor supply elasticity and the CRRA coefficient of the social welfare function.

By contrast, with linear Pareto welfare weights, set to match welfare weights at the CRS optimum, the "naive" correction is correct. To isolate the role of non-homotheticities independently of the curvature of the social welfare function, we take the specification with Pareto weights as our baseline in the next subsections. The Pareto weights are set as $\lambda(\theta) \equiv (U_{optim}(\theta))^{-\widetilde{\sigma}}$, where $U_{optim}(\theta)$ is the solution of the optimal taxation problem with homothetic utility, constant returns to scale ($\alpha = 0$), and the CRRA parameter $\widetilde{\sigma}$ for the social welfare function.

5.3.2 The Role of Non-Homotheticities

We now turn to a specification with non-homothetic utility, using non-homothetic CES (nhCES) preferences as in Hanoch (1975), Matsuyama (2019), and Comin et al. (2021).

Parametric assumptions on non-homothetic preferences. The indirect utility function $v(z^*, \boldsymbol{p})$ is given by $v \equiv v(z^*, \boldsymbol{p}) \equiv F(\mathbf{Q})$, where $\mathbf{Q}$ is the consumption vector of the agent over the set of products $i \in \mathcal{I}$. Indirect utility v is implicitly defined by:

$$\sum_{i \in \mathcal{I}} (\Omega_i v^{\varepsilon_i})^{\frac{1}{\sigma}} Q_i^{\frac{\sigma-1}{\sigma}} = 1.$$

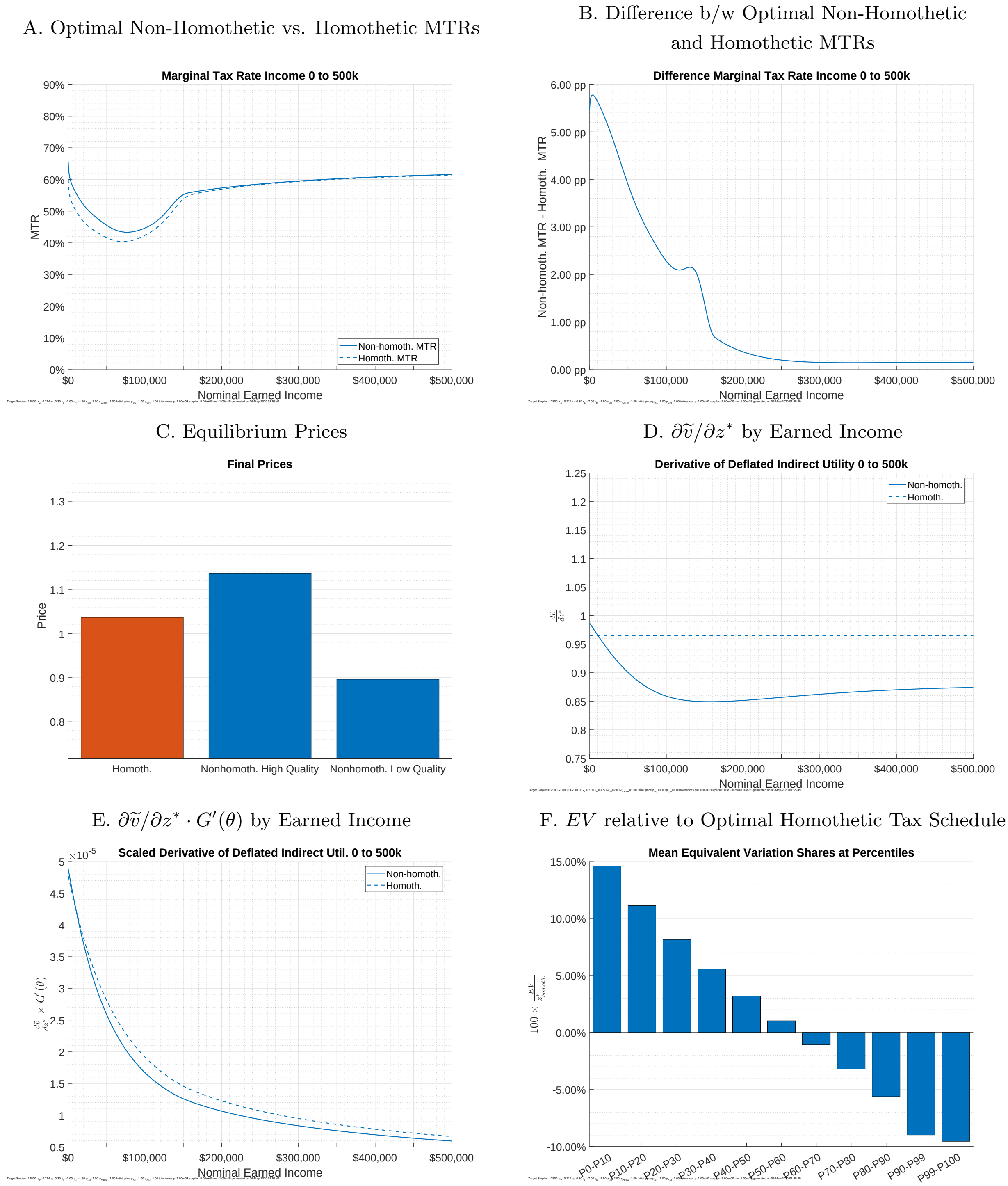
NhCES preferences have convenient features, in particular $\frac{\partial \log(Q_i/Q_j)}{\partial \log(v)} = (\varepsilon_i - \varepsilon_j)$ and $\frac{\partial \log(Q_i/Q_j)}{\partial \log(p_j/p_i)} = \sigma \quad \forall i, j \in \mathcal{I}$. This tractable specification allows us to separately examine the impact on the tax schedule of the "utility elasticities" $\{\varepsilon\}_{i \in \mathcal{I}}$, which govern non-homothetic spending patterns, and the elasticity of substitution σ.

For tractability, in our calibration we consider two products, labeled "high quality" and "low quality" products. In line with evidence on the substitutability between products within the same detailed product category (Broda and Weinstein (2006), Broda and Weinstein (2010), DellaVigna and Gentzkow (2019), and Handbury (2019)), we set $\sigma = 2$. We then specify the elasticities $\{\varepsilon\}_{H,L}$ to match the dissimilarity index of consumption shares observed across the income distribution in the Consumer Expenditure Survey in 2014. We compute the dissimilarity index at the level of the product categories available in the CEX interview files, called universal classification codes (UCC). We focus on 2014 as the data on the observed tax schedule from Hendren (2020) is available for that year. We obtain $\varepsilon_L = -7$ and $\varepsilon_H = -1.5$, implying that low-income households have a large marginal propensity to spend on the low-quality goods.

Main results. Figure 5 characterizes the impact of non-homotheticities in our baseline specification relative to the homothetic case, with $\alpha = 0.3$ and Pareto weights from the logarithmic social welfare function. Panels A and B show the effect of introducing non-homotheticities on optimal marginal tax rates. Due to non-homotheticities, marginal tax rates increase over the full range of the income distribution. The increase is larger at the bottom of the income distribution, with an increase in marginal tax rates of about 6pp for levels of earned income below \$20,000. The increase is about 2pp at an income level of \$100,000, and then gradually decreases, reaching levels close to zero above \$300,000. Non-homoheticities thus have a significant quantitative impact on optimal marginal tax rates.

Panels C through E of Figure 5 investigate the mechanism explaining the change in marginal tax rates, which operates through the change in equilibrium prices and in the marginal utility of redistribution across the skill distribution. Panel C reports the equilibrium prices, normalized to one at the observed

Figure 5 The Response of the Optimal Tax Schedule to Non-Homotheticities
($\alpha = 0.3$, $\varepsilon_z = 0.21$, Pareto weights from SWF CRRA=1)



Notes: The quantitative model uses Pareto weights computed at the optimal homothetic tax schedule obtained under a social welfare function with CRRA=1. Panel A reports the optimal tax schedule with homothetic and non-homothetic preferences. Panel B plots the difference between these two tax schedules across the income distribution. Panel C reports the equilibrium prices in the homothetic and non-homothetic models. Panel D shows the derivative of (deflated) indirect utility, $d\widetilde{v}(z^*, \boldsymbol{p})/dz^*$, across the income distribution. Panel E reports this derivative scaled by social preferences for redistribution. Panel F report the equivalent variation (EV) giving the willingness to pay of agents for the non-homothetic optimal tax schedule instead of the homothetic optimal tax schedule, expressed as a percentage of their disposable income under the homothetic optimal tax schedule.

tax schedule. In the homothetic specification with increasing returns, the price index increases by about 3.6% at the optimal tax schedule, because preferences for redistribution induce higher taxes than at the observed schedule, which reduces labor supply and market size and thus drives an increase in the price. With non-homotheticities, prices of the high-quality and low-quality products diverge: the price of the high-quality good increases by 14%, while the low-quality product becomes 10% cheaper. Indeed, additional redistribution (relative to the observed schedule) leads to an increase in the relative market size of the product which has a higher marginal propensity to consume from low-income households, i.e. the low-quality product in our specification. This result shows that the response of the optimal tax schedule to non-homotheticities lead to large endogenous price changes in equilibrium.

Panel D shows that the induced change in the marginal utility of disposable income across agents is substantial. While under homothetic utility the marginal utility is about 0.965 ($= 1/p$) throughout the distribution, with non-homotheticities the marginal utility is 0.99 at the bottom, falls gradually to 0.85 around \$150,000, and then increases slightly. The fall in marginal utility is largest for the agents with the highest marginal propensity to consume on the high-quality good, which in equilibrium occurs for earned income levels around \$150,000 in our simulation. Panel E combines each agent's marginal utility of disposable income with Pareto weights and shows a steeper decline in welfare weights across the distribution with the non-homothetic specification, because of the price effects.

Finally, panel F summarizes the willingness to pay of agents for the optimal tax schedule under non-homothetic preferences, relative to the optimal schedule under homothetic preferences.[56] The equivalent variation is close to 15% in the bottom decile of the income distribution and decreases monotonically throughout the distribution, turning negative in the seventh income decile. In the top decile, the welfare loss from the new schedule, and its induced price effects, is about 9%. These estimates show that adjusting the tax schedule for non-homotheticities generates substantial distributional effects, with large welfare gains at the bottom of the distribution. Although panels A and B depicted an increase in marginal tax rates at the bottom of the distribution, overall the change in the tax schedule benefits low-income households more. Indeed, setting higher marginal tax rates at the bottom of the income distribution raises the overall amount of redistribution in a more efficient way than increasing marginal tax rates at the top, and the induced price effects benefit agents with a high average spending share on the low-quality product.

Thus, the baseline simulation shows that non-homotheticities can have meaningful quantitative implications for optimal taxation. The results account for all feedback loops between the desirability of redistributing more and the induced price changes in general equilibrium (Proposition 4). As the relative price of the low-quality product decreases, it is optimal to redistribute more to those with a higher marginal propensity to consume, which induces further tax changes and changes in labor supply, etc. The strength of these feedback loops depend on the parameters governing increasing returns and social preferences for redistribution, which we turn to next.

[56] We study the equivalent variation defined by:

$$\widetilde{v}\left(z_H^*(\theta) + EV(\theta), \mathbf{p}_H\right) - \psi\left(\frac{z_H(\theta)}{\theta}\right) = u_{NH}(\theta),$$

where H denotes the equilibrium under the optimal tax schedule with homothetic preferences, while NH corresponds to the equilibrium with non-homothetic preferences.

Sensitivity to increasing returns. Figure 6 reports the simulation results with larger increasing returns, setting $\alpha = 0.4$, close to the baseline estimate of 0.42 in Jaravel (2019). The results and channels described for the baseline specification are all amplified by the larger increasing returns. Optimal marginal tax rates increase by 11.5 percentage points at the bottom of the income distribution (panel B). The price of the high quality good increases by 22%, while the price of the low-quality good falls by 18% (panel C). The new tax schedule and the induced price effects create welfare gains of 35% at the bottom of the skill distribution, and welfare losses of 16% at the top (panel F).

Sensitivity to preferences for redistribution. With $\alpha = 0.30$, Figure 7 investigates the effects of nonhomotheticicities when preferences for redistribution are weaker. The Pareto weights are taken from the optimal schedule with constant returns to scale and a social welfare function with a CRRA coefficient of 0.5, rather than 1 as previously.

With this specification, the impact of non-homotheticicities on the optimal tax schedule is muted. The marginal tax rate increases by 3.75pp at the bottom of the distribution (panel B), the price of the high quality product increases by about 3.75%, while the price of the low-quality product falls by about 3.5% (panel C). The willingness to pay for the tax schedule accounting for non-homotheticicities remains meaningful, especially at the bottom of the income distribution, with a welfare gain of 12% in the bottom decile and a welfare loss of about 3% in the top decile (panel F).

These results illustrate the interplay between social preferences for redistribution and endogenous prices. A weaker taste for redistribution endogenously leads to smaller changes in market size, hence smaller price changes in equilibrium and a smaller adjustment to optimal marginal tax rates.

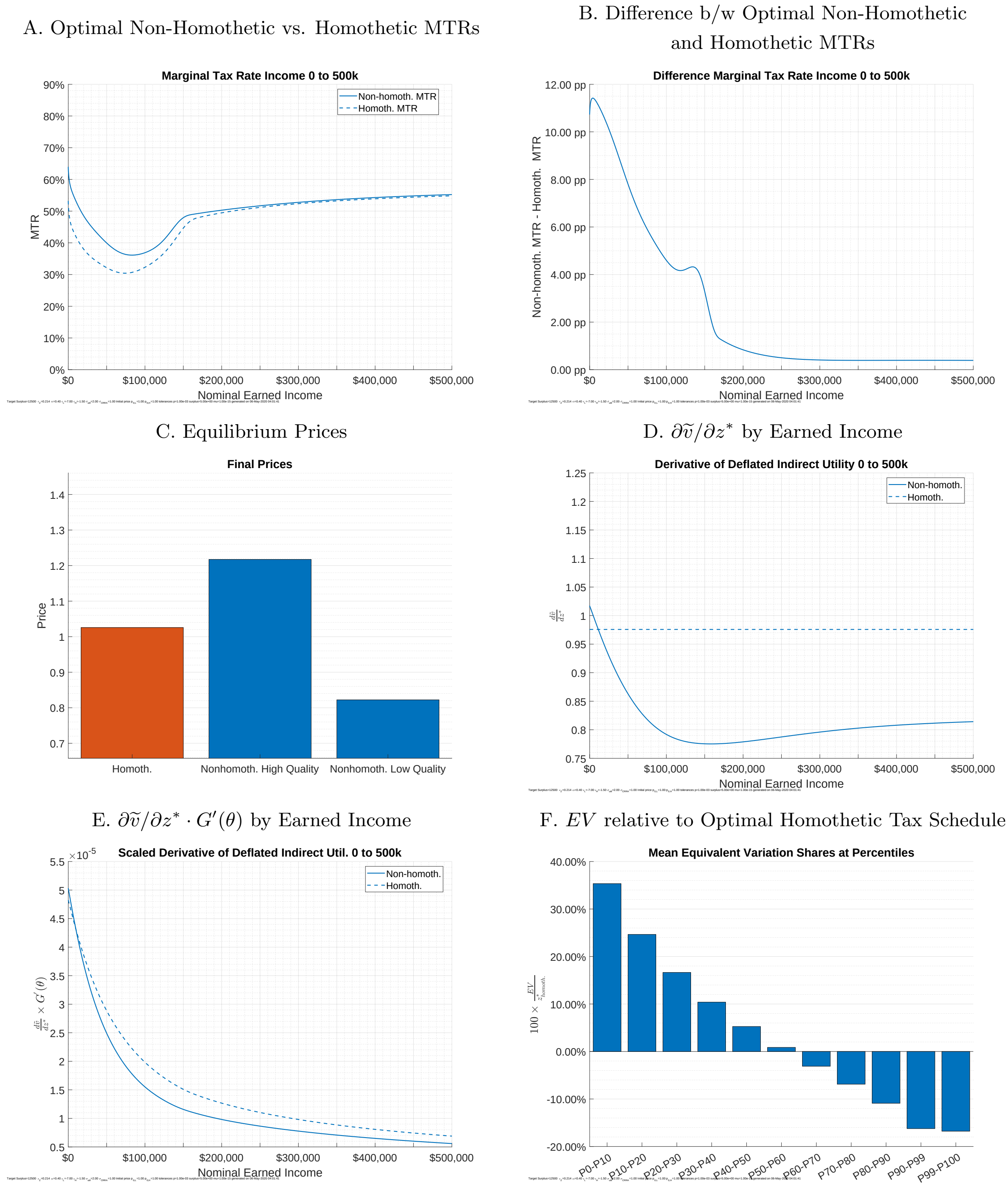
The impact of exogenous price shocks. Finally, we analyze how price shocks (e.g., from productivity shocks) affect the optimal tax schedule. We already characterized this response using a first-order approximation in Section 5.2. In Online Appendix B.1, we present complementary results with no approximation, accounting for feedback loops created by large price changes. We again find that exogenous price shocks can have a large impact on the optimal tax schedule, and that there are important amplification effects through increasing returns and the endogenous social value of redistribution.

Specifically, we consider an exogenous 5% change in the relative price of the high-quality and low-quality bundles. We find that this price shock leads to a fall in marginal tax rates of 3.25 percentage points at the bottom of the income distribution. The equivalent variation, capturing the willingness to pay of agents for the revised optimal tax schedule, ranges from -6% at the bottom to +9% at the top of the income distribution. Finally, in equilibrium the relative price of the high-quality bundle falls by 13%, more than twice the exogenous relative price shock. We thus find that the amplification effects and their welfare implications are sizable.

5.4 Extension: The Response of the Optimal Tax Schedule to Exogenous Shifts in the Skill Distribution

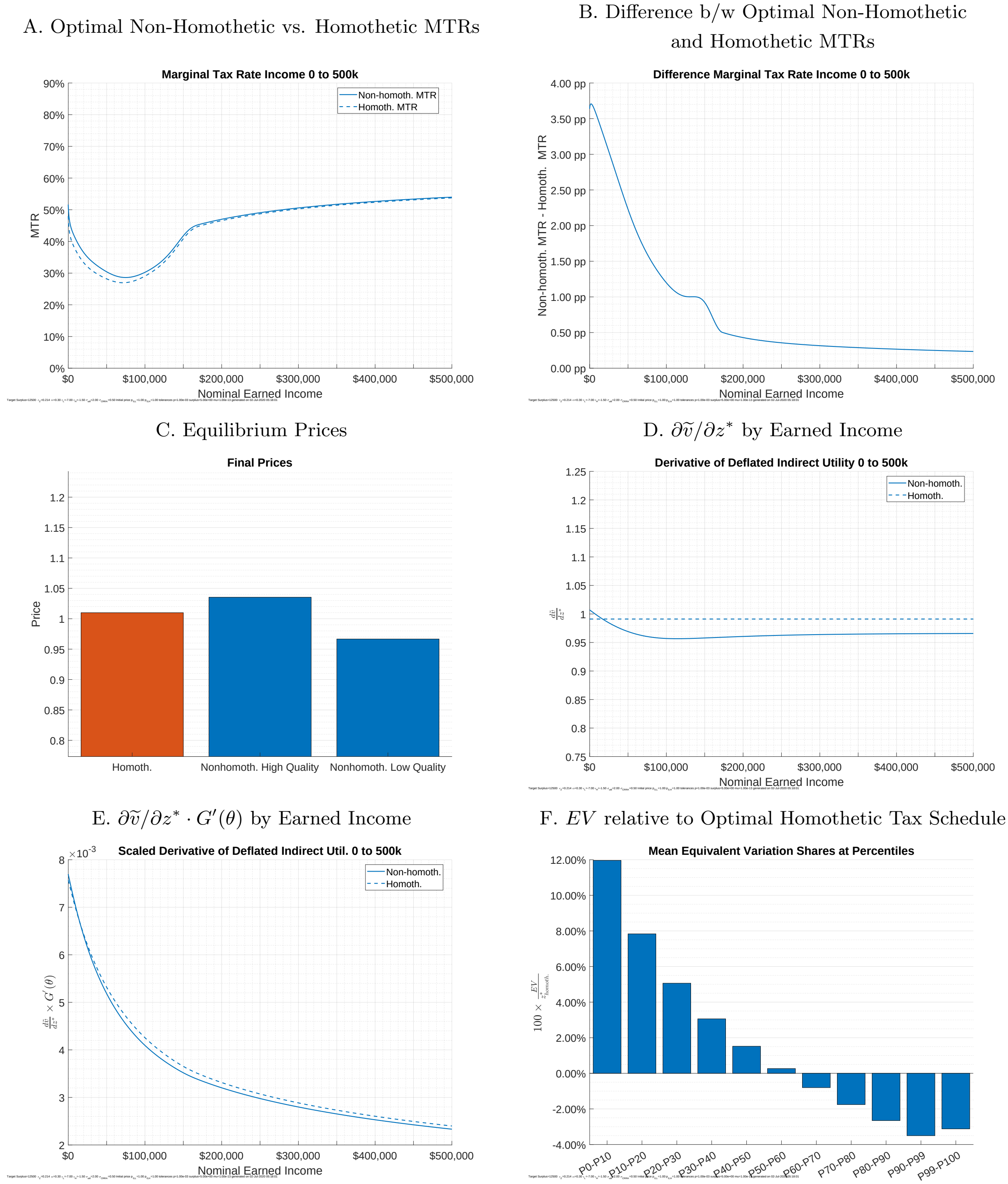
The previous section shows that the optimal tax schedule is sensitive to non-homotheticicities because redistribution induces changes in relative prices and hence in the value of further redistribution. We now present an extension analyzing shifts in the income distribution. Specifically, using the comparative statics

Figure 6 Higher Returns to Scale Magnify the Impact of Non-Homotheticicities
($\alpha = 0.4$, $\varepsilon_z = 0.21$, Pareto weights from SWF CRRA=1)



Notes: The quantitative model uses Pareto weights computed at the optimal homothetic tax schedule obtained under a social welfare function with CRRA=1. Panel A reports the optimal tax schedule with homothetic and non-homothetic preferences. Panel B plots the difference between these two tax schedules across the income distribution. Panel C reports the equilibrium prices in the homothetic and non-homothetic models. Panel D shows the derivative of (deflated) indirect utility, $d\widetilde{v}(z^*, \boldsymbol{p})/dz^*$, across the income distribution. Panel E reports this derivative scaled by social preferences for redistribution. Panel F report the equivalent variation (EV) giving the willingness to pay of agents for the non-homothetic optimal tax schedule instead of the homothetic optimal tax schedule, expressed as a percentage of their disposable income under the homothetic optimal tax schedule.

Figure 7 Lower Social Preferences for Redistribution Reduce the Impact of Non-Homotheticities ($\alpha = 0.3$, $\varepsilon_z = 0.21$, Pareto weights from SWF CRRA=0.5)



Notes: The quantitative model uses Pareto weights computed at the optimal homothetic tax schedule obtained under a social welfare function with CRRA=1. Panel A reports the optimal tax schedule with homothetic and non-homothetic preferences. Panel B plots the difference between these two tax schedules across the income distribution. Panel C reports the equilibrium prices in the homothetic and non-homothetic models. Panel D shows the derivative of (deflated) indirect utility, $d\widetilde{v}(z^*, \boldsymbol{p})/dz^*$, across the income distribution. Panel E reports this derivative scaled by social preferences for redistribution. Panel F report the equivalent variation (EV) giving the willingness to pay of agents for the non-homothetic optimal tax schedule instead of the homothetic optimal tax schedule, expressed as a percentage of their disposable income under the homothetic optimal tax schedule.

approach from Section 5.2, in Online Appendix B.2 we characterize quantitatively the optimal response of the tax schedule to exogenous shifts in the income distribution, accounting for the endogenous response of prices.

We first consider the direct, partial equilibrium response to the change in the skill distribution, with fixed prices. In this case, the rise in inequality in the United States in recent years makes it optimal to increase redistribution. Because of the shifts in the skill distribution, there is relatively more mass at the top and bottom of the skill distribution, hence the distortionary cost of taxation is higher in this range, while it is reduced in the middle of the distribution. To increase redistribution efficiently, it is therefore optimal to raise marginal tax rates especially in the middle of the income distribution.

Furthermore, general equilibrium effects are at play through prices. The direct effects on prices of the shifts in inequality is amplified through income and substitution effects, as well as changes in optimal tax rates. These effects tend to reduce optimal tax rates, because the observed shift in the income distribution lowers the price of products with a higher income elasticity. Because higher-income agents have a higher marginal propensity to spend on these goods, it is optimal to redistribute more toward them by lowering marginal tax rates, through the same channels as in Proposition 2.

Quantitatively, we find that the direct price effects, which imply more redistribution toward higher-skill agents, more than offset the motive for increased redistribution toward low-skill agents from the shift in the skill distribution. Taking into account all effects, the optimal tax schedule becomes *less* redistributive. These results show that it is important to jointly study shifts in the skill distribution and price shocks.

6 Conclusion

In this paper, we have shown that optimal commodity and income taxation is sensitive to exogenous price shocks, the elasticity of prices to market size, and non-homothetic preferences. We provided an explicit analytical characterization of the response of the optimal tax schedule to price shocks, in both partial and general equilibrium. Using simulations based on observed spending patterns and the empirical elasticity of prices to market size, we found that these channels have a sizable quantitative impact on optimal marginal tax rates and welfare across the skill distribution.

Our analysis was motivated by the fact that observed price changes are heterogeneous across product categories and across the income distribution, and that empirically prices are endogenous to market size. Going forward, our framework could be used to study the response of optimal taxation to a variety of supply shocks that could affect prices, for example due to changes in technology, trade, immigration, or market concentration. Although we considered a closed economy, we conjecture that the mechanisms we highlighted might become even richer in a model with trade. Changes in domestic demand can be even more important in an open economy than in a closed economy (Matsuyama (2019)) because they have an impact on the equilibrium patterns of specialization, which in turn have an impact on the direction of productivity growth through market size effects. Analyzing optimal taxation in an open economy model with non-homothetic preferences and endogenous prices is thus a promising direction for future research. Another interesting avenue for future work would be to estimate heterogeneity in returns to scale across sectors, which we have abstracted from in our quantitative analysis, and the consequences for the optimal tax schedule.

References

Acemoglu, Daron, "Directed technical change," *The review of economic studies*, 2002, *69* (4), 781–809. 12

_ and Joshua Linn, "Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry*," *The Quarterly Journal of Economics*, 08 2004, *119* (3), 1049–1090. 12, 41

Aghion, Philippe and Peter Howitt, "A Model of Growth through Creative Destruction," *Econometrica*, 1992, *60* (2), 323–51. 5, 12

_ and _ , "A Model of Growth Through Creative Destruction," *Econometrica: Journal of the Econometric Society*, 1992, pp. 323–351.

_ , Antonin Bergeaud, Timo Boppart, Peter J Klenow, and Huiyu Li, "A theory of falling growth and rising rents," *National Bureau of Economic Research Working Paper*, 2019.

_ , Nick Bloom, Richard Blundell, Rachel Griffith, and Peter Howitt, "Competition and innovation: An inverted-U relationship," *The quarterly journal of economics*, 2005, *120* (2), 701–728.

_ , Ufuk Akcigit, and Jesús Fernández-Villaverde, "Optimal capital versus labor taxation with innovation-led growth," *Working Paper*, 2013.

Akcigit, Ufuk, Douglas Hanley, and Stefanie Stantcheva, "Optimal Taxation and R&D Policies," NBER Working Papers 22908, National Bureau of Economic Research, Inc 2016.

Allcott, Hunt, Benjamin B Lockwood, and Dmitry Taubinsky, "Regressive Sin Taxes, with an Application to the Optimal Soda Tax," *The Quarterly Journal of Economics*, 2019, *134* (3), 1557–1626. 1

Argente, David and Munseob Lee, "Cost of Living Inequality During the Great Recession," *Journal of the European Economic Association*, 2020, *19* (2), 913–952. 1

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

Auer, Raphael, Ariel Burstein, Sarah Lein, and Jonathan Vogel, "Unequal expenditure switching: Evidence from Switzerland," *Review of Economic Studies*, 2024, *91* (5), 2572–2603. 44

Bartelme, Dominick G, Arnaud Costinot, Dave Donaldson, and Andres Rodriguez-Clare, "The Textbook Case for Industrial Policy: Theory Meets Data," Working Paper 26193, National Bureau of Economic Research August 2019. 12, 5.1

Beerli, Andreas, Franziska J. Weiss, Fabrizio Zilibotti, and Josef Zweimuller, "Demand forces of technical change evidence from the Chinese manufacturing industry," *China Economic Review*, 2020, *60*, 101157. 12, 5.1

Bell, Alex, Raj Chetty, Xavier Jaravel, Neviana Petkova, and John Van Reenen, "Who Becomes an Inventor in America? The Importance of Exposure to Innovation*," *The Quarterly Journal of Economics*, 11 2018, *134* (2), 647–713. 1

_ , _ , _ , _ , and _ , "Joseph Schumpeter Lecture, EEA Annual Congress 2017: Do Tax Cuts Produce more Einsteins? The Impacts of Financial Incentives VerSus Exposure to Innovation on the Supply of Inventors," *Journal of the European Economic Association*, 04 2019, *17* (3), 651–677. 1

Boar, Corina and Virgiliu Midrigan, "Markups and Inequality," Working Paper 25952, National Bureau of Economic Research June 2019. 1

Breinlich, Holger, Elsa Leromain, Dennis Novy, and Thomas Sampson, "The Brexit vote, inflation and UK living standards," *International Economic Review*, 2022, *63* (1), 63–93. 44

Broda, Christian and David E. Weinstein, "Globalization and the Gains From Variety*," *The Quarterly Journal of Economics*, 05 2006, *121* (2), 541–585. 5.1, 5.3.2

_ and _ , "Product Creation and Destruction: Evidence and Price Implications," *American Economic Review*, June 2010, *100* (3), 691–723. 5.1, 5.2, 5.3.2

Bustos, Paula, "Trade Liberalization, Exports, and Technology Upgrading: Evidence on the Impact of MERCOSUR on Argentinian Firms," *American Economic Review*, 2011, *101* (1), 304–40. 1

Cesarini, David, Erik Lindqvist, Matthew J Notowidigdo, and Robert Östling, "The effect of wealth on individual and household labor supply: evidence from Swedish lotteries," *American Economic Review*, 2017, *107* (12), 3917–3946. 22

Chaney, Thomas, "Distorted gravity: the intensive and extensive margins of international trade," *American Economic Review*, 2008, *98* (4), 1707–1721. 2

Chetty, Raj, "Bounds on Elasticities With Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply," *Econometrica*, 2012, *80* (3), 969–1018. 5.1

Comin, Diego, Danial Lashkari, and Marti Mestieri, "Structural Change With Long-Run Income and Price Effects," *Econometrica*, 2021, *89* (1), 311–374. 1, 5.1, 5.3.2

Costinot, Arnaud, Dave Donaldson, Margaret Kyle, and Heidi Williams, "The More We Die, The More We Sell? A Simple Test of the Home-Market Effect," *The Quarterly Journal of Economics*, 2019, *134* (2), 843–894. 5, 2

Couture, Victor, Cecile Gaubert, Jessie Handbury, and Erik Hurst, "Income growth and the distributional effects of urban spatial sorting," *National Bureau of Economic Research Working paper 26142*, 2020. 12, 41

Cravino, Javier and Andrei A Levchenko, "The distributional consequences of large devaluations," *American Economic Review*, 2017, *107* (11), 3477–3509. 44

_ **and Sebastian Sotelo**, "Trade-Induced Structural Change and the Skill Premium," *American Economic Journal: Macroeconomics*, July 2019, *11* (3), 289–326. 5.1

DellaVigna, Stefano and Matthew Gentzkow, "Uniform Pricing in U.S. Retail Chains*," *The Quarterly Journal of Economics*, 06 2019, *134* (4), 2011–2084. 5.1, 5.3.2

Dhingra, Swati and John Morrow, "Monopolistic competition and optimum product diversity under firm heterogeneity," *Journal of Political Economy*, 2019, *127* (1), 196–232.

Diamond, Peter A, "Optimal income taxation: an example with a U-shaped pattern of optimal marginal tax rates," *American Economic Review*, 1998, pp. 83–95. 4.1

Diamond, Peter and James Mirrlees, "Optimal Taxation and Public Production: I–Production Efficiency," *American Economic Review*, 1971, *61* (1), 8–27.

_ **and** _ , "Optimal Taxation and Public Production II: Tax Rules," *American Economic Review*, 1971, *61* (3), 261–78. 1, 8

_ **and Johannes Spinnewijn**, "Capital Income Taxes with Heterogeneous Discount Rates," *American Economic Journal: Economic Policy*, 2011, *3* (4), 52–76. 1

Diamond, Rebecca, "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000," *American Economic Review*, March 2016, *106* (3), 479–524. 12, 41

Dixit, Avinash K and Joseph E Stiglitz, "Monopolistic competition and optimum product diversity," *The American economic review*, 1977, *67* (3), 297–308. 12

Eeckhout, Jan, Chunyang Fu, Wenjian Li, and Xi Weng, "Optimal Taxation and Market Power," Technical Report, UPF mimeo 2021. 1, 21

Faber, Benjamin and Thibault Fally, "Firm Heterogeneity in Consumption Baskets: Evidence from Home and Store Scanner Data," *The Review of Economic Studies*, 09 2021. 5, 2, 12

Feenstra, Robert, "New Product Varieties and the Measurement of International Prices," *American Economic Review*, 1994, *84* (1), 157–77. 5.2

_ **and David Weinstein**, "Globalization, Markups, and US Welfare," *Journal of Political Economy*, 2017, *125* (4), 1040 – 1074. 1

Ferey, Antoine, Benjamin Lockwood, and Dmitry Taubinsky, "Sufficient statistics for nonlinear tax systems with general across-income heterogeneity," *NBER Working Paper*, 2023. 11

Ferriere, Axelle, Philipp Grübener, and Dominik Sachs, "Optimal Redistribution: Rising Inequality vs. Rising Living Standards," 2024.

Garner, Thesia I, David S Johnson, and Mary F Kokoski, "An experimental consumer price index for the poor," *Monthly Lab. Rev.*, 1996, *119*, 32. 44

Giupponi, Giulia, "When income effects are large: Labor supply responses and the value of welfare transfers," *Working Paper*, 2024. 22

Golosov, Mikhail, Michael Graber, Magne Mogstad, and David Novgorodsky, "How Americans respond to idiosyncratic and exogenous changes in household wealth and unearned income," *The Quarterly Journal of Economics*, 2024, *139* (2), 1321–1395. 22

Grigsby, John, Erik Hurst, and Ahu Yildirmaz, "Aggregate Nominal Wage Adjustments: New Evidence from Administrative Payroll Data," *American Economic Review*, February 2021, *111* (2), 428–71.

Handbury, Jessie, "Are Poor Cities Cheap for Everyone? Non-Homotheticity and the Cost of Living Across U.S. Cities," Working Paper 26574, National Bureau of Economic Research December 2019. 41, 5.1, 5.3.2

Hanoch, Giora, "Production and Demand Models with Direct or Indirect Implicit Additivity," *Econometrica*, 1975, *43* (3), 395–419. 1, 5.1, 5.3.2

Hendren, Nathaniel, "Measuring economic efficiency using inverse-optimum weights," *Journal of Public Economics*, 2020, *187* (C), S0047272720300621. 5.1, 1, 5.2, 47, 2, 5.2, 5.3.2, C, A11, A12, D.1.3

Imbens, Guido W, Donald B Rubin, and Bruce I Sacerdote, "Estimating the effect of unearned income on labor earnings, savings, and consumption: Evidence from a survey of lottery players," *American economic review*, 2001, *91* (4), 778–794. 22

Jaimovich, Nir and Sergio Rebelo, "Nonlinear effects of taxation on growth," *Journal of Political Economy*, 2017, *125* (1), 265–291. 1

Jaravel, Xavier, "The Unequal Gains from Product Innovations: Evidence from the U.S. Retail Sector," *The Quarterly Journal of Economics*, 2019, *134* (2), 715–783. 1, 5, 2, 12, 5.1, 42, 5.3.2, B.1

_ , "Distributional consumer price indices," *Working Paper*, 2024. 1, 44

_ **and Alan Olivi**, "Prices, non-homotheticities, and optimal taxation," *Working Paper, available at SSRN 3383693*, 2024. 39

_ **and Danial Lashkari**, "Measuring Growth in Consumer Welfare with Income-Dependent Preferences Nonparametric Methods and Estimates for the United States," *The Quarterly Journal of Economics*, 2023, p. qjad039. 1, 44

Jones, Charles I, "Taxing Top Incomes in a World of Ideas," Working Paper 25725, National Bureau of Economic Research 2019. 1

Kaplan, Greg and Sam Schulhofer-Wohl, "Inflation at the household level," *Journal of Monetary Economics*, 2017, *91* (C), 19–38. 1

Kleven, Henrik Jacobsen and Esben Anton Schultz, "Estimating Taxable Income Responses Using Danish Tax Reforms," *American Economic Journal: Economic Policy*, November 2014, *6* (4), 271–301. 5.1

Klick, Josh and Anya Stockburger, "Experimental CPI for lower and higher income households," *Bureau of Labor Statistics Working Paper*, 2021. 1

Krugman, Paul R, "Increasing returns, monopolistic competition, and international trade," *Journal of international Economics*, 1979, *9* (4), 469–479. 12

Kushnir, Alexey and Robertas Zubrickas, "Optimal Income Taxation with Endogenous Prices," Working Paper 2020. 1, 9

Lashkari, Danial, Arthur Bauer, and Jocelyn Boussard, "Information technology and returns to scale," *Working Paper*, 2018.

Linder, Staffan Burenstam, "An essay on trade and transformation," *Stockholm School of Economics PhD Dissertation*, 1961. 12

Matsuyama, Kiminori, "Engel's Law in the Global Economy: Demand-Induced Patterns of Structural Change, Innovation, and Trade," *Econometrica*, 2019, *87* (2), 497–528. 1, 5.1, 5.3.2, 6

McGranahan, Leslie and Anna Paulson, "The incidence of inflation: inflation experiences by demographic group: 1981-2004," Working Paper Series, Federal Reserve Bank of Chicago 2005. 1

Melitz, Marc, "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 2003, *71* (6), 1695–1725. 1, 5, 2, 12

Melitz, Marc J and Stephen J Redding, "Trade and Innovation," *National Bureau of Economic Research Working Paper*, 2021.

Mirrlees, James, "An Exploration in the Theory of Optimum Income Taxation," *Review of Economic Studies*, 1971, *38* (2), 175–208. 12, 5.1

Naito, Hisahiro, "Re-examination of uniform commodity taxes under a non-linear income tax system and its implication for production efficiency," *Journal of Public Economics*, 1999, *71* (2), 165–188. 8

Ridder, Maarten De, "Market power and innovation in the intangible economy," *Working Paper*, 2019.

Romer, Paul, "Endogenous Technological Change," *Journal of Political Economy*, 1990, *98* (5), S71–102. 5, 12

Romer, Paul M, "Endogenous technological change," *Journal of political Economy*, 1990, *98* (5, Part 2), S71–S102.

Rothschild, Casey and Florian Scheuer, "Redistributive taxation in the roy model," *The Quarterly Journal of Economics*, 2013, *128* (2), 623–668. 1

_ and _ , "A theory of income taxation under multidimensional skill heterogeneity," *Working Paper*, 2014.

Sachs, Dominik, Aleh Tsyvinski, and Nicolas Werquin, "Nonlinear Tax Incidence and Optimal Taxation in General Equilibrium," *Econometrica*, 2020, *88* (2), 469–493. 1

Saez, Emmanuel, "Using Elasticities to Derive Optimal Income Tax Rates," *Review of Economic Studies*, 2001, *68* (1), 205–229. 12, 3, 4.1, 5.1, 5.2, D.1.3, E.3

_ , "The desirability of commodity taxation under non-linear income taxation and heterogeneous tastes," *Journal of Public Economics*, 2002, *83* (2), 217–230. 1, 11, 3

Scheuer, Florian and Iván Werning, "Mirrlees meets diamond-mirrlees," *National Bureau of Economic Research*, 2016. 16

Schmookler, Jacob, "Invention and economic growth," *Harvard University Press*, 1966. 12

Shleifer, Andrei, "Implementation cycles," *Journal of Political Economy*, 1986, *94* (6), 1163–1190. 12

Vivalt, Eva, Elizabeth Rhodes, Alexander W Bartik, David E Broockman, Patrick Krause, and Sarah Miller, "The employment effects of a guaranteed income: Experimental evidence from two US states," *National Bureau of Economic Research*, 2024. 22

Weiss, Franziska and Timo Boppart, "Non-homothetic preferences and industry directed technical change," 2013 Meeting Papers 916, Society for Economic Dynamics 2013. 12, 41

Weyl, E. Glen and Michal Fabinger, "Pass-Through as an Economic Tool: Principles of Incidence under Imperfect Competition," *Journal of Political Economy*, 2013, *121* (3), 528–583.

For Online Publication

Appendix to "Prices, Nonhomotheticities, and Optimal Taxation"

Xavier Jaravel, *LSE & CEPR*
Alan Olivi, *UCL*
August 2025

Contents

A Proofs

In this section, we provide proofs for the theoretical results of Sections 3 and 4. We also derive the comparative statics formulas that underpin the quantitative results of Section 5.2, with many sectors. Additional results – including the derivation of optimal tax formulas, comparative statics formulas, and their qualitative characterization under general household preferences and a general specification of the supply side – are collected in Appendix E, which also provides micro-foundations for the supply-side specification.

In our quantitative analysis, we consider an economy with n sectors. Here, we formulate a simple extension of the model of Section 3, which allows us to generalize the results of Proposition 1, 2 and 4 and provide the theoretical ground for the quantitative results of Section 5.2.

The economy has n sectors indexed by k. There is a mass 1 of households with different productivity types θ distributed according to $\pi(\theta)$.

Households. Households' preferences over goods and hours worked z/θ are given by:

$$u(c_1, ..., c_n) - \frac{1}{1+\frac{1}{\varepsilon}}\left(\frac{z}{\theta}\right)^{1+\frac{1}{\varepsilon}},$$

with u concave, increasing and $\mathcal{C}^3$, and $\varepsilon \leq 1$. Given separability of preferences between consumption and labor, the household problem, under consumer prices $\mathbf{q} = \{q_1, ..., q_n\}$ and the income tax schedule $\mathbf{T}$, can be written as:

$$\begin{aligned}
V(\theta; \mathbf{T}, \mathbf{q}) &= \sup_{z, z^*} v(z^*, \mathbf{q}) - \frac{1}{1+\frac{1}{\varepsilon}}\left(\frac{z}{\theta}\right)^{1+\frac{1}{\varepsilon}}, \\
\text{such that} z^* &= z - T(z) \\
v(z^*, \mathbf{q}) &= \sup_{\mathbf{c}} u(c_1, ..., c_N), \quad \mathbf{q}\cdot\mathbf{c} = z^*.
\end{aligned}$$

The consumption problem on the third line defines an indirect utility of consumption $v(z^*, \mathbf{q})$ and a Marshallian demand function $c_k(z^*, \mathbf{q})$. Since u is concave and $\mathcal{C}^3$, the implicit function theorem directly shows that v and c_k are $\mathcal{C}^2$. The labor supply function is $z(\theta; \mathbf{T}, \mathbf{q})$ and post tax income is $z^*(\theta; \mathbf{T}, \mathbf{q}) = z(\theta; \mathbf{T}, \mathbf{q}) - T(z(\theta; \mathbf{T}, \mathbf{q}))$.

Notations. For individual consumption, we denote $e_k(z^*, \mathbf{q}) = q_k c_k(z^*, \mathbf{q})$ and $s_k(z^*, \mathbf{q}) = e_k(z^*, \mathbf{q}) / z^*$ the expenditure on good k and the budget share of k, respectively. $\partial_{z^*} e_k(z^*, \mathbf{q})$ is the marginal propensity to spend on good k.

For aggregate consumption, we denote $C_k = \int c_k(z^*(\theta; \mathbf{T}, \mathbf{q}), \mathbf{q})\pi(\theta)d\theta$, $E_k = q_k C_k$ and $\bar{s}_k = E_k / \int z^*(\theta; \mathbf{T}, \mathbf{q})\pi(\theta)d\theta$, aggregate demand for good k, aggregate spending on k, and the aggregate share of k in total expenditure. $\partial_{z^*} E_k = \int \partial_{z^*} e_k(z^*(\theta; \mathbf{T}, \mathbf{q}), \mathbf{q})\pi(\theta)d\theta$ is the average marginal propensity to spend on k. Finally, we denote the matrix of compensated cross price elasticities as $\mathcal{S}_{jk} =$

$q_k \int \left(\partial_{q_k} c_j \left(z^* \left(\theta; \mathbf{T}, \mathbf{q}\right), \mathbf{q}\right) + \partial_{z^*} c_j \left(z^* \left(\theta; \mathbf{T}, \mathbf{q}\right), \mathbf{q}\right) c_k \left(z^* \left(\theta; \mathbf{T}, \mathbf{q}\right), \mathbf{q}\right)\right) \pi\left(\theta\right) d\theta / C_j$.

For individual labor supply, as in the main text, we define $\zeta \equiv \varepsilon / \left(1 - \varepsilon \partial_{\ln(z)} \ln\left(v_{z^*}\right)\right)$, the compensated labor supply elasticity. $\tilde{\zeta}$ is the compensated labor supply elasticity corrected for non-linearities in the budget constraint: $\tilde{\zeta} = \zeta / (1 + z\zeta T'' / (1 - T'))$. Similarly, $\eta = \zeta \partial_{\ln(z)} \log\left(v_{z^*}\right)$ is the income effect with a linear budget constraint and $\tilde{\eta}$ the corrected income effect.

Firms. We adopt the same supply-side specification as in the main text. In each sector, goods are produced using labor as the sole input. The cost of producing C_k units of good k is $\chi_k\left(C_k, \xi_k\right)$ and the price of k is given by $\phi_k\left(C_k, \xi_k\right)$, where ξ_k is an exogenous supply shifter. We consider two cases: in the competitive case, we have $\phi_k\left(C_k, \xi_k\right) = \partial_{C_k} \chi_k\left(C_k, \xi_k\right)$; in the monopolistic case, we have $\chi_k\left(C_k, \xi_k\right) = C_k \phi_k\left(C_k, \xi_k\right)$. We assume, as in the main text, that the elasticity of price with respect to market size, $\alpha = -C_k \partial_{C_k} \phi_k\left(C_k, \xi_k\right) / \phi_k\left(C_k, \xi_k\right)$, is constant, equal across sectors and independent from C_k and ξ_k.

Planning Problem. The government maximizes a social welfare function $\int G(V(\theta), \theta)\pi(\theta)d\theta$ with G increasing and concave in V using a nonlinear income tax $\mathbf{T}$, and commodity taxes, $q_k - p_k$, and a profit tax. The planning problem can be written as:

$$
\begin{aligned}
\mathcal{W} = & \sup_{\mathbf{T}, \mathbf{q}} \int G(V(\theta), \theta)\pi(\theta)d\theta \\
\text{s.t} \quad & V\left(\theta\right) = \sup_{z} v\left(z - T\left(z\right), \boldsymbol{q}\right) - \frac{1}{1+\frac{1}{\epsilon}} \left(\frac{z}{\theta}\right)^{1+\frac{1}{\epsilon}} \\
& z\left(\theta\right) = \text{argmax} v\left(z - T\left(z\right), \boldsymbol{q}\right) - \frac{1}{1+\frac{1}{\epsilon}} \left(\frac{z}{\theta}\right)^{1+\frac{1}{\epsilon}} \\
& C_k = \int c_k\left(z\left(\theta\right) - T\left(z\left(\theta\right)\right), \boldsymbol{q}\right)\pi(\theta)d\theta \\
& \int T\left(z\left(\theta\right)\right)\pi(\theta)d\theta + \sum_{k=1}^{n} \left(q_k C_k - \chi_k\left(C_k, \xi_k\right)\right) = 0,
\end{aligned}
$$

where the consumption function solves $\boldsymbol{c}\left(z^*, \boldsymbol{q}\right) = \text{argmax}_{\boldsymbol{c}} u\left(\boldsymbol{c}\right)\ s.t.\ \boldsymbol{q} \cdot \boldsymbol{c} = z^*$ and $v\left(z^*, \boldsymbol{q}\right) = u\left(\boldsymbol{c}\left(z^*, \boldsymbol{q}\right)\right)$.

Preferences of the household satisfy the single crossing property: $\left(z/\theta\right)^{\epsilon^{-1}} / \left(\theta \partial_{z^*} v\left(z^*, \boldsymbol{q}\right)\right)$ is decreasing in type θ. Therefore, the planning problem can be re-expressed as a direct mechanism where global incentive compatibility constraints are replaced with a local constraint and a monotonicity condition on

$z(\theta)$.

$$\begin{aligned}
\mathcal{W} = & \sup_{V(\theta),z(\theta),\boldsymbol{q}} \int G(V(\theta),\theta)\pi(\theta)d\theta \\
s.t. \quad & V'(\theta) = \frac{1}{\theta}\left(\frac{z(\theta)}{\theta}\right)^{1+\frac{1}{\epsilon}} \quad \text{and} \quad z(\theta) \text{ is non-decreasing} \\
with \quad & V(\theta) = v\left(z^*\left(\theta\right),\boldsymbol{q}\right) - \frac{1}{1+\frac{1}{\epsilon}}\left(\frac{z}{\theta}\right)^{1+\frac{1}{\epsilon}} \\
& C_k = \int c_k(z^*\left(\theta\right),\boldsymbol{q})\pi(\theta)d\theta \\
& \int\left(z\left(\theta\right) - z^*\left(\theta\right)\right)\pi(\theta)d\theta + \sum_{k=1}^{n}\left(q_k C_k - \chi_k\left(C_k,\xi_k\right)\right) = 0
\end{aligned} \tag{A1}$$

A.1 Proof of Proposition 1

In this subsection we prove Proposition 1 in an economy with N goods; the two-good economy of the main text is included as a special case.

Proposition 1. Commodity taxes are not used at the optimum. The optimal non-linear income tax schedule is characterized by:

$$\frac{T'}{1-T'} = -t_w + \frac{1-t_w}{z\tilde{\zeta}f(z)}\left\{\mathbb{E}_{z'>z}\left(1-g\right) - \frac{1}{1-t_w}\mathbb{E}_{z'>z}\left(\left(t_w + \frac{T'}{1-T'}\right)\tilde{\eta}\right)\right\}, \tag{A2}$$

where $t_w = \alpha$ in the monopolistic case and $t_w = 0$ in the competitive case. When $\alpha = 0$, we obtain the standard formulas in both cases.

Proof: After integration by parts of the planning problem A1, the corresponding Lagrangian is:

$$\begin{aligned}
\mathcal{L} = & \int G(V(\theta),\theta)\pi(\theta)d\theta - \int\left(\mu'(\theta)V(\theta) + \mu(\theta)\frac{1}{\theta}\left(\frac{z(\theta)}{\theta}\right)^{1+\frac{1}{\epsilon}}\right)d\theta \\
& -\lambda\left(\int\left(z^*(\theta) - z(\theta)\right)\pi(\theta)d\theta - \sum_{k=1}^{n}\left(q_k C_k - \chi_k\left(C_k,\xi_k\right)\right)\right),
\end{aligned}$$

where $\mu(\theta)$ are the multipliers on the incentive constraints and λ is the multiplier on the resource constraint.

