Food subsidies in general equilibrium*

Albert Jan Hummel† Vinzenz Ziesemer‡

March 26, 2023

Abstract

The Atkinson-Stiglitz theorem on uniform consumption taxation breaks down if prices are endogenous. This paper investigates the implications for optimal food subsidies in China. To do so, we build a general equilibrium model where low-skilled workers have a comparative advantage in the production of food. Food subsidies raise the relative demand for low-skilled workers, which reduces the skill premium and indirectly redistributes income from high-skilled to low-skilled workers. The optimal food subsidy balances these distributional gains against the costs of distorting consumption decisions. We calibrate our model to match key moments from the Chinese economy, including sectoral production and spending patterns that we obtain from micro-level survey data. Our results suggest that general equilibrium effects rationalize only very modest food subsidies that generate tiny welfare gains.

JEL classification: E64, H21, Q18
Keywords: uniform consumption taxes, general equilibrium effects, food subsidies

*We would like to thank the editor Jonathan Heathcote, two anonymous referees, Bas Jacobs, Marcelo Pedroni, Dominik Sachs, Uwe Thümmel and seminar participants at EEA 2020, IIPF 2020, the University of Amsterdam, and the University of Cologne for useful comments and suggestions.

†University of Amsterdam, Tinbergen Institute and CESifo. E-mail: a.j.hummel@uva.nl.

‡Erasmus University Rotterdam. E-mail: ziesemer@euc.eur.nl.

1 Introduction

Food subsidies are a widely used policy tool, both in developed and developing countries. Despite their widespread usage, these policies remain controversial. Subsidies on agricultural products put a significant burden on government finances.[1] Moreover, they are often argued to give an unfair advantage to recipients of these subsidies compared to their foreign competitors. Finally, subsidies or low VAT rates on food and other necessities are not generally considered to be well-targeted toward low-income households (IMF, 2008).

Public economic theory provides arguments for uniform consumption taxation and hence, against the use of food subsidies. Most famously, Atkinson and Stiglitz (1976) show that consumption tax differentiation is undesirable if the government can use a nonlinear income tax and preferences are weakly separable between consumption and leisure. Intuitively, the government can use labor income taxes to achieve the same redistribution, but without distorting consumption choices. Deaton (1979) shows that this result extends to cases where income taxes are linear, provided the demand for goods increases linearly in income (i.e., provided Engel curves are linear). Partly because of these findings, uniform consumption taxation is an often-voiced piece of policy advice (cf. Ahmad and Stern, 1989).

Notwithstanding these results, it is known since at least Naito (1999) that when factor prices are endogenous, differential consumption taxation can improve welfare even if preferences are weakly separable between consumption and leisure. To illustrate, suppose the government introduces a food subsidy. The reduction in the price faced by consumers raises the demand for food, which in turn raises the demand for the labor input that is used relatively intensively in its production, presumably low-skilled labor. The increased demand for food raises the wage of low-skilled workers relative to that of high-skilled workers. A government that cares about redistribution can exploit these general equilibrium effects when designing tax policy.

What is not known, however, is how much this matters for tax policy. What is the order of magnitude by which consumption taxes should be differentiated, and what are the welfare gains from doing so? We study this form of indirect redistribution, focusing on a case where it may be particularly salient: food subsidies in China. The Chinese economy features a large agricultural sector that is relatively low-skilled labor intensive. In addition, the Chinese government is the world's largest subsidizer of agriculture (OECD, 2017). Finally, many developing countries face significant tax capacity constraints, especially when it comes to taxing personal income (Besley and Persson, 2014).[2] When nonlinear income taxes are not

[1] Specifically for China, the OECD (2017) estimates that in 2016 agricultural subsidies amounted to \$212 billion, which is close to 9.1% of total fiscal revenues.

[2] See also "Why only 2% of Chinese pay any income tax" from The Economist (December, 2018).

available, consumption taxes become particularly important.

We study the implications of general equilibrium effects for optimal food subsidies both theoretically and quantitatively. We follow the setup of Deaton (1979), in which a Ramsey planner redistributes using linear income and consumption taxes. We build a general equilibrium model with two commodities, agricultural and non-agricultural products, and two factors of production, low-skilled and high-skilled labor.[3] Within each skill type, individuals differ in their productivity. Preferences are separable between consumption and leisure and sub-utility over the two consumption goods is of the Stone-Geary form, with food entering as a necessity. This gives rise to linear Engel curves for food with a positive intercept. Low-income individuals thus devote a larger share of their total spending to food, but the marginal propensity to spend on food is the same for everyone. An immediate implication is that optimal food subsidies are zero if prices are fixed, cf. Deaton (1979). Any rationale for subsidizing food therefore must come from general equilibrium effects.

Theoretically, we start by characterizing the welfare impact of a small reduction in the food subsidy, which has both utility and budgetary effects. We show how the total welfare impact can be decomposed into 'direct effects' associated with transfers from individuals to the government, 'behavioral effects' that capture the budgetary impact from changes in consumption and labor supply, and 'general equilibrium effects' from changes in the price of food and wages. Changes in the price of food and wages affect the tax base, which generates a budgetary effect, and individual purchasing power, which has a direct impact on individual utilities. Hence, standard formulas for the welfare impact of linear instruments (Sheshinski, 1972 and Diamond, 1975) are modified to account for price responses, as in Allen (1982).

Using the property that Engel curves are linear, we then derive an expression for the optimal food subsidy. We do this by studying a tax reform that – holding prices fixed – leaves the utility and labor supply of *all* individuals unaffected.[4] This reform combines a reduction in the food subsidy with changes in the labor income tax and the lump-sum transfer that, in the absence of general equilibrium effects, offset the negative impact of a lower food subsidy on each individual's budget. We use this reform to derive an expression that links the optimal tax on agricultural goods (i.e., the negative of the food subsidy) to the welfare impact from changes in the price of food and wages. This formula demonstrates that food subsidies are only useful insofar they generate indirect distributional benefits through changes in equi-

[3]In our setup, the shares of the two skill types are considered fixed, while workers can freely move across sectors. We also assume that the economy is closed to trade with the outside world. We discuss the consequences of these choices in Section 6.

[4]In a framework with a pollution externality and fixed prices, Jacobs and van der Ploeg (2019) use the same reform to study under what conditions it is possible to construct a Pareto improvement.

librium prices. An immediate implication is that the optimal food subsidy is zero if prices are fixed, which confirms the result from Deaton (1979). In the general case where prices are endogenous, the optimal food subsidy balances the costs from distorting consumption decisions against the indirect distributional benefits from general equilibrium effects on the price of food and wages.

To study the quantitative implications of general equilibrium effects for the optimal design of food subsidies, we calibrate our model to the Chinese economy. Our main data source is the 2008 wave of the Chinese Household Income Project (CHIP), which provides micro-level survey data on earnings, personal characteristics and expenditures disaggregated by spending categories. We classify individuals as high-skilled if they completed a college degree, which is the case for roughly 10% of the individuals in our sample. In addition, we classify individuals as working in agriculture versus non-agriculture based on where they work most hours. In line with our theoretical framework, we find that agricultural production is low-skilled labor intensive: 13% of workers outside agriculture are college educated, whereas in agriculture this figure is around 1%. Combined with an estimate of the skill premium, these statistics discipline the parameters of the production function. We fit a shifted lognormal distribution to capture the substantial heterogeneity in earnings within our broad education groups, and use data on spending disaggregated by categories to obtain an estimate of the slope of the Engel curve, which is used as an input to parameterize the utility function. Data on other elasticities and government policies, including a food subsidy rate of 10% (OECD, 2017), complete our parameterization.

Turning to the results, we find that general equilibrium effect rationalize only very modest food subsidies that generate tiny welfare gains. Under a utilitarian criterion, the optimal food subsidy is 1.29%, much smaller than in the baseline economy. Not surprisingly, the small optimal food subsidy generates welfare gains close to zero compared to uniform consumption taxes. Starting from a uniform consumption tax setting without food subsidies, a government is indifferent between optimizing food subsidies and increasing everyone's consumption aggregate by less than 0.01%. The welfare gains from optimizing food subsidies are larger if they are more severely mis-optimized, though the gains remain modest.

We investigate the robustness of our results by changing the government's preferences for redistribution, some of the key parameters in the calibration, and by imposing restrictions on the government's instrument set. The optimal food subsidies and welfare gains that can be reaped from them do not turn out to be sensitive to the level of subsistence consumption. The same is true when we vary the elasticity of substitution between skill types in the production function, which determines the strength of general equilibrium effects, or the elasticity of substitution between agricultural and non-agricultural goods in the utility function, which determines how much individuals change their consumption mix in response to a change in the food subsidy. Optimal food subsidies are somewhat larger if the government maximizes the expected utility of low-skilled workers, which generates a strong motive to indirectly redistribute income by lowering the skill premium, but the welfare gains remain very small. The only instance where we find a significant role for subsidies (or taxes) on food is if the government cannot raise the lump-sum transfer (or labor tax) directly. However, the primary goal of food subsidies is then not to indirectly redistribute through general equilibrium effects, but rather to substitute for the instrument that cannot be optimized.

Given the considerable attention paid to consumption taxes in the literature and in policy discussions, we consider the small optimal food subsidies and limited welfare effects an interesting finding in and of itself. They suggest that implementing uniform consumption taxes does little damage – at least in the setting we study. At the same time, our results suggest that the welfare gains from reforming consumption taxes will typically be small.

The remainder of this paper is organized as follows. This section finishes with a review of related literature. Section 2 presents the model. Section 3 studies theoretically the implications of general equilibrium effects for optimal food subsidies. Section 4 discusses the data and the parameterization. Section 5 contains the quantitative analysis of optimal taxes, along with several robustness checks. Section 6 discusses how the results are likely to be affected if some of the key assumptions are relaxed. Section 7 concludes. The appendices contain the proofs and derivations.

1.1 Related literature

How to optimize consumption taxes is an age-old problem in public economics. Ramsey (1927) shows that the optimal linear tax on a consumption good is inversely related to its compensated elasticity of demand. His analysis abstracts from heterogeneity (and hence, a motive for redistribution) and it is assumed that a change in the tax on a specific consumption good does not affect the demand for other goods. Corlett and Hague (1953) consider a more general environment where the cross-price elasticities are not necessarily zero. They show that it is optimal for the government to alleviate tax distortions on labor supply by taxing goods that are more complementary to leisure at higher rates.

In a seminal paper, Atkinson and Stiglitz (1976) show that, in a Mirrleesian setting where the government can levy a nonlinear tax on labor income, weak separability between consumption and leisure in the utility function is enough to render consumption taxes redundant. Deaton (1979) studies a Ramsey problem where the government levies linear taxes on income and different consumption goods. He finds that the result by Atkinson and Stiglitz (1976) stands *provided* Engel curves are linear. In both cases, taxing different goods at different rates does not generate distributional benefits beyond income taxation, nor does it alleviate distortions in labor supply. Our framework with linear tax instruments is most similar to that of Deaton (1979). The key difference is that we endogenize prices by modeling the production side of the economy. As a result, there are general equilibrium effects which the government can exploit for redistributive purposes, even if preferences are weakly separable between consumption and leisure and Engel curves are linear.

The finding that the government can use general equilibrium effects to facilitate income redistribution when designing tax policy is not new. Allen (1982) studies optimal linear taxation of labor income in a model with endogenous wages. Stiglitz (1982) analyzes a similar problem where the government can levy a nonlinear tax on labor income. His analysis has recently been extended to a continuum of skill types by Sachs et al. (2020). Rothschild and Scheuer (2013) characterize optimal nonlinear income taxes in an occupational choice model with endogenous wages and overlapping wage distributions. The optimal tax rules derived in these papers differ from those from the classic analysis by Mirrlees (1971) to account for the indirect distributional effects from changes in wages, which alleviate or tighten incentive constraints.

Naito (1999) extends the analysis from Stiglitz (1982) by including different consumption goods in a Mirrleesian setting where the government can optimize non-linear income taxes. Contrary to the finding of Atkinson and Stiglitz (1976), he shows that when factor prices are endogenous, a non-uniform consumption tax can improve welfare even if preferences are weakly separable between consumption and leisure.[5] This is the case if consumption taxes compress the (pre-tax) income distribution. Jones et al. (1997), Slavík and Yazici (2014) and Kina et al. (2020) show that a similar role can be played by taxes on capital income if different types of labor cannot be taxed separately, as in the current paper, and capital (in one or several kinds) exhibits varying degrees of complementarity with different types of labor.

Our paper differs from these analyses, in particular those of Stiglitz (1982) and Naito (1999), in two substantive ways. First, we consider a Ramsey setting where the government sets linear instruments as in Allen (1982) but with heterogeneity within skill types that generates overlapping income distributions as in Rothschild and Scheuer (2013), who both abstract from consumption taxes. This allows us to derive an intuitive expression for the optimal food subsidy, which balances the costs of distorting consumption decisions against the distributional benefits from general equilibrium effects. Second, we attempt to quantify the implications of these general equilibrium effects for the optimal subsidization of food and

[5] Linear Engel curves are not required in the analysis from Naito (1999), because the government can levy a *nonlinear* tax on labor income.

the welfare gains that result from it by calibrating our model to the Chinese economy.

Several papers investigate the effect of consumption taxes in a quantitative model. Peralta-Alva et al. (2018) study the welfare implications of different taxes in low-income countries, but do not consider the role of consumption tax differentiation. Gadenne (2020) investigates the effectiveness of the 'ration shop' system, which allows the Indian government to implement what is essentially a nonlinear tax on certain consumption categories. In this paper, we abstract from such nonlinearities. Bachas et al. (2022) and Doligalski and Rojas (2022) consider the effect of the informal sector on optimal taxation. Doligalski and Rojas (2022) study the implications for income taxation, while Bachas et al. (2022) focus on consumption taxes, as we do. They find that food subsidies are hard to justify on equity grounds, because low-income households spend a larger share of their income in informal stores. Our finding that food subsidies only bring about very modest distributional benefits through general equilibrium effects complements this literature.

2 Model

In this section, we describe the model that is used in the remainder of the analysis. The economy contains individuals of two skill types: low-skilled L and high-skilled H. Each type is present with mass μ_i, where $i \in \{L, H\}$ and the total mass equals one: $\mu_L + \mu_H = 1$. Within each skill type, individuals differ in their efficiency units of labor, or simply productivity. The latter is denoted by θ, which is distributed according to some distribution function $K_i(\theta)$ on the support $\Theta_i = [\underline{\theta}_i, \overline{\theta}_i]$.[6] Both the distribution function and the support are allowed to vary across the two skill types. Production takes place in two sectors $j \in \{a, n\}$, denoting agriculture and non-agriculture, in which returns to scale are constant.[7] Markets are perfectly competitive, so there are no profits. Each worker's skill type is fixed, but she can freely move across sectors. As a result, each type earns a wage rate w_i per efficiency unit of labor that does not depend on the sector of employment. Finally, there is a government that has a preference for redistribution. We consider a Ramsey problem where the government levies linear taxes on the consumption of agricultural goods and labor income to finance a lump-sum transfer and some exogenous spending. In what follows, we describe each of the agents in the economy in more detail. Because the goal is to connect the model to data, choices of functional forms are discussed on the go.

[6] This is in line with what Acemoglu and Autor (2011) refer to as the "canonical model".

[7] In most of what follows, we use agricultural goods and food interchangeably. This is a simplification, since food products are to some extent manufactured, and some agricultural goods are inputs to non-food products. Issues regarding the mapping from consumption to production have been much-discussed in the literature on structural transformation (see, e.g., Herrendorf et al., 2014), so that we suffice with this note.

2.1 Individuals

Individuals have identical preferences over the consumption of agricultural goods c_a, non-agricultural goods c_n and labor supply ℓ. Given taxes and given prices, an individual of skill type $i \in \{L, H\}$, whose wage is w_i, with θ efficiency units of labor solves the following maximization problem:

$$\max_{\{c_a, c_n, \ell\}} \quad V_i(\theta) = u(c_a, c_n) + v(\ell)$$
$$\text{s.t.} \quad p_a(1+\tau_a)c_a + c_n = T + w_i(1-\tau_y)\theta\ell. \tag{1}$$

Utility is separable between consumption and leisure. Sub-utility over consumption and labor supply, in turn, are given by

$$u(c_a, c_n) = \frac{C(c_a, c_n)^{1-\sigma} - 1}{1-\sigma} \quad \text{where} \quad C(c_a, c_n) = [\omega^{\frac{1}{\epsilon}}(c_a - \underline{c}_a)^{\frac{\epsilon-1}{\epsilon}} + (1-\omega)^{\frac{1}{\epsilon}} c_n^{\frac{\epsilon-1}{\epsilon}}]^{\frac{\epsilon}{\epsilon-1}}, \tag{2}$$

$$v(\ell) = \psi \frac{(1-\ell)^{1-\phi}}{1-\phi}. \tag{3}$$

The parameters σ and ϕ regulate the curvature of the consumption aggregate and labor supply in utility, while ψ governs their relative importance. The relative importance of agricultural goods in consumption is regulated by ω, while ϵ determines the degree of substitutability with non-agricultural goods. Finally, $\underline{c}_a$ denotes a subsistence level for the consumption of agricultural goods (i.e., food).

Despite that preferences are potentially non-homothetic due to the subsistence level $\underline{c}_a$, the current specification of the utility function gives rise to linear Engel curves. Due to subsistence spending, the Engel curve for food has a positive intercept. Hence, low-income individuals spend a larger share of their disposable income on food, but the *marginal* propensity to spend on food is the same for everyone. Combined with the assumption that preferences are separable between consumption and leisure, an immediate implication is that the optimal food subsidy is zero in the absence of general equilibrium effects, cf. Deaton (1979).[8] Therefore, any departure from uniform consumption taxation is necessarily driven by general equilibrium effects and not by our specification of preferences.

Turning to the budget constraint, the output price of agricultural goods is denoted by p_a and non-agricultural goods are chosen as the *numeraire*: $p_n = 1$. Furthermore, note that the before-tax wage per unit of effort $w_i\theta$ depends both on the skill type i and an individual's productivity θ. By contrast, the tax instruments are the same for all individuals: they do not depend on, say, skill or sector of employment. Hence, two individuals earning the same income pay the same income taxes – irrespective of their skill type or where they work.

[8] We confirm and explain this finding in more detail in Section 3.

Tax policy consists of a linear tax rate τ_a on agricultural goods and a linear tax rate τ_y on labor income. In addition, T denotes a lump-sum transfer, which can be positive or negative. The latter can also capture that part of labor income is tax exempt, as is the case in many countries, including China.

The assumption that non-agricultural goods are not taxed is without loss of generality. This is because decisions only depend on relative prices. Consequently, any allocation that can be implemented with linear taxes on labor income and *both* consumption goods can also be implemented with only a linear tax on labor income and a linear tax on agricultural goods. We therefore harmlessly normalize $\tau_n = 0$ and account for this choice when we connect the model to the data. A convenient implication is that a preferential tax treatment of agricultural goods versus non-agricultural goods simply corresponds to subsidizing the former: $\tau_a < 0$. Uniform consumption taxation, in turn, corresponds to setting $\tau_a = 0$.