We start with the FOC with respect to consumer prices q_i. Denote $c^h(v,\mathbf{q})$ is the Hicksian demand function at prices q for a given sub-utility v, we have:

$$\begin{aligned}
\left.\frac{dc_j}{dq_i}\right|_{z,V} &= \left.\frac{dc_j}{dq_i}\right|_{v} = \frac{\partial c_j^h}{\partial q_i} \\
\left.\frac{dz^*}{dq_i}\right|_{z,V} &= \left.\frac{dz^*}{dq_i}\right|_{v} = c_i
\end{aligned}$$

We therefore have, denoting $\partial_{q_i} C_j^h = \int \partial_{q_i} c_j^h \pi(\theta) d\theta$:

$$\frac{d\mathcal{L}}{dq_i} = \lambda \left(C_i + \sum_j \left(q_j - \partial_{C_j} \chi_j \left(C_j, \xi_j \right) \right) \partial_{q_i} C_j^h - C_i \right),$$

which gives for all i:

$$\sum_j \left(q_j C_j - (1 - t_w) \, p_j C_j \right) \mathcal{S}_{j,i} = 0,$$

with $t_w = 0$ in the competitive case ($\partial_{C_j} \chi_j \left(C_j, \xi_j \right) = p_j$), and $t_w = \alpha$ in the monopolistic case ($\partial_{C_j} \chi_j \left(C_j, \xi_j \right) = \phi_j \left(C_j, \xi_j \right) + C_j \partial_{C_j} \phi_j \left(C_j, \xi_j \right) = (1 - \alpha) \, p_j$). Recall that $\mathcal{S}$ is the matrix of cross price elasticities. Given that $\mathcal{S}$ is generically of rank $N - 1$ with left kernel $\boldsymbol{qC}$, we have:

$$\boldsymbol{q} \propto \boldsymbol{p},$$

so choosing $p = q$ is optimal.

Next, we derive the FOC associated with V. $V(\theta)$ impacts consumption and producer prices through $z^*(\theta)$ with $dz^*(\theta)/dV(\theta) = 1/v_{z^*}$. We thus have:

$$\begin{aligned}
0 &= G'(V(\theta), \theta)\pi(\theta) - \mu'(\theta) - \frac{\lambda \pi(\theta)}{v_{z^*}} \left[1 - \sum_i \left(q_i - \partial_{C_i} \chi_i \left(C_i, \xi_i \right) \right) \partial_{z^*} c_i(\theta) \right] \\
&= G'(V(\theta), \theta)\pi(\theta) - \mu'(\theta) - \frac{\lambda \pi(\theta)}{v_{z^*}} \left[1 - \sum_i \left(q_i - (1 - t_w) \, p_i \right) \partial_{z^*} c_i(\theta) \right] \\
&= G'(V(\theta), \theta)\pi(\theta) - \mu'(\theta) - \frac{\lambda \pi(\theta)}{v_{z^*}} \left[1 - t_w \sum_i q_i \, \partial_{z^*} c_i(\theta) \right] \\
\Rightarrow & \mu'(\theta) \frac{v_{z^*}}{\lambda} = - \left(1 - t_w - \frac{G'(V(\theta), \theta) v_{z^*}}{\lambda} \right) \pi(\theta).
\end{aligned}$$

Finally, defining $\tilde{\mu} = \mu \, v_{z^*}/\lambda$, we have:

$$\tilde{\mu}'(\theta) + \tilde{\mu} \, \partial_{z^*} MRS \; z'(\theta) = - \left(1 - t_w - \frac{G'(V(\theta), \theta) v_{z^*}}{\lambda} \right) \pi(\theta),$$

with $MRS = \frac{1}{\theta} \left(\frac{z(\theta)}{\theta} \right)^{\epsilon^{-1}} / v_{z^*}$ the marginal rate of substitution.

Finally, the FOC associated with z, using the same steps as above to derive the response of consumption and prices, is:

$$\tilde{\mu} \, \partial_\theta MRS = \pi(\theta) ((1 - t_w) MRS - 1).$$

Since $MRS = 1 - T'(z(\theta))$, and $z \tilde{\zeta} \partial_\theta MRS = -z'(\theta)(1 - T'(z(\theta)))$, we therefore have, denoting $f(z(\theta)) =$

$\pi(\theta)/z'(\theta)$:

$$\tilde{\mu}(\theta) = f(z)z\tilde{\zeta}\left(\frac{T'}{1-T'} + t_w\right).$$

Finally, using $-z\tilde{\zeta}\,\partial_{z^*}MRS = \tilde{\eta}$, we get:

$$f(z)z\tilde{\zeta}\left(\frac{T'}{1-T'} + t_w\right) + \int_{z(\theta)}^{z(\bar{\theta})} \tilde{\eta}\left(\frac{T'}{1-T'} + t_w\right) f(z)dz = \int_{z(\theta)}^{z(\bar{\theta})} \left(1 - t_w - \frac{G' v_{z^*}}{\lambda}\right) f(z)dz$$

Using $g = G' v_{z^*}/((1-t_w)\lambda)$, we obtain the formula of Proposition 1.□

A.2 Proofs for Section 4: Propositions 2, 3, and 4, and Corollary 1

In this section, we provide proofs for our comparative statics results of Section 4. In the last subsection, we provide the comparative statics formulas underpinning the results of Section 5.2.

A.2.1 Intermediary Lemma

To streamline the presentation of the proofs, we first present an intermediary lemma valid in an $n-$sector economy. For an exogenous supply shock ξ_k, we derive the change in welfare for agent θ, $dV(\theta)/d\xi_k$,expressed in terms of the resulting equilibrium price changes, $dp_l/d\xi_k$. Much of the algebra required for Propositions 1, 2 and 3, as well as Corollary 1 is the same. The purpose of the lemma is to consolidate these repetitive derivations into a unified result.

As in the main text, we assume that there are no income effects of labor supply at initial prices (assumption $A3$). At initial prices $\mathbf{p} = \{p_1, ..., p_N\}$, we have $\partial_{z^*} v(z^*, \mathbf{p}) = 1, \forall z^*$. Recall that the utility of the agent can be rewritten $U(c_1, ..., c_n, z, \theta) = \Psi(u(c_1, ..., c_n)) - \left(1 + \epsilon^{-1}\right)^{-1} (z/\theta)^{1+\epsilon^{-1}}$, where the function Ψ can be used to calibrate the income effect of labor supply at initial prices: here we choose $\Psi'(v(z^*, \mathbf{p})) = \partial_{z^*} v(z^*, \mathbf{p})^{-1}$. The comparative statics formulas for a general Ψ are relegated to Appendix E.

Lemma A1. Under assumption $A3$, the change in welfare for agent θ, $dV(\theta)/d\xi_k$, in response to an exogenous supply shift $d\xi_k$, conditional on the change of prices $dp_l/d\xi_k$, is given by:

$$\frac{\epsilon}{(1+\epsilon)^2}\frac{\theta\pi(\theta)}{(1-T')^2}\frac{\theta}{z(\theta)}\frac{d}{d\theta}\left\{\frac{dV}{d\xi_k}\right\} + (1-t_w)\int_{\theta}^{\bar{\theta}} g\left(\gamma(\theta')\frac{dV}{d\xi_k} - \int_{\underline{\theta}}^{\bar{\theta}} g\gamma(\theta')\frac{dV}{d\xi_k}\pi d\theta'\right)\pi d\theta' = -\frac{\epsilon}{1+\epsilon}\frac{\theta\pi(\theta)}{1-T'}\sum_{l=1}^{n}\left(\tau_l(\theta) + \partial_{z^*}E_l\right)\frac{1}{p_l}\frac{dp_l}{d\xi_k},$$

$$(1-t_w)\int_{\underline{\theta}}^{\bar{\theta}} g\frac{dV}{d\xi_k}\pi d\theta = -\partial_{\xi_k}\chi_k(\xi_k, C_k),$$

where $\gamma(\theta)$, $\tau_l(\theta)$ are given by:

$$\gamma(\theta) \equiv -\frac{G''(V(\theta),\theta)}{G'(V(\theta),\theta)},$$
$$\tau_l(\theta) \equiv (1-t_w)\left(1-T'\right)\left(\frac{1+\epsilon}{\epsilon}\frac{1}{\theta\pi(\theta)}\int_{\theta}^{\bar{\theta}}\left(\partial_{z^*}e_l - \partial_{z^*}E_l\right)\pi d\theta' + \left(\partial_{z^*}e_l - \partial_{z^*}E_l\right)\right).$$

Proof: Recall from the proof of Proposition 1 that $p_l = q_l$. In addition, the optimal tax system is determined by the following envelope conditions and first order condition:

$$\begin{aligned}
\frac{d\hat{\mu}(\theta)}{d\theta} &= -\left(\frac{1-t_w}{v_{z^*}(z^*(\theta),\mathbf{p})} - \frac{G'(V(\theta),\theta)}{\lambda}\right)\pi(\theta)\\
\frac{dV(\theta)}{d\theta} &= \frac{1}{\theta}\left(\frac{z(\theta)}{\theta}\right)^{1+\frac{1}{\epsilon}}\\
\hat{\mu}(\theta)\left(1+\frac{1}{\epsilon}\right)\frac{1}{\theta^2}\left(\frac{z(\theta)}{\theta}\right)^{\frac{1}{\epsilon}} &= \pi(\theta)\left(1-(1-t_w)\frac{1}{v_{z^*}(z^*(\theta),\mathbf{p})}\frac{1}{\theta}\left(\frac{z(\theta)}{\theta}\right)^{\frac{1}{\epsilon}}\right),
\end{aligned} \tag{A3}$$

with $\hat{\mu}(\theta) = \mu(\theta)/\lambda$ and $\hat{\mu}(\underline{\theta}) = \hat{\mu}(\bar{\theta}) = 0$. Finally, the budget constraint needs to be satisfied:

$$\int (z(\theta) - z^*(\theta))\pi(\theta)d\theta + \sum_{l=1}^{n}(p_l C_l - \chi_l(\xi_l, C_l)) = 0.$$

We first start by differentiating the marginal value of income $v_{z^*}(z^*(\theta),\mathbf{p})$:

$$\begin{aligned}
\frac{d}{d\xi_k}\{v_{z^*}(z^*(\theta),\mathbf{p})\} &= v_{z^*z^*}(z^*(\theta),\mathbf{p})\frac{dz^*}{d\xi_k} + \sum_{l=1}^{n}\frac{\partial}{\partial p_l}\{v_{z^*}(z^*(\theta),\mathbf{p})\}\frac{dp_l}{d\xi_k}\\
&= v_{z^*z^*}(z^*(\theta),\mathbf{p})\frac{dz^*}{d\xi_k} + \sum_{l=1}^{n}\frac{\partial}{\partial z^*}\{v_{p_l}(z^*(\theta),\mathbf{p})\}\frac{dp_l}{d\xi_k}\\
&= v_{z^*z^*}(z^*(\theta),\mathbf{p})\frac{dz^*}{d\xi_k} - \sum_{l=1}^{n}\frac{\partial}{\partial z^*}\{v_{z^*}(z^*(\theta),\mathbf{p})c_l\}\frac{dp_l}{d\xi_k}\\
&= v_{z^*z^*}(z^*(\theta),\mathbf{p})\left(\frac{dz^*}{d\xi_k} - \sum_{l=1}^{n}c_l\frac{dp_l}{d\xi_k}\right) - v_{z^*}(z^*(\theta),\mathbf{p})\sum_{l=1}^{N}\partial_{z^*}c_l\frac{dp_l}{d\xi_k}\\
&= -v_{z^*}(z^*(\theta),\mathbf{p})\sum_{l=1}^{n}\partial_{z^*}c_l\frac{dp_l}{d\xi_k},
\end{aligned}$$

where the second line uses Schwarz's identity, the third Roy's identity, and the fifth uses $v_{z^*z^*}(z^*(\theta),\mathbf{p}) = 0$. Using the fact that t_w is constant (equal to α in the monopolistic case, equal to 0 in the competitive case), differentiating the first equation of system A3, we obtain:

$$\frac{d}{d\theta}\left\{\frac{d\hat{\mu}(\theta)}{d\xi_k}\right\} = -\left(\frac{1-t_w}{v_{z^*}(z^*(\theta),\mathbf{p})}\sum_{l=1}^{n}\partial_{z^*}c_l\frac{dp_l}{d\xi_k} + \frac{G'(V(\theta),\theta)}{\lambda}\left(\gamma(\theta)\frac{dV(\theta)}{d\xi_k} + \frac{1}{\lambda}\frac{d\lambda}{d\xi_k}\right)\right)\pi(\theta),$$

with $\gamma(\theta) = -G''(V(\theta),\theta)/G'(V(\theta),\theta)$. Using the fact that $d_{\xi_k}\hat{\mu}(\underline{\theta}) = d_{\xi_k}\hat{\mu}(\bar{\theta}) = 0$, we have in addition:

$$\begin{aligned}
\frac{1}{\lambda}\frac{d\lambda}{d\xi_k} &= -\int\left(\frac{1-t_w}{v_{z^*}\left(z^*\left(\theta\right),\mathbf{p}\right)}\sum_{l=1}^{n}\partial_{z^*}c_l\frac{dp_l}{d\xi_k} + \frac{G'(V(\theta),\theta)}{\lambda}\gamma\left(\theta\right)\frac{dV(\theta)}{d\xi_k}\right)\pi(\theta)d\theta \\
&= -\left(1-t_w\right)\int\left(\sum_{l=1}^{n}\partial_{z^*}c_l\frac{dp_l}{d\xi_k} + g\gamma\left(\theta\right)\frac{dV(\theta)}{d\xi_k}\right)\pi(\theta)d\theta,
\end{aligned}$$

where the second line uses $v_{z^*}\left(z^*\left(\theta\right),\mathbf{p}\right) = 1$ and $G'(V(\theta),\theta)//\lambda = \left(1-t_w\right)g$ (by definition of g). We therefore have:

$$\frac{d}{d\theta}\left\{\frac{d\hat{\mu}(\theta)}{d\xi_k}\right\} = -\left(1-t_w\right)\left(\sum_{l=1}^{n}\left(\partial_{z^*}c_l - g\left(\theta\right)\int\partial_{z^*}c_l\pi(\theta')d\theta'\right)\frac{dp_l}{d\xi_k} + g\left(\theta\right)\left(\gamma\left(\theta\right)\frac{dV(\theta)}{d\xi_k} - \int g\gamma\left(\theta'\right)\frac{dV(\theta')}{d\xi_k}\pi(\theta')d\theta'\right)\right)\pi(\theta).$$

Differentiating the second equation of system A3, we obtain:

$$\frac{d}{d\theta}\left\{\frac{dV(\theta)}{d\xi_k}\right\} = \left(1+\frac{1}{\epsilon}\right)\frac{1}{\theta}\left(\frac{z(\theta)}{\theta}\right)^{1+\frac{1}{\epsilon}}\frac{1}{z\left(\theta\right)}\frac{dz(\theta)}{d\xi_k}.$$

Differentiating the last equation of system A3, we obtain:

$$\begin{aligned}
\frac{d\hat{\mu}(\theta)}{d\xi_k} &= -\frac{\epsilon}{1+\epsilon}\theta\pi(\theta)\left(\left(1-t_w\right)\sum_{l=1}^{n}\partial_{z^*}c_l\frac{dp_l}{d\xi_k} + \frac{1}{\epsilon}\theta\left(\frac{\theta}{z(\theta)}\right)^{\frac{1}{\epsilon}}\frac{1}{z\left(\theta\right)}\frac{dz(\theta)}{d\xi_k}\right) \\
&= -\frac{\epsilon}{1+\epsilon}\theta\pi(\theta)\left(\left(1-t_w\right)\sum_{l=1}^{n}\partial_{z^*}c_l\frac{dp_l}{d\xi_k} + \frac{1}{1+\epsilon}\theta^2\left(\frac{\theta}{z(\theta)}\right)^{1+2\frac{1}{\epsilon}}\frac{d}{d\theta}\left\{\frac{dV(\theta)}{d\xi_k}\right\}\right).
\end{aligned}$$

Putting everything together, we obtain:

$$\begin{aligned}
&\frac{\epsilon}{1+\epsilon}\theta_0\pi(\theta_0)\left(\left(1-t_w\right)\sum_{l=1}^{n}\partial_{z^*}e_l\frac{1}{p_l}\frac{dp_l}{d\xi_k} + \frac{1}{1+\epsilon}\theta_0^2\left(\frac{\theta_0}{z(\theta_0)}\right)^{1+2\frac{1}{\epsilon}}\frac{d}{d\theta}\left\{\frac{dV(\theta_0)}{d\xi_k}\right\}\right) \\
= &-\left(1-t_w\right)\int_{\theta_0}^{\bar{\theta}}\left(\sum_{l=1}^{n}\left(\partial_{z^*}e_l - g\left(\theta\right)\partial_{z^*}E_l\right)\frac{1}{p_l}\frac{dp_l}{d\xi_k} + g\left(\theta\right)\left(\gamma\left(\theta\right)\frac{dV(\theta)}{d\xi_k} - \int g\gamma\left(\theta'\right)\frac{dV(\theta')}{d\xi_k}\pi(\theta')d\theta'\right)\right)\pi(\theta)d\theta.
\end{aligned}$$

Using the optimality of the initial schedule, we have:

$$\begin{aligned}
\frac{\epsilon}{1+\epsilon}\theta_0\pi(\theta_0)\frac{1}{1+\epsilon}\theta_0^2\left(\frac{\theta_0}{z(\theta_0)}\right)^{1+2\frac{1}{\epsilon}}\frac{d}{d\theta}\left\{\frac{dV(\theta_0)}{d\xi_k}\right\} &= -\left(1-t_w\right)\int_{\theta_0}^{\bar{\theta}}g\left(\theta\right)\left(\gamma\left(\theta\right)\frac{dV(\theta)}{d\xi_k} - \int g\gamma\left(\theta'\right)\frac{dV(\theta')}{d\xi_k}\pi(\theta')d\theta'\right)\pi(\theta)d\theta \\
&\quad -\left(1-t_w\right)\sum_{l=1}^{n}\left(\int_{\theta_0}^{\bar{\theta}}\left(\left(\partial_{z^*}e_l - g\left(\theta\right)\partial_{z^*}E_l\right)\right)\pi(\theta)d\theta + \frac{\epsilon}{1+\epsilon}\theta_0\pi(\theta_0)\partial_{z^*}e_l\right)\frac{1}{p_l}\frac{dp_l}{d\xi_k} \\
&= -\left(1-t_w\right)\int_{\theta_0}^{\bar{\theta}}g\left(\theta\right)\left(\gamma\left(\theta\right)\frac{dV(\theta)}{d\xi_k} - \int g\gamma\left(\theta'\right)\frac{dV(\theta')}{d\xi_k}\pi(\theta')d\theta'\right)\pi(\theta)d\theta \\
&\quad -\left(1-t_w\right)\sum_{l=1}^{n}\left(\int_{\theta_0}^{\bar{\theta}}\left(\left(\partial_{z^*}e_l - \partial_{z^*}E_l\right)\right)\pi(\theta)d\theta + \frac{\epsilon}{1+\epsilon}\theta_0\pi(\theta_0)\left(\partial_{z^*}e_l - \partial_{z^*}E_l\right)\right)\frac{1}{p_l}\frac{dp_l}{d\xi_k} \\
&\quad -\frac{\epsilon}{1+\epsilon}\theta_0\pi(\theta_0)\frac{1}{1-T'\left(z\left(\theta_0\right)\right)}\sum_{l=1}^{n}\partial_{z^*}E_l\frac{1}{p_l}\frac{dp_l}{d\xi_k}
\end{aligned}$$

The last line uses the fact that the initial schedule is optimal, so that:

$$\left(1-t_w\right)\int_{\theta_0}^{\bar{\theta}}g\left(\theta\right)\pi(\theta)d\theta = \left(1-t_w\right)\int_{\theta_0}^{\bar{\theta}}\pi(\theta)d\theta - \frac{\epsilon}{1+\epsilon}\theta_0\pi(\theta_0)\left(\frac{1}{1-T'\left(z\left(\theta_0\right)\right)} - \left(1-t_w\right)\right).$$

Using the definition of τ_l, we therefore have

$$\frac{\epsilon}{(1+\epsilon)^2}\frac{\theta_0\pi(\theta_0)}{\left(1-T'\left(z\left(\theta_0\right)\right)\right)^2}\left(\frac{\theta_0}{z(\theta_0)}\right)\frac{d}{d\theta}\left\{\frac{dV(\theta_0)}{d\xi_k}\right\}=-\left(1-t_w\right)\int_{\theta_0}^{\bar{\theta}}g\left(\theta\right)\left(\gamma\left(\theta\right)\frac{dV(\theta)}{d\xi_k}-\int g\gamma\left(\theta'\right)\frac{dV(\theta')}{d\xi_k}\pi(\theta')d\theta'\right)\pi(\theta)d\theta$$
$$-\frac{\epsilon}{1+\epsilon}\frac{\theta_0\pi(\theta_0)}{1-T'\left(z\left(\theta_0\right)\right)}\sum_{l=1}^{n}\left(\tau_l\left(\theta_0\right)+\partial_{z^*}E_l\right)\frac{1}{p_l}\frac{dp_l}{d\xi_k},$$

which proves the first formula of the Lemma.
Next, we differentiate the budget constraint:

$$\int_{\underline{\theta}}^{\bar{\theta}}\left(\frac{dz(\theta)}{d\xi_k}-\frac{dz^*(\theta)}{d\xi_k}\right)\pi\left(\theta\right)d\theta+\sum_{l=1}^{n}\left(p_l-\partial_{C_l}\chi_l\left(\xi_l,C_l\right)\right)\frac{dC_l}{d\xi_k}+\sum_{l=1}^{n}\frac{dp_l}{d\xi_k}C_l-\partial_{\xi_k}\chi_k\left(\xi_k,C_k\right)=0.$$

Recall that we have $C_l=\int c_l\left(z^*\left(\theta\right),\mathbf{p}\right)\pi\left(\theta\right)d\theta$ and that $c_h\left(v,\mathbf{p}\right)$ is the Hicksian demand function. Using the standard Slutsky decomposition, we have:

$$\frac{dC_l}{d\xi_k}=\int_{\underline{\theta}}^{\bar{\theta}}\partial_{z^*}c_l\left(\theta\right)\left(\frac{dz^*(\theta)}{d\xi_k}-\sum_{m=1}^{n}c_m\left(\theta\right)\frac{dp_m}{d\xi_k}\right)\pi\left(\theta\right)d\theta+\sum_{m=1}^{n}\int_{\underline{\theta}}^{\bar{\theta}}\partial_{p_m}c_l^h\left(\theta\right)\pi\left(\theta\right)d\theta\frac{dp_m}{d\xi_k}.$$

Using $p_l-\partial_{C_l}\chi_l\left(\xi_l,C_l\right)=t_wp_l$, $\sum_{l=1}^{N}p_l\partial_{p_m}c_l^h\left(\theta\right)=0$ and $\sum_{l=1}^{N}p_l\partial_{z^*}c_l\left(\theta\right)=1$, we obtain:

$$\int_{\underline{\theta}}^{\bar{\theta}}\left(\frac{dz(\theta)}{d\xi_k}-\left(1-t_w\right)\left(\frac{dz^*(\theta)}{d\xi_k}-\sum_{l=1}^{N}c_l\left(\theta\right)\frac{dp_l}{d\xi_k}\right)\right)\pi\left(\theta\right)d\theta-\partial_{\xi_k}\chi_k\left(\xi_k,C_k\right)=0.$$

Using Roy's identity and the envelope condition, we have

$$\frac{dV(\theta)}{d\xi_k}=v_{z^*}\left(\frac{dz^*(\theta)}{d\xi_k}-\sum_{l=1}^{n}c_l\left(\theta\right)\frac{dp_l}{d\xi_k}-\left(1-T'\right)\frac{dz(\theta)}{d\xi_k}\right),$$
$$\frac{d}{d\theta}\left\{\frac{dV(\theta)}{d\xi_k}\right\}=\left(1+\frac{1}{\epsilon}\right)\frac{1}{\theta}\left(1-T'\left(z(\theta)\right)\right)\frac{dz(\theta)}{d\xi_k},$$

so we can re-express the government budget constraint in terms of household's welfare:

$$\int_{\underline{\theta}}^{\bar{\theta}}\left(\left(\frac{1}{1-T'}-\left(1-t_w\right)\right)\frac{\epsilon}{1+\epsilon}\theta\frac{d}{d\theta}\left\{\frac{dV(\theta)}{d\xi_k}\right\}-\left(1-t_w\right)\frac{1}{v_{z^*}}\frac{dV(\theta)}{d\xi_k}\right)\pi\left(\theta\right)d\theta-\partial_{\xi_k}\chi_k\left(\xi_k,C_k\right)=0.$$

Finally, using the optimality of the initial schedule we have

$$\int_{\underline{\theta}}^{\bar{\theta}}\left(\left(\frac{1}{1-T'}-\left(1-t_w\right)\right)\frac{\epsilon}{1+\epsilon}\theta\frac{d}{d\theta}\left\{\frac{dV(\theta)}{d\xi_k}\right\}-\left(1-t_w\right)\frac{1}{v_{z^*}}\frac{dV(\theta)}{d\xi_k}\right)\pi\left(\theta\right)d\theta=-\left(1-t_w\right)\int_{\underline{\theta}}^{\bar{\theta}}g\left(\theta\right)\frac{dV(\theta)}{d\xi_k}\pi\left(\theta\right)d\theta,$$

which implies

$$\left(1-t_w\right)\int_{\underline{\theta}}^{\bar{\theta}}g\left(\theta\right)\frac{dV(\theta)}{d\xi_k}\pi\left(\theta\right)d\theta=-\partial_{\xi_k}\chi_k\left(\xi_k,C_k\right),$$

which proves the lemma.□

A.2.2 Proofs for Section 4.2

Here, we provide a proof of our results when the production function is linear. In that case, prices are fully exogenous $p_k = \phi_k(\xi_k)$ and $\chi_k = \phi_k(\xi_k) C_k$. Re-normalizing the shock ($\tilde{\xi}_k = \phi_k(\xi_k)$), we can directly re-write the system of Lemma A1 as:

$$\frac{\epsilon}{(1+\epsilon)^2}\frac{\theta\pi(\theta)}{(1-T')^2}\frac{\theta}{z(\theta)}\frac{d}{d\theta}\left\{\frac{dV}{dlnp_k}\right\}+\int_{\theta}^{\bar{\theta}} g\left(\gamma(\theta')\frac{dV}{dlnp_k}-\int_{\underline{\theta}}^{\bar{\theta}} g\gamma(\theta')\frac{dV}{dlnp_k}\pi d\theta'\right)\pi d\theta' = -\frac{\epsilon}{1+\epsilon}\frac{\theta\pi(\theta)}{1-T'}\left(\tau_k(\theta)+\partial_{z^*}E_k\right),$$

$$\int_{\underline{\theta}}^{\bar{\theta}} g\frac{dV}{dlnp_k}\pi d\theta = -p_k C_k.$$

As in the main text, we define an increase in the relative price of the necessity, keeping the average price level constant, as $dln\bar{p}_l$, such that $dlnp_l = \bar{s}_h dln\bar{p}_l$ and $dlnp_h = -\bar{s}_l dln\bar{p}_l$. We also define a homogeneous increase in price $dln\bar{p}$, such that $dlnp_l = dlnp_h = dln\bar{p}$. In a two-good economy, a relative increase in the price of necessity and an homogeneous increase in price summarizes all the price changes.

Linear Social Welfare Function (Proposition 2 and Corollary 1) We first provide the proof of Proposition 2:

Proposition 2.
Under $A3-A4$, the response of the optimal tax rate at θ to an increase in the price of k when $\alpha=0$ is:

$$\frac{p_k d}{dp_k}\left\{\frac{T'}{1-T'}\right\} = \frac{1}{z\tilde{\zeta}f(z(\theta))}\mathbb{E}_{z>z(\theta)}\left(\partial_{z^*}e_k-\partial_{z^*}E_k\right)-\frac{T'}{1-T'}\left(\partial_{z^*}e_k-\partial_{z^*}E_k\right).$$

With homothetic preferences ($\partial_{z^*}e_h = s_h$), we have $d_{p_k}T' = 0$. With non-homothetic preferences, under $A1$ the change in tax schedule in response to change in the price of the necessity ($k=l$) and luxury ($k=h$) good satisfies :

$$\frac{p_l d}{dp_l}\left\{\frac{T'}{1-T'}\right\} < 0 \quad and \quad \frac{p_h d}{dp_h}\left\{\frac{T'}{1-T'}\right\} = -\frac{p_l d}{dp_l}\left\{\frac{T'}{1-T'}\right\} > 0 \quad \forall\theta.$$

<u>**Proof:**</u> Under Assumption $A.4$, we have $\gamma(\theta) = 0\ \forall\theta$ (since $G''(V,\theta)=0$ when $G(V,\theta)=\lambda_\theta V$). We therefore have, using Lemma A1:

$$\frac{\epsilon}{(1+\epsilon)^2}\frac{\theta\pi(\theta)}{(1-T')^2}\frac{\theta}{z(\theta)}\frac{d}{d\theta}\left\{\frac{dV}{dlnp_k}\right\} = -\frac{\epsilon}{1+\epsilon}\frac{\theta\pi(\theta)}{1-T'}\left(\tau_k(\theta)+\partial_{z^*}E_k\right).$$

In addition, we have from the envelope condition:

$$\frac{d}{d\theta}\left\{\frac{dV(\theta)}{dlnp_k}\right\} = \left(1+\frac{1}{\epsilon}\right)\frac{1}{\theta}\left(1-T'\right)\frac{dz(\theta)}{dlnp_k}.$$

From the optimality of labor supply ($\frac{1}{\theta}\left(\frac{z(\theta)}{\theta}\right)^{\epsilon^{-1}}/v_{z^*}(z^*(\theta),\mathbf{p}) = 1-T'$), we have:

$$\frac{1}{z}\frac{dz(\theta)}{dlnp_k} = -\frac{\epsilon}{1-T'}\frac{dT'}{dlnp_k}-\epsilon\partial_{z^*}e_k.$$

Plugging these expressions in our formula, we obtain:

$$\begin{aligned}\frac{\epsilon}{1+\epsilon}\frac{\theta\pi(\theta)}{(1-T')}\left(-\frac{1}{1-T'}\frac{dT'}{dlnp_k}-\partial_{z^*}e_k\right) &= -\frac{\epsilon}{1+\epsilon}\frac{\theta\pi(\theta)}{1-T'}\left(\tau_k\left(\theta\right)+\partial_{z^*}E_k\right)\\ \frac{p_k d}{dp_k}\left\{\frac{T'}{1-T'}\right\} &= \frac{1}{1-T'}\left(\tau_k\left(\theta\right)-\left(\partial_{z^*}e_k-\partial_{z^*}E_k\right)\right).\end{aligned}$$

Finally, using the definition of $\tau_k\left(\theta\right)$, we have:

$$\begin{aligned}\frac{p_k d}{dp_k}\left\{\frac{T'}{1-T'}\right\} &= \left(\frac{1+\epsilon}{\epsilon}\frac{1}{\theta\pi(\theta)}\int_{\theta}^{\bar{\theta}}\left(\partial_{z^*}e_l-\partial_{z^*}E_l\right)\pi d\theta'-\frac{T'}{1-T'}\left(\partial_{z^*}e_l-\partial_{z^*}E_l\right)\right)\\ &= \frac{1}{z\tilde{\zeta}f(z(\theta))}\int_{z(\theta)}^{z(\bar{\theta})}\left(\partial_{z^*}e_k-\partial_{z^*}E_k\right)f(z)dz-\frac{T'}{1-T'}\left(\partial_{z^*}e_k-\partial_{z^*}E_k\right),\end{aligned}$$

which proves our formula. Since $z\tilde{\zeta}f(z(\theta))\geq 0$, the expression has the same sign as $\mathcal{F}\left(z\left(\theta\right)\right)$, defined below:

$$\begin{aligned}\mathcal{F}\left(z\left(\theta\right)\right)\equiv\int_{z(\theta)}^{z(\bar{\theta})}\left(\partial_{z^*}e_k-\partial_{z^*}E_k\right)f(z)dz-z\tilde{\zeta}f(z(\theta))\frac{T'}{1-T'}\left(\partial_{z^*}e_k-\partial_{z^*}E_k\right)=\\ \int_{z(\theta)}^{z(\bar{\theta})}\left(\partial_{z^*}e_k-\partial_{z^*}E_k\right)f(z)dz-\left(\partial_{z^*}e_k-\partial_{z^*}E_k\right)\int_{z(\theta)}^{z(\bar{\theta})}\left(1-g\right)f(z)dz.\end{aligned}$$

Inspecting the expression in the second line, we have $\mathcal{F}\left(z\left(\underline{\theta}\right)\right)=\mathcal{F}\left(z\left(\bar{\theta}\right)\right)=0$. Assume now that $\partial_{z^*}e_k$ is decreasing (k is a necessity good) and define θ^* such that $\partial_{z^*}e_k\left(z^*\left(\theta^*\right),\mathbf{p}\right)=\partial_{z^*}E_k$ (note that since $z\left(\theta\right)$is increasing due to the single crossing property and $dV/d\theta>0$ from the envelope condition, $z^*\left(\theta\right)$is increasing in θ). We have:

$$\mathcal{F}'\left(z\left(\theta\right)\right)=-\left(\partial_{z^*}e_k-\partial_{z^*}E_k\right)gf(z)-\left(1-T'\right)\partial_{z^*z^*}e_k\int_{z(\theta)}^{z(\bar{\theta})}\left(1-g\right)f(z)dz.$$

For $\theta\geq\theta^*$, we have $\partial_{z^*}e_k<\partial_{z^*}E_k$ and since $\partial_{z^*z^*}e_k\leq 0$, we have $\mathcal{F}'\left(z\left(\theta\right)\right)>0$ for $\theta\geq\theta^*$ ($\int_{z(\theta)}^{z(\bar{\theta})}\left(1-g\right)f(z)dz\geq 0$, as g is non increasing. Since $\mathcal{F}\left(z\left(\bar{\theta}\right)\right)=0$, this implies $\mathcal{F}\left(z\left(\theta\right)\right)<0$ for $\theta\geq\theta^*$.

Note in addition that we can rewrite $\mathcal{F}\left(z\left(\theta\right)\right)$ as:

$$\mathcal{F}\left(z\left(\theta\right)\right)=-\int_{z(\underline{\theta})}^{z(\theta)}\left(\partial_{z^*}e_k-\partial_{z^*}E_k\right)f(z)dz+\left(\partial_{z^*}e_k-\partial_{z^*}E_k\right)\int_{z(\underline{\theta})}^{z(\theta)}\left(1-g\right)f(z)dz.$$

For $\theta\leq\theta^*$, we have $\partial_{z^*}e_k-\partial_{z^*}E_k>0$ and decreasing in θ, so that:

$$\begin{aligned}\mathcal{F}\left(z\left(\theta\right)\right) &< -\left(\partial_{z^*}e_k-\partial_{z^*}E_k\right)\int_{z(\underline{\theta})}^{z(\theta)}f(z)dz+\left(\partial_{z^*}e_k-\partial_{z^*}E_k\right)\int_{z(\underline{\theta})}^{z(\theta)}\left(1-g\right)f(z)dz\\ &= -\left(\partial_{z^*}e_k-\partial_{z^*}E_k\right)\int_{z(\underline{\theta})}^{z(\theta)}gf(z)dz<0.\end{aligned}$$

Thus, $\mathcal{F}\left(z\left(\theta\right)\right)<0$ for $\theta<\theta^*$, which implies $\frac{p_l d}{dp_l}\left\{\frac{T'}{1-T'}\right\}<0$. By direct inspection, since $\partial_{z^*}e_l-\partial_{z^*}E_l=$

$-\left(\partial_{z^*}e_h - \partial_{z^*}E_h\right)$, we have $\frac{p_h d}{dp_h}\left\{\frac{T'}{1-T'}\right\} = -\frac{p_l d}{dp_l}\left\{\frac{T'}{1-T'}\right\} > 0$. $\square$

We now turn to our welfare analysis and provide a proof of Corollary 1. We also show that welfare decreases after an homogeneous price increase and that the decrease in welfare does not depend on consumption patterns. In particular the welfare decrease is the same when household have homothetic or non-homothetic preferences. The result is intuitive: an homogeneous price increase is equivalent to an homogeneous reduction in households' real wage independently from their consumption preferences. Households reduce their labor supply and, from Proposition 2, tax rates are left unchanged: the real income of all households falls.

Corollary 1. For an increase in the relative price of necessities $d\ln\bar{p}_l$, with $d\ln p_l = \bar{s}_h d\ln\bar{p}_l$ and $d\ln p_h = -\bar{s}_l d\ln\bar{p}_l$, the compensating scheme $dT\left(z\left(\theta\right)\right) = -\left(s_l - \bar{s}_l\right)z^*\left(\theta\right)dln\bar{p}_l$ is feasible but only optimal when preferences are homothetic. With non-homothetic preferences, under $A1-A4$ we have $dV\left(\underline{\theta}\right)/d\bar{p}_l < 0$ and $dV\left(\bar{\theta}\right)/d\bar{p}_l > 0$; $dV\left(\theta\right)/d\bar{p}_l$ is increasing in θ and $\mathbb{E}\left(gdV\left(\theta\right)/d\bar{p}_l\right) = 0$.
For an homogeneous increase in price, $d\ln\bar{p}$, such that $d\ln p_l = d\ln p_h = d\ln\bar{p}$, we have $dV\left(\theta\right)/d\bar{p}$ is negative, decreasing, and independent from consumption preferences.
Proof: Using the formula of lemma A1, we have that for an increase in the relative price of necessities, the change in welfare satisfies:

$$\frac{\epsilon}{(1+\epsilon)^2}\frac{\theta\pi(\theta)}{(1-T')^2}\frac{\theta}{z(\theta)}\frac{d}{d\theta}\left\{\frac{dV}{dln\bar{p}_l}\right\} = -\frac{\epsilon}{1+\epsilon}\frac{\theta\pi(\theta)}{1-T'}\left(\tau_l\left(\theta\right) + \partial_{z^*}E_l - \bar{s}_l\right),$$
$$\int_{\underline{\theta}}^{\bar{\theta}} g\frac{dV}{dln\bar{p}_l}\pi d\theta = 0,$$

where we used $\gamma = 0$ and $dV\left(\theta\right)/d\bar{p}_l = \bar{s}_h dV\left(\theta\right)/dp_l + \bar{s}_l dV\left(\theta\right)/dp_h$. The second equation directly shows that $\mathbb{E}\left(gdV\left(\theta\right)/d\bar{p}_l\right) = 0$. Since $\partial_{z^*}E_l \leq \bar{s}_l$, to show that $dV\left(\theta\right)/d\bar{p}_l$ is increasing it is enough to show that $\tau_l\left(\theta\right)$ is negative. We will then have $d_\theta\left\{dV\left(\theta\right)/d\bar{p}_l\right\} > 0$, which implies, given $\mathbb{E}\left(gdV\left(\theta\right)/d\bar{p}_l\right) = 0$, that $dV\left(\underline{\theta}\right)/d\bar{p}_l < 0$ and $dV\left(\bar{\theta}\right)/d\bar{p}_l > 0$.
Recall that we have:

$$\tau_l\left(\theta\right) = \left(1-T'\right)\left(\frac{1+\epsilon}{\epsilon}\frac{1}{\theta\pi(\theta)}\int_{\theta}^{\bar{\theta}}\left(\partial_{z^*}e_l - \partial_{z^*}E_l\right)\pi d\theta' + \left(\partial_{z^*}e_l - \partial_{z^*}E_l\right)\right).$$

Since $\left(1-T'\right) \geq 0$ and $\theta\pi(\theta) \geq 0$, $\tau_l\left(\theta\right)$ has the same sign as $\tilde{\tau}_l\left(\theta\right) = \int_{\theta}^{\bar{\theta}}\left(\partial_{z^*}e_l - \partial_{z^*}E_l\right)\pi d\theta' + \frac{\epsilon}{1+\epsilon}\theta\pi(\theta)\left(\partial_{z^*}e_l - \partial_{z^*}E_l\right)$.
As in the proof of Proposition 2, define θ^* such that $\partial_{z^*}e_l\left(z^*\left(\theta^*\right), \mathbf{p}\right) = \partial_{z^*}E_l$. For $\theta > \theta^*$, we have $\partial_{z^*}e_l < \partial_{z^*}E_l$, which implies $\tilde{\tau}_l\left(\theta\right) < 0$. In addition, we have $\tilde{\tau}_l'\left(\theta\right) = -\left(1 - \frac{\epsilon}{1+\epsilon}\frac{\theta\pi'(\theta)}{\pi(\theta)}\right)\pi(\theta)\left(\partial_{z^*}e_l - \partial_{z^*}E_l\right) + \left(1-T'\right)z'\left(\theta\right)\partial_{z^*z^*}e_l < 0$ for $\theta \leq \theta^*$under $A4$. Indeed, $\partial_{z^*}e_l - \partial_{z^*}E_l \geq 0$, $\partial_{z^*z^*}e_l \leq 0$ (and $z'\left(\theta\right) \geq 0$ for the tax schedule to be incentive compatible). Since $\tilde{\tau}_l\left(\underline{\theta}\right) = 0$ is decreasing on $\left(\underline{\theta}, \theta^*\right)$ and negative for $\theta > \theta^*$, we have $\tilde{\tau}_l\left(\theta\right)$ negative everywhere, which implies $d_\theta\left\{dV\left(\theta\right)/d\bar{p}_l\right\} > 0$.

For an homogeneous price increase and simplifying the formulas from Lemma A1, we have:

$$\frac{d}{d\theta}\left\{\frac{dV}{dln\bar{p}}\right\}=-\left(1+\epsilon\right)\left(1-T'\right)\frac{z(\theta)}{\theta}$$
$$\int_{\underline{\theta}}^{\bar{\theta}}g\frac{dV}{dln\bar{p}}\pi d\theta=-p_lC_l-p_hC_h=-\int_{\underline{\theta}}^{\bar{\theta}}z\left(\theta\right)\pi d\theta.$$

This implies that $dV\left(\theta\right)/d\bar{p}$ is decreasing and independent from consumption preferences. Next, we have:

$$\begin{aligned}\frac{d}{d\theta}\left\{\frac{dV}{dln\bar{p}}\right\}&=-\left(1+\epsilon\right)\left(1-T'\right)\frac{z(\theta)}{\theta}\\&=-\left(1-T'+\epsilon zT''\right)z'\left(\theta\right)\\\Rightarrow\frac{dV\left(\theta\right)}{dln\bar{p}}&=\frac{dV\left(\underline{\theta}\right)}{dln\bar{p}}-\left(z\left(\theta\right)-\left(1+\epsilon\right)T\left(z\left(\theta\right)\right)+\epsilon z\left(\theta\right)T'\left(z\left(\theta\right)\right)\right).\end{aligned}$$

Finally, recall from the proof of Lemma A1 that we have:

$$\begin{aligned}\int_{\underline{\theta}}^{\bar{\theta}}g\frac{dV}{ddln\bar{p}}\pi d\theta&=-\int_{\underline{\theta}}^{\bar{\theta}}\left(\frac{1}{1-T'}\frac{\epsilon}{1+\epsilon}\theta\frac{d}{d\theta}\left\{\frac{dV(\theta)}{d\xi_k}\right\}-\frac{dV(\theta)}{d\xi_k}\right)\pi\left(\theta\right)d\theta\\&=\int_{\underline{\theta}}^{\bar{\theta}}\left(\epsilon\left(1-T'\right)z(\theta)+z\left(\theta\right)\right)\pi\left(\theta\right)d\theta+\frac{dV\left(\underline{\theta}\right)}{dln\bar{p}}\\\Rightarrow\frac{dV\left(\underline{\theta}\right)}{dln\bar{p}}&=-\int_{\underline{\theta}}^{\bar{\theta}}\epsilon\left(1-T'\right)z(\theta)\pi\left(\theta\right)d\theta<0.\end{aligned}$$

The second line uses $\int_{\underline{\theta}}^{\bar{\theta}}T\left(z\left(\theta\right)\right)\pi d\theta=0$. We therefore have $dV\left(\underline{\theta}\right)/d\bar{p}<0$ and $dV\left(\theta\right)/d\bar{p}$decreasing, which implies that $dV\left(\theta\right)/d\bar{p}$ is everywhere negative.$\square$

Top and bottom tax rates. To conclude this section, we provide formulas for the top and bottom tax rates that we use in our discussion of the quantitative results of Section 5.2. We assume that the Pareto weights g are decreasing and denote $g\left(\underline{\theta}\right)>1>g\left(\bar{\theta}\right)$ their limit at the bottom and top of the distribution. Similarly, $\partial_{z^*}e_l\left(\underline{\theta}\right)>\partial_{z^*}E_l>\partial_{z^*}e_l\left(\bar{\theta}\right)$ are the marginal propensity to spend on the necessity good at the bottom and top of the distribution. For an increase in the price of the necessity good, our tax formula is:

$$\begin{aligned}\frac{p_ld}{dp_l}\left\{\frac{T'}{1-T'}\right\}&=\frac{1+\epsilon}{\epsilon}\frac{1}{\theta\pi(\theta)}\int_{\theta}^{\bar{\theta}}\left(\partial_{z^*}e_l-\partial_{z^*}E_l\right)\pi(\theta)d\theta-\frac{T'}{1-T'}\left(\partial_{z^*}e_l-\partial_{z^*}E_l\right),\\&=\frac{T'}{1-T'}\left(\frac{\int_{\theta}^{\bar{\theta}}\left(\partial_{z^*}e_l-\partial_{z^*}E_l\right)\pi(\theta)d\theta}{\int_{\theta}^{\bar{\theta}}\left(1-g\right)\pi(\theta)d\theta}-\left(\partial_{z^*}e_l-\partial_{z^*}E_l\right)\right).\end{aligned}$$

Using l'Hopital's rule, we have:

$$\frac{p_l d}{dp_l}\left\{\frac{T'}{1-T'}\right\}(\underline{\theta}) = -\frac{g(\underline{\theta})}{g(\underline{\theta})-1}\frac{T'}{1-T'}\left(\partial_{z^*}e_l(\underline{\theta}) - \partial_{z^*}E_l\right)$$
$$\frac{p_l d}{dp_l}\left\{\frac{T'}{1-T'}\right\}(\bar{\theta}) = -\frac{g(\bar{\theta})}{1-g(\bar{\theta})}\frac{T'}{1-T'}\left(\partial_{z^*}E_l - \partial_{z^*}e_l(\bar{\theta})\right),$$

which gives the formulas for the top and bottom tax rates.

Non-Linear Social Welfare Function (Proposition 3) Before proving Proposition 3, let us briefly discuss the tax formula when the social welfare function is non-linear. From Lemma A1, we have:

$$\frac{\epsilon}{(1+\epsilon)^2}\frac{\theta\pi(\theta)}{(1-T')^2}\frac{\theta}{z(\theta)}\frac{d}{d\theta}\left\{\frac{dV}{dlnp_k}\right\} + \int_\theta^{\bar{\theta}} g\left(\gamma(\theta')\frac{dV}{dlnp_k} - \int_{\underline{\theta}}^{\bar{\theta}} g\gamma(\theta')\frac{dV}{dlnp_k}\pi d\theta'\right)\pi d\theta' = -\frac{\epsilon}{1+\epsilon}\frac{\theta\pi(\theta)}{1-T'}\left(\tau_k(\theta) + \partial_{z^*}E_k\right),$$
$$\int_{\underline{\theta}}^{\bar{\theta}} g\frac{dV}{dlnp_k}\pi d\theta = -p_k C_k.$$

This formula implicitly determines the optimal tax rate. To see this more precisely, recall that we have:

$$\frac{d}{d\theta}\left\{\frac{dV}{dlnp_k}\right\} = -(1+\epsilon)\frac{z}{\theta}\left(1-T'\right)\left(\frac{1}{1-T'}\frac{dT'}{dlnp_k} + \partial_{z^*}e_k\right)$$
$$\frac{dV}{dlnp_k} = -v_{z^*}\left(\frac{dT}{dlnp_k} + e_k\right).$$

We can therefore rewrite the formula as:

$$z\tilde{\zeta}f(z(\theta))\frac{p_k d}{dp_k}\left\{\frac{T'}{1-T'}\right\} = \int_{z(\theta)}^{z(\bar{\theta})} g\left(\gamma v_{z^*}\left(\frac{dT}{dlnp_k} + e_k\right) - \int_{z(\underline{\theta})}^{z(\bar{\theta})} g\gamma v_{z^*}\left(\frac{dT}{dlnp_k} + e_k\right) f dz\right) f dz$$
$$+ z\tilde{\zeta}f(z(\theta))\frac{p_k d}{dp_k}\left\{\frac{T'_{lin}}{1-T'_{lin}}\right\}$$
$$z\tilde{\zeta}f(z(\theta))\frac{p_k d}{dp_k}\left\{\frac{T'_{lin}}{1-T'_{lin}}\right\} = \mathbb{E}_{z>z(\theta)}\left(\partial_{z^*}e_k - \partial_{z^*}E_k\right) - z\tilde{\zeta}f(z(\theta))\frac{T'}{1-T'}\left(\partial_{z^*}e_k - \partial_{z^*}E_k\right).$$

The tax formula is the sum of two terms. The second one is the change in tax rate when the social welfare function is linear, $\frac{p_k d}{dp_k}\left\{\frac{T'_{lin}}{1-T'_{lin}}\right\}$. As before, through this term and Channel #1 and #2, an increase in the price of necessities induces more redistribution towards the rich. The first term captures an income effect of redistribution on Pareto weights. If the planner implements the tax reform arising with a linear social welfare function, the tax burden increases for low income households, which decreases their utility and raises their Pareto weight. If the Pareto weights of lower income households increases more than average, then the first term is positive, which counteracts the impact of Channel #1 and #2 and pushes for more redistribution towards low-income households. We show however in Proposition 3 that this counterbalancing effect does not fully offset the impact of Channel #1 and #2, even when it is feasible to compensate all households.

We now turn to the proof of Proposition 3. As in the linear case, we also show that the welfare change after an homogeneous price increase does not depend on consumption patterns. The argument is the same as in the linear case.

Proposition 3. Assume that $-G''(V,\theta)/G'(V,\theta)$ is positive and non increasing (e.g., a CARA or CRRA function) and that $A1-A3$ are satisfied. For an increase in the relative price of necessities, the compensating scheme $dT(z(\theta)) = -(s_l - \bar{s}_l) z^* d\ln\bar{p}_l$ is feasible but only optimal when preferences are homothetic. With non-homothetic preferences, the change in welfare of agent θ, $dV^G/d\bar{p}_l(\theta)$, satisfies $dV/d\bar{p}_l(\underline{\theta}) < dV^G/d\bar{p}_l(\underline{\theta}) < 0$, $dV^G/d\bar{p}_l(\theta) - dV^G/d\bar{p}_l(\underline{\theta}) < dV/d\bar{p}_l(\theta) - dV/d\bar{p}_l(\underline{\theta})$, and $\mathbb{E}\left(g dV^G(\theta)/d\bar{p}_l\right) = 0$, where $dV/d\bar{p}_l(\theta)$ is the welfare impact of price change with a linear social welfare function satisfying $\lambda_\theta \propto G'(V(\theta),\theta)$.

If in addition, $\bar{\theta} = \infty$, the distribution of types is bounded by a Pareto distribution, $\theta\pi'(\theta)/\pi(\theta) \leq -1-\omega$ for θ large enough, and $G(V,\theta)$ is either CARA or CRRA with a relative risk aversion coefficient strictly higher than 0, then we have $dV^G/d\bar{p}_l(\theta) \backsim dV/d\bar{p}_l(\theta)$ at infinity.

For an homogeneous increase in price, $d\ln\bar{p}$, such that $d\ln p_l = d\ln p_h = d\ln\bar{p}$, we have that $dV^G(\theta)/d\bar{p}$ is independent from consumption preferences.

Proof: Using the formula of lemma A1, we have for an increase in the relative price of the necessity good:

$$\frac{\epsilon}{(1+\epsilon)^2}\frac{\theta\pi(\theta)}{(1-T')^2}\frac{\theta}{z(\theta)}\frac{d}{d\theta}\left\{\frac{dV}{dln\bar{p}_l}\right\} = -\int_\theta^{\bar{\theta}} g\left(\gamma(\theta')\frac{dV}{dln\bar{p}_l} - \int_{\underline{\theta}}^{\bar{\theta}} g\gamma(\theta')\frac{dV}{dln\bar{p}_l}\pi d\theta'\right)\pi d\theta' + \frac{\epsilon}{(1+\epsilon)^2}\frac{\theta\pi(\theta)}{(1-T')^2}\frac{\theta}{z(\theta)}\frac{d}{d\theta}\left\{\frac{dV^{lin}}{dln\bar{p}_l}\right\},$$

$$\int_{\underline{\theta}}^{\bar{\theta}} g\frac{dV}{dln\bar{p}_l}\pi d\theta = 0,$$

$$\frac{\epsilon}{(1+\epsilon)^2}\frac{\theta\pi(\theta)}{(1-T')^2}\frac{\theta}{z(\theta)}\frac{d}{d\theta}\left\{\frac{dV^{lin}}{dln\bar{p}_l}\right\} = -\frac{\epsilon}{1+\epsilon}\frac{\theta\pi(\theta)}{1-T'}\left(\tau_l(\theta) + \partial_{z^*}E_l - \bar{s}_l\right),$$

where $\frac{dV^{lin}}{dln\bar{p}_l}$ is the welfare change with a linear social welfare function described in Corollary 1.
First, note that implementing $\frac{dV}{dln\bar{p}_l} = 0$ implies $-dT(z(\theta)) - (s_l - \bar{s}_l) z^* d\ln\bar{p}_l = 0$. With this tax change, we have $dz(\theta) = -z\tilde{\zeta}\left(dT'(z(\theta))/(1-T') + (\partial_{z^*}e_l - \bar{s}_l)\right) = 0$ and the total cost of the reform is:

$$-\int (s_l - \bar{s}_l) z^* d\ln\bar{p}_l \pi d\theta = 0,$$

so compensating all households is feasible (as it is budget neutral). Such compensation is however not optimal as it does not solve the first equation of the system. From the proof of Corollary 1, we know that

$\frac{dV^{lin}}{dln\bar{p}_l} > 0$ for all θ. We consider the auxiliary system:

$$\frac{\epsilon}{(1+\epsilon)^2}\frac{\theta\pi(\theta)}{(1-T')^2}\frac{\theta}{z(\theta)}\frac{d}{d\theta}\left\{\frac{dV_0}{dln\bar{p}_l}\right\} = -\int_\theta^{\bar{\theta}} g\gamma\left(\theta'\right)\frac{dV_0}{dln\bar{p}_l}\pi d\theta' + \frac{\epsilon}{(1+\epsilon)^2}\frac{\theta\pi(\theta)}{(1-T')^2}\frac{\theta}{z(\theta)}\frac{d}{d\theta}\left\{\frac{dV^{lin}}{dln\bar{p}_l}\right\},$$

$$\int_{\underline{\theta}}^{\bar{\theta}} g\gamma\left(\theta\right)\frac{dV_0}{dln\bar{p}_l}\pi d\theta = 0,$$

$$\frac{\epsilon}{(1+\epsilon)^2}\frac{\theta\pi(\theta)}{(1-T')^2}\frac{\theta}{z(\theta)}\frac{d}{d\theta}\left\{\frac{dV_1}{dln\bar{p}_l}\right\} = -\int_\theta^{\bar{\theta}} g\left(\gamma\left(\theta'\right)\frac{dV_1}{dln\bar{p}_l} - \int_{\underline{\theta}}^{\bar{\theta}} g\gamma\left(\theta'\right)\frac{dV_1}{dln\bar{p}_l}\pi d\theta'\right)\pi d\theta',$$

$$\int_{\underline{\theta}}^{\bar{\theta}} g\frac{dV_1}{dln\bar{p}_l}\pi d\theta = -\int_{\underline{\theta}}^{\bar{\theta}} g\frac{dV_0}{dln\bar{p}_l}\pi d\theta.$$

We then have: $\frac{dV}{dln\bar{p}_l} = \frac{dV_0}{dln\bar{p}_l} + \frac{dV_1}{dln\bar{p}_l}$.
We first consider the term $\frac{dV_0}{dln\bar{p}_l}$. Since $\frac{dV^{lin}}{dln\bar{p}_l} > 0$, we necessarily have that $dV/d\bar{p}_l$ is strictly negative at $\underline{\theta}$. If not, $-\int_\theta^{\bar{\theta}} g\gamma\left(\theta'\right)\frac{dV_0}{dln\bar{p}_l}\pi d\theta'$ is non decreasing and therefore non negative at $\underline{\theta}$. Therefore, $\frac{d}{d\theta}\left\{\frac{dV_0}{dln\bar{p}_l}\right\}(\underline{\theta}) \geq \frac{d}{d\theta}\left\{\frac{dV^{lin}}{dln\bar{p}_l}\right\}(\underline{\theta}) > 0$, so $dV_0/dln\bar{p}_l$ is positive in a neighborhood around $\underline{\theta}$.
We show that this leads to a contradiction. Define θ_0 the first θ such that $dV_0/dln\bar{p}_l$ is 0. θ_0 must exist since $\int_{\underline{\theta}}^{\bar{\theta}} g\gamma\left(\theta\right)\frac{dV_0}{dln\bar{p}_l}\pi d\theta = 0$, so there exists $dV_0/dln\bar{p}_l < 0$. Then, since $\int_{\underline{\theta}}^{\bar{\theta}} g\gamma\left(\theta'\right)\frac{dV_0}{dln\bar{p}_l}\pi d\theta' = 0$, $\int_{\theta_0}^{\bar{\theta}} g\gamma\left(\theta'\right)\frac{dV_0}{dln\bar{p}_l}\pi d\theta' < 0$. Since in addition $\frac{d}{d\theta}\left\{\frac{dV^{lin}}{dln\bar{p}_l}\right\}$ is positive at θ_0, we have $\frac{d}{d\theta}\left\{\frac{dV_0}{dln\bar{p}_l}\right\}$ strictly positive at θ_0. Since $dV_0/dln\bar{p}_l\left(\theta_0\right) = 0$, by definition, this implies $dV_0/dln\bar{p}_l\left(\theta\right) < 0$ in $(\theta_0 - \epsilon, \theta_0)$, which contradicts the fact that $dV_0/dln\bar{p}_l\left(\theta\right) > 0$ on (θ, θ_0). Therefore, we have $dV/d\bar{p}_l$ is strictly negative at $\underline{\theta}$. Using the same logic, $dV_0/dln\bar{p}_l$ cannot be negative around $\bar{\theta}$.

In addition, $\int_\theta^{\bar{\theta}} g\gamma\left(\theta'\right)\frac{dV_0}{dln\bar{p}_l}\pi d\theta' > 0$ on $(\underline{\theta}, \bar{\theta})$. Indeed, suppose not and denote again θ_0 the smallest θ_0 such that $\int_{\theta_0}^{\bar{\theta}} g\gamma\left(\theta'\right)\frac{dV_0}{dln\bar{p}_l}\pi d\theta' = 0$. $dV_0/dln\bar{p}_l$ cannot be negative at θ_0 since then $\int_{\theta_0}^{\bar{\theta}} g\gamma\left(\theta'\right)\frac{dV_0}{dln\bar{p}_l}\pi d\theta'$ would be increasing at θ_0, contradicting the fact that it is positive for $\theta < \theta_0$. Therefore, $dV_0/dln\bar{p}_l\left(\theta\right)$ is non negative at θ_0 – however, using the same reasoning as before, this would imply $dV_0/dln\bar{p}_l\left(\theta\right) > 0$ for $\theta > \theta_0$. As before, suppose not and consider θ_1 such that $dV_0/dln\bar{p}_l\left(\theta_1\right) = 0$. We have $\int_{\theta_0}^{\bar{\theta}} g\gamma\left(\theta'\right)\frac{dV_0}{dln\bar{p}_l}\pi d\theta' = 0$ and $\int_{\theta_0}^{\theta_1} g\gamma\left(\theta'\right)\frac{dV_0}{dln\bar{p}_l}\pi d\theta' > 0$, so $\int_{\theta_1}^{\bar{\theta}} g\gamma\left(\theta'\right)\frac{dV_0}{dln\bar{p}_l}\pi d\theta' < 0$. This implies $\frac{d}{d\theta}\left\{\frac{dV_0}{dln\bar{p}_l}\right\}(\theta_1) > 0$ so $dV_0/dln\bar{p}_l\left(\theta_1\right)$ is negative below θ_1, which contradicts that θ_1 exists.

Therefore, $dV_0/dln\bar{p}_l\left(\theta\right) > 0$ for $\theta > \theta_0$, so that $\int_{\theta_0}^{\bar{\theta}} g\gamma\left(\theta'\right)\frac{dV_0}{dln\bar{p}_l}\pi d\theta' > 0$, which contradicts that $\theta_0$exists.

Next, we show that $\int_\theta^{\bar{\theta}} g\gamma\left(\theta'\right)\frac{dV_0}{dln\bar{p}_l}\pi d\theta' > 0$, for all θ, implies $\int_\theta^{\bar{\theta}} g\frac{dV_0}{dln\bar{p}_l}\pi d\theta' \geq 0$. Since $dV_0/dln\bar{p}_l$ is positive $\bar{\theta}$, we have $\int_\theta^{\bar{\theta}} g\frac{dV_0}{dln\bar{p}_l}\pi d\theta' > 0$ for θ high enough. Suppose that there exists θ_0 such that $\int_{\theta_0}^{\bar{\theta}} g\frac{dV_0}{dln\bar{p}_l}\pi d\theta' = 0$ and consider the highest $\theta_0$such that it is the case. We therefore have $\int_\theta^{\bar{\theta}} g\frac{dV_0}{dln\bar{p}_l}\pi d\theta' > 0$ for $\theta > \theta_0$ and:

$$\begin{aligned}\int_{\theta_0}^{\bar{\theta}} g\gamma\left(\theta\right)\frac{dV_0}{dln\bar{p}_l}\pi d\theta &= \int_{\theta_0}^{\bar{\theta}}\gamma'\left(\theta\right)\int_{\theta'}^{\bar{\theta}} g\frac{dV_0}{dln\bar{p}_l}\pi d\theta' d\theta + \gamma\left(\theta_0\right)\int_{\theta_0}^{\bar{\theta}} g\frac{dV_0}{dln\bar{p}_l}\pi d\theta \\ &= \int_{\theta_0}^{\bar{\theta}}\gamma'\left(\theta\right)\int_{\theta'}^{\bar{\theta}} g\frac{dV_0}{dln\bar{p}_l}\pi d\theta' d\theta \leq 0,\end{aligned}$$

where the second line uses the fact that $\gamma'\left(\theta\right) \leq 0$ and $\int_\theta^{\bar{\theta}} g\frac{dV_0}{dln\bar{p}_l}\pi d\theta' > 0$ for $\theta > \theta_0$. This contradicts the fact that $\int_{\theta_0}^{\bar{\theta}} g\gamma(\theta) \frac{dV_0}{dln\bar{p}_l}\pi d\theta > 0$, so $\int_{\theta}^{\bar{\theta}} g\frac{dV_0}{dln\bar{p}_l}\pi d\theta' \geq 0$ for all θ.

Next, we analyze $dV_1/dln\bar{p}_l$. First, if $\int_{\underline{\theta}}^{\bar{\theta}} g\frac{dV_0}{dln\bar{p}_l}\pi d\theta' = 0$, then $\int_{\underline{\theta}}^{\bar{\theta}} g\frac{dV_1}{dln\bar{p}_l}\pi d\theta' = 0$ and $dV_1/dln\bar{p}_l = 0$ solves the system's equations. Therefore $dV/dln\bar{p}_l = dV_0/dln\bar{p}_l$. Suppose $\int_{\underline{\theta}}^{\bar{\theta}} g\frac{dV_0}{dln\bar{p}_l}\pi d\theta' > 0$, which implies $\int_{\underline{\theta}}^{\bar{\theta}} g\frac{dV_1}{dln\bar{p}_l}\pi d\theta' < 0$. Suppose further $\int_{\underline{\theta}}^{\bar{\theta}} g\gamma\frac{dV_1}{dln\bar{p}_l}\pi d\theta' < 0$. Then we necessarily have that $\gamma\frac{dV_1}{dln\bar{p}_l} < \int_{\underline{\theta}}^{\bar{\theta}} g\gamma\frac{dV_1}{dln\bar{p}_l}\pi d\theta'$ in a neighborhood of $\underline{\theta}$. Indeed, using the same reasoning as above, if $\gamma\frac{dV_1}{dln\bar{p}_l}(\underline{\theta}) > \int_{\underline{\theta}}^{\bar{\theta}} g\gamma\frac{dV_1}{dln\bar{p}_l}\pi d\theta'$, then we have $\gamma\frac{dV_1}{dln\bar{p}_l} > \int_{\underline{\theta}}^{\bar{\theta}} g\gamma\frac{dV_1}{dln\bar{p}_l}\pi d\theta'$ for all θ, which would be a contradiction. We can use the same reasoning as before, if not there is a smallest θ_0 such that $\gamma\frac{dV_1}{dln\bar{p}_l}(\theta_0) = \int_{\underline{\theta}}^{\bar{\theta}} g\gamma\frac{dV_1}{dln\bar{p}_l}\pi d\theta'$; but at θ_0 we necessarily have $\frac{d}{d\theta}\left\{\frac{dV_1}{dln\bar{p}_l}\right\}(\theta_0) > 0$. Indeed, since $\gamma\frac{dV_1}{dln\bar{p}_l}(\theta) > \int_{\underline{\theta}}^{\bar{\theta}} g\gamma\frac{dV_1}{dln\bar{p}_l}\pi d\theta'$ in $(\underline{\theta},\theta_1)$, we have

$$\int_{\theta_0}^{\bar{\theta}} g\left(\gamma(\theta')\frac{dV_1}{dln\bar{p}_l} - \int_{\underline{\theta}}^{\bar{\theta}} g\gamma(\theta')\frac{dV_1}{dln\bar{p}_l}\pi d\theta'\right)\pi d\theta' < 0,$$

so $\frac{d}{d\theta}\left\{\frac{dV_1}{dln\bar{p}_l}\right\}(\theta_0) > 0$.

Since γ is positive decreasing and $\frac{dV_1}{dln\bar{p}_l}(\theta_0) < 0$, $\frac{d}{d\theta}\left\{\frac{dV_1}{dln\bar{p}_l}\right\}(\theta_0) > 0$ at z_0, this means that $\gamma\frac{dV_1}{dln\bar{p}_l}$ is increasing at θ_0, which is a contradicts the fact that $\gamma\frac{dV_1}{dln\bar{p}_l} > \int_{\underline{\theta}}^{\bar{\theta}} g\gamma\frac{dV_1}{dln\bar{p}_l}\pi d\theta'$ below θ_0. Finally, if $\gamma\frac{dV_1}{dln\bar{p}_l}(\underline{\theta}) = \int_{\underline{\theta}}^{\bar{\theta}} g\gamma\frac{dV_1}{dln\bar{p}_l}\pi d\theta'$, at $\underline{\theta}$ we would have that $\frac{d}{d\theta}\left\{\frac{dV_1}{dln\bar{p}_l}\right\} = 0$ and $\frac{dV_1}{dln\bar{p}_l} < 0$. Since γ is positive decreasing, $\gamma\frac{dV_1}{dln\bar{p}_l} > \int_{\underline{\theta}}^{\bar{\theta}} g\gamma\frac{dV_1}{dln\bar{p}_l}\pi d\theta'$ in a neighborhood of $\underline{\theta}$ and we can use the same reasoning to get $\gamma\frac{dV_1}{dln\bar{p}_l} > \int_{\underline{\theta}}^{\bar{\theta}} g\gamma\frac{dV_1}{dln\bar{p}_l}\pi d\theta'$ for all θ, which is a contradiction.

Therefore, we have $\gamma\frac{dV_1}{dln\bar{p}_l} < \int_{\underline{\theta}}^{\bar{\theta}} g\gamma\frac{dV_1}{dln\bar{p}_l}\pi d\theta'$ in a neighborhood of $\underline{\theta}$.

Next, we have that $D(\theta) = \int_{\theta}^{\bar{\theta}} g\left(\gamma(\theta')\frac{dV_1}{dln\bar{p}_l} - \int_{\underline{\theta}}^{\bar{\theta}} g\gamma(\theta')\frac{dV_1}{dln\bar{p}_l}\pi d\theta'\right)\pi d\theta'$ is positive on the interval $(\underline{\theta},\bar{\theta})$. Consider the smallest θ_0 such that it is 0 at θ_0 and negative in a neighborhood above. First, note we cannot have $\gamma\frac{dV_1}{dln\bar{p}_l} < \int_{\underline{\theta}}^{\bar{\theta}} g\gamma(\theta')\frac{dV_1}{dln\bar{p}_l}\pi d\theta'$ at θ_0 or in a neighborhood above, since $D(\theta_0)$ would be locally increasing. This means that $\gamma\frac{dV_1}{dln\bar{p}_l} \geq \int_{\underline{\theta}}^{\bar{\theta}} g\gamma(\theta')\frac{dV_1}{dln\bar{p}_l}\pi d\theta'$ in the neighborhood above θ_0, which implies by the same reasoning as at $\underline{\theta}$ that $\gamma\frac{dV_1}{dln\bar{p}_l} \geq \int_{\underline{\theta}}^{\bar{\theta}} g\gamma(\theta')\frac{dV_1}{dln\bar{p}_l}\pi d\theta'$ for all $\theta > z$ and implies (since $\gamma\frac{dV_1}{dln\bar{p}_l}$ cannot be constant) that $D(\theta)$ is positive everywhere above θ_0, a contradiction.

Therefore, when $\int_{\underline{\theta}}^{\bar{\theta}} g\gamma\frac{dV_1}{dln\bar{p}_l}\pi d\theta' < 0$, we have that $\frac{dV_1}{dln\bar{p}_l}$ is non positive and decreasing. Since the equation determining $\frac{dV_1}{dln\bar{p}_l}$ is linear, when $\int_{\underline{\theta}}^{\bar{\theta}} g\gamma\frac{dV_1}{dln\bar{p}_l}\pi d\theta' > 0$, $\frac{dV_1}{dln\bar{p}_l}$ would be non negative and increasing. Thus, to have $\int_{\underline{\theta}}^{\bar{\theta}} g\frac{dV_1}{dln\bar{p}_l}\pi d\theta' < 0$, we need $\frac{dV_1}{dln\bar{p}_l}$ non positive and decreasing and

$$\int_{\theta}^{\bar{\theta}} g\left(\gamma(\theta')\frac{dV_1}{dln\bar{p}_l} - \int_{\underline{\theta}}^{\bar{\theta}} g\gamma(\theta')\frac{dV_1}{dln\bar{p}_l}\pi d\theta'\right)\pi d\theta' > 0.$$

Since $\frac{dV}{dln\bar{p}_l} = \frac{dV_0}{dln\bar{p}_l} + \frac{dV_1}{dln\bar{p}_l}$ and $\frac{dV_1}{dln\bar{p}_l}(\underline{\theta}) < 0$, $\frac{dV_0}{dln\bar{p}_l}(\underline{\theta}) < 0$ then $\frac{dV}{dln\bar{p}_l}(\underline{\theta}) < 0$. Furthermore, since $\int_{\theta}^{\bar{\theta}} g\left(\gamma(\theta')\frac{dV_1}{dln\bar{p}_l} - \int_{\underline{\theta}}^{\bar{\theta}} g\gamma(\theta')\frac{dV_1}{dln\bar{p}_l}\pi d\theta'\right)\pi d\theta' > 0$, $\int_{\theta}^{\bar{\theta}} g\frac{dV_0}{dln\bar{p}_l}\pi d\theta' > 0$ and $\int_{\underline{\theta}}^{\bar{\theta}} g\frac{dV_0}{dln\bar{p}_l}\pi d\theta' = 0$, we have $\int_{\theta}^{\bar{\theta}} g\left(\gamma(\theta')\frac{dV}{dln\bar{p}_l} - \int_{\underline{\theta}}^{\bar{\theta}} g\gamma(\theta')\frac{dV}{dln\bar{p}_l}\pi d\theta'\right)\pi d\theta' > 0$, so $\frac{d}{d\theta}\frac{dV}{dln\bar{p}_l} < \frac{d}{d\theta}\frac{dV^{lin}}{dln\bar{p}_l}$. Since $0 = \int_{\underline{\theta}}^{\bar{\theta}} g\frac{dV}{dln\bar{p}_l}\pi d\theta = g\frac{dV}{dln\bar{p}_l}(\underline{\theta}) + \int_{\underline{\theta}}^{\bar{\theta}} \frac{d}{d\theta}\frac{dV}{dln\bar{p}_l}\int_{\theta}^{\bar{\theta}} g\pi d\theta = g\frac{dV^{lin}}{dln\bar{p}_l}(\underline{\theta}) + \int_{\underline{\theta}}^{\bar{\theta}} \frac{d}{d\theta}\frac{dV^{lin}}{dln\bar{p}_l}\int_{\theta}^{\bar{\theta}} g\pi d\theta$, we have $\frac{dV}{dln\bar{p}_l}(\underline{\theta}) > \frac{dV^{lin}}{dln\bar{p}_l}(\underline{\theta})$. In addition since $\frac{d}{d\theta}\frac{dV}{dln\bar{p}_l} < \frac{d}{d\theta}\frac{dV^{lin}}{dln\bar{p}_l}$ we have $\frac{dV}{dln\bar{p}_l}(\theta) - \frac{dV}{dln\bar{p}_l}(\underline{\theta}) < \frac{dV^{lin}}{dln\bar{p}_l}(\theta) - \frac{dV^{lin}}{dln\bar{p}_l}(\underline{\theta})$, which proves the first set of result of Proposition 3.

Next, since $\theta\pi'(\theta)/\pi(\theta) \leq -1-\omega$, we have for any $\theta > \theta_0$, $\theta_0$large enough, $\pi(\theta) \leq \pi(\theta_0)\left(\theta/\theta_0\right)^{-1-\omega}$. This implies, since g is strictly decreasing when $G(V)$ is either CARA or CRRA, that $g < \delta < 1$ for θ large enough. Therefore, $0 \leq \frac{T'}{1-T'} \leq \frac{1}{1+\gamma}\left(1-\delta\right)\frac{1+\epsilon}{\epsilon}$, which implies that there is some $1-\bar{T}' > 0$ such that $1-T' > 1-\bar{T}'$ for all θ.

Note that we have:

$$\frac{d}{d\theta}\left\{\frac{dV^{lin}}{dln\bar{p}_l}\right\} = -\left(1+\epsilon\right)\left(1-T'\right)\frac{z(\theta)}{\theta}\left(\tau_l\left(\theta\right)+\partial_{z^*}E_l-\bar{s}_l\right)$$
$$= -\left(1+\epsilon\right)\left(1-T'\right)^{1+\epsilon}\theta^{\epsilon}\left(\tau_l\left(\theta\right)+\partial_{z^*}E_l-\bar{s}_l\right).$$

Using again the fact that $\pi(\theta) \leq \pi(\theta_0)\left(\theta/\theta_0\right)^{-1-\gamma}$, it is direct to show that $0 < m < \tau_l\left(\theta\right)+\partial_{z^*}E_l-\bar{s}_l < M$, so we have, since $1 \geq 1-T' > 1-\bar{T}' > 0$,$0 < \left(1+\epsilon\right)\bar{m}\theta^{\epsilon} < \frac{d}{d\theta}\left\{\frac{dV^{lin}}{dln\bar{p}_l}\right\} < \left(1+\epsilon\right)\bar{M}\theta^{\epsilon}$. This result implies that for any $\theta > \theta_0$, $\theta_0$large enough, $m\bar{\theta}^{1+\epsilon} < \frac{dV^{lin}}{dln\bar{p}_l}(\theta) - \frac{dV^{lin}}{dln\bar{p}_l}(\theta_0) < \bar{M}\theta^{1+\epsilon}$,$\frac{d}{d\theta}\left\{\frac{dV^{lin}}{dln\bar{p}_l}\right\}$ and $\frac{dV^{lin}}{dln\bar{p}_l}(\theta)$ grows at the same rate as θ^{ϵ} and $\theta^{1+\epsilon}$ respectively.

We now show that $F(\theta) = \frac{(1+\epsilon)^2}{\epsilon}\frac{(1-T')^2}{\theta\pi(\theta)}\frac{z(\theta)}{\theta}\int_{\theta}^{\bar{\theta}} g\left(\gamma\left(\theta'\right)\frac{dV_1}{dln\bar{p}_l} - \int_{\underline{\theta}}^{\bar{\theta}} g\gamma\left(\theta'\right)\frac{dV_1}{dln\bar{p}_l}\pi d\theta'\right)\pi d\theta'$ grows at a smaller rate than θ^{ϵ}, which implies that $\frac{d}{d\theta}\frac{dV}{dln\bar{p}_l}(\theta) \sim \frac{d}{d\theta}\frac{dV^{lin}}{dln\bar{p}_l}$ and $\frac{dV}{dln\bar{p}_l}(\theta) \sim \frac{dV^{lin}}{dln\bar{p}_l}$ for θ large enough. Indeed, we have:

$$\frac{d}{d\theta}\left\{\frac{dV}{dln\bar{p}_l}\right\} = -\frac{(1+\epsilon)^2}{\epsilon}\frac{(1-T')^2}{\theta\pi(\theta)}\frac{z(\theta)}{\theta}\int_{\theta}^{\bar{\theta}} g\left(\gamma\left(\theta'\right)\frac{dV}{dln\bar{p}_l} - \int_{\underline{\theta}}^{\bar{\theta}} g\gamma\left(\theta'\right)\frac{dV}{dln\bar{p}_l}\pi d\theta'\right)\pi d\theta' + \frac{d}{d\theta}\left\{\frac{dV^{lin}}{dln\bar{p}_l}\right\},$$
$$= -F(\theta) + \frac{d}{d\theta}\left\{\frac{dV^{lin}}{dln\bar{p}_l}\right\},$$

so if $F(\theta)$ grows at a lower rate than $\frac{d}{d\theta}\left\{\frac{dV^{lin}}{dln\bar{p}_l}\right\}$ (e.g. grows at rate $ln(\theta)$), then $\frac{d}{d\theta}\frac{dV}{dln\bar{p}_l}(\theta) \sim \frac{d}{d\theta}\frac{dV^{lin}}{dln\bar{p}_l}$. This implies also that tax rate are the same for high income households. Denote $X \equiv -\int_{\underline{\theta}}^{\bar{\theta}} g\gamma\left(\theta'\right)\frac{dV_1}{dln\bar{p}_l}\pi d\theta' \geq 0$, we know from the previous result that $\frac{d}{d\theta}\frac{dV}{dln\bar{p}_l} < \frac{d}{d\theta}\frac{dV^{lin}}{dln\bar{p}_l}$ so $\frac{dV_1}{dln\bar{p}_l}(\theta) < \frac{dV_1}{dln\bar{p}_l}(\theta_0) + \bar{M}\left(\theta^{1+\epsilon} - \theta_0^{1+\epsilon}\right)$. We now consider θ_1 such that $\frac{dV_1}{dln\bar{p}_l}(\theta_0) + \bar{M}\left(\theta_1^{1+\epsilon} - \theta_0^{1+\epsilon}\right) > 0$ for $\theta > \theta_1$: we have

$$F(\theta) = \frac{(1+\epsilon)^2}{\epsilon}\frac{(1-T')^2}{\theta\pi(\theta)}\frac{z(\theta)}{\theta}\int_{\theta}^{\bar{\theta}} g\left(\gamma\left(\theta'\right)\frac{dV}{dln\bar{p}_l} - \int_{\underline{\theta}}^{\bar{\theta}} g\gamma\left(\theta'\right)\frac{dV}{dln\bar{p}_l}\pi d\theta'\right)\pi d\theta'$$
$$= \frac{(1+\epsilon)^2}{\epsilon}\frac{(1-T')^{2+\epsilon}\theta^{\epsilon}}{\theta\pi(\theta)}\int_{\theta}^{\bar{\theta}}\left(g\gamma\left(\theta'\right)\frac{dV}{dln\bar{p}_l} + X\right)\pi d\theta'$$
$$\leq \frac{(1+\epsilon)^2}{\epsilon}\left(1-T'\right)^{2+\epsilon}\theta^{\epsilon-1}\int_{\theta}^{\bar{\theta}}\left(\theta'/\theta\right)^{-1-\omega}\left(-G''(V\left(\theta'\right))\left(\frac{dV_1}{dln\bar{p}_l}(\theta_0) + \bar{M}\left(\theta'^{1+\epsilon} - \theta_0^{1+\epsilon}\right)\right) + G'(V\left(\theta'\right))X\right)d\theta',$$

where we used again the fact that $\pi(\theta') \leq \pi(\theta)\left(\theta'/\theta\right)^{-1-\omega}$.

Now, we have $V'\left(\theta\right) = \frac{1}{\theta}\left(\frac{z(\theta)}{\theta}\right)^{1+\frac{1}{\epsilon}} = \left(1-T'\right)^{1+\epsilon}\theta^{\epsilon}$, which implies $0 < m_V < V\left(\theta\right)/\theta^{1+\epsilon} < M_V$, for θ large enough $V\left(\theta\right)$ grows at rate $\theta^{1+\epsilon}$. Now, if G is CARA with coefficient β, we have:

$$F(\theta) \leq \frac{(1+\epsilon)^2}{\epsilon}\theta^{\epsilon-1}\int_{\theta}^{\bar{\theta}}\left(\theta'/\theta\right)^{-1-\omega}e^{-\beta m_V\theta'^{1+\epsilon}}\left(\beta\left(\frac{dV_1}{dln\bar{p}_l}(\theta_0) + \bar{M}\left(\theta'^{1+\epsilon} - \theta_0^{1+\epsilon}\right)\right) + X\right)d\theta'$$
$$\leq C\theta^{1+2\epsilon}e^{-\beta m_V\theta^{1+\epsilon}}$$

Since $\theta'^{1+2\epsilon}e^{-\beta m_V \theta^{1+\epsilon}} = o\left(\theta^{\epsilon}\right)$, we have $\frac{d}{d\theta}\frac{dV}{dln\bar{p}_l}(\theta) \sim \frac{d}{d\theta}\frac{dV^{lin}}{dln\bar{p}_l}$.

Next, if G is CRRA with coefficient β, we have:

$$F(\theta) \leq \frac{(1+\epsilon)^2}{\epsilon}\theta^{\epsilon-1}\int_{\theta}^{\bar{\theta}} \left(\theta'/\theta\right)^{-1-\omega} e^{-\beta m_V \theta'^{1+\epsilon}} \left(\beta\left(m_V\theta'\right)^{-(1+\beta)(1+\epsilon)}\left(\frac{dV_1}{dln\bar{p}_l}(\theta_0) + \bar{M}\left(\theta'^{1+\epsilon} - \theta_0^{1+\epsilon}\right)\right) + \left(m_V\theta'\right)^{-\beta(1+\epsilon)}X\right)d\theta'$$
$$\leq C\theta^{\epsilon-\beta(1+\epsilon)}.$$

Since $\theta^{\epsilon-\beta(1+\epsilon)} = o\left(\theta^{\epsilon}\right)$, we have $\frac{d}{d\theta}\frac{dV}{dln\bar{p}_l}(\theta) \sim \frac{d}{d\theta}\frac{dV^{lin}}{dln\bar{p}_l}$. In both cases this directly implies $\frac{dV}{dln\bar{p}_l}(\theta) \sim \frac{dV^{lin}}{dln\bar{p}_l}$.

For an homogeneous price change, we have:

$$\frac{\epsilon}{(1+\epsilon)^2}\frac{\theta\pi(\theta)}{(1-T')^2}\frac{\theta}{z(\theta)}\frac{d}{d\theta}\left\{\frac{dV}{dln\bar{p}_l}\right\} = -\int_{\theta}^{\bar{\theta}} g\left(\gamma\left(\theta'\right)\frac{dV}{dln\bar{p}_l} - \int_{\underline{\theta}}^{\bar{\theta}} g\gamma\left(\theta'\right)\frac{dV}{dln\bar{p}_l}\pi d\theta'\right)\pi d\theta'$$
$$+ \frac{\epsilon}{(1+\epsilon)^2}\frac{\theta\pi(\theta)}{(1-T')^2}\frac{\theta}{z(\theta)}\frac{d}{d\theta}\left\{\frac{dV^{lin}}{dln\bar{p}_l}\right\},$$
$$\int_{\underline{\theta}}^{\bar{\theta}} g\frac{dV}{dln\bar{p}_l}\pi d\theta = -\int_{\underline{\theta}}^{\bar{\theta}} z\pi d\theta,$$
$$\frac{\epsilon}{(1+\epsilon)^2}\frac{\theta\pi(\theta)}{(1-T')^2}\frac{\theta}{z(\theta)}\frac{d}{d\theta}\left\{\frac{dV^{lin}}{dln\bar{p}_l}\right\} = -\frac{\epsilon}{1+\epsilon}\frac{\theta\pi(\theta)}{1-T'}$$

Note that $\frac{dV^{lin}}{dln\bar{p}_l}$ is now independent from consumption preferences, so $\frac{dV}{dln\bar{p}_l}$ is independent from consumption preferences and $\int_{\underline{\theta}}^{\bar{\theta}} g\frac{dV}{dln\bar{p}_l}\pi d\theta < 0$. $\square$

Non-Linear Social Welfare Function in a Simple Example Finally we provide the derivations in our simple three-agent example. We have $0 = \theta_p < \theta_m < \theta_r$ and agent preferences are given by

$$V_i = u(c_l, c_h) - \frac{1}{1+\frac{1}{\varepsilon}}\left(\frac{z}{\theta_i}\right)^{1+\frac{1}{\varepsilon}}.$$

Given the single crossing property, the IC constraint are local and downward binding and are:

$$V_m = V_p$$
$$V_r = V_m + \frac{1}{1+\frac{1}{\varepsilon}}\left(\frac{z_m}{\theta_m}\right)^{1+\frac{1}{\varepsilon}} - \frac{1}{1+\frac{1}{\varepsilon}}\left(\frac{z_m}{\theta_r}\right)^{1+\frac{1}{\varepsilon}}.$$

The planner's problem is to solve:

$$\sup_{z_i, V_i}\sum_i G\left(V_i, \theta_i\right)\pi_i$$
$$s.t. V_m = V_p$$
$$V_r = V_m + \frac{1}{1+\frac{1}{\varepsilon}}\left(\frac{z_m}{\theta_m}\right)^{1+\frac{1}{\varepsilon}} - \frac{1}{1+\frac{1}{\varepsilon}}\left(\frac{z_m}{\theta_r}\right)^{1+\frac{1}{\varepsilon}}$$
$$\sum_i\left(z_i^* - z_i\right)\pi_i + \sum_k\left(p_k C_k - \chi_k\left(\xi_k, C_k\right)\right) = 0$$

We obtain that the tax rate is 0 at the top (no distortion at the top), 0 at the bottom (since $\theta_p = 0$), while the tax rate for $\theta_m$satisfies:

$$(g_m - 1)\,\pi_m + (g_p - 1)\,\pi_p = \pi_m \frac{T'_m}{1 - T'_m} \frac{1}{1 - (\theta_m/\theta_r)^{1+\frac{1}{\epsilon}}}.$$

We assume that at initial price $\partial_{z^*} e_{l,p} > \partial_{z^*} e_{l,m} = \partial_{z^*} E_l > \partial_{z^*} e_{l,r}$ and $\partial_{z^*} E_l < \bar{s}_l$. As before, we assume $v_{z^*} = 1$ for all three agents and we consider an increase in the relative price of the necessity good, $d\bar{p}_l$. We first show that we can express all the welfare changes, $dV_i/d\bar{p}_l$, in terms of the labor supply change of the middle income households, $dz_m/d\bar{p}_l$.

Differentiating the IC constraints and the budget constraint, we obtain:

$$\begin{aligned}
\frac{dV_p}{d\bar{p}_l} &= \frac{dV_m}{d\bar{p}_l} \\
\frac{dV_r}{d\bar{p}_l} &= \frac{dV_m}{d\bar{p}_l} + \left(1 - T'_m\right)\left(1 - \left(\frac{\theta_m}{\theta_r}\right)^{1+\frac{1}{\varepsilon}}\right)\frac{dz_m}{d\bar{p}_l} \\
0 &= \frac{dV_p}{d\bar{p}_l}\pi_p + \frac{dV_m}{d\bar{p}_l}\pi_m + \frac{dV_r}{d\bar{p}_l}\pi_r - T'_m \pi_m \frac{dz_m}{d\bar{p}_l}.
\end{aligned}$$

Solving for these equations, we obtain:

$$\begin{aligned}
\frac{dV_p}{d\bar{p}_l} = \frac{dV_m}{d\bar{p}_l} &= -\left(1 - T'_m\right)\left(1 - \left(\frac{\theta_m}{\theta_r}\right)^{1+\frac{1}{\varepsilon}}\right) g_r \pi_r \frac{dz_m}{d\bar{p}_l} \\
\frac{dV_r}{d\bar{p}_l} &= \left(1 - T'_m\right)\left(1 - \left(\frac{\theta_m}{\theta_r}\right)^{1+\frac{1}{\varepsilon}}\right)\left(1 - g_r \pi_r\right)\frac{dz_m}{d\bar{p}_l}
\end{aligned}$$

Note that $\frac{dV_p}{d\bar{p}_l}$ and $\frac{dV_m}{d\bar{p}_l}$ have the opposite signs to $\frac{dz_m}{d\bar{p}_l}$, while $\frac{dV_r}{d\bar{p}_l}$has the same sign.

Differentiating the tax formula, we obtain, defining $\gamma_i = -\frac{G''_i}{G'_i} > 0$:

$$\begin{aligned}
\frac{1}{1 - T'_m}\frac{1}{1 - (\theta_m/\theta_r)^{1+\frac{1}{\epsilon}}}\frac{1}{\epsilon}\pi_m \frac{dlnz_m}{dln\bar{p}_l} &= -\left(\left(\partial_{z^*} e_{l,r} - \partial_{z^*} E_l\right)\pi_r + \left(\partial_{z^*} E - \bar{s}_l\right)\pi_m \frac{1}{1 - T'_m}\frac{1}{1 - (\theta_m/\theta_r)^{1+\frac{1}{\epsilon}}}\right) \\
&\quad - \left(\left(1 - \pi_r g_r\right)\gamma_r \frac{dV_r}{dln\bar{p}_l} - \left(\pi_p g_p \gamma_p \frac{dV_p}{dln\bar{p}_l} + \pi_m g_m \gamma_m \frac{dV_m}{dln\bar{p}_l}\right)\right) g_r \pi_r \\
&= -\left(\left(\partial_{z^*} e_{l,r} - \partial_{z^*} E_l\right)\pi_r + \left(\partial_{z^*} E - \bar{s}_l\right)\pi_m \frac{1}{1 - T'_m}\frac{1}{1 - (\theta_m/\theta_r)^{1+\frac{1}{\epsilon}}}\right) - \mathcal{G}\frac{dlnz_m}{dln\bar{p}_l},
\end{aligned}$$

with $\mathcal{G} = z_m\left(1 - T'_m\right)\left(1 - \left(\frac{\theta_m}{\theta_r}\right)^{1+\frac{1}{\varepsilon}}\right)\left(\left(1 - g_r\pi_r\right)^2\gamma_r + g_r\pi_r\left(\pi_p g_p \gamma_p + \pi_m g_m \gamma_m\right)\right) g_r \pi_r > 0$. When $\gamma_i = 0$ (G is linear), $\mathcal{G} = 0$, When $\gamma_i > 0$ (G is concave), $\mathcal{G} > 0$. Note in addition that, since l is a necessity,$\left(\partial_{z^*} e_{l,r} - \partial_{z^*} E_l\right)\pi_r + \left(\partial_{z^*} E - \bar{s}_l\right)\pi_m \frac{1}{1-T'_m}\frac{1}{1-(\theta_m/\theta_r)^{1+\frac{1}{\epsilon}}} < 0$. Therefore, we have

$$\frac{dV_p}{d\bar{p}_l} = \frac{dV_m}{d\bar{p}_l} < 0 < \frac{dV_r}{d\bar{p}_l},$$

and denoting $\frac{dV_i^{lin}}{d\bar{p}_l}$ the tax rate with a linear social welfare function, $\frac{dV_i}{d\bar{p}_l} = \frac{1}{1+\mathcal{G}}\frac{dV_i^{lin}}{d\bar{p}_l}$. Using $\left(1 - T'_m\right)\frac{dlnz_m}{dln\bar{p}_l} =$

$-\epsilon\left(\frac{dT'^{lin}_m}{d\bar{p}_l}+\left(1-T'_m\right)\left(\partial_{z^*}E_l-\bar{s}_l\right)\right)$, it is direct to show:

$$\frac{dT'_m}{d\bar{p}_l}=\frac{1}{1+\mathcal{G}}\left(\frac{dT'^{lin}_m}{d\bar{p}_l}-\mathcal{G}\left(1-T'_m\right)\left(\partial_{z^*}E_l-\bar{s}_l\right)\right)$$

$$\frac{1}{1-\left(\theta_m/\theta_h\right)^{1+\frac{1}{\epsilon}}}\pi_m\frac{d}{dln\bar{p}_l}\left\{\frac{T'^{lin}_m}{1-T'^{lin}_m}\right\}=\left(\partial_{z^*}e_{l,h}-\partial_{z^*}E_l\right)\pi_h<0.$$

This proves our formulas for the three-agent example.

A.2.3 Proofs for Section 4.3

Non-Linear Production Function (Proposition 4) We now prove Proposition 4, our main result with non-linear production functions. Recall that we consider a cost shifter, $p^*_k=1/\partial_{\xi_k}\phi_k$, which implies $\partial_{p^*_k}\phi_k=1$ and $\partial_{p^*_k}\chi_k=\left(1-\alpha+t_w\right)^{-1}C_k$.[1] As before, we define an increase int the relative price of the necessity $d\mathrm{ln}\bar{p}_l$, such that $d\mathrm{ln}p^*_l=\bar{s}_h d\mathrm{ln}\bar{p}_l$ and $d\mathrm{ln}p^*_h=-\bar{s}_l d\mathrm{ln}\bar{p}_l$. We will also provide formulas for an homogeneous price change $d\mathrm{ln}\bar{p}$, such that $d\mathrm{ln}p^*_l=d\mathrm{ln}p^*_h=d\mathrm{ln}\bar{p}$.

We first prove the following Lemma, which characterizes the response of the tax rate to an increase in the price of k in partial equilibrium. Much of the derivation is similar to the derivation of Proposition 2.

Lemma A2. Under $A3-A4$, the partial equilibrium response of the income tax to a change in the relative price of necessities is:

$$\frac{\partial}{\partial\mathrm{ln}p^*_k}\left\{\frac{T'}{1-T'}\right\}=\frac{1-t_w}{z\tilde{\zeta}f(z(\theta))}\mathbb{E}_{z>z(\theta)}\left(\partial_{z^*}e_k-\partial_{z^*}E_k\right)-\left(\frac{T'}{1-T'}+t_w\right)\left(\partial_{z^*}e_k-\partial_{z^*}E_k\right).$$

Under $A1$, $\partial_{lnp^*_l}T'=-\partial_{lnp^*_h}T'$ is negative for all θ. The response to an increase in the relative price of necessities $\partial_{\bar{p}_l}T'=\partial_{p^*_l}T$ is also negative.
<u>Proof:</u> Using the formulas of Lemma A1, we have for an arbitrary shifter ξ_k:

$$\frac{\epsilon}{(1+\epsilon)^2}\frac{\theta\pi(\theta)}{(1-T')^2}\frac{\theta}{z(\theta)}\frac{d}{d\theta}\left\{\frac{dV}{d\xi_k}\right\}=-\frac{\epsilon}{1+\epsilon}\frac{\theta\pi(\theta)}{1-T'}\sum_{m=l,h}\left(\tau_m\left(\theta\right)+\partial_{z^*}E_m\right)\frac{1}{p_m}\frac{dp_m}{d\xi_k},$$

$$(1-t_w)\int_{\underline{\theta}}^{\bar{\theta}}g\frac{dV}{d\xi_k}\pi d\theta=-\partial_{\xi_k}\chi_k\left(\xi_k,C_k\right).$$

Since we consider the partial equilibrium response where the shifter is p^*_k and price do not endogenously respond, we have $\frac{dp_k}{d\xi_k}=1$, $\frac{dp_m}{d\xi_k}=0$ for $m\neq k$. Using as before:

$$\frac{d}{d\theta}\left\{\frac{\partial V(\theta)}{\partial lnp^*_k}\right\}=-\epsilon z\left(1+\frac{1}{\epsilon}\right)\frac{1}{\theta}\left(1-T'\right)\left(\frac{1}{1-T'}\frac{\partial T'}{\partial lnp_k}+\partial_{z^*}e_k\right),$$

[1] With monopolistic competition ($\tau_w=\alpha$), this is obvious since $\chi_k=C_k\phi_k$ so $\partial_{p^*_k}\chi_k=C_k$. With competitive firms ($\tau_w=0$), we can rewrite the pricing function as $\phi_k(\xi_k,C_k)=\tilde{\phi}_k(\xi_k)C_k^{-\alpha}=\partial_{C_k}\chi_k(\xi_k,C_k)$ so $\chi_k(\xi_k,C_k)=\phi_k(\xi_k,C_k)C_k/(1-\alpha)+\underline{\chi_k}$, where the potential fixed cost $\underline{\chi_k}$is assumed to be independent from ξ_k.

we obtain after the same algebra as in Proposition 2:

$$\frac{\partial}{\partial \ln p_k^*}\left\{\frac{T'}{1-T'}\right\}=\frac{1-t_w}{z\tilde{\zeta} f(z(\theta))}\mathbb{E}_{z>z(\theta)}\left(\partial_{z^*}e_k-\partial_{z^*}E_k\right)-\left(\frac{T'}{1-T'}+t_w\right)\left(\partial_{z^*}e_k-\partial_{z^*}E_k\right).$$

We now want to show that $\mathcal{F}\left(z\left(\theta\right)\right)=\left(1-t_w\right)\mathbb{E}_{z>z(\theta)}\left(\partial_{z^*}e_k-\partial_{z^*}E_k\right)-z\tilde{\zeta} f(z(\theta))\left(\frac{T'}{1-T'}+t_w\right)\left(\partial_{z^*}e_k-\partial_{z^*}E_k\right)$ is everywhere negative for a necessity good. Since it has the opposite sign for a luxury, it will show that it is everywhere positive for a luxury. Assume that $\partial_{z^*}e_k$ is decreasing (k is a necessity good) and define θ^* such that $\partial_{z^*}e_k\left(z^*\left(\theta^*\right),\mathbf{p}\right)=\partial_{z^*}E_k$. As before, we have:

$$\mathcal{F}'\left(z\left(\theta\right)\right)=-\left(1-t_w\right)\left(\partial_{z^*}e_k-\partial_{z^*}E_k\right)gf(z)-\left(1-t_w\right)\left(1-T'\right)\partial_{z^*z^*}e_k\int_{z(\theta)}^{z(\bar{\theta})}\left(1-g\right)f(z)dz.$$

For $\theta\geq\theta^*$, we have $\partial_{z^*}e_k<\partial_{z^*}E_k$ and since $\partial_{z^*z^*}e_k\leq 0$, we have $\mathcal{F}'\left(z\left(\theta\right)\right)>0$ for $\theta\geq\theta^*$. Since $\mathcal{F}\left(z\left(\bar{\theta}\right)\right)=0$, this implies $\mathcal{F}\left(z\left(\theta\right)\right)<0$ for $\theta\geq\theta^*$.
Note in addition that we can rewrite $\mathcal{F}\left(z\left(\theta\right)\right)$ as:

$$\mathcal{F}\left(z\left(\theta\right)\right)=-\left(1-t_w\right)\int_{z(\underline{\theta})}^{z(\theta)}\left(\partial_{z^*}e_k-\partial_{z^*}E_k\right)f(z)dz+\left(1-t_w\right)\left(\partial_{z^*}e_k-\partial_{z^*}E_k\right)\int_{z(\underline{\theta})}^{z(\theta)}\left(1-g\right)f(z)dz.$$

For $\theta\leq\theta^*$, we have $\partial_{z^*}e_k-\partial_{z^*}E_k>0$ and decreasing in θ so:

$$\begin{aligned}\mathcal{F}\left(z\left(\theta\right)\right)&<-\left(1-t_w\right)\left(\partial_{z^*}e_k-\partial_{z^*}E_k\right)\int_{z(\underline{\theta})}^{z(\theta)}f(z)dz+\left(1-t_w\right)\left(\partial_{z^*}e_k-\partial_{z^*}E_k\right)\int_{z(\underline{\theta})}^{z(\theta)}\left(1-g\right)f(z)dz\\&=-\left(1-t_w\right)\left(\partial_{z^*}e_k-\partial_{z^*}E_k\right)\int_{z(\underline{\theta})}^{z(\theta)}gf(z)dz<0.\end{aligned}$$

So $\mathcal{F}\left(z\left(\theta\right)\right)<0$ for $\theta<\theta^*$, which implies $\frac{p_l d}{dp_l}\left\{\frac{T'}{1-T'}\right\}<0$. By direct inspection, since $\partial_{z^*}e_l-\partial_{z^*}E_l=-\left(\partial_{z^*}e_h-\partial_{z^*}E_h\right)$, we have $\frac{p_h d}{dp_h}\left\{\frac{T'}{1-T'}\right\}=-\frac{p_l d}{dp_l}\left\{\frac{T'}{1-T'}\right\}>0$. $\square$

We now turn to the proof of Proposition 4. We will consider both a change in the relative price of necessity and a homogeneous price increase, which together span the entire space of price changes.
Proposition 4. Under $A3-A4$, the partial equilibrium response of the income tax to an increase in the price of k is:

$$\frac{\partial}{\partial \ln p_k^*}\left\{\frac{T'}{1-T'}\right\}=\frac{1-t_w}{z\tilde{\zeta} f(z(\theta))}\mathbb{E}_{z>z(\theta)}\left(\partial_{z^*}e_k-\partial_{z^*}E_k\right)-\left(\frac{T'}{1-T'}+t_w\right)\left(\partial_{z^*}e_k-\partial_{z^*}E_k\right).$$

Under $A1$, $\partial_{\bar{p}_l}T'=\partial_{p_l^*}T'$ is negative for all θ, $\partial_{lnp_h^*}T'=-\partial_{lnp_l^*}T'$ is positive for all θ and for an homogeneous price change $\partial_{\bar{p}}T'=0$.

In general equilibrium, response of the income tax to an increase in the relative price of necessities is:

$$\frac{dT'}{d\bar{p}_l}=\left(1-\alpha\left(\sigma+\Omega\right)\right)^{-1}\frac{\partial T'}{\partial\bar{p}_l},$$

with $\Omega = \frac{1}{1-t_w}\frac{\zeta}{\bar{s}_h\bar{s}_l}\left(\mathbb{E}_z\left((\tau_l + \partial_{z^*}E_l - \bar{s}_l)^2\right) + \frac{\alpha\zeta}{1-t_w-\alpha\zeta}\left(\mathbb{E}_z(\tau_l + \partial_{z^*}E_l - \bar{s}_l)\right)^2\right) > 0$. When $\alpha > 0$, $\frac{dT'}{d\bar{p}_l} < \frac{\partial T'}{\partial \bar{p}_l} < 0$, when $\alpha < 0$, $\frac{\partial T'}{\partial \bar{p}_l} < \frac{dT'}{d\bar{p}_l} < 0$.

The response of the income tax to an homogenous increase in prices :

$$\frac{dT'}{d\bar{p}} = (1-\alpha(\sigma+\Omega))^{-1}\frac{1}{\bar{s}_h\bar{s}_l}\frac{\alpha}{1-\alpha}\left((\partial_{z^*}E_l - \bar{s}_l) + \frac{\zeta}{1-t_w-\alpha\zeta}\mathbb{E}_z(\tau_l + \partial_{z^*}E_l - \bar{s}_l)\right)\frac{\partial T'}{\partial \bar{p}_l}.$$

With $\frac{dT'}{d\bar{p}} > 0$ if $\alpha > 0$, $\frac{dT'}{d\bar{p}} < 0$ if $\alpha < 0$.

Before proving Proposition 4, let us briefly discuss the change in tax rate for a homogeneous increase in prices y: $dlnp_l^* = dlnp_h^* = dln\bar{p}$. In partial equilibrium, this price change has no effect on tax rates. Indeed, per Lemma $A2$, we have $\partial_{lnp_h^*}T' + \partial_{lnp_l^*}T' = 0$. In general equilibrium, as the price increase is homogeneous, there are no direct substitution effects. With homothetic preferences, there is no change in relative prices and relative quantities. With non-homothetic preference, a homogeneous increase in prices endogenously increases the relative price of luxuries. In the proof of Proposition 4, we show that, when $\alpha \geq 0$, the increase in the relative price of h is given by:

$$\frac{d\log(p_h/p_l)}{d\log p^*} = -(1-\alpha(\sigma+\Omega))^{-1}\frac{1}{\bar{s}_h\bar{s}_l}\frac{\alpha}{1-\alpha}\left((\partial_{z^*}E_l - \bar{s}_l) + \frac{\zeta}{1-t_w-\alpha\zeta}\mathbb{E}_z(\tau_l + \partial_{z^*}E_l - \bar{s}_l)\right) \geq 0.$$

An increase in inflation reduces real income and therefore decreases the share of h. As a result, the relative price of h increases through market size effects ($\alpha > 0$), and this increase is amplified through the substitution and income effects described in the main text.