Whenever the solution to the utility maximization problem is interior, the following conditions together with the budget constraint pin down the optimal choices given prices and given tax policy for an individual of type (i, θ):

$$\frac{C(c_{a,i}, c_{n,i})^{-\sigma}}{P} = \frac{\psi(1-\ell_i)^{-\phi}}{w_i(1-\tau_y)\theta}, \tag{4}$$

$$c_{n,i} = \frac{1-\omega}{\omega}(c_{a,i} - \underline{c}_a)(p_a(1+\tau_a))^{\epsilon}, \tag{5}$$

where P is a price aggregator, or index:

$$P = \left(\omega(p_a(1+\tau_a))^{1-\epsilon} + (1-\omega)\right)^{\frac{1}{1-\epsilon}}. \tag{6}$$

The first of these equates the marginal rate of substitution between leisure and the consumption aggregate to the relative price, which depends on both the wage w_i and individual productivity θ. The second determines the optimal mix between agricultural and non-agricultural goods in the consumption aggregate.

2.2 Firms

Production in sector $j \in \{a, n\}$ takes place according to a constant returns to scale production function:

$$Y_j = F_j(L_j, H_j) = A_j\left[\gamma_j L_j^{\frac{\rho-1}{\rho}} + (1-\gamma_j)H_j^{\frac{\rho-1}{\rho}}\right]^{\frac{\rho}{\rho-1}}, \tag{7}$$

where L_j and H_j denote the total amount of low-skilled and high-skilled labor used in sector j. Note that total factor productivity A_j and the share parameter γ_j are allowed to vary across sectors. The first determines differences in productivity between sectors and the latter governs differences in the skill intensities used in the production of agricultural and non-agricultural goods. In what follows, we harmlessly normalize $A_n = 1$. For simplicity, the constant elasticity of substitution ρ between the two skill types is assumed to be the same in both sectors.[9]

A representative firm in sector $j \in \{a, n\}$ maximizes profits by choosing the total amounts of low-skilled and high-skilled labor, taking wages and output prices as given. Formally, it solves

$$\max_{\{L_j, H_j\}} \quad \Pi_j = p_j F_j(L_j, H_j) - w_L L_j - w_H H_j. \tag{8}$$

Because there is perfect competition with constant returns to scale, firms make zero profits in equilibrium: $\Pi_j = 0$. As can be seen from equation (8), the government does not levy any taxes on firms.[10] The first-order conditions determine the demand for low-skilled and high-skilled labor in each sector j:

$$w_L = p_j F_{L,j} = p_j A_j \left(\frac{Y_j}{A_j}\right)^{\frac{1}{\rho}} \gamma_j L_j^{-\frac{1}{\rho}}, \tag{9}$$

$$w_H = p_j F_{H,j} = p_j A_j \left(\frac{Y_j}{A_j}\right)^{\frac{1}{\rho}} (1-\gamma_j) H_j^{-\frac{1}{\rho}}. \tag{10}$$

Because workers can freely move between sectors, the wages w_L and w_H and hence the skill premium are the same in both sectors. The latter, in turn, is given by

$$\frac{w_H}{w_L} = \frac{1-\gamma_j}{\gamma_j} \left(\frac{L_j}{H_j}\right)^{\frac{1}{\rho}}. \tag{11}$$

The above relationship makes clear that low-skilled labor is used relatively intensively in the sector where it has a comparative advantage, i.e., where γ_i is highest.

2.3 Government

The government provides a lump-sum transfer T to all individuals and levies proportional tax rates τ_a and τ_y on agricultural consumption and labor income. As explained before, a food subsidy corresponds to $\tau_a < 0$. When choosing its tax instruments, the government aims to maximize the following welfare function:

$$\mathcal{W} = \mu_L \int_{\Theta} \alpha_L(\theta) V_L(\theta) dK_L(\theta) + \mu_H \int_{\Theta_H} \alpha_H(\theta) V_H(\theta) dK_H(\theta). \tag{12}$$

---

[9] This assumption follows examples of multi-sector multi-type models in the literature. See, e.g., Hendricks (2010). We study the implications of having a sector-specific elasticity of substitution ρ_j in a robustness exercise of our quantitative analysis.

[10] This assumption is without loss of generality provided the government cannot levy sector-specific taxes on the different labor inputs.

Here, $\alpha_L(\theta) \geq 0$ and $\alpha_H(\theta) \geq 0$ are exogenous Pareto weights, which, together with concavity in the individual utility function, determine the government's preferences for redistribution. A utilitarian objective corresponds to setting $\alpha_L(\theta) = \alpha_H(\theta) = 1$ for all θ, and a Rawlsian objective corresponds to setting $\alpha_L(\theta) = \alpha_H(\theta) = 0$ for all θ except for the individuals who are worst off (presumably the low-skilled workers with $\underline{\theta}_L$ efficiency units of labor), who are assigned a positive weight. Furthermore, a government that wishes to maximize the expected utility of low-skilled workers sets $\alpha_L(\theta) = 1$ for all $\theta \in \Theta_L$ and $\alpha_H(\theta) = 0$ for all $\theta \in \Theta_H$.

The government consumes an exogenous amount of G units of non-agricultural goods. Its budget constraint is

$$\tau_a p_a \sum_i \mu_i \int_{\Theta_i} c_{a,i}(\theta) dK_i(\theta) + \tau_y \sum_i \mu_i w_i \int_{\Theta_i} \theta \ell_i(\theta) dK_i(\theta) = T + G, \tag{13}$$

where the summation is over skill types $i \in \{L, H\}$. The left-hand side captures total revenues collected from taxing food consumption and labor income, while the right-hand side captures total government spending. The latter consists of the lump-sum transfer that is paid to all individuals and public consumption of non-agricultural goods. In what follows, we first characterize equilibrium given tax policy and then turn to analyze the Ramsey problem of finding the optimal linear taxes.

2.4 Equilibrium

A competitive equilibrium is formally defined as follows:

Definition 1. *A competitive equilibrium consists of consumption and labor supply decisions* $(c_{a,i}(\theta), c_{n,i}(\theta), \ell_i(\theta))_{\theta \in \Theta_i, i \in \{L,H\}}$*, labor inputs* $\{(L_a, H_a), (L_n, H_n)\}$ *and prices* (p_a, w_L, w_H) *such that, given tax policy* (τ_a, τ_y, T)*,*

(i) individuals of each type (i, θ) *maximize utility, cf. equations (1), (4) and (5),*

(ii) firms in sector $j \in \{a, n\}$ *maximize profits, cf. equations (9)–(10),*

(iii) labor and goods markets clear:

$$L_a + L_n = \mu_L \int_{\Theta_L} \theta \ell_L(\theta) dK_L(\theta), \tag{14}$$

$$H_a + H_n = \mu_H \int_{\Theta_H} \theta \ell_H(\theta) dK_H(\theta), \tag{15}$$

$$Y_a = \sum_i \mu_i \int_{\Theta_i} c_{a,i}(\theta) dK_i(\theta), \tag{16}$$

$$Y_n = \sum_i \mu_i \int_{\Theta_i} c_{n,i}(\theta) dK_i(\theta) + G. \tag{17}$$

Equations (14)–(15) give the market-clearing conditions for low-skilled and high-skilled labor, respectively. Equation (16), in turn, gives the market-clearing condition for agricultural goods. Combined with the first-order conditions of households and firms, these market-clearing conditions pin down all equilibrium quantities and prices for a given tax policy (τ_a, τ_y, T). Given a choice of these instruments, government consumption G must then be such that the market-clearing condition (17) for non-agricultural goods holds as well. If that is the case, Walras' law implies the government budget constraint is also satisfied.

3 Optimal food subsidies

This section derives our theoretical results. We start by analyzing how, through its impact on individual utilities and the government budget, an increase in the tax τ_a on agricultural goods (i.e., a reduction in the food subsidy) affects welfare (Lemma 1). Thereafter, we use a specific tax reform (Definition 2) to derive an expression for the optimal tax on agricultural goods (Proposition 1). The details of all derivations can be found in the appendices.

3.1 Welfare effect of lowering the food subsidy

The government chooses tax policy (τ_a, τ_y, T) to maximize social welfare (12), subject to the budget constraint (13). When doing so, it has to to take into account that all equilibrium quantities and prices respond to the tax instruments.[11] Before turning to the optimal tax problem, we start by analyzing how an increase in the tax τ_a on agricultural goods, i.e., a reduction in the food subsidy, affects welfare through its impact on individual utilities and the government budget.

Lemma 1. *Suppose the tax τ_y on labor income and the lump-sum transfer T are optimal. Then, the welfare impact of raising the tax τ_a on agricultural goods is*

$$\frac{\partial \mathcal{W}(\tau_a)}{\partial \tau_a}\frac{1}{\lambda} = \underbrace{\sum_i \mu_i p_a \Big[\mathbb{E}(c_{a,i}) - \mathbb{E}(c_{a,i} g_i)\Big]}_{DE} + \underbrace{\tau_a \sum_i \mu_i p_a \mathbb{E}\left(\frac{\partial c^*_{a,i}}{\partial \tau_a}\right) + \tau_y \sum_i \mu_i w_i \mathbb{E}\left(\frac{\partial(\theta \ell^*_i)}{\partial \tau_a}\right)}_{BE}$$
$$\underbrace{+ \sum_i \mu_i \frac{\partial p_a}{\partial \tau_a}\Big[\tau_a \mathbb{E}(c_{a,i}) - (1+\tau_a)\mathbb{E}(c_{a,i} g_i)\Big] + \sum_i \mu_i \frac{\partial w_i}{\partial \tau_a}\Big[\tau_y \mathbb{E}(\theta \ell_i) + (1-\tau_y)\mathbb{E}(\theta \ell_i g_i)\Big]}_{GE}, \quad (18)$$

where λ is the multiplier on the government budget constraint, the summation is over skill types $i \in \{L, H\}$, the expectation $\mathbb{E}(\cdot)$ of a variable indexed by i is computed using the distribution $K_i(\theta)$, $g_i(\theta)$ denotes the welfare weight for an individual of type (i, θ), and the

[11] Appendix B states the conditions which can be used to determine the impact of the tax instruments on equilibrium quantities and prices.

behavioral responses capture the total *impact of a change in τ_a on agricultural consumption and labor supply (i.e., taking into account the general equilibrium effects on prices). DE stands for 'direct effects', BE for 'behavioral effects' and GE for 'general equilibrium effects'.*

Proof. See Appendix A. □

An increase in τ_a affects both individual utilities and the government budget. Lemma 1 shows that the total welfare impact can be decomposed into three effects. First, a higher tax or lower subsidy on agricultural goods transfers income from individuals to the government budget. We label these 'direct effects'. By how much government revenue increases, depends on the average consumption of agricultural goods of individuals with skill $i \in \{L, H\}$, who differ in their productivity θ. The positive impact on government revenue is captured by the term proportional to $\mathbb{E}(c_{a,i})$. The negative welfare impact of raising the tax burden, which lowers individual utilities, is captured by the term proportional to $\mathbb{E}(c_{a,i}g_i)$, where $g_i(\theta)$ is the welfare weight for an individual of type (i, θ). The latter measures by how much welfare increases if the individual receives an additional unit of after-tax income.[12] The term DE on the first line of equation (18) computes the average of the direct utility and budgetary effects over individuals within skill types and then sums the effects over the two skill types.

Second, a change in the tax τ_a on agricultural goods induces changes in consumption and labor supply. We label these 'behavioral effects'. These behavioral effects, in turn, affect the government budget through so-called fiscal externalities. To illustrate, suppose that individuals, in response to a higher tax on agricultural goods, decide to work less and purchase fewer agricultural goods. Since individuals are optimizing prior to the tax reform, these changes do not affect their utility (this is an application of the envelope theorem). However, decisions to work less or purchase fewer agricultural goods do affect government finances if labor income and agricultural goods are taxed or subsidized. These are the fiscal externalities that individuals do not take into account when making their consumption and labor supply decisions. The budgetary effects due to these fiscal externalities are proportional to τ_a and τ_y, respectively. Importantly, the behavioral responses $\partial c^*_{a,i}(\theta)/\partial\tau_a$ and $\partial\ell^*_i(\theta)/\partial\tau_a$ that show up on the first line of equation (18) capture the *total* impact of an increase in τ_a on agricultural consumption and labor supply. These consist of both the direct effect from a higher after-tax price $p_a(1+\tau_a)$ of agricultural goods, and indirect effects driven by responses

[12]Formally, the welfare weight for an individual of type (i, θ) is $g_i(\theta) = \alpha_i(\theta)u_{n,i}(\theta)/\lambda$, where $u_{n,i}(\theta)$ is the marginal utility of non-agricultural consumption, of which the price is normalized to one. The higher the Pareto weight or the marginal utility of non-agricultural consumption, the higher is an individual's welfare weight.

of the before-tax prices p_a, w_L and w_H.[13] The term BE adds the fiscal externalities over the individual skill types and over the two tax bases.

Third, a higher tax on agricultural goods affects the before-tax prices p_a, w_L and w_H. We label these 'general equilibrium effects'. To illustrate, consider the typical case where an increase in τ_a reduces the demand for agricultural goods, which lowers the before-tax price p_a. A reduction in the demand for agricultural goods also reduces the demand for the labor input that is used relatively intensively in its production, presumably low-skilled labor. Consequently, the wage w_L of low-skilled workers falls relative to the wage w_H of high-skilled workers. A change in any of these prices has two welfare-relevant effects. First, a lower price p_a of agricultural goods reduces the tax base, which generates a budgetary effect that is proportional to τ_a. Changes in equilibrium wages w_L and w_H also affect the tax base, which generate a budgetary effect that is proportional to τ_y. Second, changes in equilibrium prices affect the purchasing power of individuals, which has a direct impact on their utility. A lower price of agricultural goods raises purchasing power and hence, resources available for consumption, whereas a lower wage has the opposite effect. These effects are weighed by the welfare weights $g_i(\theta)$. The term GE on the second line of equation (18) sums the welfare-relevant utility and budgetary effects over skill types and over the price responses.

3.2 A utility-neutral tax reform

Lemma 1 characterizes the total welfare impact of raising the tax τ_a on agricultural goods in terms of population shares, welfare weights and the responses of equilibrium quantities and prices to a change in τ_a.[14] Importantly, this result holds for general utility and production functions. Using the property that our preference specification (2)–(3) gives rise to linear Engel curves, we can also derive an expression for the *optimal* tax τ_a on agricultural goods. This is achieved by studying the welfare effect of a specific tax reform, which is defined next.[15]

Definition 2. *A* utility-neutral tax reform $R = (\mathrm{d}\tau_a, \mathrm{d}\tau_y, \mathrm{d}T)$ *consists of changes in tax instruments, which* — holding prices fixed — *leaves utility and labor supply for* all *individuals unaffected. Suppose the government increases the tax τ_a on agricultural goods by an amount*

[13] As mentioned before, Appendix B states the conditions which can be used to determine the *total* impact of changes in the tax instruments on individual consumption and labor supply decisions, as well as the impact from the tax instruments on equilibrium prices. Equation (53) in Appendix C splits up the total effect into direct and indirect effects.

[14] These are what Chetty (2009) refers to as the 'sufficient statistics' that determine the welfare impact of a policy change.

[15] Jacobs and van der Ploeg (2019) study the same reform in the context of pollution taxation in a framework where prices are fixed (i.e., without general equilibrium effects).

δ, so that $\mathrm{d}\tau_a = \delta > 0$. This requires adjustments in the tax τ_y on labor income and the lump-sum transfer T according to

$$\mathrm{d}\tau_y = -\left(\frac{(1-\tau_y)\zeta}{1+\tau_a}\right)\mathrm{d}\tau_a, \qquad \mathrm{d}T = \left(\frac{\eta}{1+\tau_a} + \frac{\zeta T}{1+\tau_a}\right)\mathrm{d}\tau_a, \tag{19}$$

where η and ζ are the intercept and slope of the Engel curve for agricultural consumption (formally defined in Appendix C):

$$p_a(1+\tau_a)c_{a,i}(\theta) = \eta + \zeta M_i(\theta), \quad \textit{where} \quad M_i(\theta) = w_i(1-\tau_y)\theta\ell_i(\theta) + T. \tag{20}$$

Equation (20) describes the Engel curve for food, which relates total spending on agricultural goods to disposable income $M_i(\theta)$. It is obtained by combining the first-order condition (4) and the household budget constraint (1). Because food is a necessity, the Engel curve has a positive intercept η. The slope ζ measures the marginal propensity to spend on food out of disposable income.

Because individuals have identical, linear Engel curves, it is possible to design a tax reform that, holding prices fixed, leaves the utility of *all* individuals unaffected – irrespective of their skill type i and their productivity θ. Following an increase in the tax τ_a on agricultural goods, this requires a reduction in the tax τ_y on labor income and an increase in the lump-sum transfer T that satisfy equation (19).[16] Importantly, holding prices fixed, this reform also leaves labor supply unaffected: the positive effect from a reduction in the tax on labor income is offset by the negative effect from an increase in the tax on agricultural consumption. As such, in the absence of general equilibrium effects, the *only* impact of the reform is that it leads individuals to substitute away from agricultural toward non-agricultural consumption. This generates a fiscal externality proportional to the tax τ_a on agricultural consumption. The total welfare impact of the reform R is then given by the sum of this budgetary effect, and the utility and budgetary effects that come from general equilibrium effects on prices. We use that insight to derive an expression for the optimal tax on agricultural consumption.

[16]To see why the reform R from Definition 2 leaves utility for all individuals unaffected if prices are kept fixed, write the utility maximization problem for an individual of type (i, θ) as

$$V_i(\theta) = \max_{c_{a,i}(\theta),\ell_i(\theta)} u(c_{a,i}(\theta), T + w_i(1-\tau_y)\theta\ell_i(\theta) - p_a(1+\tau_a)c_{a,i}(\theta)) + v(\ell_i(\theta)), \tag{21}$$

where we use the budget constraint (1) to substitute out for $c_{n,i}(\theta)$. Holding prices fixed (i.e., ignoring general equilibrium effects), the impact of a tax reform on individual utility is, by the envelope theorem,

$$\mathrm{d}V_i(\theta) = u_{n,i}(\theta)\big[\mathrm{d}T - w_i\theta\ell_i(\theta)\mathrm{d}\tau_y - p_a c_{a,i}(\theta)\mathrm{d}\tau_a\big], \tag{22}$$

where $u_{n,i}(\theta)$ denotes the marginal utility of non-agricultural consumption. Next, use equation (19) to substitute out for $\mathrm{d}T$ and $\mathrm{d}\tau_y$. Equation (20) then implies the term in brackets equals zero for all $\mathrm{d}\tau_a$.

3.3 Optimal tax formula

Using Definition 2, we can derive an expression for the optimal tax τ_a on agricultural goods by equating to zero the total welfare impact of the tax reform R. This leads to the following result.