Thus, while homogeneous exogenous price increases have no impact on tax rates in partial equilibrium, we find that they lead to more redistribution in general equilibrium. Households reduce their labor supply and therefore reallocate their income towards the necessity product, which increases the relative price of the luxury product. It then becomes optimal to redistribute to lower income households.

Proof: We have already proved the first part of Proposition 4 in Lemma $A.2$. We now derive formulas for the general equilibrium response. For an arbitrary cost shifter ξ_k, we have:

$$\frac{\epsilon}{(1+\epsilon)^2}\frac{\theta\pi(\theta)}{(1-T')^2}\frac{\theta}{z(\theta)}\frac{d}{d\theta}\left\{\frac{dV}{d\xi_k}\right\} = -\frac{\epsilon}{1+\epsilon}\frac{\theta\pi(\theta)}{1-T'}\sum_{m=l,h}\left(\tau_m(\theta) + \partial_{z^*}E_m\right)\frac{1}{p_m}\frac{dp_m}{d\xi_k}.$$

$$(1-t_w)\int_{\underline{\theta}}^{\bar{\theta}} g\frac{dV}{d\xi_k}\pi d\theta = -\partial_{\xi_k}\chi_k(\xi_k, C_k).$$

Using the same algebra as in Lemma $A.2$, we obtain:

$$\frac{d}{d\xi_k}\left\{\frac{T'}{1-T'}\right\} = \sum_{m=l,h}\frac{\partial}{\partial \mathrm{ln}p_k^*}\left\{\frac{T'}{1-T'}\right\}\frac{1}{p_m}\frac{dp_m}{d\xi_k}$$

$$\frac{\partial}{\partial \mathrm{ln}p_k^*}\left\{\frac{T'}{1-T'}\right\} = \frac{1-t_w}{z\tilde{\zeta}f(z(\theta))}\mathbb{E}_{z>z(\theta)}\left(\partial_{z^*}e_k - \partial_{z^*}E_k\right) - \left(\frac{T'}{1-T'} + t_w\right)\left(\partial_{z^*}e_k - \partial_{z^*}E_k\right).$$

Using the fact, from Lemma $A.2$ that $\partial_{\ln p_h^*} T' = -\partial_{\ln p_l^*} T' = -\partial_{\ln \bar{p}_l} T'$, we have

$$\frac{d}{d\xi_k}\left\{\frac{T'}{1-T'}\right\} = \frac{\partial}{\partial \ln \bar{p}_l}\left\{\frac{T'}{1-T'}\right\}\left(\frac{1}{p_l}\frac{dp_l}{d\xi_k} - \frac{1}{p_h}\frac{dp_h}{d\xi_k}\right),$$

so the general equilibrium response of the tax rate only depends on the endogenous increase of the relative price of necessities. Next, we have using our pricing function:

$$\frac{1}{p_l}\frac{dp_l}{d\xi_k} - \frac{1}{p_h}\frac{dp_h}{d\xi_k} = \frac{1}{p_l}\frac{\partial p_l}{\partial \xi_k} - \frac{1}{p_h}\frac{\partial p_h}{\partial \xi_k} - \alpha\left(\frac{1}{C_l}\frac{dC_l}{d\xi_k} - \frac{1}{C_h}\frac{dC_h}{d\xi_k}\right)$$
$$\bar{s}_l\frac{1}{p_l}\frac{dp_l}{d\xi_k} + \bar{s}_h\frac{1}{p_h}\frac{dp_h}{d\xi_k} = \bar{s}_l\frac{1}{p_l}\frac{\partial p_l}{\partial \xi_k} + \bar{s}_h\frac{1}{p_h}\frac{\partial p_h}{\partial \xi_k} - \alpha\left(\bar{s}_l\frac{1}{C_l}\frac{dC_l}{d\xi_k} + \bar{s}_h\frac{1}{C_h}\frac{dC_h}{d\xi_k}\right).$$

To complete the proof of Proposition 4, we therefore need to determine the response of aggregate demand to price shifter ξ_k. We record the derivation in the following Lemma:

Lemma A3. The response of aggregate consumption to an arbitrary cost shift $\xi_k$is given by:

$$\begin{aligned}
\frac{1}{C_l}\frac{dC_l}{d\xi_k} - \frac{1}{C_h}\frac{dC_h}{d\xi_k} =& -\frac{\zeta}{1-t_w}\frac{1}{\bar{s}_h\bar{s}_l}\int_{\underline{\theta}}^{\bar{\theta}}\left(\tau_l + \left(\partial_{z^*}E_l - \bar{s}_l\right)\right)^2\frac{z(\theta)}{Z}\pi d\theta\left(\frac{1}{p_l}\frac{dp_l}{d\xi_k} - \frac{1}{p_h}\frac{dp_h}{d\xi_k}\right)\\
&-\frac{\zeta}{1-t_w}\frac{1}{\bar{s}_h\bar{s}_l}\int_{\underline{\theta}}^{\bar{\theta}}\left(\tau_l + \left(\partial_{z^*}E_l - \bar{s}_l\right)\right)\frac{z(\theta)}{Z}\pi d\theta\left(\bar{s}_l\frac{1}{p_l}\frac{dp_l}{d\xi_k} + \bar{s}_h\frac{1}{p_h}\frac{dp_h}{d\xi_k}\right)\\
&-\sigma\left(\frac{1}{p_l}\frac{dp_l}{d\xi_k} - \frac{1}{p_h}\frac{dp_h}{d\xi_k}\right) - \frac{1}{1-t_w}\frac{1}{\bar{s}_h\bar{s}_l}\left(\partial_{z^*}E_l - \bar{s}_l\right)\frac{1}{Z}\partial_{\xi_k}\chi_k\left(\xi_k, C_k\right),\\
\bar{s}_l\frac{1}{C_l}\frac{dC_l}{d\xi_k} + \bar{s}_h\frac{1}{C_h}\frac{dC_h}{d\xi_k} =& -\frac{\zeta}{1-t_w}\int_{\underline{\theta}}^{\bar{\theta}}\left(\tau_l\left(\theta\right) + \partial_{z^*}E_l - \bar{s}_l\right)\frac{z(\theta)}{Z}\pi d\theta\left(\frac{1}{p_l}\frac{dp_l}{d\xi_k} - \frac{1}{p_h}\frac{dp_h}{d\xi_k}\right)\\
&-\frac{\zeta}{1-t_w}\left(\bar{s}_l\frac{1}{p_l}\frac{dp_l}{d\xi_k} + \bar{s}_h\frac{1}{p_h}\frac{dp_h}{d\xi_k}\right)\\
&-\frac{1}{1-t_w}\frac{1}{Z}\partial_{\xi_k}\chi_k\left(\xi_k, C_k\right),
\end{aligned}$$

and we have $\frac{1}{C_l}\frac{\partial C_l}{\partial ln\bar{p}_l} - \frac{1}{C_h}\frac{\partial C_h}{\partial ln\bar{p}_l}, \bar{s}_l\frac{1}{C_l}\frac{\partial C_l}{\partial ln\bar{p}} + \bar{s}_h\frac{1}{C_h}\frac{\partial C_h}{\partial ln\bar{p}} \leq 0$, $\frac{1}{C_l}\frac{\partial C_l}{\partial ln\bar{p}} - \frac{1}{C_h}\frac{\partial C_h}{\partial ln\bar{p}}, \bar{s}_l\frac{1}{C_l}\frac{\partial C_l}{\partial ln\bar{p}_l} + \bar{s}_h\frac{1}{C_h}\frac{\partial C_h}{\partial ln\bar{p}_l} \geq 0$.

Proof: We have, using Slutsky's formula:

$$\begin{aligned}
\frac{1}{C_m}\frac{dC_m}{d\xi_k} &= \frac{1}{C_m}\int_{\underline{\theta}}^{\bar{\theta}}\frac{dc_m}{d\xi_k}\pi d\theta\\
&= \frac{1}{C_m}\int_{\underline{\theta}}^{\bar{\theta}}\partial_{z^*}c_m\left(\frac{dz^*}{d\xi_k} - c_l\frac{dp_l}{d\xi_k} - c_h\frac{1}{p_h}\frac{dp_h}{d\xi_k}\right)\pi d\theta + \frac{1}{C_m}\int_{\underline{\theta}}^{\bar{\theta}}\partial_{p_l}c_m^h\frac{dp_l}{d\xi_k}\pi d\theta + \frac{1}{C_m}\int_{\underline{\theta}}^{\bar{\theta}}\partial_{p_h}c_m^h\frac{dp_h}{d\xi_k}\pi d\theta\\
&= \frac{1}{C_m}\int_{\underline{\theta}}^{\bar{\theta}}\partial_{z^*}c_m\left(\frac{1}{v_{z^*}}\frac{dV}{d\xi_k} + \left(1-T'\right)\frac{dz}{d\xi_k}\right)\pi d\theta + \mathcal{S}_{ml}\frac{1}{p_l}\frac{dp_l}{d\xi_k} + \mathcal{S}_{mh}\frac{1}{p_h}\frac{dp_h}{d\xi_k}\\
&= \frac{1}{C_m}\int_{\underline{\theta}}^{\bar{\theta}}\frac{d}{d\theta}\left\{\frac{dV}{d\xi_k}\right\}\int_{\theta}^{\bar{\theta}}\left(\partial_{z^*}c_m - \partial_{z^*}C_m\right)\pi d\theta + \frac{\partial_{z^*}C_m}{C_m}\int_{\underline{\theta}}^{\bar{\theta}}\frac{1}{v_{z^*}}\frac{dV}{d\xi_k}\pi d\theta\\
&+ \frac{1}{C_m}\int_{\underline{\theta}}^{\bar{\theta}}\partial_{z^*}c_m\left(1-T'\right)\frac{dz}{d\xi_k}\pi d\theta + \mathcal{S}_{ml}\frac{1}{p_l}\frac{dp_l}{d\xi_k} + \mathcal{S}_{mh}\frac{1}{p_h}\frac{dp_h}{d\xi_k},
\end{aligned}$$

where the last line is the third line integrated by part.
Next, using from the proof of Lemma $A1$:

$$\int_{\underline{\theta}}^{\bar{\theta}}\left(\left(\frac{1}{1-T'}-(1-t_w)\right)\frac{\epsilon}{1+\epsilon}\theta\frac{d}{d\theta}\left\{\frac{dV(\theta)}{d\xi_k}\right\}-(1-t_w)\frac{1}{v_{z^*}}\frac{dV(\theta)}{d\xi_k}\right)\pi(\theta)\,d\theta=-(1-t_w)\int_{\underline{\theta}}^{\bar{\theta}}g(\theta)\frac{dV(\theta)}{d\xi_k}\pi(\theta)\,d\theta$$
$$=\partial_{\xi_k}\chi_k(\xi_k,C_k),$$

we obtain:

$$\begin{aligned}\frac{1}{C_m}\frac{dC_m}{d\xi_k}&=\frac{1}{C_m}\int_{\underline{\theta}}^{\bar{\theta}}\frac{d}{d\theta}\left\{\frac{dV}{d\xi_k}\right\}\int_{\theta}^{\bar{\theta}}(\partial_{z^*}c_m-\partial_{z^*}C_m)\,\pi d\theta' d\theta\\&+\frac{1}{C_m}\int_{\underline{\theta}}^{\bar{\theta}}\partial_{z^*}c_m\left(1-T'\right)\frac{dz}{d\xi_k}\pi d\theta+\mathcal{S}_{ml}\frac{1}{p_l}\frac{dp_l}{d\xi_k}+\mathcal{S}_{mh}\frac{1}{p_h}\frac{dp_h}{d\xi_k}\\&+\frac{1}{1-t_w}\frac{\partial_{z^*}C_m}{C_m}\int_{\underline{\theta}}^{\bar{\theta}}\frac{d}{d\theta}\left\{\frac{dV}{d\xi_k}\right\}\left(\frac{1}{1-T'}-(1-t_w)\right)\frac{\epsilon}{1+\epsilon}\theta\pi(\theta)\,d\theta-\frac{1}{1-t_w}\frac{\partial_{z^*}C_m}{C_m}\partial_{\xi_k}\chi_k(\xi_k,C_k).\end{aligned}$$

Then, using $\tau_h(\theta)=-\tau_l(\theta)$, $\partial_{z^*}E_l-\bar{s}_l=-(\partial_{z^*}E_h-\bar{s}_h)$ we can rewrite the formula of Lemma $A1$ as:

$$\begin{aligned}\frac{d}{d\theta}\left\{\frac{dV}{d\xi_k}\right\}&=-(1+\epsilon)\left(1-T'\right)\frac{z(\theta)}{\theta}\left((\tau_l(\theta)+\partial_{z^*}E_l-\bar{s}_l)\left(\frac{1}{p_l}\frac{dp_l}{d\xi_k}-\frac{1}{p_h}\frac{dp_h}{d\xi_k}\right)+\bar{s}_l\frac{1}{p_l}\frac{dp_l}{d\xi_k}+\bar{s}_h\frac{1}{p_h}\frac{dp_h}{d\xi_k}\right)\\\frac{dz}{d\xi_k}&=\frac{\epsilon}{1+\epsilon}\frac{\theta}{1-T'}\frac{d}{d\theta}\left\{\frac{dV(\theta)}{d\xi_k}\right\}=-\epsilon z(\theta)\left((\tau_l(\theta)+\partial_{z^*}E_l-\bar{s}_l)\left(\frac{1}{p_l}\frac{dp_l}{d\xi_k}-\frac{1}{p_h}\frac{dp_h}{d\xi_k}\right)+\bar{s}_l\frac{1}{p_l}\frac{dp_l}{d\xi_k}+\bar{s}_h\frac{1}{p_h}\frac{dp_h}{d\xi_k}\right),\end{aligned}$$

so we have:

$$\begin{aligned}\frac{1}{C_m}\frac{dC_m}{d\xi_k}&=\frac{1}{E_m}\int_{\underline{\theta}}^{\bar{\theta}}\frac{d}{d\theta}\left\{\frac{dV}{d\xi_k}\right\}\left\{\frac{1+\epsilon}{\epsilon}\frac{1}{\theta\pi}\int_{\theta}^{\bar{\theta}}(\partial_{z^*}e_m-\partial_{z^*}E_m)\,\pi d\theta'+(\partial_{z^*}e_m-\partial_{z^*}E_m)+\frac{1}{1-T'}\frac{1}{1-t_w}\partial_{z^*}E_m\right\}\frac{\epsilon}{1+\epsilon}\theta\pi(\theta)\,d\theta\\&+\mathcal{S}_{ml}\frac{1}{p_l}\frac{dp_l}{d\xi_k}+\mathcal{S}_{mh}\frac{1}{p_h}\frac{dp_h}{d\xi_k}-\frac{1}{1-t_w}\frac{\partial_{z^*}E_m}{E_m}\partial_{\xi_k}\chi_k(\xi_k,C_k),\\&=\frac{1}{1-t_w}\frac{1}{E_m}\int_{\underline{\theta}}^{\bar{\theta}}\frac{d}{d\theta}\left\{\frac{dV}{d\xi_k}\right\}\{\tau_m+(\partial_{z^*}E_m-\bar{s}_m)\}\frac{1}{1-T'}\frac{\epsilon}{1+\epsilon}\theta\pi(\theta)\,d\theta\\&+\frac{\bar{s}_m}{E_m}\frac{1}{1-t_w}\int_{\underline{\theta}}^{\bar{\theta}}\frac{d}{d\theta}\left\{\frac{dV}{d\xi_k}\right\}\frac{1}{1-T'}\frac{\epsilon}{1+\epsilon}\theta\pi(\theta)\,d\theta\\&+\mathcal{S}_{ml}\frac{1}{p_l}\frac{dp_l}{d\xi_k}+\mathcal{S}_{mh}\frac{1}{p_h}\frac{dp_h}{d\xi_k}-\frac{1}{1-t_w}\frac{\partial_{z^*}E_m}{E_m}\partial_{\xi_k}\chi_k(\xi_k,C_k).\end{aligned}$$

We therefore obtain, using $\tau_l + \frac{1}{1-T'}\left(\partial_{z^*}E_l - \bar{s}_l\right) = -\tau_h - \frac{1}{1-T'}\left(\partial_{z^*}E_h - \bar{s}_h\right)$:

$$\begin{aligned}
\frac{1}{C_l}\frac{dC_l}{d\xi_k} - \frac{1}{C_h}\frac{dC_h}{d\xi_k} =& \frac{1}{1-t_w}\frac{E_l+E_h}{E_lE_h}\int_{\underline{\theta}}^{\bar{\theta}}\frac{d}{d\theta}\left\{\frac{dV}{d\xi_k}\right\}\left\{\tau_l + \left(\partial_{z^*}E_l - \bar{s}_l\right)\right\}\frac{1}{1-T'}\frac{\epsilon}{1+\epsilon}\theta\pi\left(\theta\right)d\theta \\
&+ \left(\mathcal{S}_{ll} - \mathcal{S}_{hl}\right)\frac{1}{p_l}\frac{dp_l}{d\xi_k} + \left(\mathcal{S}_{lh} - \mathcal{S}_{hh}\right)\frac{1}{p_h}\frac{dp_h}{d\xi_k} - \frac{1}{1-t_w}\left(\frac{\partial_{z^*}E_l}{E_l} - \frac{\partial_{z^*}E_h}{E_h}\right)\partial_{\xi_k}\chi_k\left(\xi_k, C_k\right), \\
=& -\frac{1}{1-t_w}\frac{1}{\bar{s}_h\bar{s}_l}\int_{\underline{\theta}}^{\bar{\theta}}\left(\tau_l + \left(\partial_{z^*}E_l - \bar{s}_l\right)\right)^2\epsilon\frac{z(\theta)}{Z}\pi d\theta\left(\frac{1}{p_l}\frac{dp_l}{d\xi_k} - \frac{1}{p_h}\frac{dp_h}{d\xi_k}\right) \\
&- \frac{1}{1-t_w}\frac{1}{\bar{s}_h\bar{s}_l}\int_{\underline{\theta}}^{\bar{\theta}}\left(\tau_l + \left(\partial_{z^*}E_l - \bar{s}_l\right)\right)\epsilon\frac{z(\theta)}{Z}\pi d\theta\left(\bar{s}_l\frac{1}{p_l}\frac{dp_l}{d\xi_k} + \bar{s}_h\frac{1}{p_h}\frac{dp_h}{d\xi_k}\right) \\
&- \sigma\left(\frac{1}{p_l}\frac{dp_l}{d\xi_k} - \frac{1}{p_h}\frac{dp_h}{d\xi_k}\right) - \frac{1}{1-t_w}\frac{1}{\bar{s}_h\bar{s}_l}\left(\partial_{z^*}E_l - \bar{s}_l\right)\frac{1}{Z}\partial_{\xi_k}\chi_k\left(\xi_k, C_k\right), \\
\bar{s}_l\frac{1}{C_l}\frac{dC_l}{d\xi_k} + \bar{s}_h\frac{1}{C_h}\frac{dC_h}{d\xi_k} =& \frac{1}{Z}\frac{1}{1-t_w}\int_{\underline{\theta}}^{\bar{\theta}}\frac{d}{d\theta}\left\{\frac{dV}{d\xi_k}\right\}\frac{1}{1-T'}\frac{\epsilon}{1+\epsilon}\theta\pi\left(\theta\right)d\theta - \frac{1}{1-t_w}\frac{1}{Z}\partial_{\xi_k}\chi_k\left(\xi_k, C_k\right) \\
=& -\frac{1}{1-t_w}\int_{\underline{\theta}}^{\bar{\theta}}\left(\tau_l\left(\theta\right) + \partial_{z^*}E_l - \bar{s}_l\right)\epsilon\frac{z(\theta)}{Z}\pi d\theta\left(\frac{1}{p_l}\frac{dp_l}{d\xi_k} - \frac{1}{p_h}\frac{dp_h}{d\xi_k}\right) \\
&- \frac{\epsilon}{1-t_w}\left(\bar{s}_l\frac{1}{p_l}\frac{dp_l}{d\xi_k} + \bar{s}_h\frac{1}{p_h}\frac{dp_h}{d\xi_k}\right) - \frac{1}{1-t_w}\frac{1}{Z}\partial_{\xi_k}\chi_k\left(\xi_k, C_k\right),
\end{aligned}$$

with $Z = E_l + E_h = \int_{\underline{\theta}}^{\bar{\theta}} z(\theta)\pi d\theta$. This proves the formulas of the Lemma using $\epsilon = \zeta$. The sign of the response is a direct consequence of Corollary 1 which shows $\tau_l + \left(\partial_{z^*}E_l - \bar{s}_l\right) \leq 0$. $\square$
Coming back to the proof of Proposition 4 and using the formulas of Lemma $A3$, we have:

$$\begin{aligned}
\frac{1}{p_l}\frac{dp_l}{d\xi_k} - \frac{1}{p_h}\frac{dp_h}{d\xi_k} =& \left(1-\alpha\left(\sigma+\Omega\right)\right)^{-1}\left(\frac{1}{p_l}\frac{\partial p_l}{\partial\xi_k} - \frac{1}{p_h}\frac{\partial p_h}{\partial\xi_k} + \alpha\frac{1}{1-t_w}\frac{1}{\bar{s}_h\bar{s}_l}\left(\partial_{z^*}E_l - \bar{s}_l\right)\frac{1}{Z}\partial_{\xi_k}\chi_k\left(\xi_k, C_k\right)\right) \\
&+ \left(1-\alpha\left(\sigma+\Omega\right)\right)^{-1}\frac{\frac{\alpha\zeta}{1-t_w}}{1-\frac{\alpha\zeta}{1-t_w}}\frac{1}{\bar{s}_h\bar{s}_l}\int_{\underline{\theta}}^{\bar{\theta}}\left(\tau_l + \left(\partial_{z^*}E_l - \bar{s}_l\right)\right)\frac{z(\theta)}{Z}\pi d\theta\left(\bar{s}_l\frac{1}{p_l}\frac{\partial p_l}{\partial\xi_k} + \bar{s}_h\frac{1}{p_h}\frac{\partial p_h}{\partial\xi_k} + \alpha\frac{1}{1-t_w}\frac{1}{Z}\partial_{\xi_k}\chi_k\left(\xi_k, C_k\right)\right),
\end{aligned}$$

with $\Omega = \frac{\zeta}{1-t_w}\frac{1}{\bar{s}_h\bar{s}_l}\left(\int_{\underline{\theta}}^{\bar{\theta}}\left(\tau_l + \left(\partial_{z^*}E_l - \bar{s}_l\right)\right)^2\frac{z(\theta)}{Z}\pi d\theta + \frac{\alpha\zeta}{1-t_w-\alpha\zeta}\left(\alpha\frac{\zeta}{1-t_w}\int_{\underline{\theta}}^{\bar{\theta}}\left(\tau_l + \left(\partial_{z^*}E_l - \bar{s}_l\right)\right)\frac{z(\theta)}{Z}\pi d\theta\right)^2\right) \geq 0$, with strict inequality if preferences are non-homothetic.
For an increase in the relative price of necessities, we have $\frac{\partial lnp_l}{\partial ln\bar{p}_l} = \bar{s}_h$, $\frac{\partial lnp_h}{\partial ln\bar{p}_l} = -\bar{s}_l$ $\frac{\partial lnp_l}{\partial ln\bar{p}_l} = \bar{s}_h$ and $\partial_{ln\bar{p}_l}\chi_l = \left(1-\alpha+t_w\right)^{-1}\bar{s}_hE_l$, $\partial_{ln\bar{p}_l}\chi_h = \left(1-\alpha+t_w\right)^{-1}\bar{s}_lE_h$, so we have:

$$\begin{aligned}
\frac{1}{p_l}\frac{dp_l}{dln\bar{p}_l} - \frac{1}{p_h}\frac{dp_h}{dln\bar{p}_l} &= \left(1-\alpha\left(\sigma+\Omega\right)\right)^{-1} \\
\frac{d}{dln\bar{p}_l}\left\{\frac{T'}{1-T'}\right\} &= \left(1-\alpha\left(\sigma+\Omega\right)\right)^{-1}\frac{\partial}{\partial\ln\bar{p}_l}\left\{\frac{T'}{1-T'}\right\} < \frac{\partial}{\partial\ln\bar{p}_l}\left\{\frac{T'}{1-T'}\right\} < 0 \quad if\ \alpha > 0,
\end{aligned}$$

which shows the first formula of Proposition 4 and that partial equilibrium effects are amplified in general equilibrium.
Finally for a homogenous price increase $\frac{\partial lnp_l}{\partial ln\bar{p}} = \frac{\partial lnp_h}{\partial ln\bar{p}} = 1$, $\partial_{ln\bar{p}_l}\chi_l = \left(1-\alpha+t_w\right)^{-1}E_l$, $\partial_{ln\bar{p}_l}\chi_h =$

$(1-\alpha+t_w)^{-1} E_h$, so we have:

$$\frac{1}{p_l}\frac{dp_l}{dln\bar{p}}-\frac{1}{p_h}\frac{dp_h}{dln\bar{p}}=(1-\alpha(\sigma+\Omega))^{-1}\frac{\alpha}{1-\alpha}\frac{1}{\bar{s}_h\bar{s}_l}\left((\partial_{z^*}E_l-\bar{s}_l)+\frac{\zeta}{1-t_w-\alpha\zeta}\frac{1}{\bar{s}_h\bar{s}_l}\int_{\underline{\theta}}^{\bar{\theta}}(\tau_l+(\partial_{z^*}E_l-\bar{s}_l))\frac{z(\theta)}{Z}\pi d\theta\right)$$

$$\frac{d}{dln\bar{p}}\left\{\frac{T'}{1-T'}\right\}=(1-\alpha(\sigma+\Omega))^{-1}\frac{\alpha}{1-\alpha}\frac{1}{\bar{s}_h\bar{s}_l}\left((\partial_{z^*}E_l-\bar{s}_l)+\frac{\zeta}{1-t_w-\alpha\zeta}\frac{1}{\bar{s}_h\bar{s}_l}\int_{\underline{\theta}}^{\bar{\theta}}(\tau_l+(\partial_{z^*}E_l-\bar{s}_l))\frac{z(\theta)}{Z}\pi d\theta\right)\frac{\partial}{\partial\ln\bar{p}_l}\left\{\frac{T'}{1-T'}\right\}.$$

Since l is a necessity, we have $(\partial_{z^*}E_l-\bar{s}_l)+\frac{\zeta}{1-t_w-\alpha\zeta}\frac{1}{\bar{s}_h\bar{s}_l}\int_{\underline{\theta}}^{\bar{\theta}}(\tau_l+(\partial_{z^*}E_l-\bar{s}_l))\frac{z(\theta)}{Z}\pi d\theta<0$, $\partial_{\ln\bar{p}_l}T'<0$ so $\frac{d}{dln\bar{p}}\left\{\frac{T'}{1-T'}\right\}$ has the same sign as α.

A.2.4 Formulas for Section 5.2

In this section, we provide the formulas that underpin the quantitative results of Section 5.2. The formulas allow us to compute the change in taxes in response to an arbitrary cost shock ξ_k in an n-sector economy. While the formulas are expressed in terms of $\frac{dV}{d\xi_k}$, they allow us to recover tax rates using:

$$\frac{d}{d\theta}\left\{\frac{dV}{d\xi_k}\right\}=-(1+\epsilon)\frac{z}{\theta}(1-T')\left(\frac{1}{1-T'}\frac{dT'}{d\xi_k}+\sum_{m=1}^{n}\partial_{z^*}e_m\frac{1}{p_m}\frac{dp_m}{d\xi_k}\right),$$

and we can rewrite the formula of Proposition A1 below as:

$$z\tilde{\zeta}f(z(\theta))\frac{d}{d\xi_k}\left\{\frac{T'}{1-T'}\right\}=\int_{z(\theta)}^{z(\bar{\theta})}g\left(\gamma v_{z^*}\left(\frac{dT}{d\xi_k}+\sum_{m=1}^{N}e_m\frac{1}{p_m}\frac{dp_m}{d\xi_k}\right)-\int_{z(\underline{\theta})}^{z(\bar{\theta})}g\gamma v_{z^*}\left(\frac{dT}{d\xi_k}+\sum_{m=1}^{n}e_m\frac{1}{p_m}\frac{dp_m}{d\xi_k}\right)fdz\right)fdz$$
$$+z\tilde{\zeta}f(z(\theta))\frac{d}{d\xi_k}\left\{\frac{T'_{lin}}{1-T'_{lin}}\right\}$$
$$z\tilde{\zeta}f(z(\theta))\frac{d}{d\xi_k}\left\{\frac{T'_{lin}}{1-T'_{lin}}\right\}=\sum_{m=1}^{N}\left(\mathbb{E}_{z>z(\theta)}(\partial_{z^*}e_m-\partial_{z^*}E_m)-z\tilde{\zeta}f(z(\theta))\frac{T'}{1-T'}(\partial_{z^*}e_m-\partial_{z^*}E_m)\right)\frac{1}{p_m}\frac{dp_m}{d\xi_k}.$$

The main advantage of the result below is that it provides a simple procedure to compute the change in the tax schedule in response to supply curve shifters. As a first step, we can compute $N+1$ "partial equilibrium" responses which do not depend on the endogenous adjustment of prices $\frac{dp_m}{d\xi_k}$. Once they are computed, the general equilibrium response of prices and the full response of the tax schedule is simply computed by linear algebra.

Proposition A1. Under assumption $A3$, the change in welfare for agent θ, $dV(\theta)/d\xi_k$ in response to an exogenous supply shift $d\xi_k$ is given by:

$$\frac{dV}{d\xi_k}=\sum_{m=1}^{n}\frac{\partial V}{\partial p_m}\frac{1}{p_m}\frac{dp_m}{d\xi_k}-\frac{\partial V}{\partial B}\partial_{\xi_k}\chi_k(\xi_k,C_k),$$

where $\frac{\partial V}{\partial p_m}$and $\frac{\partial V}{\partial B}$ solve:

$$\frac{\epsilon}{(1+\epsilon)^2}\frac{\theta\pi(\theta)}{(1-T')^2}\frac{\theta}{z(\theta)}\frac{d}{d\theta}\left\{\frac{\partial V}{\partial p_m}\right\}+(1-t_w)\int_\theta^{\bar{\theta}}g\left(\gamma\left(\theta'\right)\frac{\partial V}{\partial p_m}-\int_{\underline{\theta}}^{\bar{\theta}}g\gamma\left(\theta'\right)\frac{\partial V}{\partial p_m}\pi d\theta'\right)\pi d\theta'$$

$$=-\frac{\epsilon}{1+\epsilon}\frac{\theta\pi(\theta)}{1-T'}\left(\tau_m\left(\theta\right)+\partial_{z^*}E_m\right),\ (1-t_w)\int_{\underline{\theta}}^{\bar{\theta}}g\frac{\partial V}{\partial p_m}\pi d\theta=0.$$

$$\frac{\epsilon}{(1+\epsilon)^2}\frac{\theta\pi(\theta)}{(1-T')^2}\frac{\theta}{z(\theta)}\frac{d}{d\theta}\left\{\frac{\partial V}{\partial p_m}\right\}+(1-t_w)\int_\theta^{\bar{\theta}}g\left(\gamma\left(\theta'\right)\frac{\partial V}{\partial p_m}-\int_{\underline{\theta}}^{\bar{\theta}}g\gamma\left(\theta'\right)\frac{\partial V}{\partial p_m}\pi d\theta'\right)\pi d\theta'=0,\ (1-t_w)\int_{\underline{\theta}}^{\bar{\theta}}g\frac{\partial V}{\partial p_m}\pi d\theta=1.$$

To express the equilibrium change in prices, define the vector $\mathbf{C^B}$and the matrix $\mathcal{C}$ with elements:

$$C_l^B=\frac{1}{1-t_w}\frac{\partial_{z^*}E_l}{E_l}+\frac{1}{E_l}\int_{\underline{\theta}}^{\bar{\theta}}\left\{\tau_l\left(\theta\right)+\partial_{z^*}E_l\right\}\frac{1}{1-T'}\frac{1}{1-t_w}\frac{d}{d\theta}\left\{\frac{\partial V}{\partial B}\right\}\frac{\epsilon}{1+\epsilon}\theta\pi\left(\theta\right)d\theta$$

$$\mathcal{C}_{l,m}=\mathcal{S}_{l,m}+\frac{1}{E_l}\int_{\underline{\theta}}^{\bar{\theta}}\left\{\tau_l\left(\theta\right)+\partial_{z^*}E_l\right\}\frac{1}{1-T'}\frac{1}{1-t_w}\frac{d}{d\theta}\left\{\frac{\partial V}{\partial p_m}\right\}\frac{\epsilon}{1+\epsilon}\theta\pi\left(\theta\right)d\theta.$$

Then the equilibrium change in prices $\frac{\mathbf{dlnp}}{\mathbf{d\xi_k}}=\left\{\frac{1}{p_1}\frac{dp_1}{d\xi_k},...,\frac{1}{p_N}\frac{dp_N}{d\xi_k}\right\}'$ solves:

$$\frac{\mathbf{dlnp}}{\mathbf{d\xi_k}}=\left(Id+\alpha\mathcal{C}\right)^{-1}\left(\delta_{\mathbf{k}}\partial_{\xi_k}ln\phi_k\left(\xi_k,C_k\right)+\alpha\mathbf{C^B}\partial_{\xi_k}\chi_k\left(\xi_k,C_k\right)\right),$$

where $\delta_{\mathbf{k}}$ is the column vector with 1 on its k^{th} row and 0 otherwise.

Proof: Recall from Lemma $A1$ that $\frac{dV}{d\xi_k}$ solves equation:

$$\frac{\epsilon}{(1+\epsilon)^2}\frac{\theta\pi(\theta)}{(1-T')^2}\frac{\theta}{z(\theta)}\frac{d}{d\theta}\left\{\frac{dV}{d\xi_k}\right\}+(1-t_w)\int_\theta^{\bar{\theta}}g\left(\gamma\left(\theta'\right)\frac{dV}{d\xi_k}-\int_{\underline{\theta}}^{\bar{\theta}}g\gamma\left(\theta'\right)\frac{dV}{d\xi_k}\pi d\theta'\right)\pi d\theta'=-\frac{\epsilon}{1+\epsilon}\frac{\theta\pi(\theta)}{1-T'}\sum_{l=1}^{n}\left(\tau_l\left(\theta\right)+\partial_{z^*}E_l\right)\frac{1}{p_l}\frac{dp_l}{d\xi_k},$$

$$(1-t_w)\int_{\underline{\theta}}^{\bar{\theta}}g\frac{dV}{d\xi_k}\pi d\theta=-\partial_{\xi_k}\chi_k\left(\xi_k,C_k\right).$$

Since the equation is linear in $\frac{dV}{d\xi_k}$, it is direct that$\frac{dV}{d\xi_k}=\sum_{m=1}^{N}\frac{\partial V}{\partial p_m}\frac{1}{p_m}\frac{dp_m}{d\xi_k}-\frac{\partial V}{\partial B}\partial_{\xi_k}\chi_k\left(\xi_k,C_k\right)$ with $\frac{\partial V}{\partial p_m}$, $\frac{\partial V}{\partial B}$ defined in the Proposition. Next, we have, using the pricing equation:

$$\frac{dlnp_m}{d\xi_k}=\frac{\partial ln\phi_m}{\partial\xi_k}-\alpha\frac{1}{C_m}\frac{dC_m}{d\xi_k},$$

with $\frac{\partial ln\phi_m}{\partial\xi_k}=0$ if $m\neq k$. Using the same steps in Lemma $A3$, we have:

$$\frac{1}{C_m}\frac{dC_m}{d\xi_k}=\frac{1}{1-t_w}\frac{1}{E_m}\int_{\underline{\theta}}^{\bar{\theta}}\left\{\tau_m+\partial_{z^*}E_m\right\}\frac{1}{1-T'}\frac{d}{d\theta}\left\{\frac{dV}{d\xi_k}\right\}\frac{\epsilon}{1+\epsilon}\theta\pi\left(\theta\right)d\theta$$
$$+\sum_{l=1}^{N}\mathcal{S}_{ml}\frac{1}{p_l}\frac{dp_l}{d\xi_k}-\frac{1}{1-t_w}\frac{\partial_{z^*}E_m}{E_m}\partial_{\xi_k}\chi_k\left(\xi_k,C_k\right).$$

Using $\frac{dV}{d\xi_k}=\sum_{m=1}^{N}\frac{\partial V}{\partial p_m}\frac{1}{p_m}\frac{dp_m}{d\xi_k}-\frac{\partial V}{\partial B}\partial_{\xi_k}\chi_k\left(\xi_k,C_k\right)$, we obtain:

$$\frac{1}{C_m}\frac{dC_m}{d\xi_k}=\sum_{l=1}^{n}\mathcal{C}_{ml}\frac{1}{p_l}\frac{dp_l}{d\xi_k}-C_m^B\partial_{\xi_k}\chi_k\left(\xi_k,C_k\right).$$

So we have:

$$\frac{\mathbf{dlnp}}{\mathbf{d\xi_k}} = (Id + \alpha\mathcal{C})^{-1}\left(\delta_{\mathbf{k}}\partial_{\xi_k} ln\phi_k\left(\xi_k, C_k\right) + \alpha\mathbf{C^B}\partial_{\xi_k}\chi_k\left(\xi_k, C_k\right)\right),$$

which proves the result. $\square$

B Additional Quantitative Results

B.1 The Impacts of Exogenous Price Shocks

We now characterize the response of the optimal tax schedule to exogenous price shocks. As in Section 5.2, this analysis is motivated by the observed heterogeneous price changes across product categories in the United States, with lower inflation in product categories purchased by high-income households. For example, Jaravel (2019) documents that, for consumer packaged goods in the United States, annual inflation is about 2.5pp lower in the top price decile (a proxy for quality), compared with the bottom price decile.

The simulations account for feedback loops created by large price changes, and thus complement the first-order approximations in Section 5.2. To assess the quantitative relevance for the optimal schedule, we consider an exogenous productivity shock (to parameters γ_i), whose direct partial equilibrium effect is to reduce the price of the high-quality good by 2.5% and to increase the price of the low-quality good by 2.5%. Given our calibration for non-homotheticities, the induced change in the price index is 3.1 pp higher in the bottom decile of the income distribution, compared with the top decile. This corresponds to the level of inflation inequality reached over 8.5 years in U.S. data (Jaravel (2019)).

Baseline simulation. Figure A7 reports the results under the baseline parametrization. The exogenous price shock leads to lower taxes: marginal tax rates fall by about 3.25pp at the bottom of the income distribution, and gradually converge back to the reference tax schedule under homothetic utility, with a fall in marginal tax rates under 0.10pp for levels of income above $300,000 (panels A and B).

To understand the mechanism, it is instructive to examine how the exogenous shocks affect equilibrium prices. Without shocks, prices are identical to the baseline non-homothetic specification studied in Figure 5 (panel C). In partial equilibrium, the shocks would reduce the price of the high-quality good from 1.14 to 1.11 and would increase the price of the low-quality good from 0.9 to 0.92. Panel C of Figure A7 shows the amplification of the price shocks through consumer demand, additional redistribution and returns to scale: the equilibrium prices are 1.055 for the high-quality good and 0.96 for the low-quality goods. In general equilibrium, the convergence of relative prices is much larger than with the partial equilibrium shocks alone. Consequently, there is a substantial increase in the value of transferring an additional dollar to high-income households, who have a higher marginal propensity to consume on the high-quality good (panels D and E).

Thus, it is desirable for the planner to redistribute toward high-income households, which can be done efficiently by reducing marginal tax rates at the bottom of the income distribution. The welfare effects are substantial, with an equivalent variation of -6% in the bottom decile and +8.5% in the top decile, reported in Panel F.

Sensitivity to increasing returns. Figure A8 shows the results with higher returns to scale ($\alpha = 0.4$), which magnifies the impact of the exogenous productivity change. The fall in marginal tax rates at the bottom of the income distribution is about 12pp (panel B). The GE amplification of price changes is much larger and flips the relative price of the high- and low-quality bundles (panel C). The distributional effects are large, with an equivalent variation ranging from -26% at the bottom to +24% at the top, shown in Panel F.

Conversely, Figure shows the results with lower returns to scale ($\alpha = 0.2$), which reduce the impact of the exogenous productivity change. The fall in marginal tax rates at the bottom of the income distribution is about 1.70pp, the GE amplification of price changes is smaller, and the distributional effects are more modest, with an equivalent variation ranging from -2.4% at the bottom to +5% at the top.

Sensitivity to preferences for redistribution. Figure A10 documents the role of social preferences for redistribution, setting the Pareto weights to match the optimal schedule with constant returns to scale and a social welfare function with a CRRA coefficient of 0.5. The impact of exogenous price shocks is much larger than in the baseline specification, with a fall in marginal tax rates of 13pp at the bottom of the income distribution (panel B), rather than about 3pp.

To understand the mechanism, Panel C reports equilibrium prices. Before the exogenous shock, equilibrium prices are 0.965 for the low-quality product and 1.0375 for the high-quality product (identical to Figure 7). After the shock, the price of the low-quality product increases substantially, reaching 1.07, while the price of the high-quality good falls to 0.955 only. The amplification of price effects is sufficiently large to flip the relative price of the high- and low-quality bundles.

When social preferences for redistribution are low, the planner puts larger weight on the change in utility out of disposable income for high-skill agents. Therefore, the planner is more responsive to the initial fall in the relative price of the high-quality good and redistributes more toward the rich, which induces a feedback loop of changes in labor supply, spending, and prices, leading to further changes in redistribution, etc. Quantitatively, this mechanism is strong enough to flip equilibrium relative prices and increase high-skill agents' utility out of disposable income above 1 (panel D). As depicted on Panel F, the shock results in a large welfare loss at the bottom of the income distribution (-32%) and substantial welfare gains at the top (+10%).

The comparison of these results with those from Figure 7 are instructive to understand the mechanism driving the interplay between endogenous prices, increasing returns, and social preferences for redistribution. In Figure 7, weaker social preferences induced an optimal tax schedule with less redistribution toward the poor, implying smaller changes in relative market size, and hence smaller endogenous price changes. In that setting, absent exogenous shocks, weaker social preferences for redistribution reduce the importance of non-homotheticities for the optimal tax schedule, because prices change less. Introducing exogenous price shocks, Figure A10 shows that the response to price shocks is magnified with weaker social preferences, which induce more redistribution toward those with a higher propensity to spend on the cheaper products, which amplifies the exogenous shocks and leads to larger changes in equilibrium prices.

Overall, the results show that exogenous price shocks can have a large impact on the optimal tax schedule, and that there are important amplification effects through increasing returns and the endogenous social value of redistribution. In all simulations, a unifying mechanism operates: changes in equilibrium prices and the distribution of marginal propensities to consume govern the change in the optimal tax schedule.

B.2 The Response of the Tax Schedule to Exogenous Shifts in the Skill Distribution

In this section, we characterize quantitatively the optimal response of the tax schedule to exogenous shifts in the income distribution, accounting for the endogenous response of prices. Using the publicly available statistics on the income distribution from the U.S. Census, we recover the shifts in the skill distribution from the observed shifts in the income distribution from 2004 to 2015.[2] Empirically, income is stagnant at the bottom of the distribution, and increases at faster and faster rates with higher incomes.

Partial and general equilibrium results. Using the theoretical results in Appendix C.2 of our working paper (Jaravel and Olivi, 2024), Figure A11 reports the optimal response of marginal tax rates. We first consider the direct, partial equilibrium response to the change in the skill distribution, with fixed prices. As characterized in our working paper, as the income distribution becomes more spread out, the value of redistribution at higher incomes falls, which pushes for a more redistributive tax schedule, with higher marginal tax rates. Because of the shifts in the skill distribution, there is relatively more mass at the top and bottom of the skill distribution, hence the distortionary cost of taxation is higher in this range, while it is reduced in the middle of the distribution. To increase redistribution efficiently, it is therefore optimal to raise marginal tax rates especially in the middle of the income distribution. Thus, Figure A11 shows that optimal marginal tax rates increase by about 2.5pp at the bottom of the distribution, by about 5pp in the middle, and by 1pp at the very top.

Furthermore, general equilibrium effects are at play through prices, as characterized in Proposition 4. The direct effects on prices of the shifts in inequality is amplified through income and substitution effects, as well as changes in optimal tax rates. These effects tend to reduce optimal tax rates, because the observed shift in the income distribution lowers the price of products with a high income elasticity. Because higher-income agents have a higher marginal propensity to spend on these goods, it is optimal to redistribute more toward them by lowering marginal tax rates, through the same channels as in Section 4.2. Quantitatively, with $\sigma = 0.6$ optimal tax rates are reduced by a few percentage points, relative to the optimum in partial equilibrium, throughout the distribution. The fall in marginal tax rates is larger with $\sigma = 2$, reaching about -4pp at the bottom of the distribution. In this case, marginal tax rates fall below the observed tax schedule at the bottom of the income distribution.

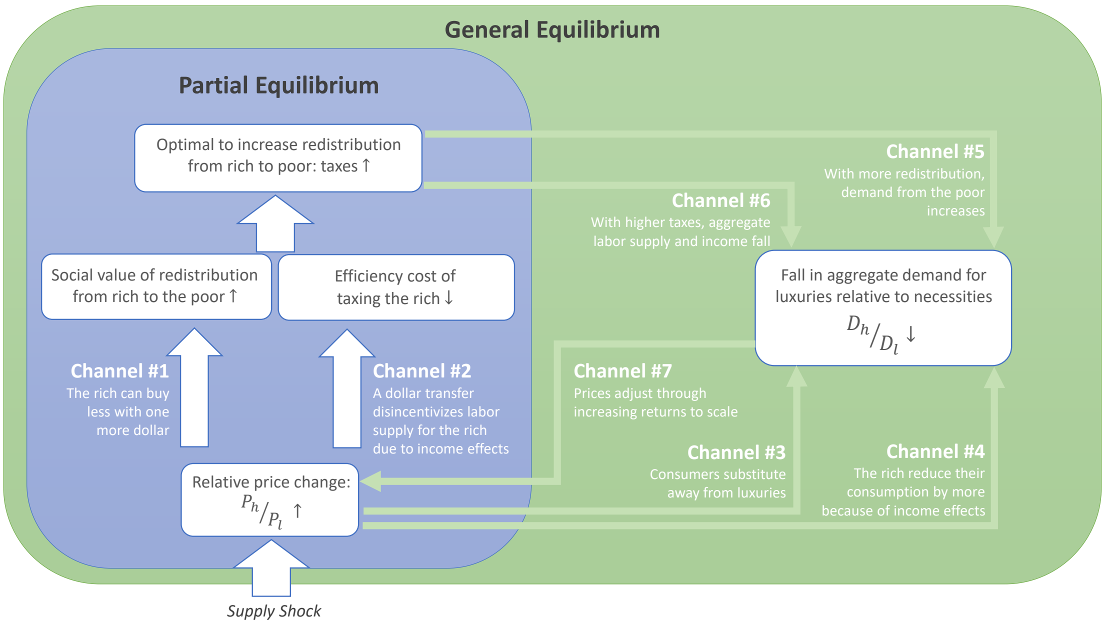
Finally, Figure A11 also shows the combined impact of the shift in the skill distribution and exogenous changes in prices. Price changes are measured from 2004 to 2015 as in Section 5.2, except that the estimation accounts at the same time for the shift in the skill distribution and the induced price changes, i.e. we estimate the "residual" price shocks to match observed price changes. Taking into account these residual price shocks leads to a substantial reduction in optimal tax rates at the bottom of the distribution. Indeed, as in Section 5.2, price shocks increase the value of redistribution at the top. Quantitatively, the direct price effects, which imply more redistribution toward higher-skill agents, more than offset the motive for increased redistribution toward low-skill agents from the shift in the skill distribution. Taking into account all effects, the optimal tax schedule becomes less redistributive. These results show that it is important to jointly study shifts in the skill distribution and price shocks.

[2] We use the historical series available at https://www.census.gov/data/tables/time-series/demo/income-poverty/historical-income-households.html (Table H-2).

The role of non-linear social preferences. Figure A12 shows the role of the curvature of the social welfare function. Both with exogenous prices (panel A) and endogenous prices (panels B and C), the response of the tax schedule is muted by additional curvature. Indeed, curvature tends to mute the motives for redistribution created either directly by the shift in the skill distribution or by the endogenous price response. We find that the changes in the optimal tax schedule remain substantial even with non-linear social welfare functions.

C Additional Figures and Tables

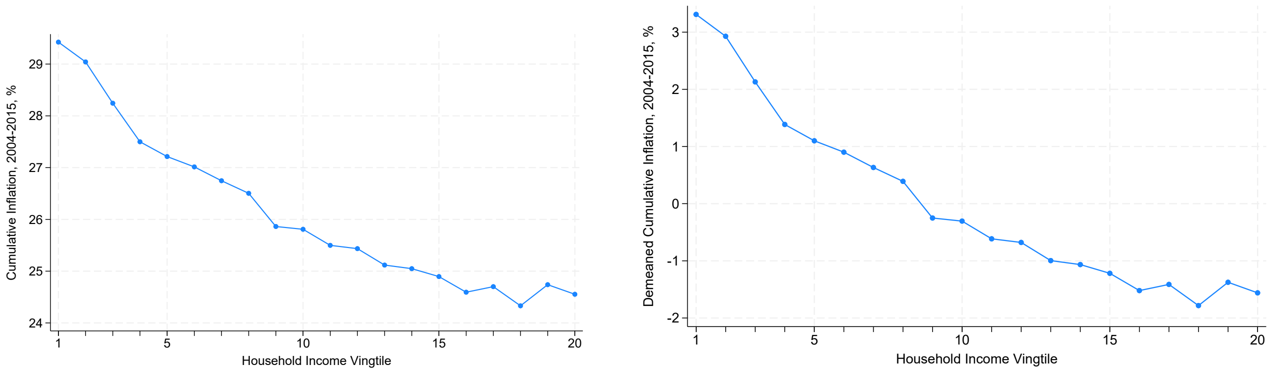
Figure A1 The Response of the Optimal Tax Schedule to a Change in the Price of Luxuries relative to Necessities, Key Channels



Notes: This table summarizes the key theoretical channels in our model whereby price shocks affect optimal taxation. We consider an increase in the price of luxuries, denoted h, relative to necessities, denoted n. The figure reports the two partial equilibrium channels characterized in the main test (Channels #1 and #2). It also reports five channels that operate in general equilibrium. First, households reallocate their spending to other products through standard substitution effects, leading to a fall in demand for luxuries (Channel #3). Second, a relative increase in the price of luxuries has a negative income effect on higher income households, as luxuries constitute a larger portion of their consumption basket. Higher income households have a higher propensity to spend on luxuries, so the aggregate share of luxuries decreases through income effects (Channel #4). Moreover, there are several changes in optimal taxes. Through the partial equilibrium channels #1 and #2, it becomes more valuable to redistribute to lower income households: tax rates increase along the income distribution. Income is reallocated to lower income households, which amplifies the decline in the share of luxuries (Channel #5). In addition, by increasing tax rates, the planner lowers labor supply in the aggregate. As households' aggregate real income decreases, they shift their consumption towards necessities, which further amplifies the fall in the share of luxuries (Channel #6). Therefore, in general equilibrium the markets for luxuries shrink relative to other markets, and the relative price of luxuries increases further through a supply side response (Channel #7). These endogenous price changes induce further rounds of changes in optimal taxes, labor supply and relative prices, as illustrated graphically with the feedback loop in the figure. The government thus amplifies both the inflation of luxuries prices and their redistributive impact through changes in optimal tax rates.

Figure A2 Inflation Inequality in the United States, 2004-2015, CEX-CPI Data

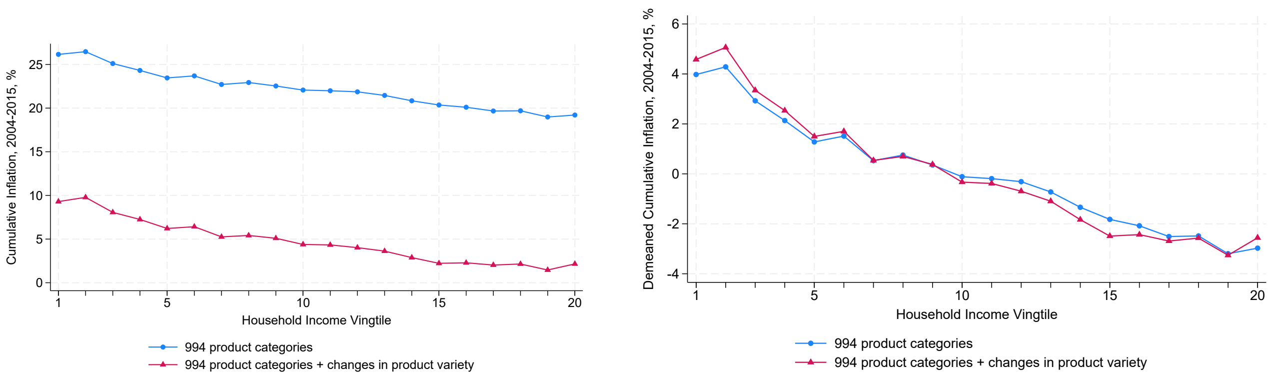
A. Raw Price Changes   B. Relative Price Changes



Notes: Panel A of this figure reports Laspeyres inflation rates across the income distribution between 2004 and 2015, considering 248 product categories observed in CEX-CPI data. Panel B reports the same patterns after demeaning price changes. We use the changes in relative prices reported in Panel B for the quantitative analysis in Section 5.2, since the relative price effect is central to our analysis while the uniform price change simply scales real wages and has limited interaction with consumption heterogeneity.

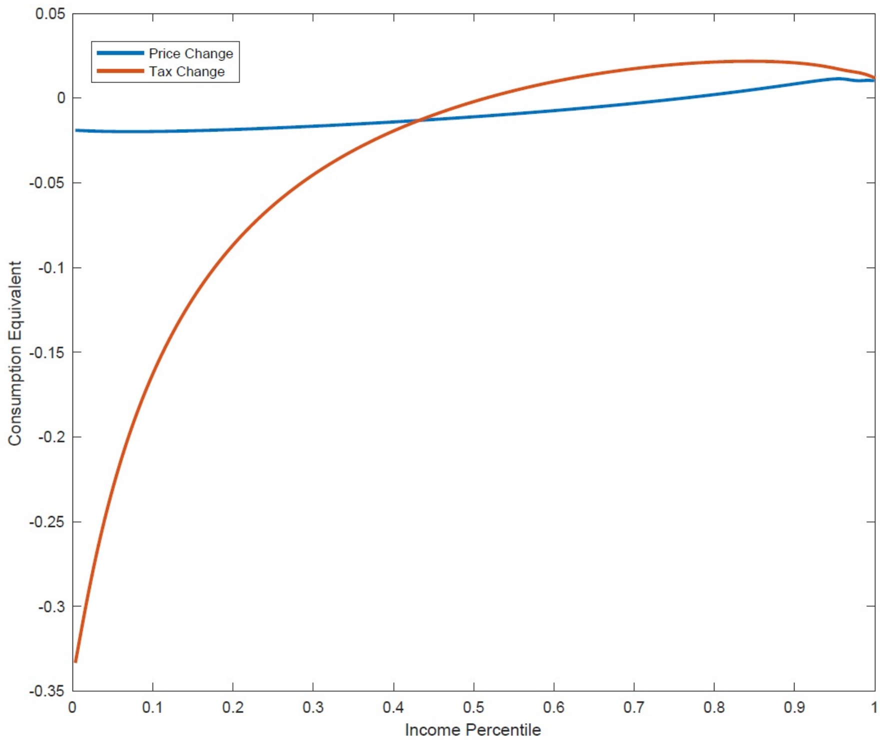
Figure A3 Inflation Inequality in the United States, 2004-2015, Nielsen Data

A. Raw Price Changes   B. Relative Price Changes



Notes: Panel A of this figure reports Laspeyres inflation rates across the income distribution between 2004 and 2015, using the Nielsen data. We use the 994 most detailed product categories, called "product modules" and report inflation rates with and without the correction for changes in product variety. Through consumers' love-of-variety, average inflation is significantly lower when allowing for changes in product variety. Panel B reports the same patterns after demeaning price changes. We use the changes in relative prices reported in Panel B for the quantitative analysis in Section 5.2, since the relative price effect is central to our analysis while the uniform price change simply scales real wages and has limited interaction with consumption heterogeneity.

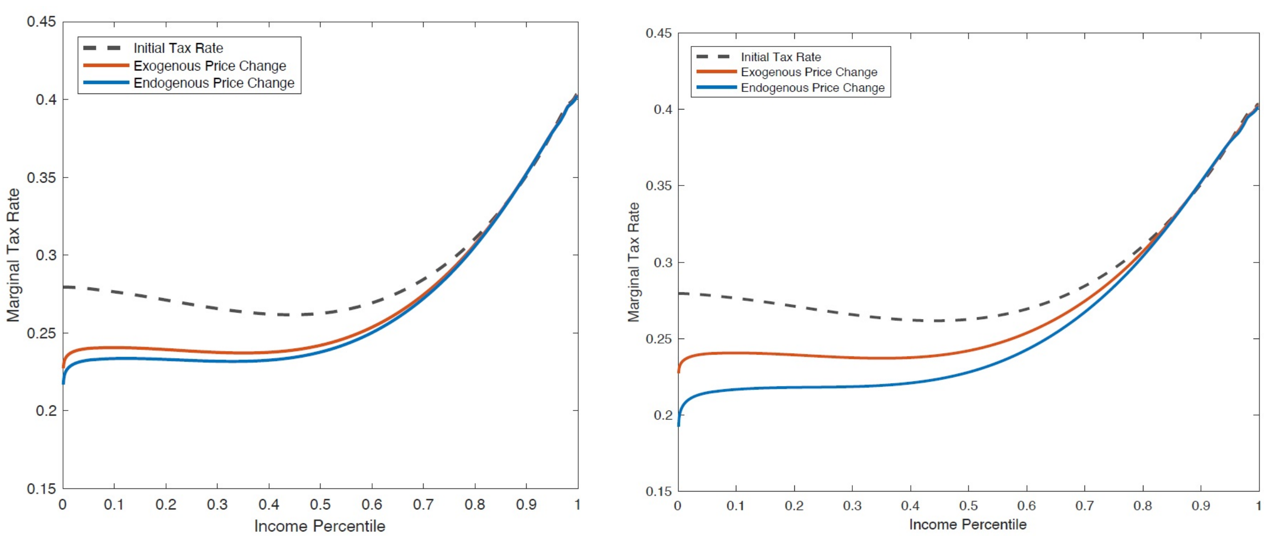
Figure A4 Willingness to Pay for Price Change vs. Optimal Tax Response



Notes: In this figure, the willingness to pay for the reform for a household with post tax income z^* is defined as $WTP\left(z^*\right) = -\sum_{k=1}^{n} s_k\left(z^*\right)\hat{p}_k - dT\left(z\right)/z^*$. $\hat{p}_k$ denotes the exogenous price change, while $dT\left(z\right)$ is the change in tax. Since a change in prices affects labor supply and therefore how much revenue is collected from the income tax, we decompose $dT\left(z\right) = dT_p\left(z\right) + dT_r\left(z\right)$. The planner rebates to the households the change in revenue arising from the change in prices, dT_p, in a lump sum fashion, and implements the optimal tax change $dT_r\left(z\right)$. We decompose the willingness to pay into $WTP_p\left(z^*\right) = -\sum_{k=1}^{n} s_k\left(z^*\right)\hat{p}_k - dT_p\left(z\right)/z^*$, the welfare impact of the reform due to prices only, and $WTP_r\left(z^*\right) = -dT_r\left(z\right)/z^*$, the welfare impact of the change in taxes.

Figure A5 Sensitivity Analysis for the Response to Observed Price Shocks (2004-2015), CEX-CPI data
Results with Labor Supply Elasticity $\varepsilon = 0.33$

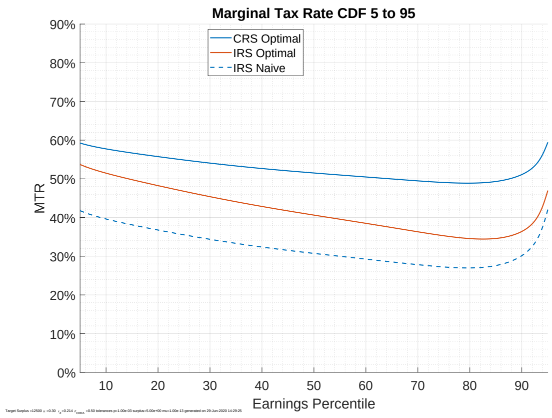
A. $\sigma = 0.6$ B. $\sigma = 2$



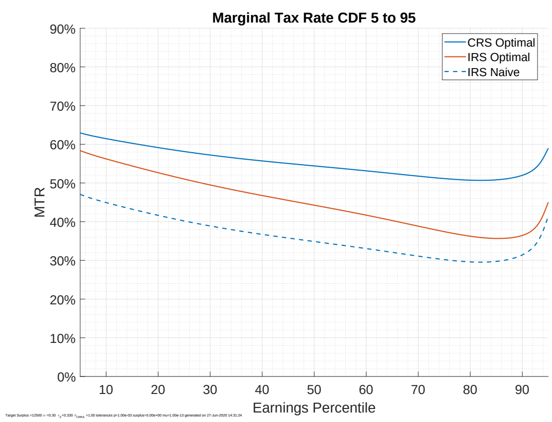
Notes: in all specifications, the IRS parameter is set to $\alpha = 0.3$ and the labor supply elasticity to $\varepsilon = 0.33$. The CEX-CPI dataset is used in both panels and the initial tax schedule is taken from Hendren (2020).

Figure A6 Returns to Scale and the Optimal Tax Schedule,
Sensitivity to Parameter Values

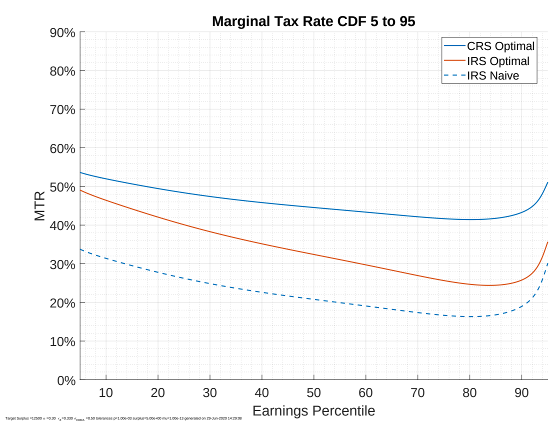
A. $\alpha = 0.3$, $\varepsilon = 0.21$, CRRA=0.5



B. $\alpha = 0.3$, $\varepsilon = 0.33$, CRRA=1



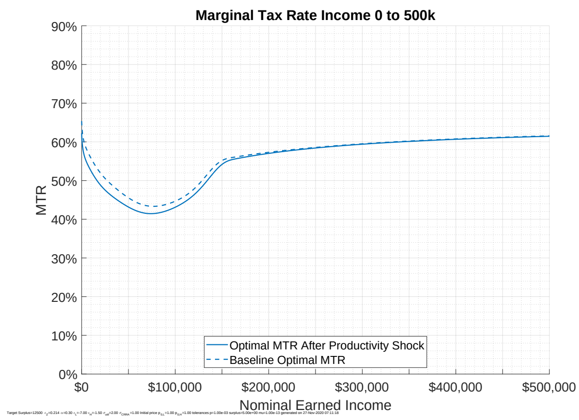
C. $\alpha = 0.3$, $\varepsilon = 0.21$, CRRA=0.5



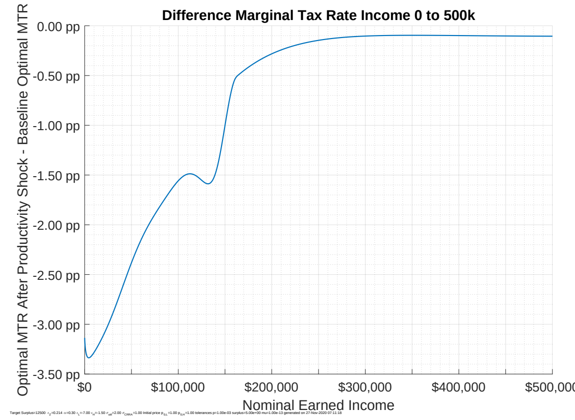
Notes: This figure plots optimal marginal tax rates under constant returns to scale (CRS, $\alpha = 0$) and increasing returns to scale (IRS, $\alpha = 0.3$). With increasing returns, the "naive" correction uses the formula $1 - T'_{NAIVE}(\theta) = \frac{1}{1-\alpha}(1 - T'_{CRS}(\theta))$. The optimal tax schedule solves the full optimization problem, accounting for endogenous changes in the value of redistribution across the income distribution. The three panels consider different values for the labor supply elasticity and social preferences for redistribution.

Figure A7 The Response of the Optimal Tax Schedule to Productivity Shocks
($\alpha = 0.3$, $\varepsilon_z = 0.21$, Pareto from SWF CRRA=1, PE price low-quality +2.5%, PE price high-quality -2.5%)

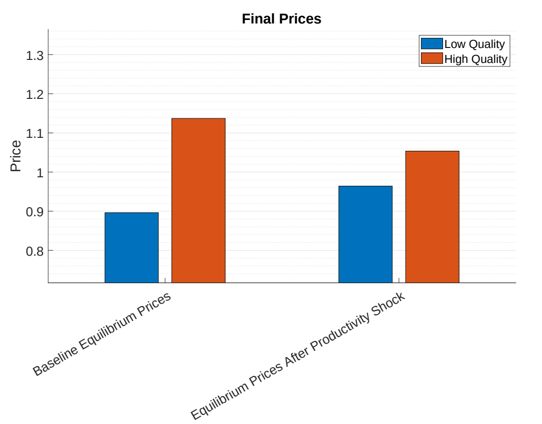
A. Optimal MTRs Before vs. After Price Shocks



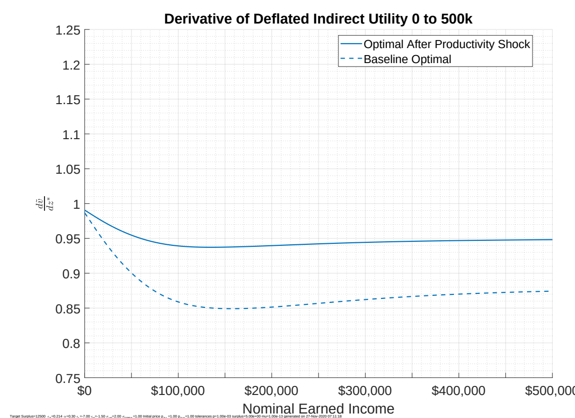
B. Difference b/w MTRs



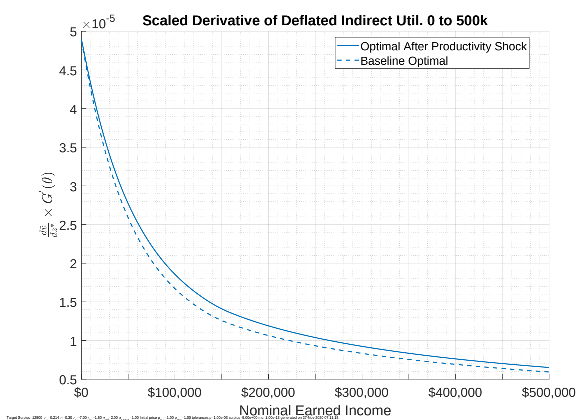
C. Equilibrium Prices



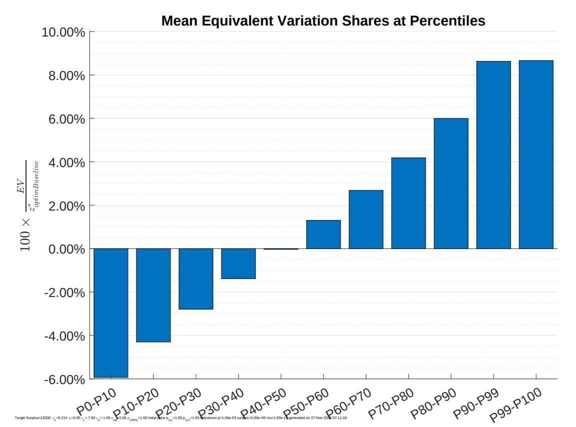
D. $\partial\widetilde{v}/\partial z^*$ by Earned Income Before vs. After Price Shocks



E. $\partial\widetilde{v}/\partial z^* \cdot G'(\theta)$ by Earned Income



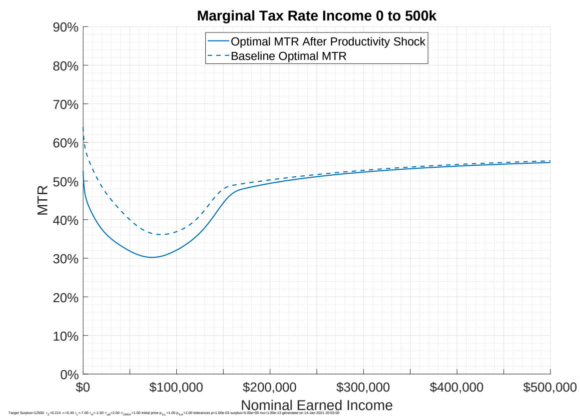
F. EV, initial vs. new schedules



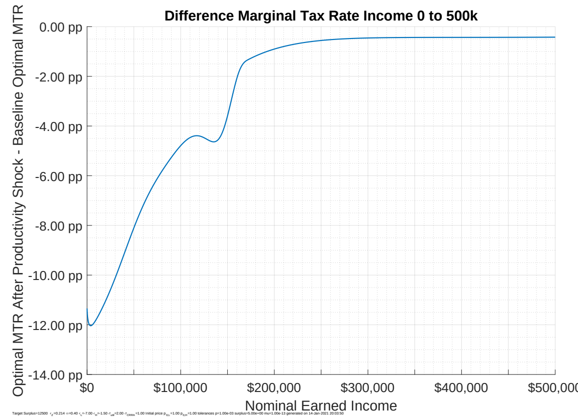
Notes: The quantitative model uses Pareto weights computed at the optimal homothetic tax schedule obtained under a social welfare function with CRRA=1. The exogenous productivity changes are such that the partial equilibrium price of the low-quality bundle increases by 2.5% while the partial equilibrium price of the high-quality bundle decreases by 2.5%.

Figure A8 Higher Returns to Scale Magnify the Impact of Productivity Shocks
($\alpha = 0.4$, $\varepsilon_z = 0.21$, Pareto weights from SWF CRRA=1, PE price low-quality +2.5%, PE price high-quality -2.5%)

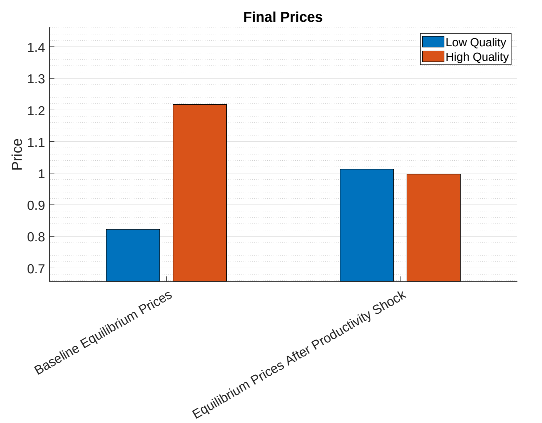
A. Optimal MTRs Before vs. After Price Shocks



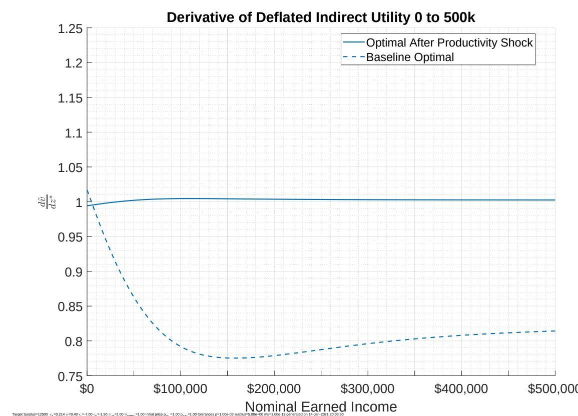
B. Change in MTRs



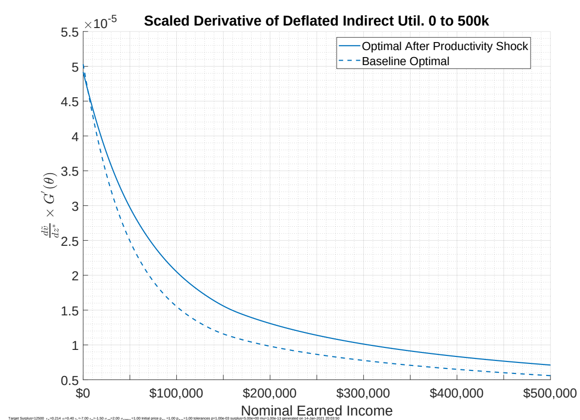
C. Equilibrium Prices



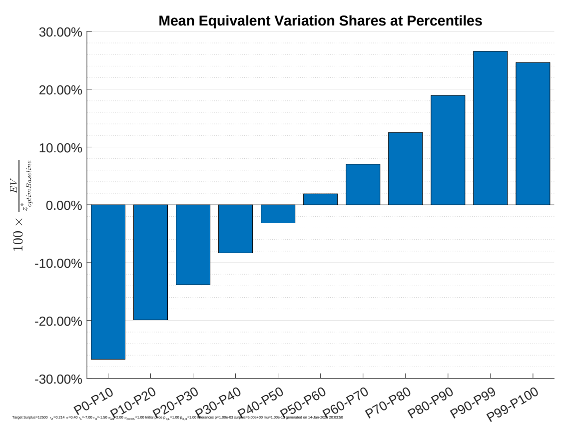
D. $\partial\widetilde{v}/\partial z^*$ by Earned Income



E. $\partial\widetilde{v}/\partial z^* \cdot G'(\theta)$ by Earned Income

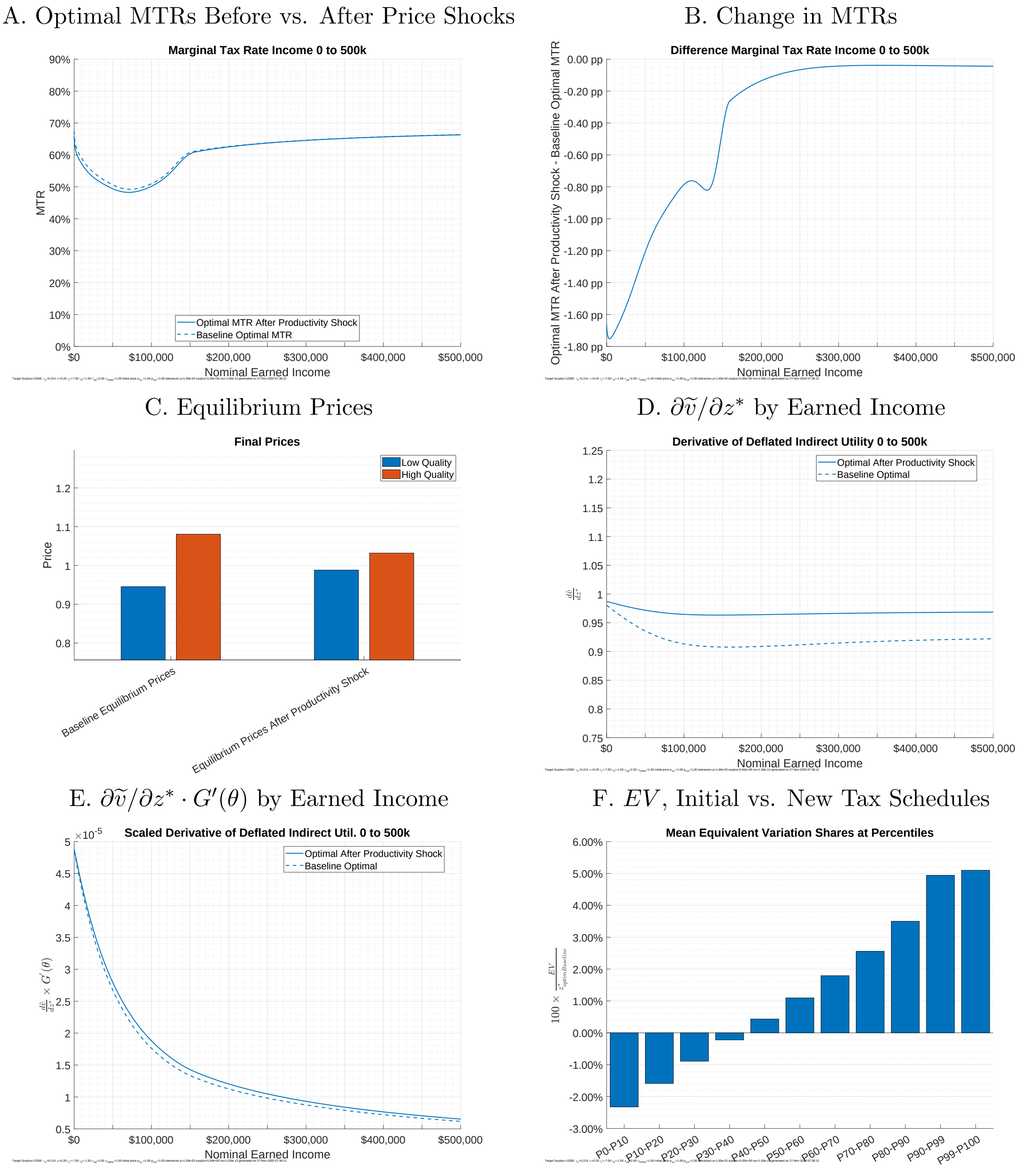


F. EV, Initial vs. New Tax Schedules



Notes: The quantitative model uses Pareto weights computed at the optimal homothetic tax schedule obtained under a social welfare function with CRRA=1. The exogenous productivity changes are such that the partial equilibrium price of the low-quality bundle increases by 2.5% while the partial equilibrium price of the high-quality bundle decreases by 2.5%.

Figure A9 Lower Returns to Scale Reduce the Impact of Productivity Shocks
($\alpha = 0.2$, $\varepsilon_z = 0.21$, Pareto weights from SWF CRRA=1, PE price low-quality +2.5%, PE price high-quality -2.5%)

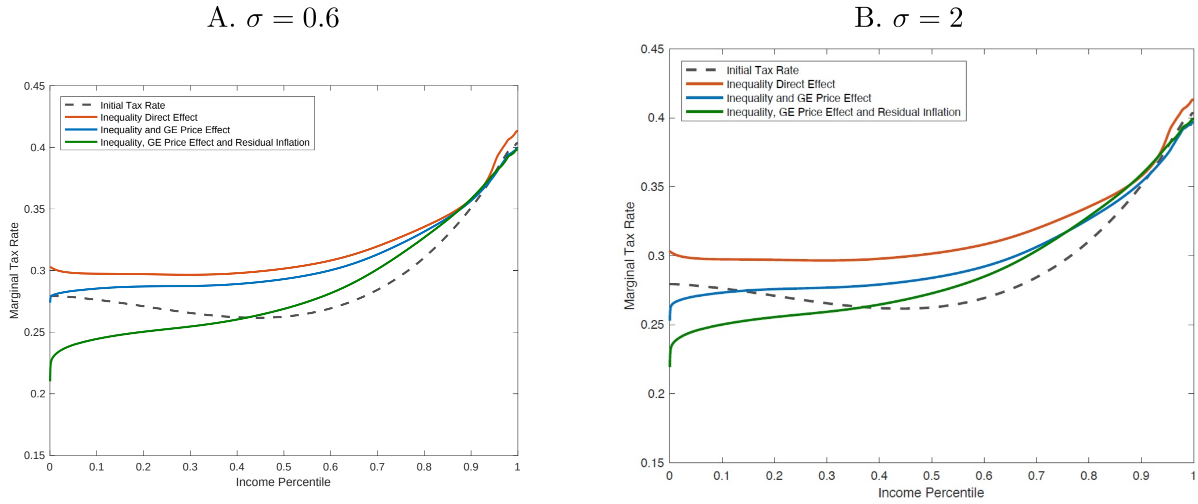


Notes: The quantitative model uses Pareto weights computed at the optimal homothetic tax schedule obtained under a social welfare function with CRRA=1. The exogenous productivity changes are such that the partial equilibrium price of the low-quality bundle increases by 2.5% while the partial equilibrium price of the high-quality bundle decreases by 2.5%.

Figure A10 Lower Social Preferences for Redistribution Magnify the Impact of Productivity Shocks ($\alpha = 0.3$, $\varepsilon_z = 0.21$, Pareto weights from SFW CRRA=0.5, PE price low-quality +2.5%, PE price high-quality -2.5%)

Notes: The model uses Pareto weights computed at the optimal homothetic tax schedule obtained under a social welfare function with CRRA=0.5. The exogenous productivity changes are such that the partial equilibrium price of the low-quality bundle increases by 2.5% while the partial equilibrium price of the high-quality bundle decreases by 2.5%.

Figure A11 The Response of the Optimal Tax Schedule to Observed Shifts in the Skill Distribution (2004-2015)

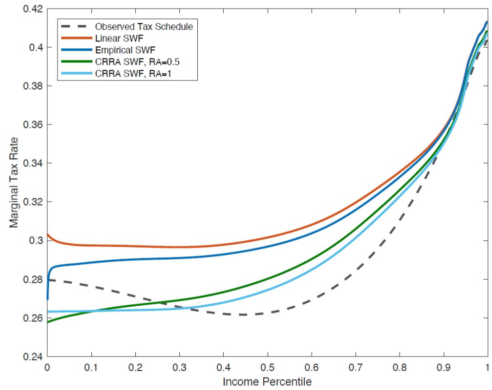


Notes: the IRS parameter is set to $\alpha = 0.3$ and the labor supply elasticity to $\varepsilon = 0.21$; the U.S. Census and CEX-CPI data sets are used in both panels and the initial tax schedule is taken from Hendren (2020).

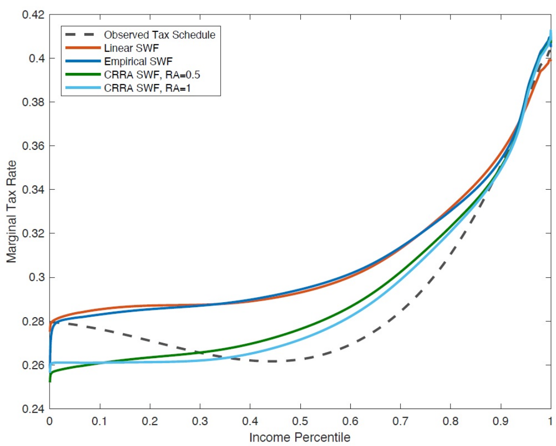
Figure A12 The Response of the Optimal Tax Schedule to Observed Shifts in the Skill Distribution (2004-2015),

The Role of the Curvature of the Social Welfare Function

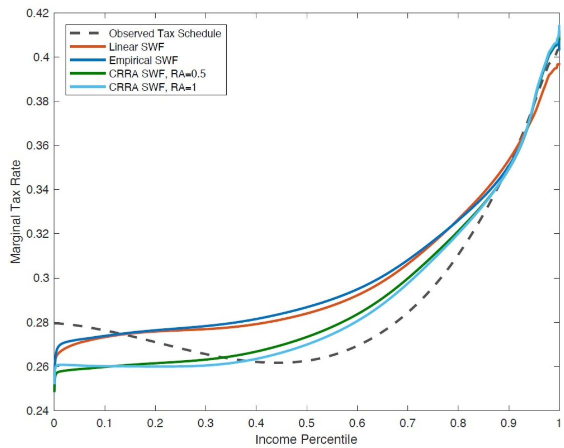
A. Exogenous prices



B. Endogenous prices, $\sigma = 0.6$



C. Endogenous prices, $\sigma = 2$



Notes: the IRS parameter is set to $\alpha = 0.3$ and the labor supply elasticity to $\varepsilon = 0.21$; the CEX-CPI data set is used in both panels and the initial tax schedule is taken from Hendren (2020).

Table A1 Notation for Model

Indices	θ	Agent productivity
	k	Sectoral index
Prices, Quantities,	p_k	Producer price index in sector k
Expenditures	ξ_k	Cost shifter for producer price index in sector k,
	q_k	Consumer price index in sector k
	C_k	Aggregate demand for sector k
	c_k	Household demand for sector k
	E	Aggregate spending
	E_k	Aggregate spending on sector k
	$\partial_{z^*} E_k$	Average marginal propensity to spend on k
	e_k	Household spending on sector k
	$\bar{s}_k$	Aggregate spending share on k
Elasticities	α_k	Elasticity of price p_k to market size C_k
	σ	Elasticity of substitution between sectors
	$\mathcal{S}$	Matrix of price elasticities
	ζ	Compensated labor supply elasticity
	$\tilde{\zeta}$	Compensated labor supply elasticity corrected for non-linearities in the tax schedule
	η	Income effect with a linear budget constraint
	$\tilde{\eta}$	Income effect corrected for nonlinearities in the tax schedule
Incomes	z	Pre-tax income
	z^*	Post-tax income
Social preferences	g	Pareto weight

Notes: This table lists the notation used for our model. The elasticity of substitution σ is relevant only for the two-sector model in the main text.

D Quantitative Model and Solution Algorithm

This section describes our quantitative model and solution algorithm. We first describe the economic environment. We then describe consumer preferences, contrasting the homothetic specification with non-homothetic preferences. Third, we describe the social planner's problem and the ordinary differential equations (ODEs) characterizing the solution. Finally, we present the solution algorithm for the ODEs.

D.1 Setting

D.1.1 Indirect Utility Function

The quantitative model uses a standard additively separable specification:

$$U\left(z^{*}, z, \boldsymbol{p}, \theta\right)=v\left(z^{*}, \boldsymbol{p}\right)-\psi\left(\frac{z}{\theta}\right) \tag{A4}$$

$$\psi\left(\frac{z}{\theta}\right)=\frac{1}{1+\frac{1}{\varepsilon_{z}}}\left(\frac{z}{\theta}\right)^{1+\frac{1}{\varepsilon_{z}}} \tag{A5}$$

where $\psi\left(\frac{z}{\theta}\right)$ is the cost of earning z given ability θ, and $v\left(z^{*}, \mathbf{p}\right)$ is the indirect utility function given prices and disposable income.

D.1.2 Pricing Function

Denoting aggregate consumption by C_i, the quantitative model is based on an isoelastic pricing function:

$$p_{i}=\gamma_{i} C_{i}^{-\alpha} \quad \forall i \in \mathcal{I} \tag{A6}$$

We calibrate γ_i to fit prices at the observed schedule, which are normalized to one without loss of generality, using the relationship:

$$\gamma_{i} \equiv p_{0, i} C_{0, i}^{\alpha} \tag{A7}$$

where $C_{0,i}$ is aggregate quantity consumed in sector i at initial prices. To obtain observed consumption, we compute disposable income at the observed schedule as defined in D.1.3, and then compute sectoral consumption given the expenditure shares described in D.2.1 and D.2.2.

D.1.3 Skill Distribution

The skill distribution $f(\theta)$ plays a key role in the shape of the optimal tax schedule. We use data from Hendren (2020) on the observed tax schedule to calibrate the skill distribution. As the data is only available for each percentile of the observed income distribution, we interpolate for marginal tax rates at income levels within the observed bounds using p-chip interpolation.

We then create a mapping from earned income at the observed tax schedule to skill type θ. Following Saez (2001), we obtain this mapping using the individual's utility function described in D.1.1, which depends on the functional form of $v\left(z^{*}, \boldsymbol{p}\right)$. We use two alternative forms of this indirect utility function - homothetic as described in D.2.1 and non-homothetic as described in D.2.2.

At the observed schedule in the homothetic case, we fit the skill distribution after setting: $p_{obs} = p_0 \equiv 1$. When computing income at the observed schedule in case of non-homothetic preferences, we use the "deflator" as defined in Definition 1 below. When we apply the deflator at any initial prices $\mathbf{p_0}$ the indirect utility of the agent will always be the same as in the homothetic case with $p_0 = 1$. This approach allows us to use the same skill distribution in the homothetic and in non-homothetic cases.

D.2 Consumer Preferences

This section describes the indirect utility function $v\left(z^*, \boldsymbol{p}\right)$ from D.1.1.

D.2.1 Homothetic Preferences

With homothetic preferences, the indirect utility function $v\left(z^*, \boldsymbol{p}\right)$ described in D.1.1 is given by:

$$v\left(z^*, \boldsymbol{p}\right) \equiv \frac{z^*}{p} \tag{A8}$$

where p is the price in the economy. The individual's utility function, per equation A4, is:

$$U\left(z^*, z, \boldsymbol{p}, \theta\right) = \frac{z^*}{p} - \psi\left(\frac{z}{\theta}\right) = \frac{z^*}{p} - \frac{1}{1+\frac{1}{\varepsilon_z}}\left(\frac{z}{\theta}\right)^{1+\frac{1}{\varepsilon_z}} \tag{A9}$$

Plugging in the definition of disposable income, the optimal $z(\theta)$ satisfies the FOC:

$$\frac{dU\left(\theta\right)}{z(\theta)} = \frac{1-T'(z(\theta))}{p} - \left(\frac{z(\theta)}{\theta}\right)^{\frac{1}{\varepsilon_z}}\frac{1}{\theta} = 0$$

We can thus express income or skill parameters as functions of observables:

$$z(\theta) = \theta^{1+\varepsilon_z}\left(\frac{1-T'(z(\theta))}{p}\right)^{\varepsilon_z} \tag{A10}$$

$$\theta = \left[z\left(\frac{p}{1-T'(z)}\right)^{\varepsilon_z}\right]^{\frac{1}{1+\varepsilon_z}} \tag{A11}$$

With $\theta = 0$, we apply the limiting case described in D.3.2.

D.2.2 Non-Homothetic CES Preferences

Definitions and Properties We use the *General Non-Homothetic CES Preferences* as defined in Appendix A.1 of Comin, Lashkari and Mestieri revision 3 (2019). The indirect utility function $v\left(z^*, \boldsymbol{p}\right)$ described in D.1.1 is given by $v \equiv v\left(z^*, \boldsymbol{p}\right) \equiv F(\mathbf{C})$, where $\mathbf{C}$ is the consumption vector of the agent.

Indirect utility v is implicitly defined by:

$$\sum_{i\in\mathcal{I}} \Omega_i^{\frac{1}{\sigma}}\left(\frac{C_i}{v^{\frac{\varepsilon_i}{1-\sigma}}}\right)^{\frac{\sigma-1}{\sigma}} = \sum_{i\in\mathcal{I}}\left(\Omega_i v^{\varepsilon_i}\right)^{\frac{1}{\sigma}} C_i^{\frac{\sigma-1}{\sigma}} = 1, \tag{A12}$$

where parameters ε_i denote the utility elasticities of each good, the elasticity of substitution between

sectors is denoted σ , and taste parameters are denoted Ω_i, for an arbitrary set of sectors $i \in \mathcal{I}$. For the quantitative analysis, we consider two sectors, labeled "high quality" (H) and "low quality" (L).
Under this specification, Marshallian Demand (spending shares) and the price index are:

$$\omega_i(z^*) = \Omega_i \left(\frac{p_i}{P}\right)^{1-\sigma} \left(\frac{z^*}{P}\right)^{\varepsilon_i-(1-\sigma)} \tag{A13}$$

$$P(\mathbf{p}, z^*) = \left[\sum_{i\in\mathcal{I}} \left(\Omega_i p_i^{1-\sigma}\right)^{\chi_i} \left(\omega_i\left(z^*\right)^{1-\sigma}\right)^{1-\chi_i}\right]^{\frac{1}{1-\sigma}} \tag{A14}$$

$$\text{where } \chi_i \equiv \frac{1-\sigma}{\varepsilon_i}$$

Using this specification, we obtain quantity consumed as:

$$C_i(z^*) = \frac{\omega_i(z^*)z^*}{p_i} \tag{A15}$$

Definition of Deflated Non-Homothetic Indirect Utility Function In the social planner's problem in D.3.1, we use a deflated indirect utility function.
Definition 1 (Deflated Indirect Utility Function) *Deflated indirect utility function $\widetilde{v}(z^*, \boldsymbol{p})$ is the inverse of the indirect utility function at initial prices, under constant returns to scale. It can be thought of as the level of "virtual disposable income" $\tilde{z}^*$ that satisfies $v(\tilde{z}^*, \boldsymbol{p}_0) = v(z^*, \boldsymbol{p})$. Formally,*

$$\widetilde{v}(z^*, \boldsymbol{p}) = v^{-1}(v(z^*, \boldsymbol{p}), \boldsymbol{p}_0) \tag{A16}$$

Properties of the deflated indirect utility function are listed below. At $\boldsymbol{p}_0$, the non-homothetic indirect utility is equivalent to the homothetic case from D.2.1:

$$\frac{d\widetilde{v}(z^*, \boldsymbol{p})}{dz^*} = \frac{dv(z^*, \boldsymbol{p})}{dz^*} \left(\frac{dv(\tilde{z}^*, \boldsymbol{p}_0)}{d\tilde{z}^*}\right)^{-1} \text{ where:} \tag{A17}$$

$$\tilde{z}^* = v^{-1}(v(z^*, \boldsymbol{p}), p), \boldsymbol{p}_0)$$

$$\widetilde{v}(z^*, \boldsymbol{p}_0) = z^* \tag{A18}$$

$$\frac{d\widetilde{v}(z^*, \boldsymbol{p}_0)}{dz^*} = 1 \tag{A19}$$

D.3 ODEs from Social Planner's Problem

D.3.1 Social Planner's Problem

The social planner chooses the optimal tax schedule to maximize total utility over the distribution of types θ, subject to budget constraint, agents' FOC and market clearing, according to an arbitrary social welfare function $G\left(U(\theta, \mathbf{p})\right)$:

$$max_{z(\theta)} \int_{\underline{\theta}}^{\overline{\theta}} G\left(U(\theta, \mathbf{p})\right) f(\theta) d\theta \quad s.t: \tag{A20}$$

$$G\left(U(\theta, \mathbf{p})\right) = G\left(\widetilde{v}\left(z^*(\theta), \boldsymbol{p}\right) - \psi\left(\frac{z(\theta)}{\theta}\right)\right) \tag{A21}$$

$$G'\left(U(\theta, \mathbf{p})\right) = \frac{dG}{dU} \tag{A22}$$

$$R \geq \int_{\underline{\theta}}^{\overline{\theta}} \left(z(\theta) - z^*(\theta)\right) f(\theta) d\theta \tag{A23}$$

$$p_i = \gamma_i \left(\overline{C_i}\right)^{-\alpha}, \forall i \tag{A24}$$

where:

1. $\overline{C_i} = \int_{\underline{\theta}}^{\overline{\theta}} C_i(\theta) dF(\theta)$ denotes aggregate consumption in sector i
2. R denotes the government surplus (government revenue requirement)
3. The state variable is $G(U(\theta, \mathbf{p}))$ and the control variable is $z(\theta)$

Using the envelope theorem and our functional form for U, we can write:

$$\dot{U}(\theta) = \frac{1}{\theta}\left(\frac{z}{\theta}\right)^{1+\frac{1}{\varepsilon_z}} \tag{A25}$$

Call $\mu(\theta)$ the co-state variable for the evolution of $G(\cdot)$. The equation for μ is

$$\dot{\mu}(\theta) = \left((1-\alpha)\frac{\lambda}{\frac{d\widetilde{v}}{dz^*}} - G'\left(U(\theta)\right)\right) f(\theta), \tag{A26}$$

where

- $\frac{d\widetilde{v}}{dz^*}$ is the derivative of deflated indirect utility $\widetilde{v}$ with respect to disposable income, evaluated at the level of disposable income z^* (which we express below as a function of θ, U and μ)
- λ is the multiplier on the government's budget constraint

The first-order condition for z gives:

$$\mu(\theta) \cdot \left(\frac{1+\frac{1}{\varepsilon_z}}{\theta^2}\left(\frac{z}{\theta}\right)^{\frac{1}{\varepsilon_z}}\right) = -\lambda\left(1 - \frac{(1-\alpha)\cdot\left(\frac{z}{\theta}\right)^{\frac{1}{\varepsilon_z}}}{\theta \cdot \frac{d\widetilde{v}}{dz^*}}\right) f(\theta) \tag{A27}$$

$$\mu(\theta) = \theta^2 \frac{\lambda\left((1-\alpha)\frac{\frac{1}{\theta}\left(\frac{z}{\theta}\right)^{\frac{1}{\varepsilon_z}}}{\frac{d\widetilde{v}}{dz^*}} - 1\right) f(\theta)}{\left(1+\frac{1}{\varepsilon_z}\right)\left(\frac{z}{\theta}\right)^{\frac{1}{\varepsilon_z}}} \tag{A28}$$

The **boundary conditions** are:

$$
\begin{aligned}
\mu(\underline{\theta}) &= 0 \\
\mu(\overline{\theta}) &= 0
\end{aligned}
$$

The government resource constraint in equation (A23) results from the fact that government revenue is distributed among the agents in the economy through a lump sum transfer such that an amount R is not redistributed. With R denoting government surplus, we have

$$
\overline{C} = \int C(\theta)\ dF(\theta) = \int C(z^*(\theta))\ dF(\theta) = \int \frac{z^*(\theta)}{p}\ dF(\theta) = \int \frac{z(\theta) - R}{p}\ dF(\theta) \tag{A29}
$$

D.3.2 Limiting Case

We need to address the case when $\theta = 0$ since several equations from D.3.1 become indeterminate. We have

$$
\zeta(\theta) \equiv \frac{1}{\theta}\left(\frac{z(\theta)}{\theta}\right)^{\frac{1}{\varepsilon_z}} \tag{A30}
$$

$$
\ell(\theta) \equiv \frac{z(\theta)}{\theta} = (\theta\zeta(\theta))^{\varepsilon_z} \tag{A31}
$$

These two functions are bounded functions of θ near 0. Using equation (A27) and the definitions above we can express:

$$
\zeta(\theta) = \left[\frac{1-\alpha}{\frac{d\widetilde{v}}{dz^*}} - \frac{\mu(\theta)}{\theta}\frac{1+\frac{1}{\varepsilon_z}}{\lambda f(\theta)}\right]^{-1} \tag{A32}
$$

$$
\dot{U}(\theta) = \zeta(\theta)\ell(\theta) \tag{A33}
$$

We still need to address the fact that $\frac{\mu(\theta)}{\theta}$ is undefined for $\theta = 0$. Given our specification we know that:

$$
\begin{aligned}
&f(0) > 0 \\
&\dot{\mu}(0) < 0 \\
&\mu(\theta) < 0 \quad \theta \to 0_+
\end{aligned}
$$

Therefore we can use:

$$
\lim_{\theta\to 0}\frac{\mu(\theta)}{\theta} = \dot{\mu}(\theta) = \left((1-\alpha)\frac{\lambda}{\frac{d\widetilde{v}}{dz^*}} - G'\left(U(\theta)\right)\right) f(\theta) \tag{A34}
$$

Using these relationships we can express several of our key variables for $\theta = 0$:

$$\dot{U}(0) = 0 \tag{A35}$$

$$z(0) = 0 \tag{A36}$$

$$\psi(\frac{0}{0}) = 0 \tag{A37}$$

$$\zeta(0) = \left[\frac{1-\alpha}{\frac{d\widetilde{v}}{dz^*}} - \dot{\mu}(0)\frac{1+\frac{1}{\varepsilon_z}}{\lambda f(0)}\right]^{-1} \tag{A38}$$

D.3.3 System of ODEs

The solution to the general case allowing for non-homotheticities in agents' utility function and an arbitrary social welfare function $G(U(\theta, \mathbf{p}))$ is given by the following system of ODEs and boundary conditions:

$$\dot{U}(\theta) = \frac{z}{\theta^2}\left(\frac{z}{\theta}\right)^{\frac{1}{\varepsilon_z}} \tag{A39}$$

$$\dot{\mu}(\theta) = \left((1-\alpha)\frac{\lambda}{\frac{d\widetilde{v}}{dz^*}} - G'\left(\widetilde{v} - \psi\right)\right) f(\theta) \tag{A40}$$

$$\mu(\theta) = \theta^2 \frac{\lambda\left((1-\alpha)\frac{\frac{1}{\theta}\left(\frac{z}{\theta}\right)^{\frac{1}{\varepsilon_z}}}{\frac{d\widetilde{v}}{dz^*}} - 1\right) f(\theta)}{\left(1+\frac{1}{\varepsilon_z}\right)\left(\frac{z}{\theta}\right)^{\frac{1}{\varepsilon_z}}} \tag{A41}$$

with the boundary conditions:

$$\mu(\underline{\theta}) = 0$$

$$\mu(\overline{\theta}) = 0$$

Furthermore, we can express incomes as a function of other variables:

$$z = \theta \cdot \left(\frac{1-\alpha}{\theta \cdot \frac{d\widetilde{v}}{dz^*}} - \frac{\mu(\theta)}{\lambda f(\theta)} \cdot \frac{1+\frac{1}{\varepsilon_z}}{\theta^2}\right)^{-\varepsilon_z} \tag{A42}$$

$$z^* = \widetilde{v}^{-1}\left(U(\theta) + \psi\left(\frac{z}{\theta}\right)\right) = \widetilde{v}^{-1}\left(U(\theta) + \frac{1}{1+\frac{1}{\varepsilon_z}}\left(\frac{z}{\theta}\right)^{1+\frac{1}{\varepsilon_z}}\right) \tag{A43}$$

This is a system of non-linear equations we can solve for to obtain z and z^* given $\theta,\ U(\theta),\ \mu(\theta),\ \lambda$. We apply the limit case as per D.3.2.

Homothetic Case With homothetic indirect utility, the system of ODEs can be expressed as:

$$
\begin{aligned}
\dot{U}(\theta) &= \frac{z}{\theta^2} \cdot \left(\frac{z}{\theta}\right)^{\frac{1}{\varepsilon_z}} \\
\dot{\mu}(\theta) &= \left((1-\alpha)\cdot\lambda\cdot p - \left(\frac{z^*}{p} - \psi\right)^{-\widetilde{\sigma}}\right)\cdot f(\theta) \\
\mu(\theta) &= \theta^2 \frac{\lambda\left(p(1-\alpha)\frac{1}{\theta}\left(\frac{z}{\theta}\right)^{\frac{1}{\varepsilon_z}} - 1\right) f(\theta)}{\left(1+\frac{1}{\varepsilon_z}\right)\left(\frac{z}{\theta}\right)^{\frac{1}{\varepsilon_z}}} \\
z &= \theta \cdot \left(\frac{-\theta^2 \lambda f(\theta)}{\mu(\theta)\left(1+\frac{1}{\varepsilon_z}\right) - \theta\lambda(1-\alpha)f(\theta)p}\right)^{\varepsilon_z} \\
z^* &= p \cdot \left(U(\theta) + \frac{1}{1+\frac{1}{\varepsilon_z}} \cdot \left(\frac{z}{\theta}\right)^{1+\frac{1}{\varepsilon_z}}\right)
\end{aligned}
$$

with boundary conditions:

$$
\begin{aligned}
\mu(\underline{\theta}) &= 0 \\
\mu(\overline{\theta}) &= 0
\end{aligned}
$$

D.3.4 Social Welfare Function

To study the role of non-linearities in the social welfare function, we consider a specification with constant relative risk aversion, where the CRRA risk parameter is denoted $\widetilde{\sigma}$. The functional form is:

$$
G'\left(U(\theta, \mathbf{p})\right) = \frac{dG}{dU} = \left(U(\theta, \mathbf{p})\right)^{-\widetilde{\sigma}}, \quad \widetilde{\sigma} \geq 0 \tag{A44}
$$

$$
G\left(U(\theta, \mathbf{p})\right) \equiv \begin{cases} \log\left(U(\theta, \mathbf{p})\right) & \text{if } \widetilde{\sigma} = 1 \\ \frac{\left(U(\theta, \mathbf{p})\right)^{1-\widetilde{\sigma}}}{1-\widetilde{\sigma}} & \text{if } \widetilde{\sigma} \geq 0 \ \wedge \ \widetilde{\sigma} \neq 1 \end{cases} \tag{A45}
$$

D.3.5 Pareto Analysis

We also perform analysis using Pareto weights, denoted $\lambda(\theta)$ and set to match the results obtained with the CRRA social welfare function $G(\cdot)$, with the CRRA risk parameter $\widetilde{\sigma}$. The Pareto weight are given by:

$$
\lambda(\theta) \equiv \left(U_{optim}(\theta)\right)^{-\widetilde{\sigma}}, \tag{A46}
$$

where $U_{optim}(\theta)$ is the solution of the optimal taxation problem with homothetic indirect utility function, $\alpha = 0$, and the CRRA parameter $\widetilde{\sigma}$.