Proposition 1. *At the optimal tax system, the tax τ_a on agricultural goods satisfies*

$$\frac{\tau_a}{1+\tau_a}\sum_i \mu_i \mathbb{E}\left[p_a c_{a,i}(-\varepsilon^c_{ap} - \zeta\varepsilon^c_{aw})\right] = \mathcal{W}^*_{p_a}\frac{\mathrm{d}p_a}{\mathrm{d}\delta} + \sum_i \mathcal{W}^*_{w_i}\frac{\mathrm{d}w_i}{\mathrm{d}\delta}, \tag{23}$$

where ε^c_{ap} and ε^c_{aw} are the compensated elasticities of agricultural consumption to changes in after-tax prices $p^*_a = p_a(1+\tau_a)$ and $w^*_i = w_i(1-\tau_y)\theta$, holding the before-tax prices fixed. Moreover,

$$\frac{\mathrm{d}p_a}{\mathrm{d}\delta} = \frac{\partial p_a}{\partial \tau_a} + \frac{\partial p_a}{\partial \tau_y}\frac{\mathrm{d}\tau_y}{\mathrm{d}\tau_a} + \frac{\partial p_a}{\partial T}\frac{\mathrm{d}T}{\mathrm{d}\tau_a}, \tag{24}$$

$$\frac{\mathrm{d}w_i}{\mathrm{d}\delta} = \frac{\partial w_i}{\partial \tau_a} + \frac{\partial w_i}{\partial \tau_y}\frac{\mathrm{d}\tau_y}{\mathrm{d}\tau_a} + \frac{\partial w_i}{\partial T}\frac{\mathrm{d}T}{\mathrm{d}\tau_a} \tag{25}$$

describe the changes in before-tax prices due to the reform R from Definition 2 that starts from an increase in τ_a by an amount δ, where $\mathrm{d}\tau_y/\mathrm{d}\tau_a$ and $\mathrm{d}T/\mathrm{d}\tau_a$ follow from equation (19), and the terms

$$\mathcal{W}^*_{p_a} = \sum_i \mu_i(1+\tau_a)\left[\frac{\tau_a}{1+\tau_a}\mathbb{E}(c_{a,i}) - \mathbb{E}(c_{a,i}g_i) + \tau_a p_a \mathbb{E}\left(\frac{\partial c_{a,i}}{\partial p^*_a}\right) + \tau_y w_i \mathbb{E}\left(\frac{\partial(\theta\ell_i)}{\partial p^*_a}\right)\right], \tag{26}$$

$$\mathcal{W}^*_{w_i} = \mu_i(1-\tau_y)\left[\frac{\tau_y}{1-\tau_y}\mathbb{E}(\theta\ell_i) + \mathbb{E}(\theta\ell_i g_i) + \tau_a p_a \mathbb{E}\left(\theta\frac{\partial c_{a,i}}{\partial w^*_i}\right) + \tau_y w_i \mathbb{E}\left(\theta\frac{\partial(\theta\ell_i)}{\partial w^*_i}\right)\right] \tag{27}$$

measure the welfare impact of an increase in prices p_a and w_i, respectively. The behavioral effects capture the uncompensated responses of agricultural consumption and labor supply to changes in after-tax prices p^*_a and w^*_i.

Proof. See Appendix C. □

Equation (23) gives an expression for the optimal tax τ_a on agricultural goods. It equates the marginal costs from distorting consumption decisions (on the left-hand side) to the marginal welfare gains from general equilibrium effects on prices (on the right-hand side). Starting with the first, by raising τ_a, the reform R from Definition 2 increases the after-tax price of agricultural goods. This leads individuals to substitute away from agricultural goods toward non-agricultural goods. As mentioned, holding prices fixed, this is the only behavioral response, as individual labor supply is not affected by the reform. A reduction in the consumption of agricultural goods generates a fiscal externality, as individuals do not internalize the impact of their consumption decisions on the government budget. The left-hand side of equation (23) captures the magnitude of this fiscal externality. It is proportional to the tax on τ_a on agricultural goods, and the reduction in agricultural consumption due to the reform R. How large this reduction is, depends on the compensated elasticity of agricultural consumption to the after-tax price of agricultural goods, ε^c_{ap}, and after-tax wages, ε^c_{aw} (recall: the reform R changes τ_a *and* τ_y). Both are formally defined in Appendix C. Given our specification of preferences, it is straightforward to verify that $-\varepsilon^c_{ap} - \zeta\varepsilon^c_{aw} > 0$. Consequently, the fiscal externality due to a reduction in agricultural consumption is positive (negative) if food is subsidized (taxed).

The right-hand side of equation (23) captures the total welfare impact associated with general equilbrium effects on prices due to the tax reform R, which consists of both utility and budgetary effects. In response to the reform, the before-tax price of agricultural goods and wages change by $\mathrm{d}p_a/\mathrm{d}\delta$ and $\mathrm{d}w_i/\mathrm{d}\delta$, cf. equations (24)–(25). To obtain the total impact on welfare, these price responses are multiplied by the sum of utility and budgetary effects of a change in the price p_a of agricultural goods and wages w_i, as given by $\mathcal{W}^*_{p_a}$ and $\mathcal{W}^*_{w_i}$. From equation (26), a change in the price of agricultural goods has a number of welfare-relevant effects. First, a change in p_a affects the tax base, which generates a budgetary effect proportional to τ_a. Second, a change in the price of agricultural goods affects individual utilities through its impact on their purchasing power. The impact on welfare is obtained by multiplying the change in purchasing power and the individual welfare weight $g_i(\theta)$. Third, a change in p_a induces (uncompensated) responses in the consumption of agricultural goods and labor supply.[17] These behavioral responses generate a fiscal externality proportional to τ_a and τ_y, respectively. Changes in equilibrium wages due to the reform R have similar welfare effects: see equation (27). Specifically, a change in w_i also affects the tax base, purchasing power, and individual consumption and labor supply decisions, which in turn have budgetary effects proportional to τ_a and τ_y.

Proposition 1 states that, at the optimum, the marginal costs from distorting consumption decisions are equal to the indirect distributional benefits from changes in prices. We now discuss the implications for optimal food subsidies, with and without price responses.

The case without price responses – If prices are fixed, i.e., $\mathrm{d}p_a/\mathrm{d}\delta = \mathrm{d}w_i/\mathrm{d}\delta = 0$, then Proposition 1 implies that the optimal tax on agricultural goods is $\tau_a = 0$. This confirms the result from Deaton (1979), who shows in partial equilibrium that uniform consumption

[17] In equation (18) from Lemma 1, these responses do not show up in the term GE. Instead they are encapsulated by the term BE, as the behavioral responses $\partial c^*_{a,i}(\theta)/\partial\tau_a$ and $\partial\ell^*_i(\theta)/\partial\tau_a$ capture the *total* impact of a higher τ_a on agricultural consumption and labor supply, which accounts for the indirect effects that go through general equilibrium responses of the tax instruments on prices. See Appendix C for details.

taxation is optimal provided Engel curves are linear. Intuitively, in the absence of general equilibrium effects, the *only* welfare-relevant effect of the reform R is that it leads individuals to substitute away from agricultural toward non-agricultural consumption, which generates a fiscal externality proportional to τ_a. By construction, the reform R is designed such that there are no other effects on individual utilities and the government budget (recall: labor supply is unaffected is by the reform). As a result, if food is taxed, it is possible to combine a *reduction* in the tax τ_a on agricultural goods with adjustments in the tax τ_y on labor income and the lump-sum transfer T, cf. equation (19). Such a reform unambiguously raises welfare, as it increases government revenue while leaving utility for all individuals unaffected. Conversely, if food is subsidized, it is possible to combine an *increase* in τ_a with changes in τ_y and T that raise government revenue without lowering individual utilities. It follows that the optimal tax on agricultural goods is $\tau_a = 0$.

The above discussion makes clear that if prices are fixed and Engel curves are linear, food subsidies cannot meaningfully complement an optimized tax on labor income and lump-sum transfer. Importantly, this is true despite the fact that the government has a preference for redistribution and low-income individuals spend a larger share of their disposable income on food (due to the subsistence level $\underline{c}_a$). *Ceteris paribus*, a higher spending share for low-income individuals calls for a lower tax on agricultural goods compared to non-agricultural goods for redistributive purposes (i.e., $\tau_a < 0$). However, because food is a necessity, the demand for it is less elastic than the demand for non-agricultural goods. *Ceteris paribus*, a lower elasticity calls for a higher tax on agricultural goods compared to non-agricultural goods (i.e., $\tau_a > 0$), cf. Ramsey (1927). As it turns out, if prices are fixed and Engel curves are linear, these effects are precisely off-setting. This explains why the optimal tax on agricultural goods is $\tau_a = 0$ in the absence of general equilibrium effects, despite the fact that low-income individuals spend a larger share of their disposable income on food.

The case with price responses – Next consider the case where prices are endogenous. As shown by Naito (1999) in a Mirrleesian setting, consumption taxes can then be used to exploit general equilibrium effects for redistributive purposes.[18] Proposition 1, in turn, derives an optimal tax formula in a Ramsey setting with linear instruments. If prices are endogenous, the sign and magnitude of the optimal tax τ_a on agricultural goods depends on the total welfare impact from general equilibrium responses to the reform R. To illustrate, consider again the typical case where an increase in τ_a lowers the demand for agricultural goods, which leads to a reduction in its before-tax price: $\mathrm{d}p_a < 0$. In Appendix D, we demonstrate

[18]Taxes on capital income can also serve this purpose if changes in the capital stock impact the skill premium. See, e.g., Jones et al. (1997), Slavík and Yazici (2014) and Kina et al. (2020).

that the corresponding impact on equilibrium wages is given by

$$\mathrm{d}w_L \left(\frac{L_a}{H_a} - \frac{L_n}{H_n} \right) = \left(\frac{Y_a}{H_a} \right) \mathrm{d}p_a, \quad \mathrm{d}w_H = -\left(\frac{L_n}{H_n} \right) \mathrm{d}w_L, \tag{28}$$

where Y_a denotes total production in the agricultural sector. If the production of food is relatively low-skilled labor intensive, i.e., if $L_a/H_a > L_n/H_n$, then a reduction in the price of agricultural goods leads to a decrease in the wage w_L for low-skilled workers and an increase in the wage w_H for high-skilled workers.[19] With our specification of the production function, this is the case if $\gamma_a > \gamma_n$, i.e., if low-skilled labor has a comparative advantage in the production of food: see equation (11). By lowering the before-tax price p_a of agricultural goods, an increase in the tax τ_a leads to a reduction in the wage w_L of low-skilled workers and an increase in the wage w_H of high-skilled workers. How large these effects are depends on how easy it is for individuals to substitute between agricultural and non-agricultural consumption, and how easy it is for firms to substitute between high-skilled and low-skilled labor.

If the combined sum of utility and budgetary effects from a reduction in the wage w_L of low-skilled workers and the price p_a of agricultural goods, and an increase in the wage w_H of high-skilled workers due to the reform R is negative, then according to equation (23) it is optimal to subsidize food: $\tau_a < 0$. While we do not have a formal result on the sign of the optimal τ_a, we expect this is the relevant case if the government wishes to redistribute on average from high-skilled to low-skilled workers, and low-skilled workers have a comparative advantage in the production of food.[20] Hence, we conjecture, and verify numerically, that the optimal $\tau_a < 0$ if $\gamma_a > \gamma_n$. By imposing a food subsidy, the government raises the demand for low-skilled labor, which reduces the skill premium and indirectly redistributes income from high-skilled to low-skilled workers. *Ceteris paribus*, this form of indirect redistribution has a positive impact on welfare. According to equation (23), these indirect distributional benefits that come from general equilibrium effects (on the right-hand side) should be weighted against the marginal costs from distorting consumption decisions (on the left-hand side). This trade-off fundamentally determines the optimal τ_a. How large the optimal tax on agricultural goods is once we account for general equilibrium effects is the topic of the following sections.

[19] This is essentially an application of the Stolper-Samuelson theorem, with the increase in the relative demand for non-agricultural goods driven by an increase in the tax τ_a rather than an opening to trade.

[20] The difficulty lies in determining the sign of the right-hand side of equation (23), which involves many terms. Despite this, if the government has a preference for redistributing from high-skilled to low-skilled worker, the negative welfare impact from a reduction in the purchasing power of low-skilled workers driven by a decrease in w_L, captured by $\mathbb{E}(\theta \ell_L g_L)$ in equation (27), is typically larger than the positive welfare impact from an increase in the purchasing power of high-skilled workers driven by an increase in w_H, as captured by $\mathbb{E}(\theta \ell_H g_H)$. This explains why we expect that the combined welfare impact from changes in the price of agricultural goods and wages due to the reform R and hence, the optimal τ_a is negative.

4 Parameterization

We aim to have our model represent the Chinese economy in the year 2008 as well as possible. To do so, we combine a variety of data sources, most importantly the 2008 sample of the Chinese Household Income Project (CHIP). Below we discuss the CHIP data and describe the moments of the data that we target in our calibration.

4.1 Data

Our main source of data is the Chinese Household Income Project (CHIP), a household micro-survey on earnings, expenditures, and personal characteristics such as education levels. We choose to work with the 2008 wave of the survey, because it is the latest at the time of writing to include questions on consumption expenditure by spending category.

Gustafsson et al. (2014) describe the income data available for China. The methodology used by the Chinese National Bureau of Statistics (NBS) to create their Annual Urban and Rural Household Surveys is the best available, even if the measurement of high incomes remains challenging. The CHIP survey is a sub-sample of this larger survey.

Two additional reasons for choosing the CHIP survey deserve mention. First, consumption-in-kind by farmers (i.e., the direct consumption of their own produce, or consumption through barter trade) makes for a measurement issue that is relevant to our research setup. Such consumption should be accounted for. Luckily, the NBS splits the survey by urban and rural populations and adjusts its methods to account for consumption-in-kind.

Second, China has a large population of migrant workers, who are registered as inhabitants of rural villages but spend most of the year living in urban areas. In addition, workers sometimes live at their place of work, for example on a construction site, rather than in a residential building. These phenomena could distort statistics on sectoral employment, but the CHIP surveys account for this as well.

4.2 Moments

We now discuss the data moments that inform our parameter choices, all of which are summarized in Table 2.

First, we obtain some statistics on employment and consumption from CHIP 2008. We drop all unemployed subjects from the rural and the urban sample and classify the remaining by sector (employed in agriculture, versus outside) and education (highest level of education completed is college or above, versus below). Sector classification is done based on where individuals work most hours.[21] According to the World Bank's World Development Indicators (December 2017 update, hereafter WDI), 53.5% of the population live in rural areas. We use this figure to weigh the CHIP's rural and urban samples.

Table 1 shows how employment is split over sectors and education levels. The share of agricultural employment in total employment is 26.6%. The total share of college educated workers is 9.8%, but they are disproportionately employed outside of agriculture. Comparing skill intensities across sectors, we find that the share of low-skilled hours in agriculture is 14% higher than in non-agriculture.[22] Matching these three figures implies matching the entire two-by-two table. In addition, we target the college premium. Wang (2012) investigates the topic in depth and finds a (non-causal) college premium of 51% using data from CHIP 2002.

	College	Non-College	Total
Agriculture	0.3%	26.4%	26.6%
Non-agriculture	9.6%	63.8%	73.4%
Total	9.8%	90.2%	100.0%

Table 1: Employment shares by sector and education level

To capture the heterogeneity in earnings within our two broad education groups, we assume that individual productivity θ in both education groups follows a shifted log-normal distribution:

$$\theta \sim \underline{\theta} + \log \mathcal{N}(\nu_i, \xi_i), \tag{29}$$

of which the average value is normalized to one for both i: $\mathbb{E}(\theta) = 1$. We set the lower bound $\underline{\theta}$, assumed common across both skill types, in such a way that all individuals could earn one percent less and still consume the subsistence level of food $\underline{c}_a$ at current equilibrium prices and labor supply.[23] This level depends on both individual choices and equilibrium prices, and is intended to bring about a feasible solution in a parsimonious way. The CHIP data allow us to obtain wages from income and hours for individuals of both educational groups. We divide individual wages by the average wage within the educational group, which yields the empirical counterpart of the individual θ's. We then calculate the standard deviations of these normalized wages, and use them as targets to pin down ν_i and ξ_i, together with the

[21] Subjects can indicate several forms of employment, but few subjects show a mixed profile of hours. This is in line with findings from Gollin et al. (2013).

[22] This figure is calculated as $\frac{26.4}{26.4+0.3} / \frac{63.8}{63.8+9.6} \approx 1.14$.

[23] Because earnings increase in productivity and $w_H > w_L$, this constraint binds for the low-skilled workers with $\underline{\theta}$ efficiency units of labor.

normalization that each θ has an expected value of one.[24]

The CHIP data also includes information on consumption expenditure for the urban population. We use the corresponding figures to determine how spending on food and non-food varies with total expenditures. To that end, we regress food expenditure on non-food expenditure. Both are normalized by average monthly urban households expenditure, so that the data can be linked to the model. The resulting coefficients are informative of the parameters ω and $\underline{c}_a$.[25] The intercept is positive and significant, indicating a minimum required spending on food, and food spending rises less than one-for-one with non-food spending. Figure 1 plots food spending against total spending. Not surprisingly, the graph shows a clear positive relationship. It also shows substantial heterogeneity in food spending conditional on total spending that we cannot match using our framework with homogeneous preferences. However, taking into account preference heterogeneity generates different motives for commodity tax differentiation (see, e.g., Saez, 2002), from which we like to abstract. Similarly, to focus exclusively on the implications of general equilibrium effects for optimal food subsidies, we restrict our attention to linear Engel curves.

Unfortunately, the data from the expenditure section of the CHIP 2008 survey do not align

[24]This normalization implies $\mathbb{E}(\theta - \underline{\theta}) = 1 - \underline{\theta}$ for both i. Using the property that the standard deviation of a shifted lognormal distribution is the same as that of a regular lognormal distribution, it is then straightforward to derive the mean and standard deviation of θ in terms of ν_i and ξ_i:

$$\mathbb{E}(\theta) = \underline{\theta} + \exp\left(\nu_i + \xi_i^2/2\right), \quad \sqrt{\mathbb{V}(\theta)} = \sqrt{(\exp\left(\xi_i^2\right) - 1)}(1 - \underline{\theta}).$$

[25]The reason we choose this specification is that we can translate the estimates from this regression into model equivalents. Specifically, we observe the following in the data: $c_a p_a(1+\tau_a)(1+\tau)$, which is total spending on food after subsidies and the general VAT rate of 17% that we denote by τ, and $c_n p_n(1+\tau)$, which is total spending on non-food (not imposing $p_n = 1$). The optimality conditions that follow from the individual's maximization problem imply the following relationship:

$$c_a p_a(1+\tau_a)(1+\tau) = \underline{c}_a p_a(1+\tau_a)(1+\tau) + c_n p_n(1+\tau)\frac{\omega}{1-\omega}(p_a(1+\tau_a))^{1-\epsilon} p_n^{\epsilon-1}.$$

To translate this into model terms, we normalize the data by $\tilde{c}_u$, which denotes the observed average total consumption of the urban population. This normalization affects the intercept, but not the slope. Let S_H^u denote the share of the urban population that is high-skilled. $\tilde{c}_u$ can be expressed as:

$$\tilde{c}_u = (1 - S_H^u)(c_{L,a} p_a(1+\tau_a)(1+\tau) + c_{L,n} p_n(1+\tau)) + S_H^u(c_{H,a} p_a(1+\tau_a)(1+\tau) + c_{H,n} p_n(1+\tau)).$$

Estimates of the intercept and slope (denoted $\hat{\beta}_0$ and $\hat{\beta}_1$) are then related to the model as follows:

$$\hat{\beta}_0 = \underline{c}_a p_a(1+\tau_a)(1+\tau)/\tilde{c}_u, \hat{\beta}_1 = \frac{\omega}{1-\omega}(p_a(1+\tau_a))^{1-\epsilon} p_n^{\epsilon-1}.$$

Hence, the results can be translated into the parameter values we are interested in given an 'extra' calibrated value of τ, which we found to be 17%, and of S_H^u, which we found to be 19.6%.

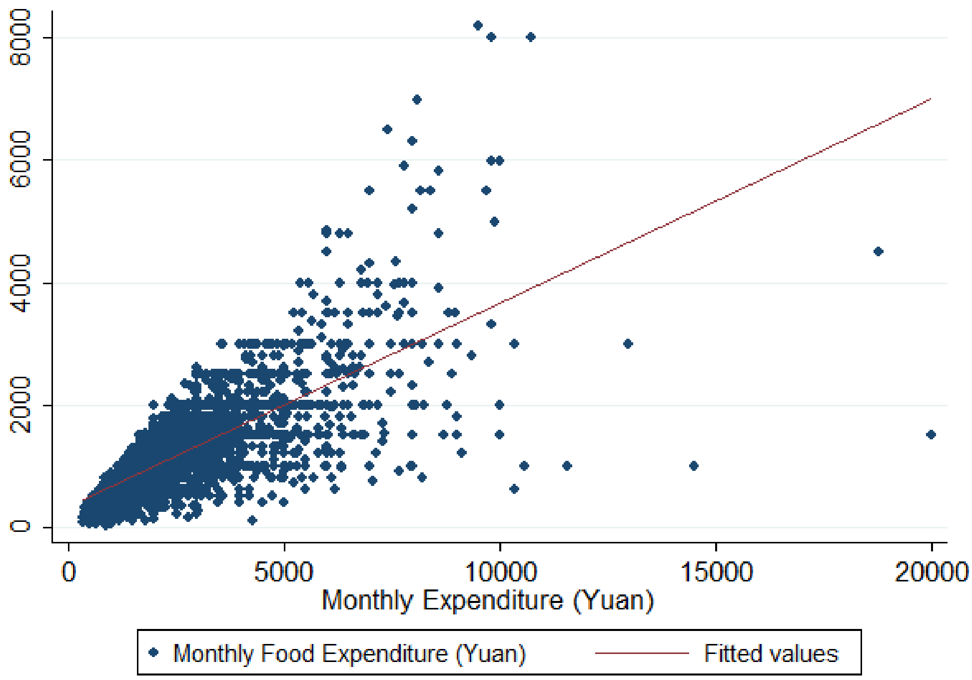

Figure 1: Engel curve based on CHIP data

with the NBS's preferred macro estimates of expenditures. In the CHIP survey, about half of aggregate expenditure goes to food products. Other data sources, such as the International Comparison Program 2011 data (by the World Bank) show an expenditure share closer to one quarter.[26] We prefer the calibrated model to be consistent with macro aggregates, and therefore choose to shift downward the linear Engel curve by reducing the intercept to match an expenditure share in line with this figure. The result is an intercept that matches the first percentile of all food spending in the CHIP data, which we consider a reasonable proxy for subsistence spending.

Next, we turn to the role of the government. Regarding the tax on agricultural goods, the OECD (2017) produces agricultural support estimates, in which it records general services support, consumer support, and producer support for agriculture by governments. General services do not include direct benefits to producers or consumers (but rather concern infrastructure), so we exclude those. Consumer support for agricultural goods is negative, but does not seem to diverge much from that of other goods (in terms of VAT and Business Tax). Producer support, instead, is sizable: it hovers around 10% of gross farm receipts.[27] We take

[26]We total food, beverages, tobacco and narcotics (which is the most common level of aggregation) and compare this to total expenditures, arriving at a share of 26.1%.

[27]The year 2008 seems to have been a downward outlier at 4.62%, so that we choose a figure that is more representative for the period.

this as the equivalent of 'food subsidies': $\tau_a = -0.1$.[28]

Chinese personal income taxes apply only after a personal allowance that is twice the average wage and make up for a small percentage of government revenues (IMF, 2018 and Lin, 2009). Instead, the standard VAT rate is significant at 17% (IMF, 2018). This figure implies an effective tax rate on labor income of $\tau_y = 1 - 1/(1 + 0.17) \approx 0.15$. Furthermore, according to the WDI, the total revenue from taxes was 10.1% of GDP in 2008. Through the lens of our model, this figure is informative about government consumption of non-agricultural goods G. The value of the lump-sum transfer T is residually determined to make sure the government runs a balanced budget.[29]

We also require values for a number of elasticities. Unfortunately, for many of these, no specific estimates are available for China. For this reason, we rely on the idea that preferences over consumption and labor supply and substitutability between high-skilled and low-skilled labor is similar across countries. We set the coefficient of relative risk-aversion σ equal to one (logarithmic utility). Moreover, we use the parameters ϕ and ψ governing the disutility of labor to target an average Frisch elasticity of 0.5 and an average Hicksian elasticity of 0.33, as suggested by Chetty et al. (2011). Furthermore, we set the elasticity of substitution between agricultural and non-agricultural goods in consumption equal to 0.5, following Buera and Kaboski (2009). As this parameter determines to what extent individuals change their consumption mix following a change in the food subsidy (and hence, to what extent there will be a change in demand for low-skilled labor), we study the robustness of our findings with respect to this parameter choice in Section 5.

Lastly, the elasticity of substitution between skill types in the production function is of particular importance, as it determines the strength of general equilibrium effects on wages. We use a value of 1.4 in our baseline, based on estimates for the US by Katz and Murphy (1992) and Ciccone and Peri (2005). Section 5 analyzes the robustness of our results with

[28]While data on agricultural subsidies are available from the OECD, less is known about subsidies in other sectors of the Chinese economy. There is, however, evidence that suggests these may be significant in certain industries (Barwick et al., 2019). We therefore performed an additional parameterization of our model economy as a robustness check, where we set subsidies to the agricultural sector to zero (instead of the 10% we use in our baseline). Using this this alternative parameterization does not affect our conclusions up to minor quantitative differences.

[29]We interpret the figure of 10.1% as revenues after transfers, so that it equals government expenditures on non-agricultural goods: $G/\text{gdp} = 0.101$. In line with our theoretical model, we assume government spending is financed by labor and consumption taxes. In reality, government spending is also financed by other taxes, e.g., on corporate income. Actual government expenditures are less informative, because of significant other sources of government revenue (in particular, income from state-owned enterprises). As a result, actual government expenditures are significantly larger than tax revenues suggest. On balance, we believe our approach is the most appropriate, parsimonious way of capturing the overall tax burden.

respect to this parameter choice as well, and to having a different elasticity of substitution between skill types in agriculture and non-agriculture.

Parameter	Value	Moment	Model	Data
μ_L	**0.90**	Share of low-skill workers	90.2%	90.2%
σ	**1.01**	Coefficient of relative risk aversion	1.01	1.01
ϕ	1.76	Average Frisch elasticity	0.50	0.50
ψ	0.56	Average Hicksian elasticity	0.32	0.33
ω	0.39	Slope of expenditure regression	0.25	0.25
ϵ	**0.50**	Consumption elasticity of substitution	0.50	0.50
$\underline{c}_a$	0.13	Food share of expenditures	26.3%	26.1%
A_a	3.71	Agriculture share of employment	26.4%	26.6%
γ_a	0.94	Relative low-skill intensity across sectors	1.14	1.14
γ_n	0.72	College premium	1.51	1.51
ρ	**1.40**	Elasticity of substitution between types	1.40	1.40
τ_a	**-0.10**	Agricultural producer support as % of receipts	10.0%	10.0%
τ_y	**0.15**	Effective tax burden on labor income	14.5%	14.5%
T	0.00	Tax revenue as % of GDP	10.5%	10.1%
$\underline{\theta}$	0.05	1% above subsistence level	1.01%	1.00%
ξ_L	0.81	Standard deviation of normalized low-skill wages	0.91	0.91
ξ_H	0.71	Standard deviation of normalized high-skill wages	0.76	0.76
ν_L	-0.38	Expected value of individual productivity	1.00	1.00
ν_H	-0.30	Expected value of individual productivity	1.00	1.00

The table informally groups the data moments with the parameters they are considered informative of, although each parameter influences many moments. Of the 19 parameters, 6 are set directly to match their empirical counterpart. These 'outside' parameters are indicated in boldface. (The last four parameters can 'almost' be set directly, as they depend only on the 'inside' parameter $\underline{\theta}$).

Table 2: Parameters and moments

Table 2 shows the parameters we use to produce the results in the next sections. For these values, our model closely matches the data moments.

5 Quantitative analysis

This section presents the quantitative results from our optimal tax analysis. It begins with our main result: general equilibrium effects rationalize very modest food subsidies that generate tiny welfare gains. We then investigate the robustness of this result by changing the government's preferences for redistribution, by changing some of the key parameters, and by imposing restrictions on the government's instrument set.

5.1 Main results

Our optimal tax analysis attempts to shed light on two questions. The first is: for a given specification of the welfare function, what is the optimal food subsidy? Because in our model any departure from uniform consumption taxes is driven by general equilibrium effects, the answer to this question gives an indication of the size of food subsidies that can be rationalized by such effects. The second question is: what are the welfare costs of setting uniform consumption taxes, i.e., of setting $\tau_a = 0$? The answer to this question gives an indication of how costly it is to abandon food subsidies or, alternatively, the welfare gains that can be reaped from optimizing them.

Optimal tax policy

We numerically solve the government's optimization problem, as formally described in Appendix C.[30] Table 3 shows the optimal taxes along some statistics on the resulting allocation. 'Baseline' refers to the calibrated model, which serves as a comparison to two sets of results under a utilitarian criterion, obtained by setting $\alpha_L(\theta) = \alpha_H(\theta) = 1$ for all θ. The column 'Optimal' shows the results if no restrictions are imposed on food subsidies. By contrast, the results under 'Uniform' are obtained by imposing that food subsidies are abandoned: $\tau_a = 0$, i.e., all goods are taxed at the same rate.

	Baseline	Utilitarian	
		Optimal	Uniform
Tax on agricultural goods (τ_a)	–10.00%	–1.29%	0.00%
Tax on labor income (τ_y)	14.53%	50.43%	50.27%
Transfer/average income ($T/\bar{y}$)	1.61%	37.40%	37.57%
Labor supply (% vs. baseline)	N/A	–17.86%	–17.84%
Food consumption (% vs. baseline)	N/A	–17.30%	–17.55%
Total output (% vs. baseline)	N/A	–17.28%	–17.27%
Skill premium	51.41%	44.09%	44.21%
Agriculture share in GDP	25.43%	25.71%	25.62%
Share of low-skilled hours in agriculture	29.11%	29.48%	29.39%
Consumption equivalent gain vs. Baseline	N/A	19.54%	19.54%
Consumption equivalent gain vs. Uniform	N/A	1E-3%	

Table 3: Optimal policy under a utilitarian criterion

Under the utilitarian criterion, the optimal food subsidy is 1.29%. The preferential tax treat-

[30] Additional details and Matlab codes are available upon request.

ment of food relative to the untaxed numeraire (non-agricultural goods) is therefore very modest, and much smaller than in the baseline. Labor income, in turn, is taxed at a rate of 50.43%. The proceeds from labor taxes are used to finance a sizable lump-sum transfer that corresponds to 37.40% of average income. Compared to the baseline, the planner thus sets a much higher tax on labor income to finance a significantly larger transfer. Due to higher labor taxes and a larger transfer, aggregate labor supply, food consumption and total output are approximately 17–18% smaller in the utilitarian optimum than in the calibrated economy, cf. rows 4–6.

The column 'Uniform' shows the results under the often-voiced advice that all goods should be taxed at the same rate and hence, food subsidies should be abandoned. To that end, we impose the restriction $\tau_a = 0$ and solve for the optimal labor tax and lump-sum transfer. The optimal tax rate on labor income is slightly lower and the transfer is slightly larger when this restriction is imposed. Intuitively, the government uses the savings from abandoning food subsidies to reduce labor market distortions (by reducing the tax on labor income) and to generate some additional redistribution (by raising the lump-sum transfer).

Rows 7–9 illustrate the mechanism through which food subsidies improve welfare. Compared to the policy under uniform taxation, the optimal food subsidy generates a small reduction in the skill premium of 0.12 percentage points, from 44.21% to 44.09%, which indirectly redistributes income from high-skilled to low-skilled workers. The reduction in the skill premium is driven by an increase in the share of the agricultural sector in GDP brought about by food subsidies: 25.71% versus 25.62% under uniform taxation. In line with the rise of the agricultural sector, the share of low-skilled workers who are employed in agriculture also increases slightly, from 29.39% to 29.48%. The fact that all these differences are small reflects that a utilitarian planner only sets a small subsidy on food of 1.29%.

The skill premium is significantly smaller in the utilitarian optimum than in the baseline: 44.09% versus 51.41%. Importantly, this is true *despite* the fact that food subsidies are larger in the baseline. The reason is that a utilitarian planner sets a much higher tax on labor income to finance a substantially larger transfer than in the calibrated economy. The larger transfer, in turn, generates negative income effects on labor supply for both high-skilled and low-skilled workers. The reduction in hours worked is larger for low-skilled workers. This is because, in relative terms, the increase in the lump-sum transfer has a larger effect on their disposable incomes. A larger reduction in labor supply for low-skilled workers, in turn, raises their marginal productivity, which lowers the skill premium. Quantitatively, this effect turns out to be stronger than the positive effect from smaller food subsidies.

To isolate the effect of food subsidies, Figure 2 shows the skill premium for different values of τ_a, starting from the baseline (indicated by the vertical line). For each of these, the tax on labor income is held constant, while the transfer clears the government's budget constraint. In line with the mechanism outlined in Section 3, a reduction in food subsidies, i.e., an increase in τ_a, increases the skill premium. The effect, however, is not large: reducing the food subsidy from the baseline value of 10% to the optimal value of 1.29% generates an increase in the skill premium of less than one percentage point.

This effect is more than offset by the negative effect of the large increase in the labor tax and lump-sum transfer that occurs if the government simultaneously optimizes labor taxes. In that case, the skill premium is *reduced* by 7.32 percentage points: see Table 3. This reduction is partly brought about by a higher wage of low-skilled workers, which increases by 0.92%. Equation (28) then implies the wage of high-skilled workers is reduced by 3.96%. The increase in the wage of low-skilled workers also leads to a higher before-tax price of agricultural goods (of about 0.84%), which are primarily produced using low-skilled labor, again cf. Equation (28). As explained, these general equilibrium effects are driven primarily by the increase in the labor tax and the lump-sum transfer, which more than offset the effect of a reduction in food subsidies that works in the opposite direction.

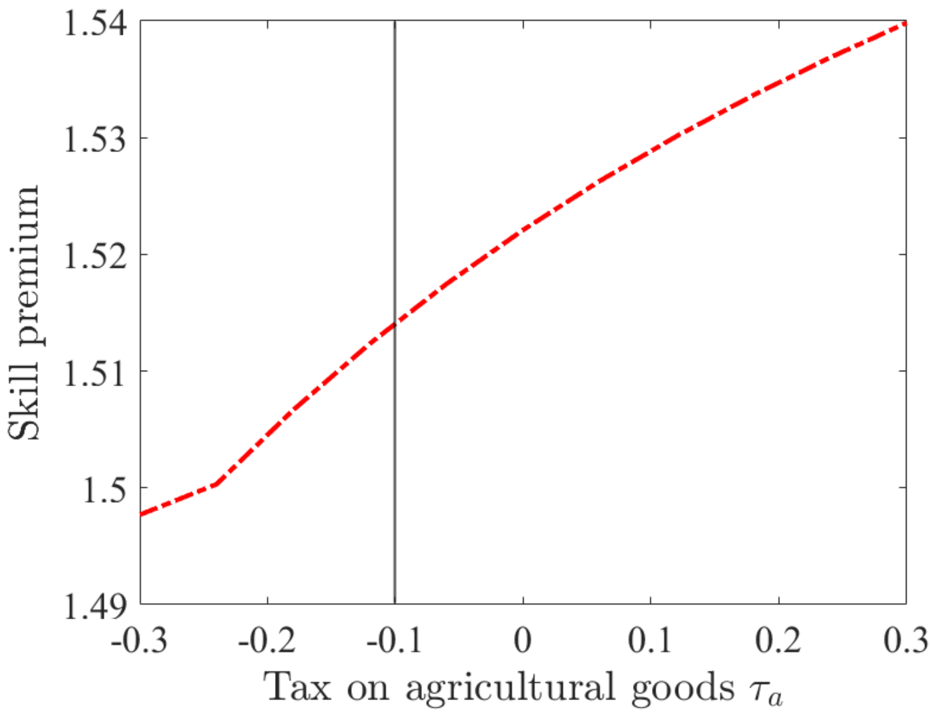

Figure 2: Skill premium at different levels of the food subsidy

Welfare gains

The last two rows of Table 3 show the welfare gains that can be obtained from optimizing taxes. Starting from the baseline, the government is indifferent between optimizing all tax instruments and increasing the consumption aggregate $C(c_{a,i}(\theta), c_{n,i}(\theta))$ of all individuals by 19.54%.[31] These gains come almost entirely from jointly optimizing the tax on labor income and the lump-sum transfer, rather than from optimizing food subsidies. In particular, fixing the food subsidy at its baseline $\tau_a = -0.1$ and letting the government jointly optimize over τ_y and T generates a welfare gain of 19.50% (not displayed). The final row compares the unconstrained optimum to the constrained optimum under uniform consumption taxes, i.e. under the restriction $\tau_a = 0$. As it turns out, the welfare costs of following a uniform consumption tax policy are negligible. In particular, starting from a setting where income taxes and the lump-sum transfer are optimized but $\tau_a = 0$, optimizing food subsidies generates a welfare gain much smaller than 0.01% in consumption equivalents. This very small welfare gain reflects that the optimal food subsidy is only 1.29%.