With Pareto weights, the social welfare function and its derivative become:

$$
G(\theta) \equiv \lambda(\theta) U(\theta, \mathbf{p}), \tag{A47}
$$

$$
G'(\theta) = \frac{dG}{dU} = \lambda(\theta). \tag{A48}
$$

D.3.6 Defining the Equivalent Variation

The equivalent variation (EV) is defined by:

$$\widetilde{v}\left(z^*_{ref}(\theta) + EV(\theta), \boldsymbol{p}_{ref}\right) - \psi\left(\frac{z_{ref}(\theta)}{\theta}\right) = u_{optim}(\theta),$$

where "*ref*" denotes the reference point and "*optim*" the new equilibrium. In our main specifications, the reference point is the outcome at the optimal tax schedule with homothetic preferences, to which we compare the outcome with non-homothetic preferences.

D.4 Solution Algorithm

This section describes the algorithm to solve the problem described in D.3.1, using nested bisection in Matlab.

D.4.1 Convergence to Optimal Schedule

The algorithm relies on three nested loops. Each loop ensures that we satisfy one of the conditions in the social planner's problem (see D.3.1); we guess the value of one parameter in each loop, and then solve for all inner loops.

The loops are structured as follows, from the outer loop to the inner loop:

1. *Price loop* - ensures prices in the economy converge such that equation (A24) is satisfied by guessing $\boldsymbol{p}$. The convergence condition is base on the price change.
2. *Surplus loop* - ensures government surplus equation (A23) converges by guessing a value of λ. The convergence condition is the distance from the revenue requirement.
3. *Utility or μ loop* - ensures that the boundary condition $\mu(\overline{\theta}) = 0$ is satisfied by guessing value of $U(\underline{\theta})$. The convergence condition is the distance between $\mu(\overline{\theta})$ and the boundary condition of 0.

We set a convergence condition (tolerance) for each loop, which determines whether the variable of interest has converged. In what follows, $\epsilon_{\mathbf{p}}, \epsilon_{\lambda}, \epsilon_{\mu}$ denote tolerance for price, surplus and utility loops, respectively.

In the description of the algorithm below, for any variable the indexes represent the iteration of price, surplus and utility loop, respectively. The optimal schedule is defined by $\mathbf{p}_{optim}$, λ_{optim} and $U_{optim}(\underline{\theta})$, denoting the values under which the variable determining convergence of each loop converged.

For illustration, assume the values of counters at convergence were 3 for price, 7 for surplus and 10 for utility. Then, the optimal value of utility is denoted $U_{3,7,10}(\underline{\theta})$, which is the value used to solve the ODE in the 10th utility loop, within the 7th surplus loop, within the 3rd price loop.

Thus, the optimal schedule can be defined as:

$$\mathbf{p}_{optim} = \mathbf{p}_{10} = \mathbf{p} \quad \textbf{such that} \tag{A49}$$

$$\epsilon_{\mathbf{p}} \geq \frac{1}{\boldsymbol{p}} \left(\boldsymbol{\gamma} \int_{\underline{\theta}}^{\overline{\theta}} \boldsymbol{C} \left(z^*(\theta; \boldsymbol{p}, \lambda_{optim}, U_{optim}(\underline{\theta})) \right) dF(\theta) - \boldsymbol{p} \right) \quad \textbf{where} \tag{A50}$$

$$\lambda_{optim} = \lambda_{3,7} = \lambda \quad \textbf{such that} \tag{A51}$$

$$\epsilon_{\lambda} > \left| \int_{\underline{\theta}}^{\overline{\theta}} z(\theta; \boldsymbol{p}, \lambda, U_{optim}(\underline{\theta})) - z^*(\theta; \boldsymbol{p}, \lambda, U_{optim}(\underline{\theta})) f(\theta) d\theta - R \right| \quad \textbf{where} \tag{A52}$$

$$U_{optim}(\underline{\theta}) = U_{3,7,10}(\underline{\theta}) = U(\underline{\theta}) \quad \textbf{such that} \tag{A53}$$

$$\epsilon_{\mu} > \left| \mu(\overline{\theta}; \boldsymbol{p}, \lambda, U(\underline{\theta})) \right| \tag{A54}$$

D.4.2 Adjustment of Bounds

In the bisection algorithm, we update the bounds at the end of each iteration of loop based on the value of the variable of interest. In the case of the **utility or μ loop**, we change bounds on $U(\underline{\theta})$ according to the rule:

$$U_{UB}(\underline{\theta}) = U_{current}(\underline{\theta}) \text{ if } \mu(\overline{\theta}) < 0 \text{ or the solver failed}$$
$$U_{LB}(\underline{\theta}) = U_{current}(\underline{\theta}) \text{ if } \mu(\overline{\theta}) \geq 0$$

In the case of the **surplus loop**, we change bounds on λ according to the rule:

$$\lambda_{UB} = \lambda_{current} \text{ if } \int_{\underline{\theta}}^{\overline{\theta}} \left(z(\theta) - z^*(\theta) \right) f(\theta) d\theta \geq R$$
$$\lambda_{LB} = \lambda_{current} \text{ if } \int_{\underline{\theta}}^{\overline{\theta}} \left(z(\theta) - z^*(\theta) \right) f(\theta) d\theta < R$$

E Extensions

In this section, we present extensions of our theoretical results under general household preferences and supply side specification. We first present the general model, then the extensions of Proposition 1-4.

E.1 Model

We consider a n-sector economy, sectors are indexed by k. There is a mass 1 of households with different productivity types θ distributed according to $\pi(\theta)$.

Households. Preferences are weakly separable between goods and leisure. Given consumer prices $\{q_k\}_{1\leq k\leq n}$ and the income tax schedule $\boldsymbol{T}$, a household of type θ solves:

$$V(\theta) = \sup_{\{c_1,...,c_n,z\}} U\left(u\left(c_1,...,c_n\right), z, \theta\right), \quad \text{s.t.} \sum_{k=1}^{N} q_k c_k = z - T\left(z\right). \tag{A55}$$

Where the functions U and u are increasing, strictly concave and three times continuously differentiable. We assume that all functions are $\mathcal{C}^3$ to ensure that consumption functions are twice continuously differentiable in prices and post-tax income. Indeed, we first have, by concavity of the utility functions, that the optimal consumption choice at a given z and θ is unique and characterized by the problem first order conditions. A direct application of the implicit function theorem then directly shows that the consumption functions c_k are $\mathcal{C}^2$ functions of prices and post tax income.

We define the sub-utility of consumption $v\left(\boldsymbol{q}, z^*\right)$ as the solution of:

$$v\left(\boldsymbol{q}, z^*\right) = \sup_{\{c_1,...,c_n\}} u\left(c_1,...,c_n\right) \quad s.t. \quad \sum_{k=1}^{N} q_k c_k = z^*.$$

The unique optimal choice of consumption allocation across sector $c_k\left(\boldsymbol{q}, z^*\right)$ is independent of type conditional on $\{\mathbf{q}, z^*\}$.

To discipline how preferences depend on type, we make the standard assumption that preferences satisfy the single crossing property.

Assumption. *Single Crossing Property. For any z^*, z and $\mathbf{q}$, the marginal rates of substitutions $-U_z\left(v\left(\boldsymbol{q}, z^*\right), z, \theta\right) / \left(U_v\left(v\left(\boldsymbol{q}, z^*\right), z, \theta\right) v_{z^*}\left(\boldsymbol{q}, z^*\right)\right)$ are decreasing in type.*

It is well known in the mechanism design literature that the assumption allows us to simplify the set of incentive compatibility constraints in the direct allocation of $z\left(\theta\right)$, $z^*\left(\theta\right)$ ($U\left(v\left(\boldsymbol{q}, z^*\left(\theta\right)\right), z\left(\theta\right), \theta\right) \geq U\left(v\left(\boldsymbol{q}, z^*\left(\theta'\right)\right), z\left(\theta'\right), \theta\right)$) into a local constraint and monotonicity condition on $z\left(\theta\right)$.

We finally provide definitions for standard demand functions and labor supply elasticities.

Definition E.1. *The expenditure function of the sub-utility problem is defined as*

$$e(\boldsymbol{q}, v) = \inf_{\{c_1,\ldots,c_n\}} \sum_{k=1}^{n} q_k c_k \quad s.t. \quad u(c_1, \ldots, c_n) \geq v.$$

The corresponding Hicksian demand function is defined as $c_k^h(\boldsymbol{q}, v) = \partial_{q_k} e(\boldsymbol{q}, v)$, the aggregate substitution matrix is defined as $\mathcal{S}_{k,l} = q_l \int \partial_{q_l} c_k^h(\boldsymbol{q}, v(\boldsymbol{q}, z(\theta) - T(z(\theta))))\, \pi(\theta)\, d\theta / C_k$. The marginal rate of substitution is given by $MRS(z^*, z, \theta, \mathbf{q}) \equiv -U_z(v(\boldsymbol{q}, z^*), z, \theta) / (U_v(v(\boldsymbol{q}, z^*), z, \theta)\, v_{z^*}(\boldsymbol{q}, z^*))$. We define the linear compensated wage elasticity as $\zeta = MRS / (z \partial_z MRS + z MRS \partial_{z^*} MRS)$, and the elasticity adjusted for non linearity in the tax schedule as $\tilde{\zeta} = \zeta / (1 + z\zeta T'' / (1 - T'))$. The linear income effect is defined as $\eta = -z\zeta\, \partial_{z^*} MRS$ and the adjusted income effect as $\tilde{\eta} = -z\tilde{\zeta}\, \partial_{z^*} MRS$.

Firms. As in the main text, we summarize the supply side of the economy through a cost function and a pricing function. The cost function in sector k is $\chi_k(C_1, \ldots, C_n, \xi_k)$, while the pricing function is $p_k = \phi_k(C_1, \ldots, C_n, \xi_k)$. The main difference with our specification in the main text is that we now allow for spillovers across sectors (an increase in demand for good l affects the price of good k). As before, we will consider two cases. In the competitive case, $\phi_k(C_1, \ldots, C_n, \xi_k) = \sum_{l=1}^{n} \partial_{C_k} \chi_l(C_1, \ldots, C_n, \xi_k)$; in the monopolistic case, $\chi_k(C_1, \ldots, C_n, \xi_k) = C_k \phi_k(C_1, \ldots, C_n, \xi_k)$.

With these more general pricing functions, we need to define the elasticities of price with respect to any increase in aggregate demand. We do so in the definition below:

Definition E.2. *The price elasticity of p_k with respect to market size C_l is defined as*

$$A_{k,l} \equiv -C_l \partial_{C_l} \phi_k(C_1, \ldots, C_N, \xi) / p_k.$$

Planning Problem. The government maximizes a social welfare function $\int G(V(\theta), \theta)\pi(\theta) d\theta$ with G increasing and concave in V using a nonlinear income tax $\mathbf{T}$, taxes on consumption prices q_k and a full profit tax, subject to the firm and household problem defined above.

$$\begin{aligned}
&\sup_{\mathbf{T}, q_k} \int G(V(\theta), \theta)\pi(\theta) d\theta \\
\text{s.t.} \quad & V(\theta) = \sup_z U(v(\boldsymbol{q}, z - T(z)), z, \theta) \quad \text{and} \int T(z(\theta))\pi(\theta) d\theta + \sum_{k=1}^{n} q_k C_k - \chi_k(C_1, \ldots, C_n, \xi_k) \geq 0 \\
\text{with} \quad & V(\theta) = U(v(\theta), z(\theta), \theta), \ \ v(\theta) = u(c_1(q, z^*(\theta)), \ldots, c_n(q, z^*(\theta))), \text{ and } z^*(\theta) = z - T(z) \\
& C_i = \int c_i(q, z^*(\theta))\pi(\theta) d\theta,
\end{aligned}$$

where the consumption function solves $\boldsymbol{c}(\boldsymbol{q}, z^*) = \text{argmax}_{\boldsymbol{c}} u(\boldsymbol{c})$ *s.t.* $\boldsymbol{q} \cdot \boldsymbol{c} = z^*$. Given our single crossing condition, the planner's problem can be re-expressed as a direct mechanism, with the additional condition that $z(\theta)$ is non decreasing in types:

$$
\begin{aligned}
&\sup_{V(\theta),z(\theta),q_i} \int G(V(\theta),\theta)\pi(\theta)d\theta \\
\text{s.t.}\quad & V'(\theta) = U_\theta(v(\theta),z(\theta),\theta) \quad \text{and} \int z^*(\theta) - z(\theta)\pi(\theta)d\theta - \sum_{k=1}^{n} q_k C_k - \chi_k\left(C_1,...,C_n,\xi_k\right) \leq 0 \\
\text{with}\quad & V(\theta) = U(v(\theta),z(\theta),\theta),\;\; v(\theta) = u(c_1(q,z^*(\theta)),...,c_n(q,z^*(\theta))),\ \text{and}\ z^*(\theta) = q\cdot c(q,z^*(\theta)) \\
& C_i = \int c_i(q,z^*(\theta))\pi(\theta)d\theta,
\end{aligned}
$$

where v is the indirect sub-utility for consumption c_i denotes demand for i and z^* post-tax income.

E.1.1 Micro-foundations of the Supply Side

While our specification of the supply side of the economy through a cost and pricing function is fully general in the competitive case, it might not be obvious which monopolistic frameworks it encompasses. In this subsection, we give micro-foundations for our reduced-form specification and argue that it covers a large class of models of free entry with monopolistic competition.

In each sector k, monopolistic producers can freely enter and produce differentiated varieties $y_k(i)$ of product k. These firms are indexed by $i \in \mathcal{I}_k$ and firm i sells its variety at price $p_k(i)$. These varieties are then aggregated by competitive retailers. Competitive retailers bundle the varieties with a constant return to scale production function $\mathcal{F}_k$, which is increasing, concave, homogeneous of degree 1and symmetric in its arguments. The retailer's problem is given by:

$$
\sup_{Y_k,\{y_k(i)\}_{j\in\mathcal{I}_k}} p_k Y_k - \int_{i\in\mathcal{I}_k} p_k(i)\,y_k(i)\,di \quad \text{s.t.}\quad \mathcal{F}_k\left(\left\{\frac{y_k(i)}{Y_k}\right\}_{j\in\mathcal{I}_k}\right) = 1.
$$

Given that $\mathcal{F}_k$ is homogenous and that retailers are competitive, we have $p_k = \int_{i\in\mathcal{I}_k} p_k(i)\,y_k(i)/Y_k di$, and demand for variety i is given by $y_k(i) = d_k\left(p_k(i),\{p_k(j)\}_{j\in\mathcal{I}_k}\right)Y_k$. On the production side, firms can freely enter all markets. Upon entering, firm i pays a fixed labor cost $\xi_{e,k}$ and draws its productivity type $\gamma(i)$ from a distribution Ψ_k. To start production, firms have to pay a second fixed cost, $\xi_{p,k}$. Firms use both labor ℓ and other sectors output $\tilde{y}_l$ to produce (through Input-Output linkages). The variable cost of producing $y_k(i)$ units of variety i in market k, χ_k is defined by the following problem:

$$
\begin{aligned}
\chi_k(i) = \inf_{\{\tilde{y}_1,...,\tilde{y}_n,\ell\}} & \sum_{k=1}^{n} p_k\tilde{y}_k + \ell \\
\text{s.t.} & F_k\left(\tilde{y}_1,...,\tilde{y}_n,\ell,\gamma(i),\xi_{c,k}\right) = y_k(i).
\end{aligned}
$$

We assume that F_k is increasing in $\tilde{y}_1,...,\tilde{y}_n,\ell$ and γ. The cost function can be rewritten as:

$$
\tilde{\chi}_k\left(\boldsymbol{p},y_{i,k},\gamma(i),\xi_{c,k}\right) = \sum_{l=1}^{N} p_l\tilde{\mathcal{Y}}_{l,k}\left(\boldsymbol{p},y_k(i),\gamma(i),\xi_{c,k}\right) + \ell_k\left(\boldsymbol{p},y_k(i),\gamma(i),\xi_{c,k}\right),
$$

where ℓ_k and $\tilde{\mathcal{Y}}_{l,k}$ are demand for labor and product l. $\tilde{\chi}_k\left(\boldsymbol{p},y_k(i),\gamma(i),\xi_{c,k}\right)$ is increasing in $\boldsymbol{p}$ and $y_k(i)$, and decreasing in γ. Total demand for good k is given by:

$$Y_k = C_k + \sum_{l=1}^{n} \int_{i \in \mathcal{I}_l} \tilde{\mathcal{Y}}_{k,l}\left(\boldsymbol{p}, y_l\left(i\right), \gamma\left(i\right), \xi_{c,l}\right) di.$$

Conditional on producing, the firm's (variable) profit maximization problem is:

$$\begin{aligned}\Pi_k\left(\gamma\left(i\right)\right) &= \sup_{p_k(i)} p_k\left(i\right) y_k\left(p_k\left(i\right)\right) - \tilde{\chi}_k\left(\boldsymbol{p}, y_k\left(p_k\left(i\right)\right), \gamma\left(i\right), \xi_{c,k}\right), \\ y_k\left(p_k\left(i\right)\right) &= d_k\left(p_k\left(i\right), \left\{p_k\left(j\right)\right\}_{j \in \mathcal{I}_k}\right) Y_k.\end{aligned}$$

The price of variety i only depends on $\gamma\left(i\right)$: $p_k\left(i\right) = p_k\left(\gamma\left(i\right)\right)$. Given that $\chi_k\left(\boldsymbol{p}, y_{i,k}\left(p_{i,k}\right), \gamma\left(i\right), \xi_{c,k}\right)$ is decreasing in γ, $\Pi_k\left(\gamma\left(i\right)\right)$ is non-decreasing in γ. Denote γ_k^* the minimum level at which firms choose to produce, that is $\gamma_k^* = \inf_\gamma \Pi_k\left(\gamma\right) \; s.t. \, \Pi_k\left(\gamma\right) \geq \xi_{p,k}$. Any firm in sector k with $\gamma \geq \gamma_k^*$ produces. Denoting M_k the mass of firms producing in k, and using the fact that the aggregator $\mathcal{F}_k$ is symmetrical, the supply side is summarized by the following equations:

$$\begin{aligned}\Pi_k\left(\gamma\left(i\right)\right) &= \sup_{p_k(i)} p_k\left(i\right) y_k\left(p_k\left(i\right)\right) - \tilde{\chi}_k\left(\boldsymbol{p}, y_k\left(p_k\left(i\right)\right), \gamma\left(i\right), \xi_{c,k}\right), \\ y_{i,k}\left(p_{i,k}\right) &= d_k\left(p_k\left(i\right), \left\{p_k\left(j\right)\right\}_{j \in \mathcal{I}_k}\right) Y_k \\ Y_k &= C_k + \sum_{l=1}^{n} M_l \int_{\gamma \geq \gamma *_l} \tilde{\mathcal{Y}}_{k,l}\left(\boldsymbol{p}, y_{\gamma,l}, \gamma, \xi_{c,l}\right) \Psi_l\left(\gamma\right) d\gamma \\ \Pi_k\left(\gamma_k^*\right) &= \xi_{p,k} \\ \int_{\gamma \geq \gamma *_k} \left(\Pi_k\left(\gamma\right) - \xi_{p,k}\right) \Psi_k\left(\gamma\right) d\gamma &= \xi_{e,k}\end{aligned}$$

The firm's problem defines in each sector three equilibrium objects, M_k,γ_k^*, and $\left\{p_{\gamma,k}\right\}_{\gamma \geq \gamma *_k}$, which only depend on $\left\{C_1, ..., C_n\right\}$, the entry costs, and the exogenous cost shifters. Therefore, the price of the retailer is itself only a function of $\left\{C_1, ..., C_n\right\}$, the entry costs, and the exogenous cost shifters. In addition, since entry is free, producing firms make no profit on average, so total cost is equal to total revenue, $p_k Y_k$. Total labor cost is equal to revenue minus intermediary good cost, $p_k C_k$, which micro-founds our cost function.

Examples of Pricing Function

As our micro-foundations are abstract, we now give some concrete examples of pricing functions in seminal entry models.

Melitz-Chaney. There are no input-output linkages, so $Y_k = C_k$. Retailers have CES preferences over varieties in sector k, $C_k = \left(\int_{i \in \mathcal{I}_k} c_{k,i}^{\frac{\epsilon_k - 1}{\epsilon_k}} di\right)^{\frac{\epsilon_k}{\epsilon_k - 1}}$; the producers' productivity distribution is Pareto: $1 - \Psi_k\left(\gamma\right) = \gamma^{-\gamma_k}$; and the variable production cost is $\chi_k\left(y_{k,i}, \gamma\left(i\right), \xi_{c,k}\right) = \xi_{c,k} y_{k,i} / \gamma\left(i\right)$. We assume $1 + \gamma_k > \epsilon_k > 1$ and $\frac{1+\gamma_k - \epsilon_k}{\epsilon_k - 1} \frac{\xi_{e,k}}{\xi_{p,k}} \geq 1$. Taking first order conditions, we obtain that unit demand for variety is given by:

$$d_{k,i}\left(p_{i,k},\{p_{j,k}\}_{j\in\mathcal{I}_k}\right)=\left(\frac{p_{k,i}}{p_k}\right)^{-\epsilon_k}$$
$$p_k=\left(\int_{i\in\mathcal{I}_k}p_{k,i}^{1-\epsilon_k}di\right)^{\frac{1}{1-\epsilon_k}}$$

The producer's problem then directly gives that variety prices and firm profit (without fixed cost) are given by:

$$p_{k,i}=\text{argmax}\left\{C_k\left(\frac{p_{k,i}}{p_k}\right)^{-\epsilon_k}\left(p_{k,i}-\frac{\xi_{c,k}}{\gamma(i)}\right)\right\}$$
$$\implies p_{k,i}=\frac{\epsilon_k}{\epsilon_k-1}\frac{\xi_{c,k}}{\gamma(i)}$$
$$\Pi_{k,i}=C_k\left(\frac{1}{p_k}\frac{\epsilon_k}{\epsilon_k-1}\frac{\xi_{c,k}}{\gamma(i)}\right)^{-\epsilon_k}\frac{1}{\epsilon_k-1}\frac{\xi_{c,k}}{\gamma(i)}$$

Note that profit are increasing in $\gamma(i)$, so all firm with $\gamma(i)\geq\gamma^*$ produce, where γ^* is the cost solving $C_k\left(\frac{1}{p_k}\frac{\epsilon_k}{\epsilon_k-1}\frac{\xi_{c,k}}{\gamma^*}\right)^{-\epsilon_k}\frac{1}{\epsilon_k-1}\frac{\xi_{c,k}}{\gamma^*}=\xi_{p,k}$. Using this, we can rewrite $\Pi_{k,i}=\xi_{p,k}\left(\frac{\gamma(i)}{\gamma^*}\right)^{\epsilon_k-1}\not\Vdash(\gamma(i)\geq\gamma^*)$. Because of free entry, firms must make zero profit *ex ante* so the entry condition therefore becomes:

$$\int_{\gamma\geq\gamma^*}\xi_{p,k}\left(\left(\frac{\gamma}{\gamma^*}\right)^{\epsilon_k-1}-1\right)\gamma_k\gamma^{-1-\gamma_k}d\gamma=\xi_{e,k}$$
$$\implies\gamma^*=\left(\frac{1+\gamma_k-\epsilon_k}{\epsilon_k-1}\frac{\xi_{e,k}}{\xi_{p,k}}\right)^{-\frac{1}{\gamma_k}}$$

Next, denoting the mass of firm entering the market by M_k, the price index p_k is given by

$$\begin{aligned}p_k&=\left(\int_{i\in\mathcal{I}_k}p_{k,i}^{1-\epsilon_k}di\right)^{\frac{1}{1-\epsilon_k}}\\&=M_k^{\frac{1}{1-\epsilon_k}}\frac{\epsilon_k}{\epsilon_k-1}\xi_{c,k}\left(\int_{\gamma\geq\gamma^*}\gamma^{\epsilon_k-1}\gamma_k\gamma^{-1-\gamma_k}d\gamma\right)^{\frac{1}{1-\epsilon_k}}\\&=M_k^{\frac{1}{1-\epsilon_k}}\frac{\epsilon_k}{\epsilon_k-1}\xi_{c,k}\left(\frac{\gamma_k}{1+\gamma_k-\epsilon_k}(\gamma^*)^{\epsilon_k-\gamma_k-1}\right)^{\frac{1}{1-\epsilon_k}}\\&=M_k^{\frac{1}{1-\epsilon_k}}\frac{\epsilon_k}{\epsilon_k-1}\xi_{c,k}\left(\frac{\gamma_k}{1+\gamma_k-\epsilon_k}\left(\frac{1+\gamma_k-\epsilon_k}{\epsilon_k-1}\frac{\xi_{e,k}}{\xi_{p,k}}\right)^{\frac{1+\gamma_k-\epsilon_k}{\gamma_k}}\right)^{\frac{1}{1-\epsilon_k}},\end{aligned}$$

where the last line uses our solution for γ^*. Finally, using the our first definition of γ^*, we have:

$$C_k \left(\frac{1}{p_k}\frac{\epsilon_k}{\epsilon_k-1}\frac{\xi_{c,k}}{\gamma^*}\right)^{-\epsilon_k}\frac{1}{\epsilon_k-1}\frac{\xi_{c,k}}{\gamma^*}=\xi_{p,k}$$

$$\Leftrightarrow C_k^{\frac{-1}{\epsilon_k}}\left(\frac{\gamma_k}{1+\gamma_k-\epsilon_k}\left(\frac{1+\gamma_k-\epsilon_k}{\epsilon_k-1}\frac{\xi_{e,k}}{\xi_{p,k}}\right)^{\frac{1+\gamma_k-\epsilon_k}{\gamma_k}}\right)^{-\frac{1}{1-\epsilon_k}}\left(\frac{1+\gamma_k-\epsilon_k}{\epsilon_k-1}\frac{\xi_{e,k}}{\xi_{p,k}}\right)^{-\frac{1}{\gamma_k}\frac{1}{\epsilon_k}}\left(\frac{1}{\epsilon_k-1}\frac{\xi_{c,k}}{\xi_{p,k}}\right)^{-\frac{1}{\epsilon_k}}=M_k^{\frac{1}{1-\epsilon_k}}$$

$$\Rightarrow p_k = C_k^{\frac{-1}{\epsilon_k}}\epsilon_k\xi_{p,k}\left(\frac{1+\gamma_k-\epsilon_k}{\epsilon_k-1}\frac{\xi_{e,k}}{\xi_{p,k}}\right)^{-\frac{1}{\gamma_k}\frac{1}{\epsilon_k}}\left(\frac{1}{\epsilon_k-1}\frac{\xi_{c,k}}{\xi_{p,k}}\right)^{\frac{\epsilon_k-1}{\epsilon_k}}$$

The pricing function is therefore given by $\phi_k(C_k,\boldsymbol{\xi}_k) = C_k^{\frac{-1}{\epsilon_k}}\epsilon_k\xi_{p,k}\left(\frac{1+\gamma_k-\epsilon_k}{\epsilon_k-1}\frac{\xi_{e,k}}{\xi_{p,k}}\right)^{-\frac{1}{\gamma_k}\frac{1}{\epsilon_k}}\left(\frac{1}{\epsilon_k-1}\frac{\xi_{c,k}}{\xi_{p,k}}\right)^{\frac{\epsilon_k-1}{\epsilon_k}}$, while the negative of the elasticity with respect to market size is $\alpha_k = 1/\epsilon_k$

HARA aggregator. To showcase how demand can directly impact producers' markups, we consider a case with non-CES preferences. Retailers preferences for varieties are given by a HARA aggregator without love for variety: $\mathcal{F}_k\left(\{c_{k,i}/c_k\}_{i\in\mathcal{I}_k}\right) = M_k^{\frac{-1}{\epsilon_k-1}}\left(\int_{i\in\mathcal{I}_k}(c_{k,i}/c_k+b)^{\frac{\epsilon_k-1}{\epsilon_k}}di\right)^{\frac{\epsilon_k}{\epsilon_k-1}}$, with $b>0$. In addition, assume that Ψ_k is a mass point at $\gamma=1$, that $\xi_{p,k}=0$ and that $\chi_k\left(y_{k,i},\gamma\left(i\right),\xi_{c,k}\right)=\xi_{c,k}y_{k,i}$. From the retailer first order condition, we obtain:

$$\frac{c_{k,i}}{c_k}+b=\left(\frac{p_{k,i}}{p_k}\right)^{-\epsilon_k}\left(\int_{i\in\mathcal{I}_k}\frac{c_{k,i}}{c_k}\left(\frac{c_{k,i}}{c_k}+b\right)^{\frac{-1}{\epsilon_k}}di\right)^{-\epsilon_k}.$$

Using $M_k^{\frac{-1}{\epsilon_k-1}}\left(\int_{i\in\mathcal{I}_k}(c_{k,i}/c_k+b)^{\frac{\epsilon_k-1}{\epsilon_k}}di\right)^{\frac{\epsilon_k}{\epsilon_k-1}}=1$, we obtain:

$$\left(\int_{i\in\mathcal{I}_k}\frac{c_{k,i}}{c_k}(c_{k,i}/c_k+b)^{\frac{-1}{\epsilon_k}}di\right)^{-\epsilon_k}=M_k^{\frac{1}{\epsilon_k-1}}\left(\int_{i\in\mathcal{I}_k}\left(\frac{p_{k,i}}{p_k}\right)^{1-\epsilon_k}di\right)^{-\frac{\epsilon_k}{\epsilon_k-1}}$$

So we have:

$$c_{k,i}=\left(M_k^{\frac{1}{\epsilon_k-1}}\left(\left(\int_{i\in\mathcal{I}_k}\left(\frac{p_{k,i}}{p_k}\right)^{1-\epsilon_k}di\right)^{\frac{1}{1-\epsilon_k}}\right)^{\epsilon_k}\left(\frac{p_{k,i}}{p_k}\right)^{-\epsilon_k}-b\right)c_k,$$

$$d_{k,i}\left(p_{i,k},\{p_{j,k}\}_{j\in\mathcal{I}_k}\right)=\left(M_k^{\frac{1}{\epsilon_k-1}}\left(\left(\int_{i\in\mathcal{I}_k}\left(\frac{p_{k,i}}{p_k}\right)^{1-\epsilon_k}di\right)^{\frac{1}{1-\epsilon_k}}\right)^{\epsilon_k}\left(\frac{p_{k,i}}{p_k}\right)^{-\epsilon_k}-b\right)$$

Multiplying the last expression by $p_{k,i}/p_k$ and integrating we obtain:

$$p_k=M_k^{\frac{1}{\epsilon_k-1}}\left(\int_{i\in\mathcal{I}_k}p_{k,i}^{1-\epsilon_k}di\right)^{\frac{1}{1-\epsilon_k}}-b\int_{i\in\mathcal{I}_k}p_{k,i}di.$$

Next, the producer's FOC, conditional on producing, gives:

$$\left(M_k^{\frac{1}{\epsilon_k-1}}\left(\left(\left(\int_{i\in\mathcal{I}_k}\left(\frac{p_{k,i}}{p_k}\right)^{1-\epsilon_k}di\right)^{\frac{1}{1-\epsilon_k}}\right)^{\epsilon_k}\left(\frac{p_{k,i}}{p_k}\right)^{-\epsilon_k}-b\right)p_{k,i}-\epsilon_k M_k^{\frac{1}{\epsilon_k-1}}\left(\left(\left(\int_{i\in\mathcal{I}_k}\left(\frac{p_{k,i}}{p_k}\right)^{1-\epsilon_k}di\right)^{\frac{1}{1-\epsilon_k}}\right)^{\epsilon_k}\left(\frac{p_{k,i}}{p_k}\right)^{-\epsilon_k}\left(p_{k,i}-\xi_{c,k}\right)=0.$$

Denoting p_k^* the common price chosen by firms, we have $p_k=(1-bM_k)\,p_k^*$ so we can rewrite the last expression as:

$$\begin{aligned}(1-bM_k)\,p_k^*-\epsilon_k\left(p_k^*-\xi_{c,k}\right)=0,&\\ \Rightarrow p_k^*=&\frac{\epsilon_k}{\epsilon_k-1+bM_k}\xi_{c,k}\\ \Rightarrow p_k=&\frac{\epsilon_k\left(1-bM_k\right)}{\epsilon_k-1+bM_k}\xi_{c,k}\end{aligned}$$

Finally using the free entry condition, we obtain:

$$C_k\frac{(1-bM_k)^2}{\epsilon_k-(1-bM_k)}=-\frac{1}{b}\frac{\xi_{e,k}}{\xi_{c,k}}\left((1-bM_k)-1\right)$$

Solving the quadratic equation in $1-bM_k$ yields:

$$1-bM_k=\frac{\epsilon_k}{bC_k\frac{\xi_{c,k}}{\xi_{e,k}}-1}\frac{-1+\sqrt{1+4\frac{1}{\epsilon_k}\left(bC_k\frac{\xi_{c,k}}{\xi_{e,k}}-1\right)}}{2}$$

Plugging this expression in p_k, we obtain:

$$p_k=\frac{\epsilon_k\left(\sqrt{1+4\frac{1}{\epsilon_k}\left(bC_k\frac{\xi_{c,k}}{\xi_{e,k}}-1\right)}-1\right)}{2\left(bC_k\frac{\xi_{c,k}}{\xi_{e,k}}-1\right)-\left(\sqrt{1+4\frac{1}{\epsilon_k}\left(bC_k\frac{\xi_{c,k}}{\xi_{e,k}}-1\right)}-1\right)}\xi_{c,k},$$

and direct algebra shows that the negative of elasticity of price with respect to demand is:

$$\alpha_k=\frac{E_kb}{\xi_{e,k}}/\left(\epsilon_k-1+2\frac{E_kb}{\xi_{e,k}}\right),$$

with $E_k=p_kC_k$. As demand increases, more firms enter; with HARA preferences, demand for varieties becomes more price elastic with more firms, so markups decrease and the aggregate price index decreases.

Translog aggregator. We finally consider a standard case where retailers have translog preferences. As is standard, we express preferences directly in terms of the expenditure function. Denoting $\tilde{M}_k$ the mass of potential entrants, we have:

$$ln\left(e_k\right) = ln(c_k) + \alpha_o + \frac{\tilde{M}_k - M_k}{2\gamma\tilde{M}_k M_k} + \int_{i\in\mathcal{I}_k}\frac{1}{M_k}ln(p_{i,k})di + \frac{\beta_k}{2M_k}\int_{i\in\mathcal{I}_k}\int_{j\in\mathcal{I}_k}ln(p_{i,k})(ln(p_{j,k}) - ln(p_{i,k}))didj$$

$$ln\left(p_k\right) = \alpha_o + \frac{\tilde{M}_k - M_k}{2\gamma\tilde{M}_k M_k} + \int_{i\in\mathcal{I}_k}\frac{1}{M_k}ln(p_{i,k})di + \frac{\beta_k}{2M_k}\int_{i\in\mathcal{I}_k}\int_{j\in\mathcal{I}_k}ln(p_{i,k})(ln(p_{j,k}) - ln(p_{i,k}))didj$$

Unit demand is given by

$$d_{k,i}\left(p_{i,k}, \left\{p_{j,k}\right\}_{j\in\mathcal{I}_k}\right) = \frac{p_k}{M_k p_{i,k}}\left(1 + \beta_k\int_{j\in\mathcal{I}_k}ln(p_{j,k}))dj - \beta_k M_k ln(p_{i,k}))\right).$$

We assume that Ψ_k is a mass point at $\gamma = 1$, that $\xi_{p,k} = 0$ and that $\chi_k\left(y_{k,i}, \gamma\left(i\right), \xi_{c,k}\right) = \xi_{c,k}y_{k,i}$, i.e. we consider identical firms with a linear production function. The producer first order condition gives:

$$-\gamma_k\frac{p_k}{p_{i,k}}\left(1 - \frac{\xi_{k,c}}{p_{i,k}}\right) + \frac{\xi_{k,c}}{p_{i,k}}\frac{p_k}{M_k p_{i,k}}\left(1 + \beta_k\int_{j\in\mathcal{I}_k}ln(p_{j,k}))dj - \beta_k M_k ln(p_{i,k}))\right) = 0.$$

So the common price set by producers is

$$p_k^* = \xi_{k,c}\left(1 + \frac{1}{\beta_k M_k}\right).$$

We therefore obtain that the price index is given by:

$$ln\left(p_k\right) = ln(\xi_{k,c}) + \kappa_0 + \frac{1}{2\beta_k M_k} + ln\left(1 + \frac{1}{\beta_k M_k}\right),$$

with $\kappa_0 = \alpha_o - \frac{1}{2\gamma\tilde{M}_k}$, while $M_k$is determined implicitly by the free entry condition:

$$\frac{p_k}{p_{i,k}}\left(p_{i,k} - \xi_{k,c}\right)\frac{C_k}{M_k} = \xi_{k,e}$$

$$\kappa_0 + \frac{1}{2\beta_k M_k} + ln\left(\beta_k\right) - 2ln\left(M_k\right) = ln\left(\frac{\xi_{k,e}}{\xi_{k,c}}\right) - ln\left(C_k\right).$$

We therefore obtain:

$$\alpha_k = \left(\frac{1}{2\beta_k M_k} + \frac{1}{\beta_k M_k + 1}\right)\frac{dln\left(M_k\right)}{dln\left(C_k\right)},$$

with

$$\frac{dln\left(M_k\right)}{dln\left(C_k\right)} = \left(\frac{1}{2\beta_k M_k} + 2\right)^{-1}.$$

As with HARA preferences, an increase in demand leads to more entry by producers and lower markups.

E.2 First Order Approach in the General Model

We provide a version of Proposition 1 in our extended model. Relative to the main text, the formula for the optimal income tax remains unchanged; however, commodity taxation can play a nontrivial role in the monopolistic case. Under perfect competition or when mark-ups are uniform across sectors, we recover a variant of the Atkinson-Stiglitz result: optimal tax policy relies solely on the income tax, with no role for commodity taxes. In contrast, when pricing frictions (captured by the mark-ups $\phi_k / \sum_{l=1}^n \partial_{C_k}\chi_l - 1$) differ across sectors, commodity taxes serve to correct *relative* mark-ups. Specifically, they ensure that consumers internalize the social benefit of consumption across markets by adjusting for distortions in relative prices.

To fix ideas, suppose that ϕ_k and χ_k only depends on C_k. Then, an increase in demand for good k induces a price change determined by $-A_{kk}$. In Proposition E1, we show that if this price falls by more than α (the average market size elasticity), then good k is subsidized. Conversely, if the price response is smaller than α, the good is taxed. Intuitively, consumption taxes reallocate demand toward sectors with inefficiently high mark-ups, where demand is otherwise too low. This reallocation helps restore efficiency by offsetting sector-specific pricing distortions.

Proposition E1. *The optimal commodity taxes t_k, with $q_k = (1+t_k)p_k$, are null in the competitive case. In the monopolistic case, they are given by:*

$$1 + t_i = \frac{1 - \sum_j A_{j,i} p_j C_j / p_i C_i}{1-\alpha}, \tag{A56}$$

with $\alpha \equiv \frac{\sum_i (\sum_j A_{i,j}) p_i C_i}{\sum_i p_i C_i}$. The optimal non-linear income tax schedule is characterized by:

$$\frac{T'}{1-T'} = -t_w + \frac{1-t_w}{z\tilde{\zeta} f(z)} \left\{ \mathbb{E}_{z'>z} \left(1-g\right) - \frac{1}{1-t_w} \mathbb{E}_{z'>z} \left(\left(t_w + \frac{T'}{1-T'} \right) \tilde{\eta} \right) \right\}. \tag{A57}$$

with $g = G' U_v v_{z^*} / ((1-\alpha)\lambda)$, $t_w = \alpha$ in the monopolistic case, $t_w = 0$ in the competitive case.

Proof of Proposition E1. After integration by parts of the planning problem, the corresponding Lagrangian is:

$$\begin{aligned} \mathcal{L} = &\int G(V(\theta), \theta) \pi(\theta) d\theta - \int (\mu'(\theta) V(\theta) + \mu(\theta) U_\theta(v(\theta), z(\theta), \theta)) d\theta \\ &- \lambda \left(\int z^*(\theta) - z(\theta) \pi(\theta) d\theta - \sum_{k=1}^{n} q_k C_k - \chi_k \left(C_1, ..., C_n, \xi_k\right) \right) \end{aligned}$$

where $\mu(\theta)$ are the multipliers on the incentive constraints and λ is the multiplier on the resource constraint.

We start with the FOC with respect to consumer prices q_i. Recall that $c^h(q, v)$ is the Hicksian demand function at prices q for a given sub-utility v; we have:

$$\begin{aligned} \left.\frac{dc_j}{dq_i}\right|_{z,V} &= \left.\frac{dc_j}{dq_i}\right|_{v} = \frac{\partial c_j^h}{\partial q_i} \\ \left.\frac{dz^*}{dq_i}\right|_{z,V} &= \left.\frac{dz^*}{dq_i}\right|_{v} = c_i \end{aligned}$$

We therefore have, denoting $\partial_{q_i} C_j^h = \int \partial_{q_i} c_j^h \pi(\theta) d\theta$:

$$
\begin{aligned}
\frac{d\mathcal{L}}{dq_i} &= \lambda\left(C_i + \sum_j \left(q_j - \sum_k \partial_{C_j}\chi_k\right)\partial_{q_i}C_j^h - C_i\right)\\
\Rightarrow 0 &= \sum_j \left(q_j C_j - \sum_k C_j \partial_{C_j}\chi_k\right)\mathcal{S}_{j,i},
\end{aligned}
$$

where A and $\mathcal{S}$ are as in definition E1 and E2. Given that $\mathcal{S}$ is generically of rank $N-1$ with left kernel $\boldsymbol{qC}$, we have:

$$
(1-\beta)q_k = \sum_k \partial_{C_j}\chi_k,
$$

where $1-\beta > 0$ is an arbitrary scaling constant. Therefore, since $p_k = \sum_k \partial_{C_j}\chi_k$ in the competitive case, no commodity tax is optimal. In the monopolistic case, we have $\sum_k \partial_{C_j}\chi_k = p_j - \sum_k p_k C_k A_{k,j}/C_j$. Denoting α the scaling such that $q\cdot C - p\cdot C = 0$ gives:

$$
\begin{aligned}
\boldsymbol{q}\cdot\boldsymbol{C} - \boldsymbol{p}\cdot\boldsymbol{C} &= \frac{1}{1-\beta}\left(\beta\sum_i p_i C_i - \sum_i p_i C_i \sum_j A_{i,j}\right)\\
\Rightarrow \quad \alpha &= \frac{\sum_i(\sum_j A_{i,j})p_i C_i}{\sum_i p_i C_i}.
\end{aligned}
$$

With this scaling, the ad valorem commodity taxes in the monopolistic are:

$$
1+t_i = \frac{1-\sum_j A_{j,i}p_j C_j/p_i C_i}{1-\alpha}.
$$

Next, we derive the FOC associated with V. $V(\theta)$ impacts consumption and producer prices through $z^*(\theta)$ with $dz^*(\theta)/dV(\theta) = (U_v v_{z^*})^{-1}$. Denoting $t_w = \alpha$ in the monopolistic case, $t_w = 0$ in the competitive case, and using our result on optimal commodity taxes, we have:

$$
\begin{aligned}
0 &= G'(V(\theta),\theta)\pi(\theta) - \mu'(\theta) - \mu\frac{U_{\theta,v}}{U_v} - \frac{\lambda\pi(\theta)}{U_v v_{z^*}}\left[1 - \sum_i\left(q_i - \sum_k \partial_{C_i}\chi_k\right)\partial_{z^*}c_i(\theta)\right]\\
&= G'(V(\theta),\theta)\pi(\theta) - \mu'(\theta) - \mu\frac{U_{\theta,v}}{U_v} - \frac{\lambda\pi(\theta)}{U_v v_{z^*}}\left[1 - \sum_i\left(q_i - (1-t_w)(1+t_i)\,p_i\right)\partial_{z^*}c_i(\theta)\right]\\
&= G'(V(\theta),\theta)\pi(\theta) - \mu'(\theta) - \mu\frac{U_{\theta,v}}{U_v} - \frac{\lambda\pi(\theta)}{U_v v_{z^*}}\left[1 - t_w\sum_i q_i\,\partial_{z^*}c_i(\theta)\right]\\
\Rightarrow &\mu'(\theta)\frac{U_v v_{z^*}}{\lambda} + \mu\frac{U_{\theta,v}v_{z^*}}{\lambda} = -\left(1 - t_w - \frac{G'(V(\theta),\theta)U_v v_{z^*}}{\lambda}\right)\pi(\theta).
\end{aligned}
$$

Finally, defining $\tilde{\mu} = \mu\, U_v v_{z^*}/\lambda$, we have:

$$\tilde{\mu}'(\theta) + \tilde{\mu}\,\partial_{z^*} MRS\, z'(\theta) = -\left(1 - t_w - \frac{G'(V(\theta),\theta)U_v v_{z^*}}{\lambda}\right)\pi(\theta),$$

with $MRS = -U_z/U_v v_{z^*}$ the marginal rate of substitution.

Finally, the FOC associated with z, using the same steps as above to derive the response of consumption and prices, is:

$$\begin{aligned} 0 &= \mu(-U_{\theta,z} - U_{\theta,z^*} MRS) - \lambda\pi(\theta)\,(MRS - 1 - t_w MRS) \\ \Rightarrow \quad & \tilde{\mu}\,\partial_\theta MRS = \pi(\theta)((1-t_w)MRS - 1) \end{aligned}$$

Since $MRS = 1 - T'(z(\theta))$, and $z\tilde{\zeta}\partial_\theta MRS = -z'(\theta)(1 - T'(z(\theta)))$, where $\tilde{\zeta}$ is defined in Definition E.1, we therefore have, denoting $f(z(\theta)) = \pi(\theta)/z'(\theta)$

$$\tilde{\mu}(\theta) = f(z) z\tilde{\zeta}\left(\frac{T'}{1-T'} + t_w\right)$$

Finally, using $-z\tilde{\zeta}\,\partial_{z^*} MRS = \tilde{\eta}$ we get:

$$f(z)z\tilde{\zeta}\left(\frac{T'}{1-T'} + t_w\right) + \int_{z(\theta)}^{z(\bar{\theta})} \tilde{\eta}\left(\frac{T'}{1-T'} + t_w\right) f(z)dz = \int_{z(\theta)}^{z(\bar{\theta})}\left(1 - t_w - \frac{G'U_v v_{z^*}}{\lambda}\right) f(z)dz$$

Using $g = G'U_v v_{z^*}/((1-\alpha)\lambda)$, we obtain the formula of Proposition A1.$\square$

E.3 Comparative Statics Results

In this section, we we extend the results of Section 4. We begin by deriving a general comparative statics formula for the change in the optimal tax rate in response to exogenous supply shocks (i.e. changes in the parameters ξ). This result generalizes lemma A1 from Appendix A. To obtain streamlined formulas, we assume that utility is additively separable between consumption and labor.