The finding that food subsidies generate very small welfare gains does not necessarily imply that the welfare costs of setting suboptimal food subsidies are also small. To investigate this further, we fix τ_a at a particular value and optimize with respect to the labor income tax and the lump-sum transfer. We then measure by how much welfare increases if the government could also optimize food subsidies. Figure 3 displays the results. Naturally, the welfare costs are zero if τ_a is restricted at its optimal value of –1.29%. Furthermore, the costs of following a uniform consumption tax policy with $\tau_a = 0$ are much smaller than 0.01% in consumption equivalent gains, cf. Table 3.[32] The figure shows that, unless severely mis-optimized, the welfare costs of setting suboptimal food subsidies are modest, typically well below 0.5% in consumption equivalents.

5.2 Robustness

Our optimal tax analysis indicates that general equilibrium effects rationalize very modest food subsidies that generate tiny welfare gains. We now demonstrate that this finding is robust to changing the government's preferences for redistribution and some of the key parameters, in particular the elasticity of substitution between high-skilled and low-skilled

[31]Note from equation (2) that increasing an individual's consumption aggregate $C(c_{a,i}(\theta), c_{n,i}(\theta))$ by x% is equivalent to increasing her consumption of non-agricultural goods $c_{n,i}(\theta)$ and agricultural goods *net of the subsistence level* $c_{a,i}(\theta) - \underline{c}_a$ by x%. The required increase in *gross* consumption $c_{n,i}(\theta)$ and $c_{a,i}(\theta)$ is slightly smaller.

[32]The welfare costs of setting $\tau_a = -0.1$, in turn, is approximately 0.04%, which equals the difference between the 19.54% and 19.50% consumption equivalent gains mentioned above.

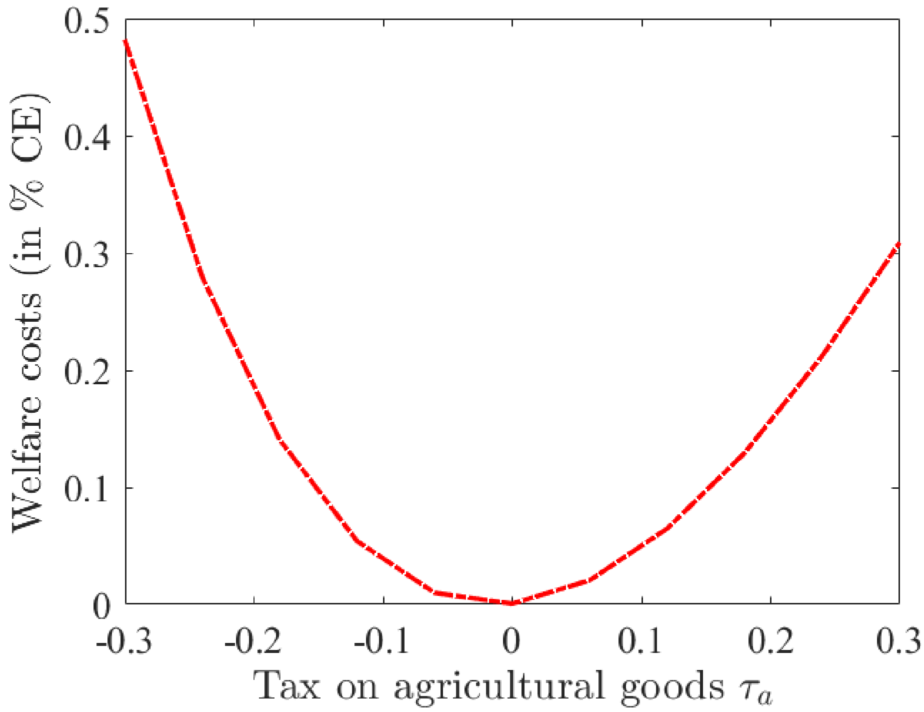

Figure 3: Welfare costs of suboptimal food subsidies

labor in production (ρ), the elasticity of substitution between agricultural goods and non-agricultural goods in consumption (ϵ), and subsistence consumption of agricultural goods ($\underline{c}_a$). We do find a significant role for subsidies (or taxes) on food if the government cannot adjust the lump-sum transfer (or labor income tax) directly. However, the government then primarily uses food subsidies as a partial substitute for the instrument that cannot be optimized, rather than as a tool to indirectly redistribute income through general equilibrium effects.

Non-utilitarian preferences

Our benchmark reports results for a utilitarian criterion. What if policymakers have stronger preferences for redistribution? We investigate this by considering the case where a Rawlsian government attaches a positive weight only to individuals with the lowest earnings, i.e., the low-skilled workers with the lowest productivity. Perhaps surprisingly, we find that optimal food subsidies are very close to zero. In fact, the welfare function is so flat around this point that it becomes numerically difficult to find an exact optimum.

The reason that a Rawlsian planner does not make use of substantial food subsidies is that low-skilled individuals with a very low productivity θ hardly benefit from an increase in their wage w_L. The main source of income for the worst-off individuals is the lump-sum transfer rather than their labor income. Using food subsidies to indirectly redistribute income by reducing the skill premium is therefore of very little use to a Rawlsian planner. Instead, the Rawlsian government aims to maximize tax revenues by setting a labor income tax of approximately 70% and uses the proceeds to finance a substantial lump-sum transfer.[33]

Optimal food subsidies are somewhat larger if the government has a very strong preference to indirectly redistribute income by lowering the skill premium. To illustrate this, we consider a government that attempts to maximize the expected utility of low-skilled workers: $\alpha_L(\theta) = 1$ for all $\theta \in \Theta_L$ and $\alpha_H(\theta) = 0$ for all $\theta \in \Theta_H$. Hence, the government has a 'Hybrid' objective with maxi-min (Rawlsian) preferences *between* skill types, but with equal (utilitarian) weights to all individuals *within* each skill type, who differ in their productivity. Table 4 shows that even with these preferences, optimal food subsidies remain quite small at 5.44%. These food subsidies generate a 0.48 percentage point reduction in the skill premium compared to the uniform tax optimum with $\tau_a = 0$, from 43.26% to 42.78%. This increase is larger than the 0.12 percentage points from the benchmark (see Table 3), though still modest. Similarly, despite generating substantially larger welfare gains than in the utilitarian benchmark (see Table 3), the optimal food subsidy still only brings about an additional 0.01% consumption equivalents compared to the uniform tax optimum.

	Baseline	Hybrid	
		Optimal	Uniform
Tax on agricultural goods (τ_a)	–10.00%	–5.44%	0.00%
Tax on labor income (τ_y)	14.53%	52.48%	51.94%
Transfer/average income ($T/\bar{y}$)	1.61%	38.54%	39.42%
Skill premium	51.41%	42.78%	43.26%
Consumption equivalent gain vs. Uniform	N/A	0.01%	

Table 4: Optimal policy when $\alpha_L(\theta) = 1$ and $\alpha_H(\theta) = 0$ for all θ

Strength of general equilibrium effects (ρ)

How wages respond to a change in the food subsidy critically depends on the elasticity of substitution between high-skilled and low-skilled labor, as captured by ρ. If it is very easy

[33] Food subsidies only become part of the optimal policy mix if the worst-off individuals are more dependent on labor income. To illustrate, suppose we move the individuals with the lowest income further from the subsistence constraint by exogenously tripling the lower bound $\underline{\theta}$ of the productivity distribution. In this case, the optimal food subsidy is approximately 1.60%, which exceeds the optimal food subsidy of 1.29% in the utilitarian case.

to substitute between skill types (i.e., if ρ is large), the skill premium hardly responds to a change in the food subsidy. In our baseline calibration, we use an elasticity of substitution of $\rho = 1.4$. This figure is based on Katz and Murphy (1992), who relate changes in the skill premium to changes in the supply of college and non-college graduates in the US. Ciccone and Peri (2005) use variation in child labor and compulsory school attendance laws across US states and arrive at a very similar estimate of around 1.5. Evidence from the immigration literature points to somewhat larger values. In a survey on the topic, Card (2009) suggests a value between 1.5 and 2.5. Findings for less-developed countries also fall within that range. For example, Angrist (1995) and Behar (2009) find a value of approximately two.

Because our results are likely to be sensitive to this elasticity, we redo our analysis with different values of ρ. The remaining parameters are then set to match the moments outlined in Table 2. Figure 4 plots (on the left axis) the optimal food subsidy for different values of $\rho \in [1.1, 3.5]$. The vertical line shows the baseline with $\rho = 1.4$. The optimal food subsidy is declining in the elasticity of substitution ρ, though the effect is not large. For plausible values of ρ up to 2.5, the optimal food subsidy does not exceed 1.5% and does not fall below 0.9%. The reason that the optimal food subsidy is declining in ρ is that the benefits of using food subsidies are smaller if it is easier to substitute between high-skilled and low-skilled workers. In that case, the skill premium is not very responsive to a change in the food subsidy. To illustrate, if $\rho = 3.5$, the optimal food subsidy generates a reduction in the skill premium of a mere 0.03 percentage points relative to the uniform tax optimum, compared to 0.12 percentage points in the baseline (see Table 3). This figure increases to 0.16 percentage points if the elasticity of substitution is $\rho = 1.1$, though the effect remains small.

Figure 4 also shows the welfare gains from using food subsidies for different values of ρ. It plots (on the right axis) the percentage increase in the consumption aggregate that makes a utilitarian planner indifferent between this increase and optimizing food subsidies, starting from a constrained optimum with $\tau_a = 0$. The corresponding figure in the baseline calibration can be found in the last row of Table 3. The graph looks similar to that of optimal food subsidies. In particular, the welfare gains from food subsidies are decreasing in the elasticity of substitution between high-skilled and low-skilled labor. Because food subsidies are only useful insofar they indirectly redistribute income through general equilibrium effects (see Proposition 1), there is less to gain from food subsidies if it is very easy to substitute between the different skill types. However, the welfare gains from using food subsidies are tiny for all values of ρ considered. This should come as no surprise, since in each case the optimal food subsidy is close to zero.

Some literature (e.g. Blankenau and Cassou, 2011) suggests that the elasticity of substitution between high-skilled and low-skilled labor may be particularly high in agriculture. Therefore,

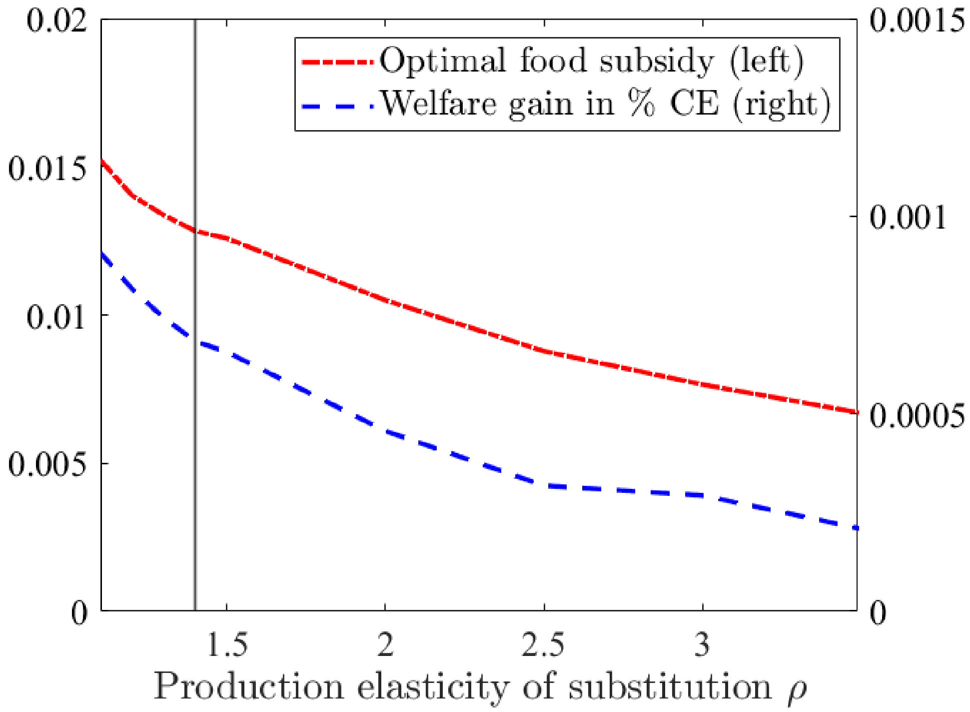

Figure 4: Optimal food subsidy and welfare gains, for different values of ρ

as an additional robustness check, we consider a version of our model where the elasticity of substitution differs across sectors. To that end, we increase the elasticity of substitution in the agricultural sector from 1.4 up to 5 and calibrate the remaining parameters to match the moments from Table 2. As the elasticity of substitution between high-skilled and low-skilled labor in the agricultural sector increases, the optimal food subsidy remains almost flat around 1.3%, while the welfare gains compared to a uniform policy with $\tau_a = 0$ remain virtually zero.

Strength of consumption responses (ϵ)

The extent to which individuals change their consumption mix in response to a change in the food subsidy critically depends on ϵ, which measures the elasticity of substitution between agricultural goods above the subsistence level and non-agricultural goods. Our baseline calibration employs a value of $\epsilon = 0.5$. The literature on structural transformation consistently finds values between zero and one, depending, among other things, on whether sectors are characterized as value added or final expenditure categories. Some recent examples are Acemoglu and Guerrieri (2008), Rogerson (2008), Buera and Kaboski (2009), Herrendorf et al. (2013), Stefanski (2014) and Moro et al. (2017). Because the change in the consumption mix is key to determining the general equilibrium effects, we calculate the optimal food subsidy and welfare gains for different values of $\epsilon \in [0.3, 1.2]$. As before, the remaining parameters are set to match the moments from Table 2.

Perhaps surprisingly, Figure 5 shows that the optimal food subsidy is not sensitive at all to the degree of substitutability between agricultural and non-agricultural goods. For each value of ϵ considered, the optimal food subsidy is around 1.3% – as in the baseline with $\epsilon = 0.5$. The reason that the optimal food subsidy hardly responds to the elasticity of substitution between agricultural and non-agricultural goods is that *both* the marginal costs and the marginal benefits of food subsidies are increasing in this elasticity. A food subsidy leads to distortions in consumption decisions that are increasing in ϵ. At the same time, a food subsidy generates indirect distributional benefits by reducing the skill premium. By how much the skill premium declines depends on the increase in the demand for agricultural goods following a rise in the subsidy, which is determined by ϵ as well. To illustrate, optimal food subsidies reduce the skill premium by 0.27 percentage points compared to the uniform tax optimum if $\epsilon = 1.2$. The corresponding figure is 0.12 percentage points in the baseline, and decreases to 0.06 percentage points if $\epsilon = 0.3$. As a result, both the indirect distributional benefits from changes in the skill premium and the costs of distorting consumption decisions are larger if individuals find it easier to substitute between agricultural and non-agricultural goods. On balance, the optimal food subsidy is not sensitive to ϵ.

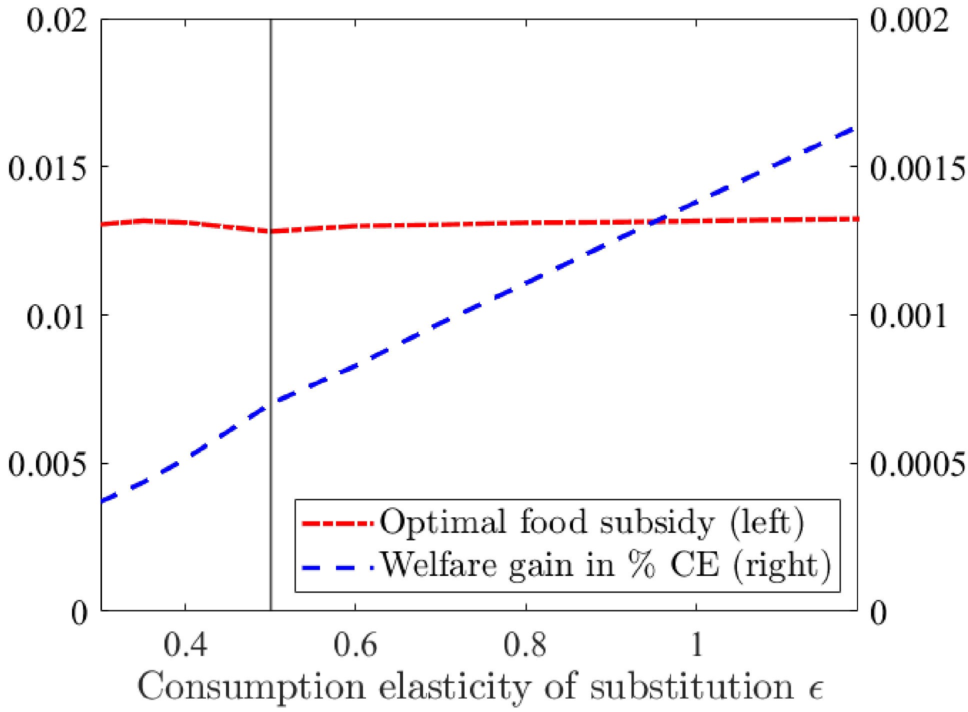

Figure 5: Optimal food subsidy and welfare gains, for different values of ϵ

According to Figure 5, the welfare gains of using food subsidies are larger if individuals find it is easier to substitute between agricultural and non-agricultural goods. Recall that the welfare gains are calculated by comparing the unconstrained optimum to a constrained optimum with $\tau_a = 0$. If the food subsidy is optimal, the *marginal* costs of distorting consumption decisions are equal to the *marginal* distributional benefits that come from general equilibrium effects, cf. equation (23). However, infra-marginally, i.e., between $\tau_a = 0$ and its optimal value, the marginal benefits of increasing the food subsidy exceed the marginal costs. Therefore, the *total* benefits of using food subsidies are larger if an increase in the food subsidy leads to a larger increase in the demand for agricultural goods, and thereby a larger decline in the skill premium. This is the case when it is easier to substitute between agricultural and non-agricultural goods, i.e., when ϵ is large. However, as can be seen from Figure 5, the welfare gains remain negligible for all values of ϵ considered.

The role of subsistence consumption ($\underline{c}_a$)

In the absence of general equilibrium effects, the optimal food subsidy is zero, even though low-income individuals spend a larger share of their disposable income on food due to subsistence consumption (which calls for a subsidy). As explained in Section 3, the reason is that subsistence consumption also makes the demand for agricultural goods less elastic (which calls for a tax). On balance, these effects are precisely off-setting if Engel curves are linear and prices are fixed, cf. Deaton (1979). However, if there are general equilibrium effects, this may be different: subsistence consumption could affect the optimal food subsidy.

To explore this further, we parameterize a version of our model in which $\underline{c}_a$ is set to zero and all other parameters are set to match the data moments as well as possible (excluding the food share of expenditures, which strongly depends on the subsistence level). In this case, the optimal food subsidy is approximately 1.09%, compared to 1.29% in the baseline. These small food subsidies generate welfare gains that are close to zero compared to the uniform tax optimum, as in the baseline. Changing the level of subsistence consumption therefore does not affect our main results, up to small quantitative differences. The reason is that food subsidies are only helpful insofar they generate indirect distributional benefits through general equilibrium effects, cf. Proposition 1. The effect of food subsidies on wages, in turn, is governed by how much individuals change their consumption mix, and how much firms change their labor input mix. These responses depend on the elasticity of substitution in the consumption aggregate (ϵ) and the production function (ρ), but not to a large degree on subsistence consumption ($\underline{c}_a$). Therefore, the optimal food subsidy is not very sensitive to the level of subsistence consumption, even if prices are endogenous.

Tax capacity constraints

Finally, we turn to study what role there is for subsidies (or taxes) on food if the government faces severe tax capacity constraints that prevent it from raising either the tax on labor income τ_y or the lump-sum transfer T.

Suppose the government cannot raise the labor tax. We investigate this case by keeping τ_y fixed at its baseline value of approximately 15% and solve for the constrained optimal policies. Because the labor income tax is kept fixed, an increase in the tax on agricultural goods is then used to finance a larger transfer.[34] In this case, it becomes optimal to *tax* (rather than to subsidize) food at an extraordinary rate of approximately 160%. Intuitively, if the government cannot raise the labor income tax (or, equivalently, the uniform consumption tax) to finance a larger transfer, it is optimal to generate significant revenues from taxing food consumption and to use the proceeds to increase the transfer. Food taxes are then used as a partial substitute for the tax on labor income, which is fixed at a level much lower than in the utilitarian optimum, where it is around 50% (see Table 3). Because total spending on food increases in income, this policy generates substantial distributional benefits. Relative to the baseline, the welfare gains are 9.11% in consumption equivalents, compared to 19.54% if the government can also optimize labor taxes directly (see Table 3). These gains are very large even though the policy generates indirect distributional losses. In particular, the very high taxes on food combined with an increase in the lump-sum transfer bring about a sizable increase in the skill premium, from 51.41% in the baseline to 57.28%.

Next, we solve for the constrained optimal policies if the government cannot adjust the lump-sum transfer T. An increase in food subsidies must then be financed by higher taxes on labor income. In this case, it is optimal to subsidize food at a rate of approximately 40%, much higher than in the baseline. Larger food subsidies financed by higher taxes on labor income generate distributional benefits, because low-income individuals spend a larger share of their disposable income on food (due to subsistence consumption). Food subsidies are therefore used a partial substitute for the lump-sum transfer, which is fixed at a level much lower than in the utilitarian optimum (see Table 3). The direct distributional benefits are complemented by indirect distributional benefits from general equilibrium effects. In particular, the increase in food subsidies financed by higher labor taxes brings about a substantial reduction in the

[34] Note that the total tax *wedge* on labor income τ_y that is kept fixed consists not only of direct taxes on labor income, but also of uniform consumption taxes. Keeping the labor wedge fixed thus refers to the case where the government can neither raise the tax rate on labor income nor the uniform consumption tax. This may be relevant when a government is capacity constrained for both income and consumption taxes but can still differentiate between goods, say using subsidies.

skill premium, from 51.41% in the baseline to 46.28%.[35] Compared to the baseline, the total welfare gains amount to 1.12% in consumption equivalents, out of the 19.54% that can be gained if the government can also optimize the transfer (see Table 3). Hence, food subsidies are a useful policy tool if the government faces severe capacity constraints that prevent it from raising the lump-sum transfer. At the same time, the benefits from relaxing this constraint are substantially larger.

6 Discussion

This section discusses how we expect the results to be affected in several extensions of the model, or if some of the key assumptions are relaxed.

A number of features would further reduce the (already very small) optimal food subsidies we find. First, the economy we analyze is closed to trade with the outside world. In a small open economy, none of the general equilibrium effects we study would be relevant, since food prices are determined on world markets. However, the vast majority of agricultural goods produced in China are not exported but instead consumed domestically (FAO, 2012) and the price of these goods appears very sensitive to country-specific shocks.[36] Hence, while weakening the case for food subsidies, we doubt that allowing for trade significantly affects our results.[37] Second, we have abstracted from nonlinear taxes on labor income. With a nonlinear income tax, the government can use marginal tax rates at different points in the income distribution to increase (decrease) the aggregate labor supply of high-skilled (low-skilled) workers, which reduces the skill premium, cf. Stiglitz (1982) and Rothschild and Scheuer (2013). If a nonlinear income tax enables the government to already exploit general equilibrium effects, we conjecture it further limits the scope for food subsidies as a means to indirectly redistribute income.[38] Third, we have treated each worker's skill type and the size

[35]This reduction is larger than what one would expect from Figure 2. The reason is that in Figure 2, a higher food subsidy is not financed by higher labor taxes, as in the current experiment, but by a lower transfer, which generates a positive income effect on labor supply. Because the latter is stronger for low-skilled workers, this mutes the reduction in the skill premium from larger food subsidies.

[36]See, for example, the article "China pork price hits 2019 low as swine fever spurs selloff" from Bloomberg News (April, 2021).

[37]Of course, how much international trade reduces general equilibrium effects depends very much on the country under investigation. There is, however, another sense in which our results carry over to an open economy setting: if a planner is interested in maximizing global welfare (e.g., by coordinating agricultural subsidies), general equilibrium effects are as relevant as in a closed economy.

[38]In a previous draft of this paper that did not include heterogeneity within skill types, we indeed found that the optimal food subsidy is smaller if the government optimizes a nonlinear income tax. Solving the optimal nonlinear tax problem with heterogeneity within skill types is beyond the scope of this paper.

of each skill group as fixed. Hence, a reduction in the skill premium driven by an increase in food subsidies does not lead to different education choices. Allowing for an educational choice margin mutes the impact of food subsidies on wages. This limits the scope for the government to exploit general equilibrium effects for redistributive purposes (Saez, 2004), though most likely only in the run.

On another adjustment margin, we have taken the opposite stance. While workers' skill type is fixed, they can freely move between sectors and their productivity is the same in both sectors. One might question that assumption on the grounds of moving frictions: the production of food mostly occurs in rural environments, and that of non-food mostly in urban ones. Introducing costs of switching between sectors likely strengthens the general equilibrium effects of food subsidies on wages and hence, the importance of the mechanism we study. However, this would be most relevant in the short run. In the long run, we do observe large-scale reallocation of employment between sectors (Herrendorf et al., 2014).

Finally, our model does not include subsistence farming, i.e., farming activity that is not part of the formal marketplace, but rather results in home production of food. This is still an important phenomenon in rural China (Prändl-Zika, 2008). Home production is not subject to a VAT, which affects the labor tax wedge, nor is it likely subject to food subsidies. However, the general equilibrium effects we focus on will be relevant for subsistence farmers if they sell part of their output.[39] In that case, subsistence farmers might benefit from food subsidies if they actually receive these subsidies. On the other hand, if they are not recipients, subsistence farmers might be harmed by food subsidies due to the negative impact of these subsidies on consumer prices. Therefore, while it cannot be excluded entirely, it seems unlikely that subsistence farming significantly strengthens the case for food subsidies.

7 Conclusion

This paper investigates to what extent food subsidies are helpful as a means to indirectly redistribute income through general equilibrium effects. To do so, we analyze an economy with two two goods (agriculture and non-agriculture) and individuals of two skill types (low-skilled and high-skilled) who, within each skill type, differ in their productivity. Low-skilled labor has a comparative advantage in the production of agricultural goods. An increase in food subsidies raises the demand for agricultural goods and thereby the demand for low-skilled labor. This reduces the skill premium and compresses the income distribution. A government that is interested in redistribution can exploit these general equilibrium effects

[39]The impact of food subsidies on the price of agricultural goods has a direct impact on the incomes of subsistence farmers if they sell part of their output. The impact of food subsidies on wages, however, is unlikely to affect them because they operate their own technology.

to indirectly redistribute income from high-skilled to low-skilled workers.

We show theoretically how the welfare impact from a reduction in food subsidies can be decomposed into 'direct effects' due to transfers from individuals to the government, 'behavioral effects' in consumption and labor supply that generate fiscal externalities, and 'general equilibrium effects' driven by changes in the price of food and wages. These price responses affect welfare by influencing the tax base and individual purchasing power. Using the property that in our model Engel curves are linear, we also derive an expression for the optimal food subsidy. This formula demonstrates that food subsidies are only useful insofar they generate indirect distributional benefits, which confirms the result from Deaton (1979) that the optimal food subsidy is zero if prices are fixed. However, if prices are endogenous, the optimal food subsidy strikes a balance between the costs of distorting consumption decisions and the indirect distributional benefits that come from changes in the price of food and wages.

We then investigate the quantitative importance of general equilibrium effects for optimal food subsidies by calibrating our model to the economy of China, the world's largest subsidizer of agriculture (OECD, 2017). We parameterize our model to match key moments on agricultural versus non-agricultural employment and individual consumption patterns that we obtain from micro-level survey data, in particular the 2008 wave of the Chinese Household Income Project, and other sources. We find that general equilibrium effects rationalize only very modest food subsides that generate tiny welfare gains. In particular, a utilitarian government sets a food subsidy of approximately 1.29%. The welfare gains of using the optimal food subsidy (or, equivalently, the welfare costs of following a uniform consumption tax policy) are typically below 0.01% in consumption equivalent gains. We show that these numbers are fairly robust to changing the government's preferences for redistribution, the elasticity of substitution between skill types, the elasticity of substitution between agricultural and non-agricultural consumption, and subsistence consumption of agricultural goods. The only instance where we find a significant subsidy (or tax) on food is if the government cannot raise the lump-sum transfer (or labor tax) directly, in which case it serves to substitute for the instrument that cannot be optimized.

To focus exclusively on the implications of general equilibrium effects for optimal food subsidies, we have maintained throughout the assumption that individuals have identical, linear Engel curves. However, there is substantial variation between households in food spending conditional on total expenditures. And even in the absence of such variation, food spending and total expenditures need not be linearly related. Furthermore, Jensen and Miller (2008) find that for the very poorest consumers, the demand for agricultural goods actually increases in price. Exploring how properties of the demand structure (such as preference heterogeneity, nonlinear Engel curves, and Giffen behavior) affects the optimal design of food subsidies is an interesting topic for future research.

References

Acemoglu, D. and D. Autor (2011). Skills, tasks and technologies: Implications for employment and earnings. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume 4, Chapter 12, pp. 1043–1171. Amsterdam: Elsevier.

Acemoglu, D. and V. Guerrieri (2008). Capital deepening and nonbalanced economic growth. *Journal of Political Economy 116*(3), 467–498.

Ahmad, E. and N. Stern (1989). Taxation for developing countries. In H. Chenery and T. Srinivasan (Eds.), *Handbook of Development Economics*, Volume 2, Chapter 20, pp. 1005–1092. Amsterdam: Elsevier.

Allen, F. (1982). Optimal linear income taxation with general equilibrium effects on wages. *Journal of Public Economics 17*(2), 135–143.

Angrist, J. (1995). The economic returns to schooling in the West Bank and Gaza Strip. *American Economic Review 85*(5), 1065–1087.

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

Bachas, P., L. Gadenne, and A. Jensen (2022). Informality, consumption taxes and redistribution. *Review of Economic Studies*. Forthcoming.

Barwick, P., M. Kalouptsidi, and N. Zahur (2019). China's industrial policy: An empirical evaluation. NBER Working Paper No. 26075, Cambridge-MA: NBER.

Behar, A. (2009). Directed technical change, the elasticity of substitution and wage inequality in developing countries. mimeo, University of Oxford.

Besley, T. and T. Persson (2014). Why do developing countries tax so little? *Journal of Economic Perspectives 28*(4), 99–120.

Blankenau, W. and S. Cassou (2011). Industry estimates of the elasticity of substitution and the rate of biased technological change between skilled and unskilled labour. *Applied Economics 43*(23), 3129–3142.

Buera, F. and J. Kaboski (2009). Can traditional theories of structural change fit the data? *Journal of the European Economic Association 7*(2-3), 469–477.

Card, D. (2009). Immigration and inequality. *American Economic Review 99*(2), 1–21.

Chetty, R. (2009). Sufficient statistics for welfare analysis: a bridge between structural and reduced-form methods. *Annual Review of Economics 1*(1), 451–488.

Chetty, R., A. Guren, D. Manoli, and A. Weber (2011). Are micro and macro labor supply elasticities consistent? A review of evidence on the intensive and extensive margins. *American Economic Review 101*(3), 471–475.

Ciccone, A. and G. Peri (2005). Long-run substitutability between more and less educated workers: evidence from US states, 1950–1990. *Review of Economics and Statistics 87*(4), 652–663.

Corlett, W. and D. Hague (1953). Complementarity and the excess burden of taxation. *Review of Economic Studies 21*(1), 21–30.

Deaton, A. (1979). Optimally uniform commodity taxes. *Economics Letters 2*(4), 357–361.

Diamond, P. (1975). A many-person Ramsey tax rule. *Journal of Public Economics 4*(4), 335–342.

Doligalski, P. and L. Rojas (2022). Optimal redistribution with a shadow economy. *Theoretical Economics*. Forthcoming.

FAO (2012). Agriculture and trade policy background note: People's Republic of China. Rome: FAO.

Gadenne, L. (2020). Can rationing increase welfare? Theory and an application to India's ration shop system. *American Economic Journal: Economic Policy 12*(4), 144–77.

Gollin, D., D. Lagakos, and M. Waugh (2013). The agricultural productivity gap. *Quarterly Journal of Economics 129*(2), 939–993.

Gustafsson, B., L. Shi, and H. Sato (2014). Data for studying earnings, the distribution of household income and poverty in China. *China Economic Review 30*, 419–431.

Hendricks, L. (2010). Cross-country variation in educational attainment: structural change or within-industry skill upgrading? *Journal of Economic Growth 15*(3), 205–233.

Herrendorf, B., R. Rogerson, and A. Valentinyi (2013). Two perspectives on preferences and structural transformation. *American Economic Review 103*(7), 2752–89.

Herrendorf, B., R. Rogerson, and A. Valentinyi (2014). Growth and structural transformation. In P. Aghion and S. Durlauf (Eds.), *Handbook of Economic Growth*, Volume 2, Chapter 6, pp. 855–941. Amsterdam: Elsevier.

IMF (2008). Fuel and food price subsidies: issues and reform options. Washington, DC: IMF.

IMF (2018). Technical assistance report - tax policy and employment creation. Washington, DC: IMF.

Jacobs, B. and F. van der Ploeg (2019). Redistribution and pollution taxes with non-linear Engel curves. *Journal of Environmental Economics and Management 95*, 198–226.

Jensen, R. and N. Miller (2008). Giffen behavior and subsistence consumption. *American Economic Review 98*(4), 1553–77.

Jones, L., R. Manuelli, and P. Rossi (1997). On the optimal taxation of capital income. *Journal of Economic Theory 73*(1), 93–117.

Katz, L. and K. Murphy (1992). Changes in relative wages, 1963–1987: supply and demand factors. *Quarterly Journal of Economics 107*(1), 35–78.

Kina, Ã., C. Slavík, and H. Yazici (2020). Redistributive capital taxation revisited. CESifo

Working Paper No. 8627, Munich: CESifo.

Lin, S. (2009). The rise and fall of China's government revenue. EAI Working Paper No. 150, Singapore: EAI.

Mirrlees, J. (1971). An exploration in the theory of optimum income taxation. *Review of Economic Studies 38*(114), 175–208.

Moro, A., S. Moslehi, and S. Tanaka (2017). Does home production drive structural transformation? *American Economic Journal: Macroeconomics 9*(3), 116–146.

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

OECD (2017). Agricultural policy monitoring and evaluation 2017. Paris: OECD.

Peralta-Alva, A., M. Mendes Tavares, X. Tam, and X. Tang (2018). The macroeconomic and distributional implications of fiscal consolidations in low-income countries. IMF Working Paper, Washington DC: IMF.

Prändl-Zika, V. (2008). From subsistence farming towards a multifunctional agriculture: Sustainability in the Chinese rural reality. *Journal of Environmental Management 87*(2), 236–248.

Ramsey, F. (1927). A contribution to the theory of taxation. *The Economic Journal 37*(145), 47–61.

Rogerson, R. (2008). Structural transformation and the deterioration of European labor market outcomes. *Journal of Political Economy 116*(2), 235–259.

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

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

Saez, E. (2002). The desirability of commodity taxation under non-linear income taxation and heterogeneous tastes. *Journal of Public Economics 83*(2), 217–230.

Saez, E. (2004). Direct or indirect tax instruments for redistribution: short-run versus long-run. *Journal of Public Economics 88*(3-4), 503–518.

Sheshinski, E. (1972). The optimal linear income-tax. *Review of Economic Studies 39*(3), 297–302.

Slavík, C. and H. Yazici (2014). Machines, buildings, and optimal dynamic taxes. *Journal of Monetary Economics 66*, 47–61.

Stefanski, R. (2014). Structural transformation and the oil price. *Review of Economic Dynamics 17*(3), 484–504.

Stiglitz, J. (1982). Self-selection and Pareto efficient taxation. *Journal of Public Economics 17*(1), 213–240.

Wang, L. (2012). Economic transition and college premium in urban China. *China Economic*

Review 23(2), 238–252.

A Proof Lemma 1

Suppose the government optimizes the tax τ_y on labor income and the lump-sum transfer T. Welfare as a function of the tax τ_a on agricultural goods can then be written as

$$\mathcal{W}(\tau_a) = \max_{\tau_y, T} \Bigg\{ \sum_i \mu_i \int_{\Theta_i} \alpha_i(\theta) V_i^*(\tau_a, \tau_y, T; \theta) dK_i(\theta) \;\; \text{s.t.} \;\; \tau_a \sum_i \mu_i p_a(\tau_a, \tau_y, T) \tag{30}$$
$$\times \int_{\Theta_i} c_{a,i}^*(\tau_a, \tau_y, T; \theta) dK_i(\theta) + \tau_y \sum_i \mu_i w_i(\tau_a, \tau_y, T) \int_{\Theta_i} \theta \ell_i^*(\tau_a, \tau_y, T; \theta) dK_i(\theta) = T + G \Bigg\},$$

where the summation is over skill types $i \in \{L, H\}$. Here, we explicitly account for the dependency of equilibrium prices, individual utility and individual decisions on the tax instruments (τ_a, τ_y, T). The conditions which pin down equilibrium prices and individual decisions as a function of the tax instruments can be found in Appendix B. The Lagrangian corresponding to the maximization problem (30) is

$$\mathcal{L} = \sum_i \mu_i \int_{\Theta_i} \alpha_i(\theta) V_i^*(\tau_a, \tau_y, T; \theta) dK_i(\theta) + \lambda \Bigg[\tau_a \sum_i \mu_i p_a(\tau_a, \tau_y, T) \int_{\Theta_i} c_{a,i}^*(\tau_a, \tau_y, T; \theta) dK_i(\theta)$$
$$+ \tau_y \sum_i \mu_i w_i(\tau_a, \tau_y, T) \int_{\Theta_i} \theta \ell_i^*(\tau_a, \tau_y, T; \theta) dK_i(\theta) - T - G \Bigg], \tag{31}$$

where λ is the multiplier on the government budget constraint. The indirect utility function, which shows up in the Lagrangian (31), is given by

$$V_i^*(\tau_a, \tau_y, T; \theta) = \max_{c_a, c_n, \ell} \Bigg\{ u(c_a, c_n) + v(\ell)$$
$$\text{s.t.} \;\; p_a(\tau_a, \tau_y, T)(1 + \tau_a) c_a + c_n = w_i(\tau_a, \tau_y, T)(1 - \tau_y) \theta \ell + T \Bigg\}, \tag{32}$$

To determine how a change in the tax instruments affects individual utility, use the budget constraint to substitute out for c_n. By the envelope theorem,

$$\frac{\partial V_i^*(\theta)}{\partial \tau_a} = u_{n,i}(\theta) \Bigg[- p_a c_{a,i}(\theta) + (1 - \tau_y) \theta \ell_i(\theta) \frac{\partial w_i}{\partial \tau_a} - (1 + \tau_a) c_{a,i}(\theta) \frac{\partial p_a}{\partial \tau_a} \Bigg], \tag{33}$$
$$\frac{\partial V_i^*(\theta)}{\partial \tau_y} = u_{n,i}(\theta) \Bigg[- w_i \theta \ell_i(\theta) + (1 - \tau_y) \theta \ell_i(\theta) \frac{\partial w_i}{\partial \tau_y} - (1 + \tau_a) c_{a,i}(\theta) \frac{\partial p_a}{\partial \tau_y} \Bigg], \tag{34}$$
$$\frac{\partial V_i^*(\theta)}{\partial T} = u_{n,i}(\theta) \Bigg[1 + (1 - \tau_y) \theta \ell_i(\theta) \frac{\partial w_i}{\partial T} - (1 + \tau_a) c_{a,i}(\theta) \frac{\partial p_a}{\partial T} \Bigg], \tag{35}$$

where $u_{n,i}(\theta)$ denotes the marginal utility of non-agricultural goods and we dropped the function arguments (τ_a, τ_y, T) to save on notation.