Assumption E3. *Additive Separability. U is additively separable and takes the form $U\left(u\left(c_1,...,c_n\right), z, \theta\right) = u\left(c_1,...,c_n\right) - \psi\left(z/\theta\right)$, with ψ and u increasing in their arguments and respectively convex and concave. We denote $\epsilon\left(z/\theta\right) = \psi'/\left(z/\theta\psi''\right)$ and assume $|\left(z/\theta\right)\epsilon'\left(z/\theta\right)/\epsilon| \leq 1$.*

Under this assumption, we derive the equation determining the change in taxes in response to an exogenous supply shift $d\xi$. As for lemma A1, the system is more easily expressed in terms of $dV/d\xi = -v'\left(dT/d\xi + \sum_{i=1}^N e_i dlnq_i/d\xi\right)$. We will use this system to extend the results of the main text with general production functions and households preferences

Proposition E2. *Under assumption E3, the change in the income tax schedule, expressed in terms of $dV/d\xi = -v'\left(dT/d\xi + \sum_{i=1}^N e_i dlnq_i/d\xi\right)$, in response to an exogenous supply shift $d\xi$, conditional on* *the change consumer prices $dq_i/d\xi$ and market size elasticity $dt_w/d\xi$ is given by:*

$$\frac{\theta\epsilon}{(1+\epsilon)^2}\tilde{\pi}\kappa(\theta)\frac{1}{(1-T')^2}\frac{\theta}{z}\frac{1}{v'}\frac{d}{d\theta}\left\{\frac{dV}{d\xi}\right\}+(1-t_w)\int_{\theta}^{\bar{\theta}}\tilde{\pi}g\left(\gamma(\theta)\frac{dV}{d\xi}-\int_{\underline{\theta}}^{\theta}\tilde{\pi}g\gamma(\theta)\frac{dV}{d\xi}d\theta\right)=-\frac{\theta\epsilon}{1+\epsilon}\frac{\tilde{\pi}(\theta)}{1-T'}\sum_{i=1}^{n}\left(\tau_i(\theta)+\partial_{z^*}\tilde{E}_i\right)\frac{1}{q_i}\frac{dq_i}{d\xi}$$
$$+\frac{\theta\epsilon}{1+\epsilon}\frac{\tilde{\pi}(\theta)}{1-T'}\frac{1}{1-t_w}\frac{dt_w}{d\xi}$$
$$(1-t_w)\int_{\underline{\theta}}^{\bar{\theta}}g\frac{1}{v'}\frac{dV}{d\xi}\pi dz=-\sum_{i=1}^{n}\frac{\partial\chi_i}{\partial\xi}$$

where $\tilde{\pi}=\frac{1}{v'}\pi/\int_{\underline{\theta}}^{\theta}\frac{1}{v'}\pi d\theta$, $\partial_{z^*}\tilde{E}_i=\int_{\underline{\theta}}^{\theta}\partial_{z^*}e_i\tilde{\pi}d\theta$, $\kappa(\theta)=1-(1-t_w)(1-T')^2\epsilon z\frac{v''}{v'}-\frac{(z/\theta)\epsilon'}{1+\epsilon}(1-(1-T')(1-t_w))$, $\gamma(\theta)=-\frac{G''}{G'}-g^{-1}\left(\frac{v''}{(v')^2}-\frac{1}{\tilde{\pi}}\frac{d}{d\theta}\left\{\frac{\theta\epsilon}{1+\epsilon}\tilde{\pi}(\theta)\frac{v''}{(v')^2}\right\}\right)$, *and*

$$\tau_l(\theta)=(1-t_w)\left(1-T'\right)\left(\frac{1+\epsilon}{\epsilon}\frac{1}{\theta\tilde{\pi}\theta)}\int_{\theta}^{\bar{\theta}}\left(\partial_{z^*}e_l-\partial_{z^*}E_l\right)\tilde{\pi}d\theta'+\left(\partial_{z^*}e_l-\partial_{z^*}E_l\right)\right).$$

Proof of Proposition E2. Recall that $\mu(\theta)=\tilde{\mu}(\theta)/v'(z(\theta)-T(z(\theta)),q)$ is the co-state on the local incentive constraint. With additive separability in consumption and labor, the income tax schedule is determined by the following system of equations in μ and V, with unknown λ:

$$\begin{aligned}
\mu'(\theta)&=-(1-t_w)\left(\frac{1}{v'(z^*(\theta),q)}-\frac{G'(V(\theta),\theta)}{\lambda}\right)\pi(\theta),\\
\mu(\theta)&=\frac{1}{v'(z^*(\theta),q)}\frac{\theta\epsilon(z/\theta)}{1+\epsilon(z/\theta)}\pi(\theta)\left(\frac{T'(z(\theta))}{1-T'(z(\theta))}+t_w\right),\\
V'(\theta)&=z(\theta)/\theta^2\psi'(z(\theta)/\theta),\quad V(\theta)=v(z^*(\theta),q)-\psi(z(\theta)/\theta),\\
0&=\int(z(\theta)-z^*(\theta))\pi(\theta)+\sum_{i=1}^{N}(q_iC_i-\chi_i),
\end{aligned}\tag{A58}$$

with $\mu(\bar{\theta})=\mu(\underline{\theta})=0$, $\epsilon(z/\theta)=\psi'(z/\theta)/(z/\theta\psi''(z/\theta))$ and $v'(z^*(\theta),q)=\partial_{z^*}v(z^*(\theta),q)$. Using Roy's identity and differentiating the IC-FOC condition, we have:

$$\begin{aligned}
\frac{\partial V(\theta)}{\partial ln(q_i)}&=\frac{\partial v(z^*(\theta),q)}{\partial ln(q_i)}=-v'(z(\theta)-T(z(\theta)),q)e_i(z^*(\theta),q),\\
\frac{dV}{d\xi}&=v'(z(\theta)-T(z(\theta)),q)\left(\frac{dz^*(\theta)}{d\xi}-\sum_{i=1}^{N}e_i(z^*(\theta),q)\frac{dq_i}{d\xi}-\left(1-T'\right)\frac{dz(\theta)}{d\xi}\right),\\
\frac{d}{d\theta}\left\{\frac{dV}{d\xi}\right\}&=1/\theta^2\psi'(z(\theta)/\theta)\left(1+\frac{1}{\epsilon(z/\theta)}\right)\frac{dz(\theta)}{d\xi}=v'\left(1-T'\right)\frac{1}{\theta}\left(1+\frac{1}{\epsilon(z/\theta)}\right)\frac{dz(\theta)}{d\xi}
\end{aligned}$$

Using Young's identity on the first of the two equations above, we have:

$$\begin{aligned}
\frac{\partial^2 v(z^*(\theta),q)}{\partial z^*\partial ln(q_i)}=\frac{\partial v'(z^*(\theta),q)}{\partial ln(q_i)}&=-\frac{\partial}{\partial z^*}\left\{v'(z^*(\theta),q)e_i(z^*(\theta),q)\right\}\\
&=-\left(v''(z^*(\theta),q)e_i(z^*(\theta),q)+v'(z^*(\theta),q)\partial_{z^*}e_i(z^*(\theta),q)\right).
\end{aligned}$$

We use these four equations to differentiate the first two equations of system A58. Differentiating the first equation of the system with respect to ξ, we obtain (omitting the arguments of the function for clarity):

$$\frac{d}{d\theta}\left\{\frac{d\mu}{d\xi}\right\}=-\frac{d\mu}{d\theta}\frac{1}{1-t_w}\frac{dt_w}{d\xi}-\left(1-t_w\right)\frac{1}{v'}\pi\sum_{i=1}^{N}\partial_{z^*}e_i\frac{1}{q_i}\frac{dq_i}{d\xi}$$
$$-\left(1-t_w\right)\left(-\frac{1}{v'}\frac{v''}{\left(v'\right)^2}\left(\frac{dV}{d\xi}+v'\left(1-T'\right)\frac{dz\left(\theta\right)}{d\xi}\right)-\frac{G'}{\lambda}\frac{G''}{G'}\frac{dV}{d\xi}+\frac{G'}{\lambda}\frac{1}{\lambda}\frac{d\lambda}{d\xi}\right)\pi.$$

Since $\mu\left(\bar{\theta}\right)=\mu\left(\underline{\theta}\right)=0$ for all ξ, we have $d\mu/d\xi\left(\bar{\theta}\right)=d\mu/d\xi\left(\underline{\theta}\right)=0$, so integrating between $\underline{\theta}$ and $\bar{\theta}$ yields:

$$\begin{aligned}\frac{1}{\lambda}\frac{d\lambda}{d\xi}&=-\sum_{i=1}^{N}\int_{\underline{\theta}}^{\bar{\theta}}\frac{\frac{1}{v'}\pi}{\int_{\underline{\theta}}^{\theta}\frac{1}{v'}\pi d\theta}\partial_{z^*}e_i d\theta\frac{1}{q_i}\frac{dq_i}{d\xi}+\int_{\underline{\theta}}^{\bar{\theta}}\frac{\frac{1}{v'}\pi}{\int_{\underline{\theta}}^{\theta}\frac{1}{v'}\pi d\theta}\frac{v''}{v'}\left(\frac{dz^*\left(\theta\right)}{d\xi}-\sum_{i=1}^{N}e_i\frac{1}{q_i}\frac{dq_i}{d\xi}\right)d\theta+\int_{\underline{\theta}}^{\bar{\theta}}\frac{G'v'\frac{1}{v'}\pi}{\int_{\underline{\theta}}^{\bar{\theta}}\frac{1}{v'}\pi d\theta}\frac{G''}{G'}\frac{dV}{d\xi}d\theta\\&=-\sum_{i=1}^{N}\int_{\underline{\theta}}^{\bar{\theta}}\frac{\frac{1}{v'}\pi}{\int_{\underline{\theta}}^{\theta}\frac{1}{v'}\pi d\theta}\partial_{z^*}e_i d\theta\frac{1}{q_i}\frac{dq_i}{d\xi}+\int_{\underline{\theta}}^{\bar{\theta}}\frac{\frac{1}{v'}\pi}{\int_{\underline{\theta}}^{\theta}\frac{1}{v'}\pi d\theta}\frac{v''}{\left(v'\right)^2}\left(\frac{dV}{d\xi}+v'\left(1-T'\right)\frac{dz\left(\theta\right)}{d\xi}\right)d\theta+\int_{\underline{\theta}}^{\bar{\theta}}\frac{G'v'\frac{1}{v'}\pi}{\int_{\underline{\theta}}^{\bar{\theta}}\frac{1}{v'}\pi d\theta}\frac{G''}{G'}\frac{dV}{d\xi}d\theta\end{aligned}$$

Defining $\tilde{\pi}=\frac{1}{v'}\pi/\int_{\underline{\theta}}^{\theta}\frac{1}{v'}\pi d\theta$, which corrects the distribution of types to take into account the income effects, we can rewrite the change in the value of public funds as:

$$\frac{1}{\lambda}\frac{d\lambda}{d\xi}=-\sum_{i=1}^{N}\int_{\underline{\theta}}^{\bar{\theta}}\partial_{z^*}e_i\tilde{\pi}d\theta\frac{dq_i}{d\xi}+\int_{\underline{\theta}}^{\bar{\theta}}\tilde{\pi}\frac{v''}{\left(v'\right)^2}\left(\frac{dV}{d\xi}+v'\left(1-T'\right)\frac{dz\left(\theta\right)}{d\xi}\right)\tilde{\pi}d\theta+\int_{\underline{\theta}}^{\theta}\frac{G'v'\tilde{\pi}}{\int_{\underline{\theta}}^{\theta}G'v'\tilde{\pi}d\theta}\frac{G''}{G'}\frac{dV}{d\xi}d\theta$$

Differentiating the system with respect to ξ, for the second equation we obtain (omitting the arguments of the function for clarity):

$$\begin{aligned}\frac{d\mu}{d\xi}=&\frac{\theta\epsilon}{1+\epsilon}\frac{1}{v'}\pi(\theta)\frac{dt_w}{d\xi}-\left(1-t_w\right)\frac{\theta\epsilon}{1+\epsilon}\frac{1}{v'}\pi(\theta)\sum_{i=1}^{N}\partial_{z^*}e_i\frac{1}{q_i}\frac{dq_i}{d\xi}+\left(1-t_w\right)\frac{\theta\epsilon}{1+\epsilon}\frac{1}{v'}\pi(\theta)\frac{v''}{\left(v'\right)^2}\left(\frac{dV}{d\xi}+v'\left(1-T'\right)\frac{dz\left(\theta\right)}{d\xi}\right)\\&+\frac{\epsilon'}{\left(1+\epsilon\right)^2}\frac{1}{v'}\pi(\theta)\left(\frac{T'}{1-T'}+t_w\right)\frac{dz}{d\xi}-\frac{1}{1+\epsilon}\frac{1}{v'}\pi(\theta)\frac{1}{1-T'}\frac{\theta}{z}\frac{dz}{d\xi}\end{aligned}$$

We therefore obtain the following equation in terms of $\frac{dV}{d\xi}$:

$$\begin{aligned}&\frac{d}{d\theta}\left\{\frac{\theta\epsilon}{1+\epsilon}\tilde{\pi}\left(\frac{\left(1-t_w\right)v''}{\left(v'\right)^2}\frac{dV}{d\xi}+\left(\left(1-t_w\right)\left(1-T'\right)\frac{v''}{v'}+\frac{\epsilon'/\left(\theta\epsilon\right)}{1+\epsilon}\left(\frac{T'}{1-T'}+t_w\right)-\frac{1}{1-T'}\frac{1}{\epsilon z}\right)\frac{dz}{d\xi}\right)\right\}\\&\qquad=-\frac{d}{d\theta}\left\{\frac{1}{1-T'}\frac{\theta\epsilon}{1+\epsilon}\tilde{\pi}\right\}\frac{1}{1-t_w}\frac{dt_w}{d\xi}\\&-\left(1-t_w\right)\sum_{i=1}^{n}\tilde{\pi}\left(\partial_{z^*}e_i-\int_{\underline{\theta}}^{\theta}\partial_{z^*}e_i\tilde{\pi}d\theta\right)\frac{1}{q_i}\frac{dq_i}{d\xi}-\frac{d}{d\theta}\left\{\frac{\theta\epsilon}{1+\epsilon}\tilde{\pi}\sum_{i=1}^{n}\left(\left(1-t_w\right)\partial_{z^*}e_i-\left(\left(1-t_w\right)-\frac{1}{1-T'}\right)\int_{\underline{\theta}}^{\bar{\theta}}\partial_{z^*}e_i\tilde{\pi}d\theta\right)\frac{1}{q_i}\frac{dq_i}{d\xi}\right\}\\&\qquad+\left(1-t_w\right)\tilde{\pi}\left(\frac{v''}{\left(v'\right)^2}\left(\frac{dV}{d\xi}+v'\left(1-T'\right)\frac{dz\left(\theta\right)}{d\xi}\right)-g\int_{\underline{\theta}}^{\theta}\tilde{\pi}\frac{v''}{\left(v'\right)^2}\left(\frac{dV}{d\xi}+v'\left(1-T'\right)\frac{dz\left(\theta\right)}{d\xi}\right)d\theta\right)\\&\qquad+\left(1-t_w\right)g\tilde{\pi}\left(\frac{G''}{G'}\frac{dV}{d\xi}-\int_{\underline{\theta}}^{\bar{\theta}}\frac{G''}{G'}\frac{dV}{d\xi}g\tilde{\pi}d\theta\right).\qquad\text{(A59)}\end{aligned}$$

Next, we simplify the formula using $v'\left(1-T'\right)\frac{dz(\theta)}{d\xi}=\frac{\theta\epsilon}{1+\epsilon}\frac{d}{d\theta}\left\{\frac{dV}{d\xi}\right\}$. Focusing on the first term in the

RHS of equation A59, we can re-express it as:

$$\frac{d}{d\theta}\left\{\frac{\theta\epsilon}{1+\epsilon}\tilde{\pi}(\theta)\frac{v''}{(v')^2}\frac{dV}{d\xi}\right\}=\frac{\theta\epsilon}{1+\epsilon}\tilde{\pi}(\theta)\frac{v''}{(v')^2}\frac{d}{d\theta}\left\{\frac{dV}{d\xi}\right\}+\frac{d}{d\theta}\left\{\frac{\theta\epsilon}{1+\epsilon}\tilde{\pi}(\theta)\frac{v''}{(v')^2}\right\}\frac{dV}{d\xi}$$
$$=\tilde{\pi}(\theta)\frac{v''}{(v')^2}v'\left(1-T'\right)\frac{dz\left(\theta\right)}{d\xi}+\frac{d}{d\theta}\left\{\frac{\theta\epsilon}{1+\epsilon}\tilde{\pi}(\theta)\frac{v''}{(v')^2}\right\}\frac{dV}{d\xi}.$$

Substituting in equation A59 yields:

$$\frac{d}{d\theta}\left\{\left(\frac{\theta\epsilon}{1+\epsilon}\right)^2\tilde{\pi}\left(\left(1-t_w\right)\left(1-T'\right)\frac{v''}{v'}+\frac{\epsilon'/\left(\theta\epsilon\right)}{1+\epsilon}\left(\frac{T'}{1-T'}+t_w\right)-\frac{1}{1-T'}\frac{1}{\epsilon z}\right)\frac{1}{1-T'}\frac{1}{v'}\frac{d}{d\theta}\left\{\frac{dV}{d\xi}\right\}\right\}$$
$$=-\frac{d}{d\theta}\left\{\frac{1}{1-T'}\frac{\theta\epsilon}{1+\epsilon}\tilde{\pi}\right\}\frac{1}{1-t_w}\frac{dt_w}{d\xi}$$
$$-\left(1-t_w\right)\sum_{i=1}^{n}\left(\tilde{\pi}\left(\partial_{z^*}e_i-\int_{\underline{\theta}}^{\theta}\partial_{z^*}e_i\tilde{\pi}d\theta\right)-\frac{d}{d\theta}\left\{\frac{\theta\epsilon}{1+\epsilon}\tilde{\pi}\left(\left(1-t_w\right)\partial_{z^*}e_i-\left(\left(1-t_w\right)-\frac{1}{1-T'}\right)\int_{\underline{\theta}}^{\bar{\theta}}\partial_{z^*}e_i\tilde{\pi}d\theta\right)\right\}\right)\frac{1}{q_i}\frac{dq_i}{d\xi}$$
$$+\left(1-t_w\right)\tilde{\pi}\left(\left(\frac{v''}{(v')^2}-\frac{d}{d\theta}\left\{\frac{\theta\epsilon}{1+\epsilon}\tilde{\pi}(\theta)\frac{v''}{(v')^2}\right\}\right)\frac{dV}{d\xi}-g\int_{\underline{\theta}}^{\theta}\tilde{\pi}\left(\frac{v''}{(v')^2}-\frac{d}{d\theta}\left\{\frac{\theta\epsilon}{1+\epsilon}\tilde{\pi}(\theta)\frac{v''}{(v')^2}\right\}\right)\frac{dV}{d\xi}d\theta\right)$$
$$+\left(1-t_w\right)g\tilde{\pi}\left(\frac{G''}{G'}\frac{dV}{d\xi}-\int_{\underline{\theta}}^{\bar{\theta}}\frac{G''}{G'}\frac{dV}{d\xi}g\tilde{\pi}d\theta\right).$$

Integrating the formula above, we obtain:

$$\left(\frac{\theta\epsilon}{1+\epsilon}\right)^2\tilde{\pi}\kappa\left(\theta\right)\frac{1}{1-T'}\frac{1}{v'}\frac{d}{d\theta}\left\{\frac{dV}{d\xi}\right\}$$
$$=\frac{1}{1-T'}\frac{\theta\epsilon}{1+\epsilon}\tilde{\pi}\frac{1}{1-t_w}\frac{dt_w}{d\xi}$$
$$-\left(1-t_w\right)\sum_{i=1}^{n}\left(\int_{\theta}^{\bar{\theta}}\tilde{\pi}\left(\partial_{z^*}e_i-\partial_{z^*}\tilde{E}_i\right)d\theta+\left\{\frac{\theta\epsilon}{1+\epsilon}\tilde{\pi}\left(\left(1-t_w\right)\left(\partial_{z^*}e_i-\partial_{z^*}\tilde{E}_i\right)+\frac{1}{1-T'}\partial_{z^*}\tilde{E}_i\right)\right\}\right)\frac{1}{q_i}\frac{dq_i}{d\xi}$$
$$-\left(1-t_w\right)\int_{\theta}^{\bar{\theta}}\tilde{\pi}g\left(\gamma\left(\theta\right)\frac{dV}{d\xi}-\int_{\underline{\theta}}^{\theta}\tilde{\pi}g\gamma\left(\theta\right)\frac{dV}{d\xi}d\theta\right),$$

where $\partial_{z^*}\tilde{E}_i=\int_{\underline{\theta}}^{\theta}\partial_{z^*}e_i\tilde{\pi}d\theta$, $\kappa\left(\theta\right)=\frac{1}{\epsilon}\frac{1}{1-T'}\frac{1}{z}-\left(1-t_w\right)\left(1-T'\right)\frac{v''}{v'}-\frac{\epsilon'/(\theta\epsilon)}{1+\epsilon}\left(\frac{T'}{1-T'}+t_w\right)$ and $\gamma\left(\theta\right)=-\frac{G''}{G'}-g^{-1}\left(\frac{v''}{(v')^2}-\frac{1}{\tilde{\pi}}\frac{d}{d\theta}\left\{\frac{\theta\epsilon}{1+\epsilon}\tilde{\pi}(\theta)\frac{v''}{(v')^2}\right\}\right)$.
Finally, using the optimality of consumption taxes, the government budget constraint becomes:

$$0=\int_{\underline{\theta}}^{\bar{\theta}}\left(\frac{dT}{d\xi}+\sum_{i=1}^{N}e_i\frac{1}{q_i}\frac{dq_i}{d\xi}+T'\frac{dz}{d\xi}\right)\pi d\theta-\sum_{i=1}^{n}\frac{\partial\chi_i}{\partial\xi}+t_w\sum_{i=1}^{n}q_i\frac{dC_i}{d\xi}$$
$$\frac{dC_i}{d\xi}=-\int_{\underline{\theta}}^{\bar{\theta}}\partial_{z^*}c_i\left(\frac{dT}{d\xi}+\sum_{i=1}^{N}e_i\frac{1}{q_i}\frac{dq_i}{d\xi}-\left(1-T'\right)\frac{dz}{d\xi}\right)\pi d\theta+\sum_{i=1}^{N}\int_{\underline{\theta}}^{\bar{\theta}}\partial_{q_j}c_i^h\pi d\theta\frac{dq_j}{d\xi}.$$

Plugging the expression for aggregate consumption, we obtain:

$$0=\int_{\underline{\theta}}^{\bar{\theta}}\left(\frac{dT}{d\xi}+\sum_{i=1}^{N}e_i\frac{1}{q_i}\frac{dq_i}{d\xi}+T'\frac{dz}{d\xi}\right)\pi d\theta-\sum_{i=1}^{n}\frac{\partial\chi_i}{\partial\xi}-t_w\int_{\underline{\theta}}^{\bar{\theta}}\left(\frac{dT}{d\xi}+\sum_{i=1}^{N}e_i\frac{1}{q_i}\frac{dq_i}{d\xi}-\left(1-T'\right)\frac{dz}{d\xi}\right)\pi d\theta$$

Using the definitions of $dV/d\xi$ and $dz/d\xi$ and the optimality of the initial income tax schedule, we obtain:

$$0=-\int_{\underline{\theta}}^{\bar{\theta}}\left(1-t_w\right)\frac{1}{v'}\pi\frac{dV}{d\xi}+\frac{d}{d\theta}\left\{\left(\frac{T'}{1-T'}+t_w\right)\frac{\theta\epsilon}{1+\epsilon}\frac{1}{v'}\pi\right\}\frac{dV}{d\xi}d\theta-\sum_{i=1}^{n}\frac{\partial\chi_i}{\partial\xi}$$
$$\Leftrightarrow\left(1-t_w\right)\int_{\underline{\theta}}^{\bar{\theta}}g\frac{1}{v'}\frac{dV}{d\xi}\pi dz=-\sum_{i=1}^{n}\frac{\partial\chi_i}{\partial\xi},$$

which completes the proof. □

The formula in Proposition E2 retains the same structure as the one in Lemma A1, indicating that the introduction of richer income effects does not substantially alter the qualitative response of the optimal tax schedule to price changes. Nonetheless, several distinctions arise.

First, that the relevant distribution with income effects is now given by $\tilde{\pi}=\frac{1}{v'}\pi/\int_{\underline{\theta}}^{\theta}\frac{1}{v'}\pi d\theta$. This re-weighted distribution corrects the original type distribution π to accounts for income effects. Such a correction is standard: for instance, Saez (2001) introduces implicitly an equivalent re-weighting of the income distribution, $\tilde{f}\equiv f\left(z\right)e^{-\int_{\underline{z}}^{z}\frac{\eta}{z\zeta}dz}fdz/\int_{\underline{z}}^{\bar{z}}f\left(z\right)e^{-\int_{\underline{z}}^{z}\frac{\eta}{z\zeta}dz}dz$. Importantly, if v is CRRA, $v=\left(z^*\right)^{1-\beta}/\left(1-\beta\right)$, and π is asymptotically Pareto with tail index ω, then $\tilde{\pi}$ also has Pareto tails with adjusted coefficient $\tilde{\omega}=\omega-\frac{1+\epsilon}{1+\beta\epsilon}\beta$, so the asymptotic properties of the distribution are preserved.

Second, the terms κ and γ differ from lemma $A2$ since they now depends on the curvature of utility $-\frac{v''}{v'}$ (which, as $-\frac{G''}{G'}$, captures the income effect on Pareto weights), on the super-elasticity of ζ (which depends on $dln(\epsilon)/dln\theta$), and on η (which depends on $dln(-\frac{v''}{v'})/dln\theta$). By construction, $\kappa\left(\theta\right)\geq0$ for all θ and, under reasonable assumptions, $\gamma\left(\theta\right)\geq0$. Indeed, if v is CRRA, we have:

$$\gamma\left(\theta\right)=-\frac{G''}{G'}-g^{-1}\frac{v''}{\left(v'\right)^2}\left(1-\frac{1}{\tilde{\pi}}\frac{d}{d\theta}\left\{\frac{\theta\epsilon}{1+\epsilon}\tilde{\pi}(\theta)\right\}+\frac{\epsilon\left(1-\beta\right)}{1+\epsilon\left(1-T'\right)\beta\frac{z}{z^*}+\frac{\epsilon zT''}{1-T'}}\frac{\left(1-T'\right)z}{z^*}\right)$$

Therefore, if $\frac{d}{d\theta}\left\{\frac{\theta\epsilon}{1+\epsilon}\tilde{\pi}(\theta)\right\}/\tilde{\pi}(\theta)\leq1$ (which parallels Assumption A2 in the main text) and ϵ small, then we have $\gamma\geq0$. To highlight the dependence of κ and γ on super-elasticities more transparently, it is helpful to express the RHS of the formula of Proposition E2 in terms of z rather than θ:

$$z\tilde{\zeta}\tilde{f}\kappa\left(z\right)\frac{1}{1-T'}\frac{1}{v'}\frac{d}{dz}\left\{\frac{dV}{d\xi}\right\}+\left(1-t_w\right)\int_{z}^{\bar{z}}g\tilde{f}\left(\gamma\left(z'\right)\frac{dV}{d\xi}-\int_{\underline{z}}^{\bar{z}}\gamma\left(z\right)\frac{dV}{d\xi}g\tilde{f}dz\right)dz',$$

with

$$\kappa\left(z\right)\equiv\left(\frac{1}{1-T'}\frac{\tilde{\zeta}}{\zeta}-\left(\frac{T'}{1-T'}+\alpha\right)\left(\tilde{\eta}+z\tilde{\zeta}\frac{\zeta'/\zeta-\eta'/\left(1+\eta\right)}{1+\eta/\zeta-zT''/\left(1-T'\right)}\right)\right),$$
$$\gamma\left(z\right)\equiv-\frac{1}{\left(1-T'\right)g\left(z\right)}\frac{1}{v'\left(z\right)}\frac{\tilde{\eta}}{z\zeta}\left(1-z\zeta\left(z\right)\frac{\tilde{f}'\left(z\right)}{\tilde{f}\left(z\right)}+\eta-z\zeta\left(z\right)\frac{\tilde{\eta}'\left(z\right)}{\tilde{\eta}\left(z\right)}\right)-\frac{G''\left(z\right)}{G'\left(z\right)}.$$

Finally, the novelty of Proposition $E2$ is that, under a more general supply-side structure, the optimal tax rate becomes sensitive to changes in the average market-size elasticity, $t_w=\alpha$ (in the monopolistic case). The intuition is straightforward: if α increases, an increase in demand for goods reduces prices more strongly. As a consequence, the planner reduces the income tax to encourage labor supply, increase total income, and thereby further reduce prices.

E.3.1 Partial equilibrium results

We now extend the results of Proposition 3, allowing for income effects on labor supply. As in the main text, we consider an increase in the relative price of the necessity, $\bar{p}_l$, in a two-sector economy with linear production functions ($\alpha = 0$). As before, good l is defined to be a necessity by assumption A1.[3] We prove our extension assuming that the coefficient γ is positive and under a restriction on the distribution of types $\tilde{\pi}$ similar to assumption $A2$. The assumption that γ is positive reflects social preferences placing some weight on reducing income inequality. Indeed, $\gamma < 0$ implies that the objective function of the planner is convex: in that case, the regressive impact of the price change discussed in the main text would be exacerbated. To avoid artificially strengthening our results, we assume that $\gamma > 0$, which ensures that the planner is inequality-averse and values redistribution.

Assumption (E3). $\frac{d}{d\theta}\left\{\frac{\theta\epsilon}{1+\epsilon}\tilde{\pi}(\theta)\right\}/\tilde{\pi}(\theta) \leq 1$*and* $\gamma(\theta) \geq 0$ *for all* θ.

As in the main text, we compare the welfare impact of a change in the relative price of necessities, $\bar{p}_l$, compared to the change in welfare that would occur if the social welfare function was linear and satisfied:

$$\frac{\theta\epsilon}{(1+\epsilon)^2}\tilde{\pi}\kappa(\theta)\frac{1}{(1-T')^2}\frac{\theta}{z}\frac{1}{v'}\frac{d}{d\theta}\left\{\frac{dV_{lin}}{d\bar{p}_l}\right\} = -\frac{\theta\epsilon}{1+\epsilon}\frac{\tilde{\pi}(\theta)}{1-T'}\sum_{i=1}^{n}\left(\tau_i(\theta) + \partial_{z^*}\tilde{E}_i - \bar{s}_l\right).$$

As before, $dV_{lin}/d\bar{p}_l$ is increasing in θ, negative at the bottom of the distribution, and positive at the top. Consistent with Proposition 3, we find that the change in welfare increases more slowly than in the linear benchmark. Households experience a strict welfare loss, while households at the top benefit – even though, in principle, it would be feasible to compensate all individuals. Hence, our qualitative conclusions from the main text remain robust when income effects on labor supply are introduced.

While Proposition 3 only considered the case where γ is decreasing (consistent with most standard social welfare functions), in Proposition E3 we also examine the case where γ is increasing. Indeed, with a linear social welfare function and v CRRA, γ increases when the risk aversion parameter is sufficiently large, and decreases when it is small.

In addition, we characterize the willingness to pay for the price change (including the optimal tax reform) for agents at the top of the income distribution. The willingness to pay for a price change and induced optimal tax reform is $\mathcal{Z} = \frac{1}{z^* v_{z^*}}\frac{dV}{d\bar{p}_l}$. Here, $\frac{1}{v_{z^*}}\frac{dV}{d\bar{p}_l}$ represents the dollar amount household θ would be willing to pay to implement the reform, while $\frac{1}{z^* v_{z^*}}\frac{dV}{d\bar{p}_l}$ expresses how much the household would be willing to pay as a share of their pre-reform income z^* . Under the assumption of Proposition 3, the willingness to pay for the reform is positive and given by $\mathcal{Z} = \partial_{z^*}e_l - \bar{s}_l$ at the top of the distribution. When v is CRRA with coefficient β, we show in Proposition E3 that the willingness to pay for the reform is $\mathcal{Z} = C_0\left(\partial_{z^*}e_l - \bar{s}_l\right)$, with $C_0 > 0$. In this sense, income effects only scale the impact of the reform on high income households by a positive scalar C_0: when $C_0 > 1$, top-income households gain more than in

[3] Note that $\partial_{z^*}E_l \leq \bar{s}_l$ directly implies $\partial_{z^*}\tilde{E}_l \leq \bar{s}_l$ since $\tilde{\pi}$ puts more weight on high income households when v is concave.

the linear case; when $C_0 < 1$, they gain less. Interestingly, when $\beta \leq 1$ and the Pareto coefficient of the distribution $\tilde{\pi}$ is not too high,[4] then $C_0 \geq 1$: there is more redistribution towards high income households than in the case without income effects. The case $\beta \leq 1$ and $\frac{1+\beta\epsilon}{1+\epsilon}\omega + \beta \leq 2$ is empirically relevant for the U.S. Indeed, if $\beta > 1$, hours worked would counterfactually decline with income. In addition, the pre-tax income distribution in the US has a Pareto coefficient of less than 2 (it ranges from 1.5 to 2, depending on how much capital income is included), which implies that the underlying distribution of type $\tilde{\pi}$ has a Pareto coefficient of around 1 for $\beta = 1$, which satisfies the sufficient condition on ω.

Therefore, under realistic calibrations, our main result is amplified: an increase in the relative price of necessities leads to even greater redistribution in favor of high-income households than in the baseline case without income effects.

Proposition E3. *Under assumption E4, for an increase in the relative price of necessities, the compensating scheme $dT(z(\theta)) = -(s_l - \bar{s}_l) z^* dln\bar{p}_l$ is feasible but only optimal when preferences are homothetic. With non-homothetic preferences, the change in welfare of agent θ, $dV^G/d\bar{p}_l(\theta) \neq 0$.*

If γ is decreasing, $dV^G/d\bar{p}_l(\theta)$ satisfies $dV_{lin}/d\bar{p}_l(\underline{\theta}) < dV^G/d\bar{p}_l(\underline{\theta}) < 0$, $dV^G/d\bar{p}_l(\theta) - dV^G/d\bar{p}_l(\underline{\theta}) < dV_{lin}/d\bar{p}_l(\theta) - dV_{lin}/d\bar{p}_l(\underline{\theta})$, and $\mathbb{E}\left(gdV^G(\theta)/d\bar{p}_l\right) = 0$.

If γ is increasing , $dV^G/d\bar{p}_l(\theta)$ satisfies $dV_{lin}/d\bar{p}_l\left(\bar{\theta}\right) > dV^G/d\bar{p}_l\left(\bar{\theta}\right) > 0$, $dV^G/d\bar{p}_l(\theta) - dV^G/d\bar{p}_l(\underline{\theta}) < dV_{lin}/d\bar{p}_l(\theta) - dV_{lin}/d\bar{p}_l(\underline{\theta})$, and $\mathbb{E}\left(gdV^G(\theta)/d\bar{p}_l\right) = 0$.
Asymptotically, if there is θ_0 such that for $\theta \geq \theta_0$, $\psi'/(z/\theta\psi'') = \epsilon < 1$, v is CRRA with coefficient β, $\tilde{\pi}$ is Pareto with coefficient ω, and the social welfare function G is CRRA, CARA or linear, then the willingness to pay $\mathcal{Z} = (z^* v_{z^*})^{-1} dV/d\bar{p}_l$ satisfies:

$$\mathcal{Z} \sim -C_0\left(\tau_l + \partial_{z^*}\tilde{E}_l - \bar{s}_l\right) > 0$$
$$C_0 = \frac{(1+\beta\epsilon)(1+\epsilon(1+\omega))}{\left((1-\beta)(1+\epsilon(1+\omega)+\beta\epsilon^2\omega) + \omega\beta\left(\beta(1+\epsilon)^2 + \epsilon\omega(1+\beta\epsilon)\right)\right)},$$

with $\tau_l = \partial_{z^*}e_l - \partial_{z^*}\tilde{E}_l$, $C_0 > 0$ and $C_0 \geq 1$ if $\beta \leq 1$, $\epsilon, \omega \leq 1$.
Proof of Proposition E3. For an increase in the price of necessities and when production functions are linear, the formula of Proposition E2 becomes:

$$\frac{\theta\epsilon}{(1+\epsilon)^2}\tilde{\pi}\kappa(\theta)\frac{1}{(1-T')^2}\frac{\theta}{z}\frac{1}{v'}\frac{d}{d\theta}\left\{\frac{dV}{d\bar{p}_l}\right\} + \int_\theta^{\bar{\theta}}\tilde{\pi}g\left(\gamma(\theta)\frac{dV}{d\bar{p}_l} - \int_{\underline{\theta}}^{\bar{\theta}}\tilde{\pi}g\gamma(\theta)\frac{dV}{d\bar{p}_l}d\theta\right) = -\frac{\theta\epsilon}{1+\epsilon}\frac{\tilde{\pi}(\theta)}{1-T'}\left(\tau_l(\theta) + \partial_{z^*}\tilde{E}_l - \bar{s}_l\right).$$

Under the assumption that $\frac{d}{d\theta}\left(\frac{\theta\epsilon}{1+\epsilon}\tilde{\pi}\right) \leq \tilde{\pi}$, following the same steps as in Corollary 1, we have

$$-\frac{\theta\epsilon}{1+\epsilon}\frac{\tilde{\pi}(\theta)}{1-T'}\left(\tau_l(\theta) + \partial_{z^*}\tilde{E}_l - \bar{s}_l\right) > 0$$

for all θ. Since $\kappa(\theta) > 0$, if $\gamma(\theta)$ is positive and decreasing, following the same step as in Proposition 3,

[4] A sufficient condition is $\frac{1+\beta\epsilon}{1+\epsilon}\omega + \beta \leq 2$ for $\epsilon = 0.21$, which implies that the Pareto tail of the income distribution is less than 2.

we have:

$$\frac{dV}{d\bar{p}_l}\left(\underline{\theta}\right) < 0$$

$$\int_{\theta}^{\bar{\theta}} \tilde{\pi} g \left(\gamma\left(\theta\right) \frac{dV}{d\bar{p}_l} - \int_{\underline{\theta}}^{\theta} \tilde{\pi} g \gamma\left(\theta\right) \frac{dV}{d\bar{p}_l} d\theta\right) > 0$$

$$\int_{\underline{\theta}}^{\theta} \tilde{\pi} g\left(\theta\right) \frac{dV}{d\bar{p}_l} d\theta = 0,$$

which implies, as before, $\frac{dV_{lin}}{d\bar{p}_l}\left(\underline{\theta}\right) < \frac{dV}{d\bar{p}_l}\left(\underline{\theta}\right) < 0$ and $\frac{dV_{lin}}{d\bar{p}_l}\left(\theta\right) - \frac{dV_{lin}}{d\bar{p}_l}\left(\underline{\theta}\right) > \frac{dV}{d\bar{p}_l}\left(\theta\right) - \frac{dV}{d\bar{p}_l}\left(\underline{\theta}\right)$.

For $\gamma\left(\theta\right)$ positive and increasing, consider the change of variable $\tilde{\theta} = 1/\theta$, we obtain:

$$-\frac{\epsilon}{\left(1+\epsilon\right)^2} \tilde{\pi} \kappa \left(\tilde{\theta}^{-1}\right) \frac{1}{\left(1-T'\right)^2} \frac{1}{z} \frac{1}{v'} \frac{d}{d\tilde{\theta}} \left\{\frac{dV}{d\bar{p}_l}\right\} - \int_{\tilde{\theta}}^{\bar{\tilde{\theta}}} \tilde{\pi} g \left(\gamma\left(\tilde{\theta}^{-1}\right) \frac{dV}{d\bar{p}_l} - \int_{\underline{\theta}}^{\theta} \tilde{\pi} g \gamma\left(\tilde{\theta}^{-1}\right) \frac{dV}{d\bar{p}_l} d\theta\right) \frac{1}{\tilde{\theta}^2} d\tilde{\theta} = -\frac{\tilde{\theta}^{-1}\epsilon}{1+\epsilon} \frac{\tilde{\pi}(\tilde{\theta}^{-1})}{1-T'} \left(\tau_l\left(\tilde{\theta}^{-1}\right) + \partial_{z^*} \tilde{E}_l - \bar{s}_l\right).$$

with $\bar{\tilde{\theta}} = 1/\underline{\theta}$. Since $\gamma\left(\tilde{\theta}^{-1}\right)$ is decreasing, we can use our previous result which implies:

$$\frac{dV}{d\bar{p}_l}\left(\underline{\tilde{\theta}}\right) = \frac{dV}{d\bar{p}_l}\left(\bar{\theta}\right) > 0$$

$$\int_{\theta}^{\bar{\theta}} \tilde{\pi} g \left(\gamma\left(\theta\right) \frac{dV}{d\bar{p}_l} - \int_{\underline{\theta}}^{\theta} \tilde{\pi} g \gamma\left(\theta\right) \frac{dV}{d\bar{p}_l} d\theta\right) = -\int_{\tilde{\theta}}^{\bar{\tilde{\theta}}} \tilde{\pi} g \left(\gamma\left(\tilde{\theta}^{-1}\right) \frac{dV}{d\bar{p}_l} - \int_{\underline{\theta}}^{\theta} \tilde{\pi} g \gamma\left(\tilde{\theta}^{-1}\right) \frac{dV}{d\bar{p}_l} d\theta\right) \frac{1}{\tilde{\theta}^2} d\tilde{\theta} > 0$$

$$\int_{\underline{\theta}}^{\theta} \tilde{\pi} g\left(\theta\right) \frac{dV}{d\bar{p}_l} d\theta = 0,$$

and $\frac{dV_{lin}}{d\bar{p}_l}\left(\theta\right) - \frac{dV_{lin}}{d\bar{p}_l}\left(\underline{\theta}\right) > \frac{dV}{d\bar{p}_l}\left(\theta\right) - \frac{dV}{d\bar{p}_l}\left(\underline{\theta}\right)$, as well as $\frac{dV_{lin}}{d\bar{p}_l}\left(\bar{\theta}\right) > \frac{dV}{d\bar{p}_l}\left(\bar{\theta}\right) > 0$.

We now prove the asymptotic results, assuming that $v\left(z^*, \mathbf{p}\right) = \left(z^*\right)^{1-\beta} / \left(1-\beta\right)$ and $\tilde{\pi} = \tilde{\pi}_0 \omega \theta^{-1-\omega}$. We first start with a case where the equation can be solved in closed form. Assume that, above some θ_0, $g = G'' = \epsilon' = 0$, T' and $\partial_{z^*} e_l$ are constant and $z^* = \left(1-T'\right) z$. In that case, we can rewrite our equation as:

$$\frac{\theta\epsilon}{\left(1+\epsilon\right)^2} \tilde{\pi} \kappa \frac{1}{\left(1-T'\right)^2} \frac{\theta}{z} \frac{1}{v'} \frac{d}{d\theta} \left\{\frac{dV}{d\bar{p}_l}\right\} + \int_{\theta}^{\bar{\theta}} \tilde{\pi} \tilde{\gamma}\left(\theta\right) \frac{dV}{d\bar{p}_l} d\theta = -\frac{\theta\epsilon}{1+\epsilon} \frac{\tilde{\pi}(\theta)}{1-T'} \left(\tau_l + \partial_{z^*} \tilde{E}_l - \bar{s}_l\right),$$

with $\kappa = 1 + \beta\epsilon\left(1-T'\right)$, $\tilde{\gamma}\left(\theta\right) = \left(z^*\right)^{\beta-1} \beta \left(\beta \frac{1+\epsilon}{1+\beta\epsilon} + \frac{\epsilon\omega}{1+\epsilon}\right)$, $\tau_l = \partial_{z^*} e_l - \partial_{z^*} \tilde{E}_l$. Next, define $\mathcal{Z} = \left(z^*\right)^{\beta-1} \frac{dV}{d\bar{p}_l}$. We have:

$$\frac{\theta\epsilon}{\left(1+\epsilon\right)^2} \tilde{\pi} \kappa \frac{1}{\left(1-T'\right)^2} \frac{\theta}{z} \frac{1}{v'} \left(\left(z^*\right)^{1-\beta} \frac{d}{d\theta} \left\{\mathcal{Z}\right\} + \left(1-T'\right) \frac{\epsilon+1}{1+\beta\epsilon} \frac{z}{\theta} \left(1-\beta\right) \left(z^*\right)^{-\beta} \mathcal{Z}\right)$$
$$+ \beta \left(\beta \frac{1+\epsilon}{1+\beta\epsilon} + \frac{\epsilon\omega}{1+\epsilon}\right) \int_{\theta}^{\bar{\theta}} \tilde{\pi} \mathcal{Z} d\theta = -\frac{\theta\epsilon}{1+\epsilon} \frac{\tilde{\pi}(\theta)}{1-T'} \left(\tau_l + \partial_{z^*} \tilde{E}_l - \bar{s}_l\right),$$

which simplifies to:

$$\frac{\theta\epsilon}{\left(1+\epsilon\right)^2} \tilde{\pi} \kappa \frac{1}{1-T'} \left(\theta \frac{d}{d\theta} \left\{\mathcal{Z}\right\} + \frac{\epsilon+1}{1+\beta\epsilon} \left(1-\beta\right) \mathcal{Z}\right)$$
$$+ \beta \left(\beta \frac{1+\epsilon}{1+\beta\epsilon} + \frac{\epsilon\omega}{1+\epsilon}\right) \int_{\theta}^{\bar{\theta}} \tilde{\pi} \mathcal{Z} d\theta = -\frac{\theta\epsilon}{1+\epsilon} \frac{\tilde{\pi}(\theta)}{1-T'} \left(\tau_l + \partial_{z^*} \tilde{E}_l - \bar{s}_l\right).$$

Differentiating the equation, we obtain:

$$\theta^2 \frac{d^2}{d\theta^2}\{\mathcal{Z}\} + D_1 \theta \frac{d}{d\theta}\{\mathcal{Z}\} - D_0 \mathcal{Z} = \frac{(1+\epsilon)\,\omega}{\kappa}\left(\tau_l + \partial_{z^*}\tilde{E}_l - \bar{s}_l\right)$$

$$D_1 = 1 - \omega + \frac{\epsilon+1}{1+\beta\epsilon}(1-\beta)$$

$$D_0 = \frac{1+\epsilon}{1+\beta\epsilon}\omega\left((1-\beta) + \frac{\omega}{1+\epsilon(1+\omega)+\beta\epsilon^2\omega}\beta\left(\beta(1+\epsilon)^2 + \epsilon\omega(1+\beta\epsilon)\right)\right) > 0$$

The solution of the homogenous equation are $\theta^{\rho_+}$and $\theta^{-\rho_-}$, where ρ_+, ρ_- are the roots of the polynomial:

$$x^2 + (D_1 - 1)\,x - D_0 = 0.$$

Since the last term of the polynomial, $-D_0$,is negative (assuming that $\omega \geq 1$ if $\beta > 1$), ρ_+, ρ_- are positive. The solution of the equation is:

$$\mathcal{Z} = -C_0\left(\tau_l + \partial_{z^*}\tilde{E}_l - \bar{s}_l\right) + C_+\theta^{\rho_+} + C_-\theta^{-\rho_-}$$

$$C_0 = \frac{(1+\beta\epsilon)(1+\epsilon(1+\omega))}{\left((1-\beta)(1+\epsilon(1+\omega)+\beta\epsilon^2\omega) + \omega\beta\left(\beta(1+\epsilon)^2 + \epsilon\omega(1+\beta\epsilon)\right)\right)} > 0.$$

Since $\int_\theta^{\bar{\theta}} \tilde{\pi}\tilde{\gamma}(\theta)\,\frac{dV}{d\bar{p}_l}d\theta = \int_\theta^{\bar{\theta}} \tilde{\pi}\mathcal{Z}d\theta$ is positive for all θ, we have $C_+ \geq 0$, since $\frac{1}{1-T'}\frac{\theta}{z}\frac{1}{v'}\frac{d}{d\theta}\left\{\frac{dV}{d\bar{p}_l}\right\} \leq -\frac{1+\epsilon}{\kappa}\left(\tau_l + \partial_{z^*}\tilde{E}_l - \bar{s}_l\right)$ for all θ, we have $C_+ \leq 0$ so necessarily, $C_+ = 0$ and $\mathcal{Z} = -C_0\left(\tau_l + \partial_{z^*}\tilde{E}_l - \bar{s}_l\right) + C_-\theta^{-\rho_-}$. When $\beta \leq 1$, $\epsilon, \omega \leq 1$, direct algebra shows that $C_0 \geq 1$.

Next, we consider the case where we simply assume that G is $CARA$,$CRRA$ or linear. As a preliminary, we characterize $1-T'$, z, $z^*$and V at the top of the distribution. First, the tax rate above $\theta_0$satisfies:

$$\frac{T'}{1-T'} = \frac{\epsilon+1}{\epsilon}\frac{1}{\theta\tilde{\pi}}\int_\theta^{\bar{\theta}}(1-g)\,\tilde{\pi}d\theta.$$

Using l'Hopital rule, we have $\lim_{\theta\to\infty} 1-T' = \epsilon\omega/(1+\epsilon(1+\omega))$. In addition, we have:

$$\frac{1}{z}\frac{dz}{d\theta} = \frac{\epsilon}{1+\epsilon(1-T')\,\beta\frac{z}{z^*} + \frac{\epsilon z T''}{1-T'}}\frac{\epsilon+1}{\epsilon}\frac{1}{\theta}$$

$$\frac{\epsilon}{1+(1-T')\,\beta\frac{z}{z^*} + \frac{\epsilon z T''}{1-T'}}\frac{zT''}{(1-T')^2} = \left(g - \frac{\omega\int_\theta^{\bar{\theta}} g\tilde{\pi}d\theta}{\theta\tilde{\pi}}\right).$$

Since g is decreasing and converges to 0, we have $\frac{zT''}{1-T'}$ is positive and converges to 0. In addition, it is direct to show that $\lim_{\theta\to\infty} z = \infty$, $\lim_{\theta\to\infty}(1-T')\,z/z^* = 1$, $(1-T')\,z/z^* - 1 = O(\theta^{-\bar{\lambda}_0})$, with $\bar{\lambda}_0 = \frac{1+\epsilon}{1+\epsilon\beta}$ and $\frac{zT''}{1-T'} = o\left((z^*)^{-\beta}\right)$. In addition, we have:

$$\frac{dV}{d\theta} = \frac{1}{\theta}\psi'\left(\frac{z}{\theta}\right)\frac{z}{\theta} = (1-T')\,(z^*)^{-\beta}\,\frac{z}{\theta}$$

$$= (z^*)^{-\beta}\,\frac{dz^*}{d\theta}\frac{1+\epsilon(1-T')\,\beta\frac{z}{z^*} + \frac{\epsilon z T''}{1-T'}}{\epsilon+1}.$$

Since $\frac{1+\epsilon(1-T')\beta\frac{z}{z^*}+\frac{\epsilon z T''}{1-T'}}{\epsilon+1}$ converges to $\frac{1+\epsilon\beta}{\epsilon+1}$, we have for $\beta \leq 1$ that $V \approx \frac{1+\epsilon\beta}{\epsilon+1}\frac{(z^*)^{1-\beta}}{1-\beta}$ at infinity, while V converges to a constant if $\beta > 1$.

Next, consider, as before, the change of variable $\mathcal{Z} = (z^*)^{\beta-1}\frac{dV}{d\bar{p}_l}$; we obtain:

$$\begin{aligned}
&\frac{\theta\epsilon}{(1+\epsilon)^2}\tilde{\pi}\kappa(\theta)\frac{1}{(1-T')^2}\frac{\theta}{z}(z^*)^{\beta}\frac{d}{d\theta}\left\{(z^*)^{1-\beta}\mathcal{Z}\right\} \\
&\qquad\qquad + \int_{\theta}^{\bar{\theta}}\tilde{\pi}g\left(\gamma(\theta)(z^*)^{1-\beta}\mathcal{Z} - \int_{\underline{\theta}}^{\theta}\tilde{\pi}g\gamma(\theta)\frac{dV}{d\bar{p}_l}d\theta\right)d\theta \\
&\qquad\qquad\qquad\qquad = -\frac{\theta\epsilon}{1+\epsilon}\frac{\tilde{\pi}(\theta)}{1-T'}\left(\tau_l(\theta) + \partial_{z^*}\tilde{E}_l - \bar{s}_l\right),
\end{aligned}$$

which simplifies to:

$$\begin{aligned}
&\frac{\theta\epsilon}{1+\epsilon}\frac{\tilde{\pi}(\theta)}{1-T'}\kappa(\theta)\left(\frac{1}{1+\epsilon}\frac{1}{1-T'}\frac{z^*}{z}\theta\frac{d}{d\theta}\{\mathcal{Z}\} + \frac{1}{1+\epsilon(1-T')\beta\frac{z}{z^*}+\frac{\epsilon z T''}{1-T'}}(1-\beta)\mathcal{Z}\right) \\
&\qquad\qquad + \int_{\theta}^{\bar{\theta}}\tilde{\pi}g\left(\gamma(\theta)(z^*)^{1-\beta}\mathcal{Z} - \int_{\underline{\theta}}^{\theta}\tilde{\pi}g\gamma(\theta)\frac{dV}{d\bar{p}_l}d\theta\right)d\theta \\
&\qquad\qquad\qquad\qquad = -\frac{\theta\epsilon}{1+\epsilon}\frac{\tilde{\pi}(\theta)}{1-T'}\left(\tau_l(\theta) + \partial_{z^*}\tilde{E}_l - \bar{s}_l\right),
\end{aligned}$$

with:

$$\begin{aligned}
\kappa(\theta) &= 1 + \epsilon\beta\left(1-T'\right)\frac{(1-T')z}{z^*} \\
g\gamma(\theta) &= -g\frac{G''}{G'} + \beta\left(\left(1+\frac{\omega\epsilon}{1+\epsilon}\right) + (1-\beta)\frac{\epsilon}{1+\epsilon(1-T')\beta\frac{z}{z^*}+\frac{\epsilon z T''}{1-T'}}\frac{(1-T')z}{z^*}\right)(z^*)^{\beta-1}.
\end{aligned}$$

We first bound $\mathcal{Z}$ and $\theta\frac{d}{d\theta}\{\mathcal{Z}\}$ using these equations. Since the integral term

$$\int_{\theta}^{\bar{\theta}}\tilde{\pi}g\left(\gamma(\theta)(z^*)^{1-\beta}\mathcal{Z} - \int_{\underline{\theta}}^{\theta}\tilde{\pi}g\gamma(\theta)\frac{dV}{d\bar{p}_l}d\theta\right)d\theta$$

is positive, we obtain by Chaplygin's theorem that $\mathcal{Z} \leq \bar{\mathcal{Z}}$, where $\bar{\mathcal{Z}}$ solves:

$$\begin{aligned}
\frac{\theta\epsilon}{1+\epsilon}\frac{\tilde{\pi}(\theta)}{1-T'}\kappa(\theta)\left(\frac{1}{1+\epsilon}\frac{1}{1-T'}\frac{z^*}{z}\theta\frac{d}{d\theta}\{\bar{\mathcal{Z}}\} + \frac{\epsilon}{1+\epsilon(1-T')\beta\frac{z}{z^*}+\frac{\epsilon z T''}{1-T'}}(1-\beta)\bar{\mathcal{Z}}\right)& \\
= -\frac{\theta\epsilon}{1+\epsilon}\frac{\tilde{\pi}(\theta)}{1-T'}\left(\tau_l(\theta) + \partial_{z^*}\tilde{E}_l - \bar{s}_l\right)& \\
\bar{\mathcal{Z}}(\theta_0) = \mathcal{Z}(\theta_0)&
\end{aligned}$$

Using the variation of parameters for first order equations, we obtain:

$$\bar{\mathcal{Z}}(\theta) = exp\left(-\int_{\theta_0}^{\theta} \lambda_0(\vartheta)\frac{1-\beta}{\vartheta}d\vartheta\right)\mathcal{Z}(\theta_0)$$
$$+\int_{\theta_0}^{\theta} exp\left(-\int_{\vartheta}^{\theta}\lambda_0(\vartheta')\frac{1-\beta}{\vartheta'}d\vartheta'\right)\frac{(1-T')\frac{z}{z^*}(1+\epsilon)}{\kappa(\vartheta)\vartheta}\left(-\left(\tau_l(\vartheta)+\partial_{z^*}\tilde{E}_l-\bar{s}_l\right)\right)d\vartheta$$
$$\lambda_0(\vartheta) = \frac{(1-T')\frac{z}{z^*}(1+\epsilon)\epsilon}{1+\epsilon\beta(1-T')\frac{z}{z^*}+\frac{\epsilon z T''}{1-T'}} > 0$$

First, consider the case $\beta < 1$. Since $\lambda_0(\vartheta)$, $\kappa(\vartheta)$, $(1-T')\frac{z}{z^*}$ and $\tau_l(\vartheta)$ converge to positive constant at infinity (respectively $\bar{\lambda}_0 = \frac{(1+\epsilon)\epsilon}{1+\epsilon\beta}, \bar{\kappa} = 1+\epsilon\beta(1-T'), 1$ and $\bar{\tau}_l = \partial_{z^*}e_l - \partial_{z^*}\tilde{E}_l$), we have for θ_0 large enough and δ small,

$$\bar{\mathcal{Z}}(\theta) \le |\mathcal{Z}(\theta_0)|\left(\frac{\theta}{\theta_0}\right)^{-\bar{\lambda}_0(1-\beta)+\delta} + (1+\delta)\frac{(1+\epsilon)}{\bar{\kappa}}\left(-\left(\bar{\tau}_l+\partial_{z^*}\tilde{E}_l-\bar{s}_l\right)\right)\frac{1}{\bar{\lambda}_0(1-\beta)}\left(1-\left(\frac{\theta}{\theta_0}\right)^{-\bar{\lambda}_0(1-\beta)+\delta}\right).$$

So we have, for θ large enough, that $\bar{\mathcal{Z}}(\theta)$ is bounded above by a positive constant, so $\mathcal{Z}(\theta)$ is bounded above by a positive constant $\mathcal{Z}^*$. With G as $CARA$,$CRRA$ or linear, we have $g\frac{G''}{G'}(z^*)^{1-\beta} = o\left(z^{*-\beta}\right)$, $g = o\left(z^{*-\beta}\right)$, so $g\gamma(\theta)(z^*)^{1-\beta} = O(1)$, which implies that:

$$\left(\frac{\theta\epsilon}{1+\epsilon}\frac{\tilde{\pi}(\theta)}{1-T'}\right)^{-1}\int_{\theta}^{\bar{\theta}}\tilde{\pi}g\left(\gamma(\theta)(z^*)^{1-\beta}\mathcal{Z}-\int_{\underline{\theta}}^{\theta}\tilde{\pi}g\gamma(\theta)\frac{dV}{d\bar{p}_l}d\theta\right)d\theta \le \left(\frac{\theta\epsilon}{1+\epsilon}\frac{\tilde{\pi}(\theta)}{1-T'}\right)^{-1}\int_{\theta}^{\bar{\theta}}\tilde{\pi}g\left(\gamma(\theta)(z^*)^{1-\beta}\bar{\mathcal{Z}}(\theta)-\int_{\underline{\theta}}^{\theta}\tilde{\pi}g\gamma(\theta)\frac{dV}{d\bar{p}_l}d\theta\right)d\theta \le M\mathcal{Z}^*.$$

We therefore have, applying again Chaplygin's theorem, that $\mathcal{Z}(\theta) \ge \underline{\mathcal{Z}}(\theta)$, where $\underline{\mathcal{Z}}(\theta)$ solves:

$$\frac{\theta\epsilon}{1+\epsilon}\frac{\tilde{\pi}(\theta)}{1-T'}\kappa(\theta)\left(\frac{1}{1+\epsilon}\frac{1}{1-T'}\frac{z^*}{z}\theta\frac{d}{d\theta}\{\underline{\mathcal{Z}}\}+\frac{\epsilon}{1+\epsilon(1-T')\beta\frac{z}{z^*}+\frac{\epsilon z T''}{1-T'}}(1-\beta)\underline{\mathcal{Z}}\right) = -\frac{\theta\epsilon}{1+\epsilon}\frac{\tilde{\pi}(\theta)}{1-T'}\left(\tau_l(\theta)+\partial_{z^*}\tilde{E}_l-\bar{s}_l-M\mathcal{Z}^*\right),$$

which implies, as before, that $\mathcal{Z}(\theta) \ge \underline{\mathcal{Z}}(\theta) \ge \mathcal{Z}_*$. We therefore have that, for $\beta < 1$, $\mathcal{Z}(\theta)$ is bounded, which directly implies that $\theta\frac{d}{d\theta}\{\mathcal{Z}\}$ is bounded.