The welfare impact of raising the tax on agricultural goods can be found by differentiating the Lagrangian (31) with respect to τ_a:

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \tau_a} =& \sum_i \mu_i \int_{\Theta_i} \alpha_i(\theta) \frac{\partial V_i^*(\theta)}{\partial \tau_a} dK_i(\theta) + \lambda \Bigg[\sum_i \mu_i p_a \int_{\Theta_i} c_{a,i}(\theta) dK_i(\theta) \\ &+ \left(\tau_a \sum_i \mu_i p_a \int_{\Theta_i} \frac{\partial c^*_{a,i}(\theta)}{\partial \tau_a} dK_i(\theta) + \tau_y \sum_i \mu_i w_i \int_{\Theta_i} \theta \frac{\partial \ell_i^*(\theta)}{\partial \tau_a} dK_i(\theta) \right) \\ &+ \left(\tau_a \sum_i \mu_i \frac{\partial p_a}{\partial \tau_a} \int_{\Theta_i} c_{a,i}(\theta) dK_i(\theta) + \tau_y \sum_i \mu_i \frac{\partial w_i}{\partial \tau_a} \int_{\Theta_i} \theta \ell_i(\theta) dK_i(\theta) \right) \Bigg]. \end{aligned} \tag{36}$$

To proceed, substitute out for $\partial V_i^*(\theta)/\partial \tau_a$ using equation (33) and collect terms to get:

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \tau_a} =& \sum_i \mu_i p_a \int_{\Theta_i} (\lambda - \alpha_i(\theta) u_{n,i}(\theta)) c_{a,i}(\theta) dK_i(\theta) \\ &+ \lambda \left(\tau_a \sum_i \mu_i p_a \int_{\Theta_i} \frac{\partial c^*_{a,i}(\theta)}{\partial \tau_a} dK_i(\theta) + \tau_y \sum_i \mu_i w_i \int_{\Theta_i} \theta \frac{\partial \ell_i^*(\theta)}{\partial \tau_a} dK_i(\theta) \right) \\ &+ \sum_i \mu_i \frac{\partial p_a}{\partial \tau_a} \int_{\Theta_i} (\lambda \tau_a - \alpha_i(\theta) u_{n,i}(\theta)(1+\tau_a)) c_{a,i}(\theta) dK_i(\theta) \\ &+ \sum_i \mu_i \frac{\partial w_i}{\partial \tau_a} \int_{\Theta_i} (\lambda \tau_y + \alpha_i(\theta) u_{n,i}(\theta)(1-\tau_y)) \theta \ell_i(\theta) dK_i(\theta). \end{aligned} \tag{37}$$

Next, divide equation (37) by λ and define the welfare weight of an individual of type (i, θ) as $g_i(\theta) = \alpha_i(\theta) u_{n,i}(\theta)/\lambda$. The latter measures by how much welfare increases if an individual of type (i, θ) receives an additional unit of non-agricultural goods or, equivalently, an additional unit of after-tax income.[40] Equation (37) then reads

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \tau_a} \frac{1}{\lambda} =& \sum_i \mu_i p_a \int_{\Theta_i} (1 - g_i(\theta)) c_{a,i}(\theta) dK_i(\theta) \\ &+ \tau_a \sum_i \mu_i p_a \int_{\Theta_i} \frac{\partial c^*_{a,i}(\theta)}{\partial \tau_a} dK_i(\theta) + \tau_y \sum_i \mu_i w_i \int_{\Theta_i} \theta \frac{\partial \ell_i^*(\theta)}{\partial \tau_a} dK_i(\theta) \\ &+ \sum_i \mu_i \frac{\partial p_a}{\partial \tau_a} \int_{\Theta_i} (\tau_a - g_i(\theta)(1+\tau_a)) c_{a,i}(\theta) dK_i(\theta) \\ &+ \sum_i \mu_i \frac{\partial w_i}{\partial \tau_a} \int_{\Theta_i} (\tau_y + g_i(\theta)(1-\tau_y)) \theta \ell_i(\theta) dK_i(\theta). \end{aligned} \tag{38}$$

Using the expectation operator $\mathbb{E}(\cdot)$, the latter can be written as

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \tau_a} \frac{1}{\lambda} =& \underbrace{\sum_i \mu_i p_a \Big[\mathbb{E}(c_{a,i}) - \mathbb{E}(c_{a,i} g_i) \Big]}_{DE} + \underbrace{\tau_a \sum_i \mu_i p_a \mathbb{E}\left(\frac{\partial c^*_{a,i}}{\partial \tau_a} \right) + \tau_y \sum_i \mu_i w_i \mathbb{E}\left(\frac{\partial (\theta \ell_i^*)}{\partial \tau_a} \right)}_{BE} \\ &\underbrace{+ \sum_i \mu_i \frac{\partial p_a}{\partial \tau_a} \Big[\tau_a \mathbb{E}(c_{a,i}) - (1+\tau_a) \mathbb{E}(c_{a,i} g_i) \Big] + \sum_i \mu_i \frac{\partial w_i}{\partial \tau_a} \Big[\tau_y \mathbb{E}(\theta \ell_i) + (1-\tau_y) \mathbb{E}(\theta \ell_i g_i) \Big]}_{GE}, \end{aligned} \tag{39}$$

[40] Recall that the price of non-agricultural goods is normalized to one.

which coincides with equation (18) from Lemma 1.

B Equilibrium given tax policy

This Appendix states the conditions which pin down the equilibrium consumption and labor supply decisions $(c_{a,i}(\theta), c_{n,i}(\theta), \ell_i(\theta))_{\theta\in\Theta_i, i\in\{L,H\}}$, labor inputs $\{(L_a, H_a), (L_n, H_n)\}$ and prices (p_a, w_L, w_H) as a function of tax policy (τ_a, τ_y, T).

From the utility maximization problem, for each $\theta \in \Theta_i$ and $i \in \{L, H\}$,

$$\frac{C(c_{a,i}(\theta), c_{n,i}(\theta))^{-\sigma}}{P} = \frac{\psi(1-\ell_i(\theta))^{-\phi}}{w_i(1-\tau_y)\theta}, \tag{40}$$

$$c_{n,i}(\theta) = \frac{1-\omega}{\omega}(c_{a,i}(\theta) - \underline{c}_a)(p_a(1+\tau_a))^{\epsilon}, \tag{41}$$

$$p_a(1+\tau_a)c_{a,i}(\theta) + c_{n,i}(\theta) = T + w_i(1-\tau_y)\theta\ell_i(\theta), \tag{42}$$

where $C(c_{a,i}(\theta), c_{n,i}(\theta))$ and P are as defined in equation (2) and (6), respectively. From the firm's problem, for each $j \in \{a, n\}$,

$$w_L = p_j A_j \left(\frac{Y_j}{A_j}\right)^{\frac{1}{\rho}} \gamma_j L_j^{-\frac{1}{\rho}}, \tag{43}$$

$$w_H = p_j A_j \left(\frac{Y_j}{A_j}\right)^{\frac{1}{\rho}} (1-\gamma_j) H_j^{-\frac{1}{\rho}}, \tag{44}$$

where outputs Y_a and Y_n are determined from equation (7) and $p_n = 1$ as a normalization. Labor and agricultural goods market clearing, in turn, requires

$$L_a + L_n = \mu_L \int_{\Theta_L} \theta_L \ell_L(\theta_L) dK_L(\theta_L), \tag{45}$$

$$H_a + H_n = \mu_H \int_{\Theta_H} \theta_H \ell_H(\theta_H) dK_H(\theta_H), \tag{46}$$

$$Y_a = \mu_L \int_{\Theta_L} c_{a,L}(\theta_L) dK_L(\theta_L) + \mu_H \int_{\Theta_H} c_{a,H}(\theta_H) dK_H(\theta_H). \tag{47}$$

Combined, equations (40)–(47) pin down the equilibrium quantities and prices for a given tax policy. We denote the equilibrium prices by $p_a(\tau_a, \tau_y, T)$, $w_L(\tau_a, \tau_y, T)$, $w_H(\tau_a, \tau_y, T)$ and individual choices by $c^*_{a,i}(\tau_a, \tau_y, T; \theta)$, $c^*_{n,i}(\tau_a, \tau_y, T; \theta)$ and $\ell^*_i(\tau_a, \tau_y, T; \theta)$. These are used in Appendix A to study the welfare impact of raising τ_a. Importantly, the functions $c^*_{n,i}(\cdot)$, $c^*_{a,i}(\cdot)$ and $\ell^*_i(\cdot)$ capture the *total* impact of a change in the tax instruments on individual consumption and labor supply. The total impact consists of both the direct effect from a change in the tax instrument on consumption and labor supply decisions, and indirect effects driven by general equilibrium effects from tax instruments on prices.

It is worth pointing out that G does not show up in the above system. Hence, for a particular choice of the tax instruments (τ_a, τ_y, T), the value of G must be such that the government budget constraint (or equivalently, by Walras' law, the market-clearing condition for non-agricultural goods) is satisfied. This is why in the Lagrangian (31) from Appendix A, the government budget constraint is explicitly taken into account, while the equilibrium conditions (40)–(47) are summarized through reduced-form relationships that highlight the dependency of the equilibrium prices and individual decisions on tax policy (τ_a, τ_y, T).

C Proof Proposition 1

The proof proceeds as follows. We start by characterizing the optimal tax system in terms of welfare weights, behavioral responses and general equilibrium effects from the tax instruments on prices. We then use our specification of preferences to obtain expressions for the behavioral responses. Substituting these in the optimal tax formulas leads to equation (23) from Proposition 1.

Appendix A derives an expression for the welfare impact of raising the tax on agricultural goods that involves behavioral responses $\partial c_{a,i}^*(\theta)/\partial\tau_a$ and $\partial\ell_i^*(\theta)/\partial\tau_a$. As mentioned before, these capture the *total* impact of a higher tax τ_a on the consumption of agricultural goods and labor supply. The total impact consists of both the direct effect from a higher after-tax price of agricultural goods (driven by an increase in τ_a) as well as the indirect effects driven by general equilibrium responses from τ_a on before-tax prices p_a, w_L and w_H. It turns out insightful to split up these effects. To that end, reconsider the utility maximization problem. Write the indirect utility function as

$$V(p_a(1+\tau_a), w_i(1-\tau_y)\theta, T) = \max_{c_a, c_n, \ell} \Big\{ u(c_a, c_n) + v(\ell) \quad \text{s.t.} \quad p_a(1+\tau_a)c_a + c_n = w_i(1-\tau_y)\theta\ell + T \Big\}, \tag{48}$$

The difference with equation (32) is that equation (48) does not account for the impact of (τ_a, τ_y, T) on equilibrium prices (p_a, w_L, w_H).[41] Therefore, the function arguments are different. Denote by $p_a^* = p_a(1+\tau_a)$ and $w_i^* = w_i(1-\tau_y)\theta$ the after-tax price of one unit of agricultural consumption and one unit of leisure for an individual of type (i, θ) in terms of

[41] If we take this dependency into account, the value functions are the same. Hence, for each (τ_a, τ_y, T):

$$V_i^*(\tau_a, \tau_y, T; \theta) = V(p_a(\tau_a, \tau_y, T)(1+\tau_a), w_i(\tau_a, \tau_y, T)(1-\tau_y)\theta, T). \tag{49}$$

the *numeraire*.[42] By the envelope theorem,

$$\frac{\partial V_i(\theta)}{\partial p_a^*} = -u_{n,i}(\theta)c_{a,i}(\theta), \tag{50}$$

$$\frac{\partial V_i(\theta)}{\partial w_i^*} = u_{n,i}(\theta)\ell_i(\theta), \tag{51}$$

$$\frac{\partial V_i(\theta)}{\partial T} = u_{n,i}(\theta), \tag{52}$$

where we ignored function arguments to save on notation.

Write the uncompensated demand functions, i.e., the solution to the maximization problem (48), as $c_a(p_a^*, w_i^*, T)$, $c_n(p_a^*, w_i^*, T)$ and $\ell(p_a^*, w_i^*, T)$.[43] The Slutsky equations for $c_a(\cdot)$ and $\ell(\cdot)$, which we use below, are:

$$\frac{\partial c_{a,i}^c(\theta)}{\partial p_a^*} = \frac{\partial c_{a,i}(\theta)}{\partial p_a^*} + \frac{\partial c_{a,i}(\theta)}{\partial T}c_{a,i}(\theta), \tag{54}$$

$$\frac{\partial c_{a,i}^c(\theta)}{\partial w_i^*} = \frac{\partial c_{a,i}(\theta)}{\partial w_i^*} - \frac{\partial c_{a,i}(\theta)}{\partial T}\ell_i(\theta), \tag{55}$$

$$\frac{\partial \ell_i^c(\theta)}{\partial p_a^*} = \frac{\partial \ell_i(\theta)}{\partial p_a^*} + \frac{\partial \ell_i(\theta)}{\partial T}c_{a,i}(\theta), \tag{56}$$

$$\frac{\partial \ell_i^c(\theta)}{\partial w_i^*} = \frac{\partial \ell_i(\theta)}{\partial w_i^*} - \frac{\partial \ell_i(\theta)}{\partial T}\ell_i(\theta), \tag{57}$$

where the terms on the left-hand side denote compensated responses (i.e., holding utility fixed) and the first (second) term on the right-hand side captures uncompensated responses (income effects). As before, we ignore function arguments to save on notation.

The government chooses tax policy (τ_a, τ_y, T) to maximize welfare (12), subject to the budget constraint (13), taking into account the impact of taxes on individual decisions $c_a(p_a(1+\tau_a), w_i(1-\tau_y)\theta, T)$, $c_n(p_a(1+\tau_a), w_i(1-\tau_y)\theta, T)$ and $\ell(p_a(1+\tau_a), w_i(1-\tau_y)\theta, T)$ *and* the impact

[42] For notational convenience, in what follows we suppress the dependence of w_i^* on θ.

[43] These differ from the functions $c_{a,i}^*(\tau_a, \tau_y, T; \theta)$, $c_{n,i}^*(\tau_a, \tau_y, T; \theta)$ and $l_i^*(\tau_a, \tau_y, T; \theta)$ characterized in Appendix B, because the latter account for the impact of tax policy on equilibrium prices. Taking this into account, the functions for agricultural consumption are linked through $c_{a,i}^*(\tau_a, \tau_y, T; \theta) = c_a(p_a(\tau_a, \tau_y, T)(1+\tau_a), w_i(\tau_a, \tau_y, T)(1-\tau_y)\theta, T)$, and similarly for non-agricultural consumption and labor supply. Differentiating both sides with respect to τ_a shows how the behavioral responses are related:

$$\frac{\partial c_{a,i}^*(\theta)}{\partial \tau_a} = \frac{\partial c_{a,i}(\theta)}{\partial p_a^*}p_a + \frac{\partial c_{a,i}(\theta)}{\partial p_a^*}(1+\tau_a)\frac{\partial p_a}{\partial \tau_a} + \frac{\partial c_{a,i}(\theta)}{\partial w_i^*}(1-\tau_y)\theta\frac{\partial w_i}{\partial \tau_a}, \tag{53}$$

where again we ignored function arguments to save on notation. The *total* impact of a higher τ_a on the left-hand side equals the sum of the *direct* effect from an increase in the after-tax price (first term on the right-hand side) and the *indirect* effects driven by changes in equilibrium prices (second and third term on the right-hand side).

of taxes on equilibrium before-tax prices (p_a, w_L, w_H).[44] The corresponding Lagrangian of the government's maximization problem is

$$\mathcal{L} = \sum_i \mu_i \int_{\Theta_i} \alpha_i(\theta) V(p_a(\tau_a, \tau_y, T)(1+\tau_a), w_i(\tau_a, \tau_y, T)(1-\tau_y)\theta, T) dK_i(\theta) \tag{58}$$
$$+ \lambda \Big[\tau_a \sum_i \mu_i p_a(\tau_a, \tau_y, T) \int_{\Theta_i} c_a(p_a(\tau_a, \tau_y, T)(1+\tau_a), w_i(\tau_a, \tau_y, T)(1-\tau_y)\theta, T) dK_i(\theta) + \tau_y$$
$$\times \sum_i \mu_i w_i(\tau_a, \tau_y, T) \int_{\Theta_i} \theta \ell(p_a(\tau_a, \tau_y, T)(1+\tau_a), w_i(\tau_a, \tau_y, T)(1-\tau_y)\theta, T) dK_i(\theta) - T - G \Big],$$

which differs from equation (31) because the function arguments of the indirect utility function and the consumption and labor supply decisions are different: see also footnotes 41 and 43. The first-order condition with respect to τ_a is, after collecting terms,

$$\frac{\partial \mathcal{L}}{\partial \tau_a} = \sum_i \mu_i \int_{\Theta_i} p_a \left(\alpha_i(\theta) \frac{\partial V_i(\theta)}{\partial p_a^*} + \lambda c_{a,i}(\theta) \right) dK_i(\theta) + \lambda \Bigg(\tau_a \sum_i \mu_i p_a \int_{\Theta_i} \frac{\partial c_{a,i}(\theta)}{\partial p_a^*} p_a dK_i(\theta)$$
$$+ \tau_y \sum_i \mu_i w_i \int_{\Theta_i} \theta \frac{\partial \ell_i(\theta)}{\partial p_a^*} p_a dK_i(\theta) \Bigg) + \sum_i \mu_i \frac{\partial p_a}{\partial \tau_a} (1+\tau_a) \int_{\Theta_i} \Bigg[\alpha_i(\theta) \frac{\partial V_i(\theta)}{\partial p_a^*} + \lambda \Bigg(\frac{\tau_a}{1+\tau_a} c_{a,i}(\theta)$$
$$+ \tau_a p_a \frac{\partial c_{a,i}(\theta)}{\partial p_a^*} + \tau_y w_i \theta \frac{\partial \ell_i(\theta)}{\partial p_a^*} \Bigg) \Bigg] dK_i(\theta) + \sum_i \mu_i \frac{\partial w_i}{\partial \tau_a} (1-\tau_y) \int_{\Theta_i} \theta \Bigg[\alpha_i(\theta) \frac{\partial V_i(\theta)}{\partial w_i^*}$$
$$+ \lambda \Bigg(\frac{\tau_y}{1-\tau_y} \ell_i(\theta) + \tau_a p_a \frac{\partial c_{a,i}(\theta)}{\partial w_i^*} + \tau_y w_i \theta \frac{\partial \ell_i(\theta)}{\partial w_i^*} \Bigg) \Bigg] dK_i(\theta) = 0. \tag{59}$$

In this equation, the term $\partial c_{a,i}(\theta)/\partial p_a^*$ shows up twice (and similarly for $\partial \ell_i(\theta)/\partial p_a^*$). The term on the first line captures the direct behavioral response due to a higher τ_a, whereas the term on the third line captures the indirect behavioral response driven by general equilibrium effects (hence, the multiplication with $\partial p_a/\partial \tau_a$). In equation (36) from Appendix A, these effects are combined in the term $\partial c_{a,i}^*(\theta)/\partial \tau_a$, cf. equation (53).