For $\beta > 1$ using the same reasoning, we obtain $\bar{\mathcal{Z}}(\theta) = O\left(\theta^{\bar{\lambda}_0(\beta-1)}\right)$, which implies, since $g\gamma(\theta)(z^*)^{1-\beta} = O(1)$:

$$\mathcal{Z}(\theta) \le M\theta^{\bar{\lambda}_0(\beta-1)}$$
$$\left(\frac{\theta\epsilon}{1+\epsilon}\frac{\tilde{\pi}(\theta)}{1-T'}\right)^{-1}\int_{\theta}^{\bar{\theta}}\tilde{\pi}g\left(\gamma(\theta)(z^*)^{1-\beta}\mathcal{Z}-\int_{\underline{\theta}}^{\theta}\tilde{\pi}g\gamma(\theta)\frac{dV}{d\bar{p}_l}d\theta\right)d\theta \le M\theta^{\bar{\lambda}_0(\beta-1)}$$

for some M large enough. Applying Chaplygin's theorem once more for a lower bound, we therefore obtain $|\mathcal{Z}(\theta)| \le M\theta^{\bar{\lambda}_0(\beta-1)}$, which implies by direct inspection of the equations $\left|\theta\frac{d}{d\theta}\{\mathcal{Z}\}\right| \le M\theta^{\bar{\lambda}_0(\beta-1)}$.

Finally, using the same steps as before in the case $\beta = 1$, we obtain $|\mathcal{Z}(\theta)|, \left|\theta\frac{d}{d\theta}\{\mathcal{Z}\}\right| \le ln(\theta)$.

Next, we differentiate the equation

$$\frac{\theta\epsilon}{1+\epsilon}\frac{\tilde{\pi}(\theta)}{1-T'}\kappa(\theta)\left(\frac{1}{1+\epsilon}\frac{1}{1-T'}\frac{z^*}{z}\theta\frac{d}{d\theta}\{\mathcal{Z}\}+\frac{1}{1+\epsilon(1-T')\beta\frac{z}{z^*}+\frac{\epsilon z T''}{1-T'}}(1-\beta)\mathcal{Z}\right)+\int_{\theta}^{\bar{\theta}}\tilde{\pi}g\left(\gamma(\theta)(z^*)^{1-\beta}\mathcal{Z}-\int_{\underline{\theta}}^{\theta}\tilde{\pi}g\gamma(\theta)\frac{dV}{d\bar{p}_l}d\theta\right)d\theta$$
$$= -\frac{\theta\epsilon}{1+\epsilon}\frac{\tilde{\pi}(\theta)}{1-T'}\left(\tau_l(\theta)+\partial_{z^*}\tilde{E}_l-\bar{s}_l\right).$$

We obtain:

$$
\begin{aligned}
\theta^2 \frac{d^2}{d\theta^2} \{\mathcal{Z}\} + (D_1 + \delta_1(\theta)) \theta \frac{d}{d\theta} \{\mathcal{Z}\} - (D_0 + \delta_0(\theta)) \mathcal{Z} &= \frac{(1+\epsilon)\omega}{\kappa} \left(\tau_l + \partial_{z^*} \tilde{E}_l - \bar{s}_l\right) + \delta(\theta) \\
\delta_1(\theta) &= \left(1 - T'\right) \frac{z}{z^*} \theta \frac{d}{d\theta} \left\{ \frac{1}{1-T'} \frac{z^*}{z} \right\} + \theta \frac{d}{d\theta} \left\{ \frac{1}{1-T'} \right\} \\
&\quad + \frac{1}{\kappa} \theta \frac{d}{d\theta} \{\kappa(\theta)\} + \epsilon \frac{\epsilon+1}{1+\beta\epsilon} (1-\beta) \frac{\beta \left(1 - (1-T') \frac{z}{z^*}\right) - \frac{\epsilon z T''}{1-T'}}{1 + \epsilon (1-T') \beta \frac{z}{z^*} + \frac{\epsilon z T''}{1-T'}} \\
\delta_0(\theta) &= \frac{(1-\beta)(1+\epsilon)\left(1-T'\right) \frac{z}{z^*} \mathcal{Z} \left(-\omega + \theta \frac{d}{d\theta} \left\{ \frac{1}{1-T'} \right\} + \frac{1}{\kappa} \theta \frac{d}{d\theta} \{\kappa(\theta)\} - \frac{\beta\theta \frac{d}{d\theta} \left\{ \frac{1}{1-T'} \frac{z^*}{z} \right\} + \theta \frac{d}{d\theta} \left\{ \frac{z T''}{1-T'} \right\}}{1 + \epsilon (1-T') \beta \frac{z}{z^*} + \frac{\epsilon z T''}{1-T'}} \right)}{1 + \epsilon (1-T') \beta \frac{z}{z^*} + \frac{\epsilon z T''}{1-T'}} \\
&\quad - \frac{(1+\epsilon)^2}{\epsilon} \frac{1-T'}{\kappa(\theta)} \left(1-T'\right) \frac{z}{z^*} g \gamma(\theta) (z^*)^{1-\beta} - D_1 \\
\delta(\theta) &= \frac{(1+\epsilon)^2}{\epsilon} \frac{1-T'}{\kappa(\theta)} \left(1-T'\right) \frac{z}{z^*} g \int_{\underline{\theta}}^{\theta} \tilde{\pi} g \gamma(\theta) \frac{dV}{d\bar{p}_l} d\theta \\
&= -\frac{1+\epsilon}{\kappa(\theta)} \left(1-T'\right) \frac{z}{z^*} \left(-\omega + \theta \frac{d}{d\theta} \left\{ \frac{1}{1-T'} \right\} \right) \left(\tau_l(\theta) + \partial_{z^*} \tilde{E}_l - \bar{s}_l \right) \\
&\quad + \frac{1+\epsilon}{\kappa(\theta)} \left(1-T'\right) \frac{z}{z^*} \theta \frac{d}{d\theta} \{\tau_l(\theta)\} - \frac{(1+\epsilon)\omega}{\kappa} \left(\tau_l + \partial_{z^*} \tilde{E}_l - \bar{s}_l \right).
\end{aligned}
$$

It is direct to show, using our asymptotic characterization of the initial tax rate, that $\delta_1(\theta) = O\left(\theta^{-\bar{\lambda}_0}\right)$, $\delta_0(\theta) = O\left(\theta^{-\bar{\lambda}_0}\right)$ and $\delta(\theta)$ converges to 0. Indeed, the dominant term of $\delta_1(\theta)$ and $\delta_0(\theta)$ is

$$
\theta \frac{d}{d\theta} \left\{ \frac{1}{1-T'} \frac{z^*}{z} \right\} = \frac{(\epsilon+1)}{1 + \epsilon (1-T') \beta \frac{z}{z^*} + \frac{\epsilon z T''}{1-T'}} \left(\frac{(1-T') z}{z^*} - 1 + \frac{z T''}{1-T'} \right),
$$

which converges to 0 at rate $\theta^{-\bar{\lambda}_0}$. Using the variation of parameters for second order equations, we therefore have:

$$
\begin{aligned}
\mathcal{Z} &= -C_0 \left(\tau_l + \partial_{z^*} \tilde{E}_l - \bar{s}_l \right) + C_+ \theta^{\rho_+} + C_- \theta^{-\rho_-} \\
&\quad - \frac{1}{\rho_+ + \rho_-} \int \left(\theta^{-\rho_-} \hat{\theta}^{\rho_- - 1} - \theta^{\rho_+} \hat{\theta}^{-\rho_+ - 1} \right) \left(\delta\left(\hat{\theta}\right) + \delta_0\left(\hat{\theta}\right) \mathcal{Z} + \delta_1\left(\hat{\theta}\right) \hat{\theta} \frac{d}{d\theta} \{\mathcal{Z}\} \right) d\hat{\theta}.
\end{aligned}
$$

First note that, since $\delta\left(\hat{\theta}\right)$ goes to 0, we have $\int \left(\theta^{-\rho_-} \hat{\theta}^{\rho_- - 1} - \theta^{\rho_+} \hat{\theta}^{-\rho_+ - 1} \right) \delta\left(\hat{\theta}\right) d\hat{\theta} = o(1)$. Next if $\beta < 1$, we have $\mathcal{Z}, \theta \frac{d}{d\theta} \{\mathcal{Z}\} = O(1)$, which implies $\int \left(\theta^{-\rho_-} \hat{\theta}^{\rho_- - 1} - \theta^{\rho_+} \hat{\theta}^{-\rho_+ - 1} \right) \left(\delta_0\left(\hat{\theta}\right) \mathcal{Z} + \delta_1\left(\hat{\theta}\right) \hat{\theta} \frac{d}{d\theta} \{\mathcal{Z}\} \right) d\hat{\theta} = o(1)$ and $\mathcal{Z}$ converges to $-C_0 \left(\tau_l + \partial_{z^*} \tilde{E}_l - \bar{s}_l \right)$, since as before C_+ is necessarily 0. If $\beta > 1$, we have $\left| \delta_0(\theta) \mathcal{Z} + \delta_1(\theta) \theta \frac{d}{d\theta} \{\mathcal{Z}\} \right| \leq M \theta^{\bar{\lambda}_0 (\beta - 2)}$, so

$$
\int \left(\theta^{-\rho_-} \hat{\theta}^{\rho_- - 1} - \theta^{\rho_+} \hat{\theta}^{-\rho_+ - 1} \right) \left(\delta\left(\hat{\theta}\right) + \delta_0\left(\hat{\theta}\right) \mathcal{Z} + \delta_1\left(\hat{\theta}\right) \hat{\theta} \frac{d}{d\theta} \{\mathcal{Z}\} \right) d\hat{\theta} \leq N \theta^{\bar{\lambda}_0 (\beta - 2)}.
$$

Direct inspection shows that $\rho_+ > \bar{\lambda}_0 (\beta - 2)$, so using the same reasoning as before, we have $C_+ = 0$. This implies $\mathcal{Z} = -C_0 \left(\tau_l + \partial_{z^*} \tilde{E}_l - \bar{s}_l \right) + C_- \theta^{-\rho_-} + O\left(\theta^{\bar{\lambda}_0 (\beta - 2)} \right)$. If $\beta < 2$, we have $\mathcal{Z}$ converges to $-C_0 \left(\tau_l + \partial_{z^*} \tilde{E}_l - \bar{s}_l \right)$, otherwise, noting that $|\mathcal{Z}| \leq M \theta^{\bar{\lambda}_0 (\beta - 2)}$ directly implies $\left| \theta \frac{d}{d\theta} \{\mathcal{Z}\} \right| \leq M \theta^{\bar{\lambda}_0 (\beta - 2)}$, we obtain $\left| \delta_0(\theta) \mathcal{Z} + \delta_1(\theta) \theta \frac{d}{d\theta} \{\mathcal{Z}\} \right| \leq M \theta^{\bar{\lambda}_0 (\beta - 3)}$. Reiterating our operation, if $\beta < 3$, we have $\mathcal{Z}$ converges to $-C_0 \left(\tau_l + \partial_{z^*} \tilde{E}_l - \bar{s}_l \right)$; otherwise by direct induction, we obtain $\mathcal{Z}$ converges to $-C_0 \left(\tau_l + \partial_{z^*} \tilde{E}_l - \bar{s}_l \right)$. □

E.3.2 General equilibrium results

In this subsection, we characterize the general equilibrium responses of the tax schedule to shifts in the supply curves when pricing and cost functions are non-linear. We follow the same approach as in Proposition A1 of Appendix A: we decompose the welfare response $dV/d\xi$ into a sum of $n+1$ partial equilibrium components, each of which can be computed independently of the supply-side response. These partial equilibrium components allow us to directly compute changes in aggregate consumption, $dC/d\xi$, in Lemma E1, and to recover the change in equilibrium prices $dq_k/d\xi$ through matrix inversion.

Due to the complexity of the formulas involved, we conclude the section with an intuitive discussion of how the optimal schedule responds to increases in the price elasticity α.

As a direct consequence of Proposition E2, we can decompose the welfare response to an exogenous supply shock ξ as:

$$\frac{dV}{d\xi}(\theta) = \sum_{k=1}^{n} \frac{\partial V}{\partial q_k}(\theta)\left(\frac{1}{q_k}\frac{dq_k}{d\xi} - \frac{1}{1-t_w}\frac{dt_w}{d\xi}\right) - \frac{\partial V}{\partial B}(\theta)\sum_{i=1}^{n}\frac{\partial \chi_i}{\partial \xi}.$$

Here, $\frac{\partial V}{\partial q_k}$ denotes the change in welfare to an increase in the price of k keeping the government budget fixed. It satisfies:

$$\frac{\theta\epsilon}{(1+\epsilon)^2}\tilde{\pi}\kappa(\theta)\frac{1}{(1-T')^2}\frac{\theta}{z}\frac{1}{v'}\frac{d}{d\theta}\left\{\frac{\partial V}{\partial q_k}\right\} + (1-t_w)\int_{\theta}^{\bar{\theta}}\tilde{\pi}g\left(\gamma(\theta)\frac{\partial V}{\partial q_k} - \int_{\underline{\theta}}^{\theta}\tilde{\pi}g\gamma(\theta)\frac{\partial V}{\partial q_k}d\theta\right) = -\frac{\epsilon}{1+\epsilon}\frac{\theta\tilde{\pi}(\theta)}{1-T'}\left(\tau_k(\theta) + \partial_{z^*}\tilde{E}_k\right)$$
$$(1-t_w)\int_{\underline{\theta}}^{\bar{\theta}} g\frac{1}{v'}\frac{\partial V}{\partial q_k}\pi dz = 0.$$

Note that defining $\frac{\partial V}{\partial t_w}$ the (partial equilibrium) response of welfare to an increase in the corrective tax t_w, as the solution of:

$$\frac{\theta\epsilon}{(1+\epsilon)^2}\tilde{\pi}\kappa(\theta)\frac{1}{(1-T')^2}\frac{\theta}{z}\frac{1}{v'}\frac{d}{d\theta}\left\{\frac{\partial V}{\partial t_w}\right\} + (1-t_w)\int_{\theta}^{\bar{\theta}}\tilde{\pi}g\left(\gamma(\theta)\frac{\partial V}{\partial t_w} - \int_{\underline{\theta}}^{\theta}\tilde{\pi}g\gamma(\theta)\frac{\partial V}{\partial t_w}d\theta\right) = \frac{\epsilon}{1+\epsilon}\frac{\theta\tilde{\pi}(\theta)}{1-T'}$$
$$(1-t_w)\int_{\underline{\theta}}^{\bar{\theta}} g\frac{1}{v'}\frac{\partial V}{\partial t_w}\pi dz = 0,$$

we have $\frac{\partial V}{\partial t_w} = -\sum_{k=1}^{n}\frac{\partial V}{\partial q_k}$, so $\frac{\partial V}{\partial t_w}$ need not be computed separately.

Finally, $\frac{\partial V}{\partial B}$ denotes the welfare sensitivity to a relaxation of the government's budget constraint. It satisfies:

$$\frac{\theta\epsilon}{(1+\epsilon)^2}\tilde{\pi}\kappa(\theta)\frac{1}{(1-T')^2}\frac{\theta}{z}\frac{1}{v'}\frac{d}{d\theta}\left\{\frac{\partial V}{\partial B}\right\} + (1-t_w)\int_{\theta}^{\bar{\theta}}\tilde{\pi}g\left(\gamma(\theta)\frac{\partial V}{\partial B} - \int_{\underline{\theta}}^{\theta}\tilde{\pi}g\gamma(\theta)\frac{\partial V}{\partial B}d\theta\right) = 0$$
$$(1-t_w)\int_{\underline{\theta}}^{\bar{\theta}} g\frac{1}{v'}\frac{\partial V}{\partial t_w}\pi dz = 1.$$

With these definitions in hand, we can characterize the response of aggregate consumption to supply shocks as a function of the partial equilibrium response of welfare to price and budget changes. We do so in the following lemma, which is a generalization of lemma A3.

Lemma E1. *The change in aggregate consumption in response to an exogenous supply change $d\xi$ is given by:*

$$
\begin{aligned}
\frac{1}{C_i}\frac{dC_i}{d\xi} &= \sum_{j=1}^{n}\frac{q_j\partial C_i}{C_i\partial q_j}\left(\frac{1}{q_k}\frac{dq_k}{d\xi} - \frac{1}{1-t_w}\frac{dt_w}{d\xi}\right) - \frac{\partial C_i}{C_i\partial B}\sum_{i=1}^{n}\frac{\partial \chi_i}{\partial \xi} \\
\frac{q_j\partial C_i}{C_i\partial q_j} &\equiv \frac{1}{1-t_w}\frac{1}{\tilde{E}_k}\int_{\underline{\theta}}^{\bar{\theta}}\frac{d}{d\theta}\left\{\frac{\partial V}{\partial q_j}\right\}\frac{1}{1-T'}\frac{\theta\epsilon}{1+\epsilon}\left(\tau_k + \partial_{z^*}\tilde{E}_k\right)\tilde{\pi}d\theta + \mathcal{S}_{i,j} \\
\frac{\partial C_i}{C_i\partial B} &\equiv \frac{1}{1-t_w}\frac{1}{\tilde{E}_k}\int_{\underline{\theta}}^{\bar{\theta}}\frac{d}{d\theta}\left\{\frac{\partial V}{\partial B}\right\}\frac{1}{1-T'}\frac{\theta\epsilon}{1+\epsilon}\left(\tau_k + \partial_{z^*}\tilde{E}_k\right)\tilde{\pi}d\theta + \frac{1}{1-t_w}\frac{\partial_{z^*}\tilde{E}_i}{E_i},
\end{aligned}
$$

with $\tilde{E}_i = \int_{\underline{\theta}}^{\theta} e_i\tilde{\pi}d\theta$.

Proof of Lemma E1. We follow the same steps as in Lemma A3. We have:

$$
\begin{aligned}
\frac{1}{C_k}\frac{dC_k}{d\xi} &= \frac{1}{C_k}\int_{\underline{\theta}}^{\bar{\theta}}\frac{dc_k}{d\xi}\pi d\theta \\
&= \frac{1}{C_k}\int_{\underline{\theta}}^{\bar{\theta}}\partial_{z^*}c_k\left(\frac{dz^*}{d\xi} - \sum_{l=1}^{n}c_l\frac{dq_l}{d\xi}\right)\pi d\theta + \frac{1}{C_k}\sum_{l=1}^{n}\int_{\underline{\theta}}^{\bar{\theta}}\partial_{p_l}c_k^h\frac{dq_l}{d\xi}\pi d\theta \\
&= \frac{1}{C_k}\int_{\underline{\theta}}^{\bar{\theta}}\partial_{z^*}c_k\left(\frac{1}{v_{z^*}}\frac{dV}{d\xi} + \left(1-T'\right)\frac{dz}{d\xi}\right)\pi d\theta + \sum_{l=1}^{n}\mathcal{S}_{kl}\frac{1}{q_l}\frac{dq_l}{d\xi} \\
&= \frac{1}{\tilde{E}_k}\int_{\underline{\theta}}^{\bar{\theta}}\frac{d}{d\theta}\left\{\frac{dV}{d\xi}\right\}\int_{\theta}^{\bar{\theta}}\left(\partial_{z^*}e_k - \partial_{z^*}\tilde{E}_k\right)\tilde{\pi}d\theta + \frac{\partial_{z^*}\tilde{E}_k}{\tilde{E}_k}\int_{\underline{\theta}}^{\bar{\theta}}\frac{dV}{d\xi_k}\tilde{\pi}d\theta + \frac{1}{\tilde{E}_k}\int_{\underline{\theta}}^{\bar{\theta}}\partial_{z^*}e_k\frac{\theta\epsilon}{1+\epsilon}\frac{d}{d\theta}\left\{\frac{dV}{d\xi}\right\}\tilde{\pi}d\theta + \sum_{l=1}^{n}\mathcal{S}_{kl}\frac{1}{q_l}\frac{dq_l}{d\xi},
\end{aligned}
$$

with $\tilde{E}_k = \int_{\theta}^{\bar{\theta}} e_k\tilde{\pi}d\theta$ and where we used, from the IC-FOC, $v'\left(1-T'\right)\frac{dz(\theta)}{d\xi} = \frac{\theta\epsilon}{1+\epsilon}\frac{d}{d\theta}\left\{\frac{dV}{d\xi}\right\}$. Next, using the government budget constraint (see the proof of Proposition E1), we have:

$$
\int_{\underline{\theta}}^{\bar{\theta}}\left(1-t_w\right)\frac{1}{v'}\pi\frac{dV}{d\xi} - \frac{d}{d\theta}\left\{\frac{dV}{d\xi}\right\}\left(\frac{T'}{1-T'} + t_w\right)\frac{\theta\epsilon}{1+\epsilon}\frac{1}{v'}\pi d\theta = -\sum_{i=1}^{n}\frac{\partial \chi_i}{\partial \xi}.
$$

Using this expression, we have:

$$
\begin{aligned}
\frac{1}{C_k}\frac{dC_k}{d\xi} = &\frac{1}{\tilde{E}_k}\int_{\underline{\theta}}^{\bar{\theta}}\frac{d}{d\theta}\left\{\frac{dV}{d\xi}\right\}\int_{\theta}^{\bar{\theta}}\left(\partial_{z^*}e_k - \partial_{z^*}\tilde{E}_k\right)\tilde{\pi}d\theta + \frac{1}{1-t_w}\frac{\partial_{z^*}\tilde{E}_k}{\tilde{E}_k}\left(\int_{\underline{\theta}}^{\bar{\theta}}\frac{d}{d\theta}\left\{\frac{dV}{d\xi}\right\}\left(\frac{T'}{1-T'} + t_w\right)\frac{\theta\epsilon}{1+\epsilon}\tilde{\pi}d\theta\right) \\
&+ \frac{1}{\tilde{E}_k}\int_{\underline{\theta}}^{\bar{\theta}}\partial_{z^*}e_k\frac{\theta\epsilon}{1+\epsilon}\frac{d}{d\theta}\left\{\frac{dV}{d\xi}\right\}\tilde{\pi}d\theta \\
&- \frac{1}{1-t_w}\frac{\partial_{z^*}\tilde{E}_k}{E_k}\sum_{i=1}^{n}\frac{\partial \chi_i}{\partial \xi} + \sum_{l=1}^{n}\mathcal{S}_{kl}\frac{1}{q_l}\frac{dq_l}{d\xi} \\
= &\frac{1}{\left(1-t_w\right)}\frac{1}{\tilde{E}_k}\int_{\underline{\theta}}^{\bar{\theta}}\frac{d}{d\theta}\left\{\frac{dV}{d\xi}\right\}\frac{1}{1-T'}\frac{\theta\epsilon}{1+\epsilon}\tilde{\pi}\left(\left(1-t_w\right)\left(1-T'\right)\left(\frac{1+\epsilon}{\epsilon}\frac{1}{\theta\tilde{\pi}\theta)}\int_{\theta}^{\bar{\theta}}\left(\partial_{z^*}e_k - \partial_{z^*}\tilde{E}_k\right)\tilde{\pi}d\theta + \partial_{z^*}e_k - \partial_{z^*}\tilde{E}_k\right) + \partial_{z^*}\tilde{E}_k\right)d\theta \\
&- \frac{1}{1-t_w}\frac{\partial_{z^*}\tilde{E}_k}{E_k}\sum_{i=1}^{n}\frac{\partial \chi_i}{\partial \xi} + \sum_{l=1}^{n}\mathcal{S}_{kl}\frac{1}{q_l}\frac{dq_l}{d\xi} \\
= &\frac{1}{\left(1-t_w\right)}\frac{1}{\tilde{E}_k}\int_{\underline{\theta}}^{\bar{\theta}}\frac{d}{d\theta}\left\{\frac{dV}{d\xi}\right\}\frac{1}{1-T'}\frac{\theta\epsilon}{1+\epsilon}\tilde{\pi}\left(\tau_k + \partial_{z^*}\tilde{E}_k\right)d\theta - \frac{1}{1-t_w}\frac{\partial_{z^*}\tilde{E}_k}{E_k}\sum_{i=1}^{n}\frac{\partial \chi_i}{\partial \xi} + \sum_{l=1}^{n}\mathcal{S}_{kl}\frac{1}{q_l}\frac{dq_l}{d\xi}.
\end{aligned}
$$

Using the fact that $\frac{dV}{d\xi}(\theta) = \sum_{k=1}^{n} \frac{\partial V}{\partial q_k}(\theta)\left(\frac{1}{q_k}\frac{dq_k}{d\xi} - \frac{1}{1-t_w}\frac{dt_w}{d\xi}\right) - \frac{\partial V}{\partial B}(\theta)\sum_{i=1}^{n}\frac{\partial \chi_i}{\partial \xi}$ and $\sum_j \mathcal{S}_{i,j} = 0$ then gives the decomposition of the lemma.$\square$

Using the partial equilibrium welfare responses and the aggregate demand adjustments derived above, we now characterize the general equilibrium response of the optimal tax rates. To derive these results, we must first define the relevant supply-side elasticities that govern the equilibrium adjustment in prices.

In particular, in addition to the standard price elasticity with respect to market size, we also need to introduce the super-elasticity of supply with respect to market size. This is necessary because demand shifts also affect the elasticity itself, altering the slope of the supply curve in equilibrium. The super-elasticity captures how the elasticity changes as the market expands, and thus play a role in determining how prices adjust in response to supply shocks.

Definition. *Recall that A is the matrix of price elasticities with respect to market sizes with $A_{i,j} = -C_j \partial_{C_j}\phi(C_1, ..., C_n, \xi)/p_i$. We define $a_{i,j,l} \equiv C_l \partial_l A_{i,j}/A_{i,j}$ the super-elasticity of prices with respect to market sizes and $\mathscr{A}$ the matrix with entries $\mathscr{A}_{i,l} = \sum_{j=1}^{N} A_{j,i} a_{j,i,l} \frac{p_j C_j}{p_i C_i}$. The matrix $\mathscr{A}$ summarizes how a change in the market size of l impact the externality of an increase in the demand for i on the average producer price. Denote $\Delta\left[(1+t_i)^{-1}\right]$ the diagonal matrix with $(1+t_i)^{-1}$ on the diagonal (and similarly $\Delta[p_i C_i]$, $\Delta\left[(p_i C_i)^{-1}\right]$) and define:*

$$\mathcal{A}_m \equiv Id - \frac{1}{1-\alpha}\Delta\left[(1+t_i)^{-1}\right]\left(\left(Id - \Delta[p_i C_i] A^T \Delta\left[(p_i C_i)^{-1}\right]\right)(Id - A) - \mathscr{A}\right)$$

$$\mathcal{A}_c \equiv A$$

$$\mathcal{C}_{i,j} \equiv \frac{q_j \partial C_i}{C_i \partial q_j}$$

The matrices $\mathcal{A}_m$ and $\mathcal{A}_c$ summarize the relevant supply side elasticity in the monopolistic and competitive cases respectively while $\mathcal{C}$ summarizes the demand size elasticities.

The equilibrium responses of prices and of t_w can then simply be obtained as a function of $\mathcal{A}_m$, $\mathcal{A}_c$ and $\mathcal{C}$.

Proposition E4. *Consider a change in the parameters ξ and define $\partial p_i/\partial \xi \equiv \partial \phi_i(C_1, ..., C_N, \xi)/\partial \xi$, $\partial A_{i,j}/\partial \xi \equiv \partial\left(\partial_{lnC_j} ln\phi_i(C_1, ..., C_N, \xi)\right)/\partial \xi$. The change in the utility of households in response to a change in ξ is given by:*

$$\frac{dV}{d\xi}(z) = \sum_{i=1}^{n} q_i \frac{\partial V}{\partial q_i}(z)\frac{1}{q_i}\frac{dq_i}{d\xi} - \frac{\partial V}{\partial B}(z)\sum_{i=1}^{n}\frac{\partial \chi_i}{\partial \xi},$$

where the endogenous price vector $\frac{\mathbf{1}}{\boldsymbol{q}}\frac{\boldsymbol{dq}}{\boldsymbol{d\xi}} = \left[\frac{1}{q_1}\frac{dq_1}{d\xi}, ..., \frac{1}{q_n}\frac{dq_n}{d\xi}\right]$ solves:

$$\frac{\mathbf{1}}{\boldsymbol{q}}\frac{\boldsymbol{dq}}{\boldsymbol{d\xi}} = (Id + \mathcal{A}\mathcal{C})^{-1}\left(\frac{\mathbf{1}}{\boldsymbol{q}}\frac{\boldsymbol{\partial q}}{\boldsymbol{\partial \xi}} + \sum_{i=1}^{n}\frac{\partial \chi_i}{\partial \xi}\mathcal{A}\frac{\boldsymbol{\partial C}}{\boldsymbol{C\partial B}}\right),$$

where $\partial q_i/\partial \xi = \partial p_i/\partial \xi$, $\mathcal{A} = \mathcal{A}_c$ in the competitive case and $\partial q_i/\partial \xi = (1+t_i)\partial p_i/\partial \xi - \frac{p_i}{(1-\alpha)}\sum_{j=1}^{N}\frac{p_j C_j}{p_i C_i}\frac{\partial A_{j,i}}{\partial \xi}$,$\mathcal{A} =$

$\mathcal{A}_m$ in the monopolistic case. The change in taxes in the competitive case is given by:

$$\frac{dT}{d\xi} = -\frac{1}{v'}\frac{dV}{d\xi} - \sum_{i=1}^{n} e_i \frac{1}{q_i}\frac{dq_i}{d\xi},$$

and in the monopolistic case by:

$$\frac{dT}{d\xi} = -\frac{1}{v'}\frac{dV}{d\xi} - \sum_{i=1}^{n} e_i \left(\frac{1}{q_i}\frac{dq_i}{d\xi} + \frac{1}{1-\alpha}\frac{d\alpha}{d\xi}\right)$$

$$\frac{1}{1-\alpha}\frac{d\alpha}{d\xi} = \frac{1}{1-\alpha}\frac{\partial\alpha}{\partial\xi} - \boldsymbol{s}^T\left((Id-\mathcal{A}) - \Delta\left[(1+t_i)^{-1}\right](Id-A)\right)\left(\mathcal{C}\frac{\mathbf{1}}{\boldsymbol{q}}\frac{\boldsymbol{dq}}{\boldsymbol{d\xi}} - \left(\sum_{j=1}^{N} p_j C_j \frac{1}{p_j}\frac{\partial p_j}{\partial\xi}\right)\frac{\boldsymbol{\partial C}}{\boldsymbol{C\partial B}}\right).$$

Proof of Proposition E4. Let us start with the competitive case. There are no consumption taxes in that case and we have $q_i = p_i = \phi_i(C_1, ..., C_N, \xi)$. In addition, $t_w = 0$ and $dt_w/d\xi$=0. We therefore have:

$$\begin{aligned}\frac{1}{q_i}\frac{dq_i}{d\xi} &= \frac{1}{p_i}\frac{\partial p_i}{\partial\xi} - \sum_{j=1}^{N} A_{i,j}\frac{1}{C_j}\frac{dC_j}{d\xi} \\ &= \frac{1}{p_i}\frac{\partial p_i}{\partial\xi} - \sum_{j=1}^{N} A_{i,j}\left(\sum_{j=1}^{n}\frac{q_j\partial C_i}{C_i\partial q_j}\frac{1}{q_k}\frac{dq_k}{d\xi} - \frac{\partial C_i}{C_i\partial B}\sum_{i=1}^{n}\frac{\partial\chi_i}{\partial\xi}\right).\end{aligned}$$

Therefore, we have

$$\frac{\mathbf{1}}{\boldsymbol{q}}\frac{\boldsymbol{dq}}{\boldsymbol{d\xi}} = (Id + \mathcal{A}\mathcal{C})^{-1}\left(\frac{\mathbf{1}}{\boldsymbol{q}}\frac{\boldsymbol{\partial q}}{\boldsymbol{\partial\xi}} + \sum_{i=1}^{n}\frac{\partial\chi_i}{\partial\xi}\mathcal{A}\frac{\boldsymbol{\partial C}}{\boldsymbol{C\partial B}}\right),$$

with $\mathcal{A} = A$ and $\frac{1}{q}\frac{\partial q}{\partial\xi} = \frac{1}{p}\frac{\partial p}{\partial\xi}$. Next, consider the monopolistic case. Recall from Proposition A.1. that we have $q_i = p_i\frac{1-\sum_j A_{j,i}p_jC_j/p_iC_i}{1-\alpha}$ with $1+t_i = \frac{1-\sum_j A_{j,i}p_jC_j/p_iC_i}{1-\alpha}$, defining $1/q_i d\tilde{q}_i/d\xi = 1/q_i dq_i/d\xi - 1/$

$(1-\alpha)\, d\alpha/d\xi$, we have:

$$\begin{aligned}
\frac{1}{q_i}\frac{d\tilde{q}_i}{d\xi} =& \frac{1}{p_i}\frac{\partial p_i}{\partial \xi} - \frac{1}{(1+t_i)(1-\alpha)}\sum_{j=1}^{n}\frac{p_j C_j}{p_i C_i}\frac{\partial A_{j,i}}{\partial \xi} \\
& - \sum_{j=1}^{n} A_{i,j}\frac{1}{C_j}\frac{dC_j}{d\xi} + \frac{1-(1+t_i)(1-\alpha)}{(1+t_i)(1-\alpha)}\left(\frac{1}{C_i}\frac{dC_i}{d\xi} - \sum_{j=1}^{n} A_{i,j}\frac{1}{C_j}\frac{dC_j}{d\xi}\right) \\
& - \frac{1}{(1+t_i)(1-\alpha)}\sum_{j=1}^{n} A_{j,i}\frac{p_j C_j}{p_i C_i}\left(\frac{1}{C_j}\frac{dC_j}{d\xi} - \sum_{l=1}^{n} A_{j,l}\frac{1}{C_l}\frac{dC_l}{d\xi}\right) \\
& - \frac{1}{(1+t_i)(1-\alpha)}\sum_{j=1}^{n} A_{j,i}\frac{p_j C_j}{p_i C_i}\sum_{l=1}^{n} a_{j,i,l}\frac{1}{C_l}\frac{dC_l}{d\xi} \\
=& \frac{1}{p_i}\frac{\partial p_i}{\partial \xi} - \frac{1}{(1+t_i)(1-\alpha)}\sum_{j=1}^{n} A_{i,j}\frac{p_j C_j}{p_i C_i}\frac{1}{A_{i,j}}\frac{\partial A_{i,j}}{\partial \xi} \\
& - \frac{1}{C_i}\frac{dC_i}{d\xi} \\
& + \frac{1}{(1+t_i)(1-\alpha)}\left(\frac{1}{C_i}\frac{dC_i}{d\xi} - \sum_{j=1}^{n} A_{i,j}\frac{1}{C_j}\frac{dC_j}{d\xi} - \sum_{j=1}^{n} A_{j,i}\frac{p_j C_j}{p_i C_i}\left(\frac{1}{C_j}\frac{dC_j}{d\xi} - \sum_{l=1}^{n} A_{j,l}\frac{1}{C_l}\frac{dC_l}{d\xi}\right) - \sum_{l=1}^{n}\mathscr{A}_{i,l}\frac{1}{C_l}\frac{dC_l}{d\xi}\right),
\end{aligned}$$

where we used $p_i = \phi(C_1, ..., C_N, \xi)$. In matrix form, denoting $\boldsymbol{\frac{1}{q}\frac{d\tilde{q}}{d\xi}} = \left[\frac{1}{q_1}\frac{d\tilde{q}_1}{d\xi}, ..., \frac{1}{q_N}\frac{d\tilde{q}_N}{d\xi}\right]$, and similarly $\boldsymbol{\frac{1}{C}\frac{dC}{d\xi}}$, $\Delta\left[(1+t_i)^{-1}\right]$ the diagonal matrix with $(1+t_i)^{-1}$ on the diagonal – and similarly $\Delta[p_i C_i]$, $\Delta\left[(p_i C_i)^{-1}\right]$and defining $\frac{1}{q_i}\frac{\partial q_i}{\partial \xi} = \frac{1}{p_i}\frac{\partial p_i}{\partial \xi} - \frac{1}{(1+t_i)(1-\alpha)}\sum_{j=1}^{N}\frac{p_j C_j}{p_i C_i}\frac{\partial A_{j,i}}{\partial \xi}$ – ,we have:

$$\boldsymbol{\frac{1}{q}\frac{d\tilde{q}}{d\xi}} = \boldsymbol{\frac{1}{q}\frac{\partial q}{\partial \xi}} - \left(Id - \frac{1}{1-\alpha}\Delta\left[(1+t_i)^{-1}\right]\left(\left(Id - \Delta[p_i C_i]\, A^T \Delta\left[(p_i C_i)^{-1}\right]\right)(Id - A) - \mathscr{A}\right)\right)\boldsymbol{\frac{1}{C}\frac{dC}{d\xi}}.$$

Next using the result of Lemma E1, we have, noting that $t_w = \alpha$:

$$\begin{aligned}
\frac{1}{C_i}\frac{dC_i}{d\xi} &= \sum_{j=1}^{n}\left(\frac{q_j \partial C_i}{C_i \partial q_j}\left(\frac{1}{q_j}\frac{dq_i}{d\xi} - \frac{1}{1-\alpha}\frac{d\alpha}{d\xi}\right) - \frac{\partial C_i}{C_i \partial B}\sum_{i=1}^{n}\frac{\partial \chi_i}{\partial \xi}\right) \\
&= \sum_{j=1}^{N}\left(\frac{q_j \partial C_i}{C_i \partial q_j}\frac{1}{q_i}\frac{d\tilde{q}_i}{d\xi} - \frac{\partial C_i}{C_i \partial B}\sum_{i=1}^{n}\frac{\partial \chi_i}{\partial \xi}\right).
\end{aligned}$$

Defining $\mathcal{C}_{i,j} \equiv \frac{q_j \partial C_i}{C_i \partial q_j}$, $\mathcal{A} \equiv Id - \frac{1}{1-\alpha}\Delta\left[(1+t_i)^{-1}\right]\left(\left(Id - \Delta[p_i C_i]\, A^T \Delta\left[(p_i C_i)^{-1}\right]\right)(Id - A) - \mathscr{A}\right)$, we obtain

$$\begin{aligned}
\boldsymbol{\frac{1}{q}\frac{d\tilde{q}}{d\xi}} &= \boldsymbol{\frac{1}{q}\frac{\partial q}{\partial \xi}} - \mathcal{A}\mathcal{C}\boldsymbol{\frac{1}{q}\frac{d\tilde{q}}{d\xi}} + \sum_{i=1}^{n}\frac{\partial \chi_i}{\partial \xi}\mathcal{A}\boldsymbol{\frac{\partial C}{C \partial B}} \\
\Rightarrow \boldsymbol{\frac{1}{q}\frac{d\tilde{q}}{d\xi}} &= (Id + \mathcal{A}\mathcal{C})^{-1}\left(\boldsymbol{\frac{1}{q}\frac{\partial q}{\partial \xi}} + \sum_{i=1}^{n}\frac{\partial \chi_i}{\partial \xi}\mathcal{A}\boldsymbol{\frac{\partial C}{C \partial B}}\right).
\end{aligned}$$

Recall that we have:

$$\begin{aligned}\frac{dV}{d\xi} &= \sum_{i=1}^{n} q_i \frac{\partial V}{\partial q_i}\left(\frac{1}{q_k}\frac{dq_k}{d\xi} - \frac{1}{1-t_w}\frac{dt_w}{d\xi}\right) - \frac{\partial V}{\partial B}\sum_{i=1}^{n}\frac{\partial \chi_i}{\partial \xi} \\ &= \sum_{i=1}^{n} q_i \frac{\partial V}{\partial \bar{q}_i}\frac{1}{q_i}\frac{d\tilde{q}_i}{d\xi} - \frac{\partial V}{\partial B}\sum_{i=1}^{n}\frac{\partial \chi_i}{\partial \xi},\end{aligned}$$

which gives the formula. Finally, we have:

$$\alpha = \frac{\sum_{i=1}^{n} p_i C_i \sum_{j=1}^{n} A_{i,j}}{\sum_{i=1}^{N} p_i C_i}.$$

So the change in average price elasticity with respect to market size is given by:

$$\begin{aligned}\frac{d\alpha}{d\xi} &= \frac{\sum_{i=1}^{N} p_i C_i \sum_{j=1}^{N}\left(A_{j,i}\frac{p_j C_j}{p_i C_i} - \alpha \Bbbk(i=j)\right)\frac{1}{p_j}\frac{\partial p_j}{\partial \xi}}{\sum_{i=1}^{N} p_i C_i} + \frac{\sum_{i=1}^{N} p_i C_i \sum_{j=1}^{N}\frac{\partial A_{j,i}}{\partial \xi}}{\sum_{i=1}^{N} p_i C_i} \\ &+ \frac{\sum_{i=1}^{N}\left(\frac{1}{C_i}\frac{dC_i}{d\xi} - \sum_{j=1}^{N} A_{i,j}\frac{1}{C_j}\frac{dC_j}{d\xi}\right) p_i C_i \sum_{j=1}^{N} A_{i,j}}{\sum_{i=1}^{N} p_i C_i} + \frac{\sum_{i=1}^{N} p_i C_i \sum_{j=1}^{N} A_{j,i}\frac{p_j C_j}{p_i C_i}\sum_{l=1}^{N} a_{j,i,l}\frac{1}{C_l}\frac{dC_l}{d\xi}}{\sum_{i=1}^{N} p_i C_i} \\ &- \alpha\frac{\sum_{i=1}^{N}\left(\frac{1}{C_i}\frac{dC_i}{d\xi} - \sum_{j=1}^{N} A_{i,j}\frac{1}{C_j}\frac{dC_j}{d\xi}\right) p_i C_i}{\sum_{i=1}^{N} p_i C_i} \\ &= \sum_{i=1}^{N}\frac{s_i}{1+t_i}\left(\sum_{j=1}^{N} A_{j,i}\frac{p_j C_j}{p_i C_i} - \alpha\Bbbk(i=j)\right)\frac{1}{p_j}\frac{\partial p_j}{\partial \xi} + \sum_{i=1}^{N}\frac{s_i}{1+t_i}\sum_{j=1}^{N}\frac{\partial A_{j,i}}{\partial \xi} \\ &+ \sum_{i=1}^{N}\frac{s_i}{1+t_i}\sum_{j=1}^{N}\left(A_{j,i}\frac{p_j C_j}{p_i C_i} - \alpha\Bbbk(i=j)\right)\left(\frac{1}{C_j}\frac{dC_j}{d\xi} - \sum_{j=1}^{N} A_{j,l}\frac{1}{C_l}\frac{dC_l}{d\xi}\right) + \sum_{i=1}^{N}\frac{s_i}{1+t_i}\sum_{l=1}^{N}\mathscr{A}_{i,l}\frac{1}{C_l}\frac{dC_l}{d\xi}.\end{aligned}$$

Defining $\partial\alpha/\partial\xi = \sum_{i=1}^{N}\left(s_i/\left(1+t_i\right)\right)\sum_{j=1}^{N}\left(\left(A_{j,i}\frac{p_j C_j}{p_i C_i} - \alpha\Bbbk(i=j)\right)\frac{1}{p_j}\frac{\partial p_j}{\partial \xi} + \frac{\partial A_{j,i}}{\partial \xi}\right)$, we therefore have in matrix form:

$$\begin{aligned}\frac{d\alpha}{d\xi} &= \frac{\partial \alpha}{\partial \xi} + \boldsymbol{s}^T\Delta\left[(1+t_i)^{-1}\right]\left(\left(\Delta\left[p_i C_i\right]A^T\Delta\left[(p_i C_i)^{-1}\right] - \alpha Id\right)(Id - A) + \mathscr{A}\right)\frac{\mathbf{1}}{\boldsymbol{C}}\frac{\boldsymbol{dC}}{\boldsymbol{d\xi}} \\ \frac{1}{1-\alpha}\frac{d\alpha}{d\xi} &= \frac{1}{1-\alpha}\frac{\partial\alpha}{\partial\xi} - \boldsymbol{s}^T\left((Id - \mathcal{A}) - \Delta\left[(1+t_i)^{-1}\right](Id - A)\right)\left(\mathcal{C}\frac{\mathbf{1}}{\boldsymbol{q}}\frac{\boldsymbol{d\tilde{q}}}{\boldsymbol{d\xi}} - \left(\sum_{j=1}^{N} p_j C_j \frac{1}{p_j}\frac{\partial p_j}{\partial \xi}\right)\frac{\boldsymbol{\partial C}}{\boldsymbol{C\partial B}}\right).\end{aligned}$$

Using the fact that $dV/d\xi = -v'\left(dT/d\xi + \sum_{i=1}^{N} c_i dq_i/d\xi\right)$ gives the final expression of the proposition.□

Proposition E4 generalizes the formulas of Proposition A1. As can be seen from the formulas, a change in the return to scale parameter α has an effect on optimal taxes. To better understand their impact, consider a simple example with two sectors and where the market size elasticity α is common across sectors. As prices become more sensitive to market size (e.g., α increases), the planner implements a larger wage subsidy. As explained in Proposition A1, the subsidy incentivizes labor supply, aggregate income increases, and all prices decrease. A naive interpretation of Proposition A1 suggests that this corrective wage subsidy can be implemented independently from redistributive policies: since the wage subsidy is $1/(1-\alpha)$, the derivative of the tax rate with respect to α would then simply be $dT'/d\alpha = -(1-T')/(1-\alpha)$. Our comparative statics results unveil a subtler interaction. A higher wage subsidy is equivalent to a homogeneous reduction in prices. As seen in section 4.3, this implies that the share of luxuries increases while the share of necessities decreases. The relative price of luxuries therefore decreases. This triggers a readjustment of optimal redistribution policies, with more redistribution toward higher income households.

To see this explicitly, consider an application of Proposition E4 in our benchmark two-good example with Assumption A1 and A2. We find that commodity taxes are optimally set to zero and that the response of the income tax to an increase in α is:

$$\frac{dT'}{d\alpha} = \overbrace{-\frac{1-T'}{1-\alpha}}^{\text{Wage Subsidy}} - \overbrace{\frac{\partial T'}{\partial p_h}\left(\frac{1}{p_h}\frac{dp_h}{d\alpha} - \frac{1}{p_l}\frac{dp_l}{d\alpha}\right)}^{\text{Redistribution Induced by Price Changes}}$$

$$\text{with} \quad \frac{1}{p_h}\frac{dp_h}{d\alpha} - \frac{1}{p_l}\frac{dp_l}{d\alpha} = \frac{1}{1-\alpha}\frac{\frac{\alpha\zeta}{1-\alpha}}{1-\frac{\alpha\zeta}{1-\alpha}}\frac{\mathbb{E}_z(\partial_{z^*}E_l - s_l + \tau_l)}{s_h s_l(1-\alpha\sigma)\left(1-\frac{\alpha\zeta}{1-\alpha}\Omega\right)}.$$

Thus, when prices become more elastic to demand, the tax schedule becomes more regressive not only because of the corrective wage subsidy, but also because of the change in redistribution policies it induces. When $\partial_{z^*}E_l - s_l \leq 0$,[5] the wage subsidy reduces the relative price of the luxury good. Note that this reduction is exactly the same as the one generated by an exogenous decrease in all producer prices of $1/(1-\alpha)$ percent, which corresponds to the increase in the wage subsidy. The price decrease is then amplified by general equilibrium effects[6] and further decreases tax rates according to $\partial_{q_h}T'$, as it becomes more valuable to redistribute towards higher income households. The interaction between corrective and redistributive taxation is therefore non trivial and works through prices. When prices instead become less sensitive to demand, the wage subsidy is lowered, the relative price of the necessity good decreases, which makes redistribution towards lower income households socially more valuable. A general lesson is that corrective and redistributive taxation cannot be conducted independently when prices are elastic.

[5] Or, alternatively, when the share of luxuries $e_h(z^*)/z^*$ increases along the income distribution.

[6] As before, this amplification is larger when the elasticity of substitution σ, the initial price elasticity α and non-homothecities are stronger.