To simplify the above expression, substitute $\partial V_i(\theta)/\partial p_a^*$ and $\partial V_i(\theta)/\partial w_i^*$ from equations (50)–(51), divide the resulting expression by λ and use the definition of the welfare weight $g_i(\theta) = \alpha_i(\theta) u_{n,i}(\theta)/\lambda$. This gives

$$0 = \sum_i \mu_i p_a \int_{\Theta_i} (1 - g_i(\theta)) c_{a,i}(\theta) dK_i(\theta) \tag{60}$$
$$+ \tau_a \sum_i \mu_i p_a \int_{\Theta_i} \frac{\partial c_{a,i}(\theta)}{\partial p_a^*} p_a dK_i(\theta) + \tau_y \sum_i \mu_i w_i \int_{\Theta_i} \theta \frac{\partial \ell_i(\theta)}{\partial p_a^*} p_a dK_i(\theta)$$
$$+ \sum_i \mu_i \frac{\partial p_a}{\partial \tau_a} (1+\tau_a) \int_{\Theta_i} \Bigg[- g_i(\theta) c_{a,i}(\theta) + \frac{\tau_a}{1+\tau_a} c_{a,i}(\theta) + \tau_a p_a \frac{\partial c_{a,i}(\theta)}{\partial p_a^*} + \tau_y w_i \theta \frac{\partial \ell_i(\theta)}{\partial p_a^*} \Bigg] dK_i(\theta)$$
$$+ \sum_i \mu_i \frac{\partial w_i}{\partial \tau_a} (1-\tau_y) \int_{\Theta_i} \theta \Bigg[g_i(\theta) \ell_i(\theta) + \frac{\tau_y}{1-\tau_y} \ell_i(\theta) + \tau_a p_a \frac{\partial c_{a,i}(\theta)}{\partial w_i^*} + \tau_y w_i \theta \frac{\partial \ell_i(\theta)}{\partial w_i^*} \Bigg] dK_i(\theta).$$

[44] As mentioned before, Appendix B states the conditions which can be used to determine the impact of the tax instruments on equilibrium prices.

The first line captures the direct welfare effect of redistributing income from individuals to the government due to higher tax τ_a. The second line, in turn, captures the fiscal externalities associated with changes in the consumption of agricultural goods (first term) and labor supply (second term) following a rise in τ_a. Importantly, the behavioral responses only capture the *direct* impact due to a higher after-tax price of agricultural goods and not the indirect impact driven by general equilibrium effects. The third and fourth line capture the welfare impacts from general equilibrium effects on the price of agricultural goods and wages. Starting with the first, a change in p_a due to a higher τ_a affects individual purchasing power (the term proportional to $-g_i(\theta)$) and the tax base (the term proportional to $\tau_a c_{a,i}(\theta)$). Moreover, a change in the before-tax price p_a also induces (indirect) behavioral responses in agricultural consumption and labor supply, which generates fiscal externalities (the terms proportional to $\tau_a \partial c_{a,i}(\theta)/\partial p_a^*$ and $\tau_y \partial \ell_i(\theta)/\partial p_a^*$). Similarly, changes in equilibrium wages due to a higher τ_a affect purchasing power, the tax base and induces fiscal externalities from changes in agricultural consumption and labor supply. Equation (60) states that at the optimum, the sum of all these welfare effects should be zero.

We can rearrange equation (60) to obtain

$$-\frac{\partial p_a}{\partial \tau_a} \times \mathcal{W}_{p_a}^* - \sum_i \frac{\partial w_i}{\partial \tau_a} \times \mathcal{W}_{w_i}^* = \sum_i \mu_i p_a \int_{\Theta_i} (1 - g_i(\theta)) c_{a,i}(\theta) dK_i(\theta) \tag{61}$$

$$+ \tau_a \sum_i \mu_i p_a \int_{\Theta_i} \frac{\partial c_{a,i}(\theta)}{\partial p_a^*} p_a dK_i(\theta) + \tau_y \sum_i \mu_i w_i \int_{\Theta_i} \theta \frac{\partial \ell_i(\theta)}{\partial p_a^*} p_a dK_i(\theta),$$

where the terms

$$\mathcal{W}_{p_a}^* = \frac{\partial \mathcal{W}}{\partial p_a} \frac{1}{\lambda} = \tag{62}$$

$$\sum_i \mu_i (1 + \tau_a) \int_{\Theta_i} \left[- g_i(\theta) c_{a,i}(\theta) + \frac{\tau_a}{1 + \tau_a} c_{a,i}(\theta) + \tau_a p_a \frac{\partial c_{a,i}(\theta)}{\partial p_a^*} + \tau_y w_i \theta \frac{\partial \ell_i(\theta)}{\partial p_a^*} \right] dK_i(\theta),$$

$$\mathcal{W}_{w_i}^* = \frac{\partial \mathcal{W}}{\partial w_i} \frac{1}{\lambda} = \tag{63}$$

$$\mu_i (1 - \tau_y) \int_{\Theta_i} \theta \left[g_i(\theta) \ell_i(\theta) + \frac{\tau_y}{1 - \tau_y} \ell_i(\theta) + \tau_a p_a \frac{\partial c_{a,i}(\theta)}{\partial w_i^*} + \tau_y w_i \theta \frac{\partial \ell_i(\theta)}{\partial w_i^*} \right] dK_i(\theta)$$

measure the welfare effect driven by a change in p_a and w_i, respectively, scaled by the multiplier on the government budget constraint. These capture the welfare impact from changes in prices on purchasing power (first term), the tax base (second term) and fiscal externalities due to behavioral responses in agricultural consumption (third term) and labor supply (fourth term).

The (uncompensated) behavioral responses that show up on the right-hand side of equation

(61) can be split up using the Slutsky equations. This gives

$$-\frac{\partial p_a}{\partial \tau_a} \times \mathcal{W}^*_{p_a} - \sum_i \frac{\partial w_i}{\partial \tau_a} \times \mathcal{W}^*_{w_i} = \sum_i \mu_i p_a \int_{\Theta_i} (1 - g^*_i(\theta)) c_{a,i}(\theta) dK_i(\theta) \tag{64}$$

$$+ \tau_a \sum_i \mu_i p_a \int_{\Theta_i} \frac{\partial c^c_{a,i}(\theta)}{\partial p^*_a} p_a dK_i(\theta) + \tau_y \sum_i \mu_i w_i \int_{\Theta_i} \theta \frac{\partial \ell^c_i(\theta)}{\partial p^*_a} p_a dK_i(\theta),$$

where the Diamond (1975) based social welfare weight is defined as

$$g^*_i(\theta) = g_i(\theta) + \tau_a p_a \frac{\partial c_{a,i}(\theta)}{\partial T} + \tau_y w_i \theta \frac{\partial \ell_i(\theta)}{\partial T}. \tag{65}$$

In words, $g^*_i(\theta)$ measures the welfare impact of giving one unit of income to an individual of type (i, θ) by raising the lump-sum transfer, holding prices fixed. It captures both the direct utility benefit (first term) as well as the budgetary impact driven by income effects in agricultural consumption (second term) and labor supply (third term).

The first-order conditions for τ_y and T can be simplified in analogous fashion. The one for the labor income tax τ_y is given by

$$-\frac{\partial p_a}{\partial \tau_y} \times \mathcal{W}^*_{p_a} - \sum_i \frac{\partial w_i}{\partial \tau_y} \times \mathcal{W}^*_{w_i} = \sum_i \mu_i w_i \int_{\Theta_i} (1 - g^*_i(\theta)) \theta \ell_i(\theta) dK_i(\theta) \tag{66}$$

$$- \tau_a \sum_i \mu_i p_a \int_{\Theta_i} \frac{\partial c^c_{a,i}(\theta)}{\partial w^*_i} w_i \theta dK_i(\theta) - \tau_y \sum_i \mu_i w_i \int_{\Theta_i} \theta \frac{\partial \ell^c_i(\theta)}{\partial w^*_i} w_i \theta dK_i(\theta).$$

Moreover, the one for the lump-sum transfer T is

$$-\frac{\partial p_a}{\partial T} \times \mathcal{W}^*_{p_a} - \sum_i \frac{\partial w_i}{\partial T} \times \mathcal{W}^*_{w_i} = \sum_i \mu_i \int_{\Theta_i} (g^*_i(\theta) - 1) dK_i(\theta). \tag{67}$$

Because a change in T does not generate substitution effects, there are no compensated responses in equation (67). Combined, equations (64), (66) and (67) together with the government budget constraint pin down the optimal tax policy (τ_a, τ_y, T) and the multiplier λ on the government budget constraint.

To derive an expression for the optimal tax τ_a on agricultural goods using the optimal tax formulas (64), (66) and (67), we closely follow Jacobs and van der Ploeg (2019). The main differences with their framework is that (i) we do not have an environmental block, but (ii) prices in our model are endogenous (and show up in the optimal tax formulas through multiplication with $\mathcal{W}^*_{p_a}$ and $\mathcal{W}^*_{w_i}$). The idea is to use our preference specification to obtain expressions for the behavioral responses that show up in equations (64), (66) and (67). The resulting expressions can then be used to derive equation (23) from Proposition 1.

Consider again the utility maximization problem. From the first-order conditions (41) and (42), we can derive the following relationships for each individual (i, θ)

$$p^*_a c_{a,i}(\theta) = \eta(p^*_a) + \zeta(p^*_a) M_i(\theta), \tag{68}$$

$$C(c_{a,i}(\theta), c_{n,i}(\theta)) P(p^*_a) = M_i(\theta) - \varphi(p^*_a), \tag{69}$$

where $p_a^* = p_a(1+\tau_a)$ and the intercept $\eta(\cdot)$ and slope $\zeta(\cdot)$ of the Engel curve for agricultural consumption are given by

$$\eta(p_a^*) = \frac{p_a^* \underline{c}_a}{1 + \frac{\omega}{1-\omega}(p_a^*)^{1-\epsilon}}, \quad \zeta(p_a^*) = \frac{1}{1 + \frac{1-\omega}{\omega}(p_a^*)^{\epsilon-1}} \tag{70}$$

and $M_i(\theta)$ denotes disposable income, i.e.,

$$M_i(\theta) = T + w_i^* \ell_i(\theta) \tag{71}$$

with $w_i^* = w_i(1-\tau_y)\theta$. Moreover, the aggregate price index $P(\cdot)$ is

$$P(p_a^*) = (\omega(p_a^*)^{1-\epsilon} + (1-\omega))^{\frac{1}{1-\epsilon}} \tag{72}$$

and the function $\varphi(\cdot)$ captures subsistence spending:

$$\varphi(p_a^*) = p_a^* \underline{c}_a. \tag{73}$$

To derive the compensated responses that show up in the optimal tax formulas (64), (66) and (67), first take changes on both sides of the first-order condition for labor supply (40):

$$-\sigma \tilde{C}(c_{a,i}(\theta), c_{n,i}(\theta)) + \tilde{w}_i^* - \tilde{P}(p_a^*) = \phi \frac{\ell_i(\theta)}{1-\ell_i(\theta)} \tilde{\ell}_i(\theta), \tag{74}$$

where the tildes denote relative changes (i.e., $\tilde{C} = \mathrm{d}C/C$). Using the definition of $C(\cdot)$,

$$\tilde{C}(c_{a,i}(\theta), c_{n,i}(\theta)) = \kappa_a \tilde{c}_{a,i}(\theta) + \kappa_n \tilde{c}_{n,i}(\theta), \tag{75}$$

where the share parameters of the consumption aggregate are

$$\kappa_a = \frac{C_a(c_{a,i}(\theta), c_{n,i}(\theta)) c_{a,i}(\theta)}{C(c_{a,i}(\theta), c_{n,i}(\theta))} = C(c_{a,i}(\theta), c_{n,i}(\theta))^{\frac{1-\epsilon}{\epsilon}} \omega^{\frac{1}{\epsilon}} (c_{a,i}(\theta) - \underline{c}_a)^{-\frac{1}{\epsilon}} c_{a,i}(\theta), \tag{76}$$

$$\kappa_n = \frac{C_n(c_{a,i}(\theta), c_{n,i}(\theta)) c_{n,i}(\theta)}{C(c_{a,i}(\theta), c_{n,i}(\theta))} = C(c_{a,i}(\theta), c_{n,i}(\theta))^{\frac{1-\epsilon}{\epsilon}} (1-\omega)^{\frac{1}{\epsilon}} c_{n,i}(\theta)^{-\frac{1}{\epsilon}} c_{n,i}(\theta), \tag{77}$$

and the subscripts a and n refer to the partial derivatives with respect to the first and second argument, respectively. Moreover, using equation (72), we can write

$$\tilde{P}(p_a^*) = \zeta(p_a^*) \tilde{p}_a^*. \tag{78}$$

Next, take changes on both sides of the first-order condition (41) to get

$$\chi_a \tilde{c}_{a,i}(\theta) - \chi_n \tilde{c}_{n,i}(\theta) = \tilde{p}_a^*, \tag{79}$$

where the χ's are given by

$$\chi_a = \frac{C_{aa}(c_{a,i}(\theta), c_{n,i}(\theta)) c_{a,i}(\theta)}{C_a(c_{a,i}(\theta), c_{n,i}(\theta))} - \frac{C_{na}(c_{a,i}(\theta), c_{n,i}(\theta)) c_{a,i}(\theta)}{C_n(c_{a,i}(\theta), c_{n,i}(\theta))} = -\frac{1}{\epsilon} \frac{c_{a,i}(\theta)}{c_{a,i}(\theta) - \underline{c}_a}, \tag{80}$$

$$\chi_n = \frac{C_{nn}(c_{a,i}(\theta), c_{n,i}(\theta)) c_{n,i}(\theta)}{C_n(c_{a,i}(\theta), c_{n,i}(\theta))} - \frac{C_{an}(c_{a,i}(\theta), c_{n,i}(\theta)) c_{n,i}(\theta)}{C_a(c_{a,i}(\theta), c_{n,i}(\theta))} = -\frac{1}{\epsilon}. \tag{81}$$

The second step in both equations uses the definition of $C(\cdot)$ from equation (2). Lastly, set the change in the utility function equal to zero (recall: we derive compensated responses). This gives

$$C(c_{a,i}(\theta), c_{n,i}(\theta))^{1-\sigma}\tilde{C}(c_{a,i}(\theta), c_{n,i}(\theta)) - \psi(1-\ell_i(\theta))^{-\phi}\ell_i(\theta)\tilde{\ell}_i(\theta) = 0. \tag{82}$$

Combining the latter with the first-order condition (40) for labor supply:

$$\tilde{C}(c_{a,i}(\theta), c_{n,i}(\theta)) = \vartheta\tilde{\ell}_i(\theta), \quad \vartheta = \frac{w_i^*\ell_i(\theta)}{P(p_a^*)C(c_{a,i}(\theta), c_{n,i}(\theta))}. \tag{83}$$

We are then left with the following system of linear equations in the unknowns $\tilde{C}(\cdot)$, $\tilde{c}_{a,i}(\cdot)$, $\tilde{c}_{a,i}(\cdot)$ and $\tilde{l}_i(\cdot)$ as a function of price changes $\tilde{p}_a^*$ and $\tilde{w}_i^*$:

$$-\sigma\tilde{C}(c_{a,i}(\theta), c_{n,i}(\theta)) + \tilde{w}_i^* - \zeta(p_a^*)\tilde{p}_a^* = \phi\frac{\ell_i(\theta)}{1-\ell_i(\theta)}\tilde{\ell}_i(\theta), \tag{84}$$

$$\tilde{C}(c_{a,i}(\theta), c_{n,i}(\theta)) = \kappa_a\tilde{c}_{a,i}(\theta) + \kappa_n\tilde{c}_{n,i}(\theta), \tag{85}$$

$$\chi_a\tilde{c}_{a,i}(\theta) - \chi_n\tilde{c}_{n,i}(\theta) = \tilde{p}_a^*, \tag{86}$$

$$\tilde{C}(c_{a,i}(\theta), c_{n,i}(\theta)) = \vartheta\tilde{\ell}_i(\theta). \tag{87}$$

Solving for the relative changes:

$$\begin{aligned}\tilde{c}_{a,i}(\theta) = {} & \frac{\chi_n}{\kappa_a\chi_n+\kappa_n\chi_a}\frac{\vartheta}{\sigma\vartheta+\phi\ell_i(\theta)/(1-\ell_i(\theta))}\tilde{w}_i^* \\ & - \frac{\chi_n}{\kappa_a\chi_n+\kappa_n\chi_a}\left(\frac{\vartheta\zeta(p_a^*)}{\sigma\vartheta+\phi\ell_i(\theta)/(1-\ell_i(\theta))} - \frac{\kappa_n}{\chi_n}\right)\tilde{p}_a^*, \end{aligned} \tag{88}$$

$$\begin{aligned}\tilde{c}_{n,i}(\theta) = {} & \frac{\chi_a}{\kappa_a\chi_n+\kappa_n\chi_a}\frac{\vartheta}{\sigma\vartheta+\phi\ell_i(\theta)/(1-\ell_i(\theta))}\tilde{w}_i^* \\ & - \frac{\chi_a}{\kappa_a\chi_n+\kappa_n\chi_a}\left(\frac{\vartheta\zeta(p_a^*)}{\sigma\vartheta+\phi\ell_i(\theta)/(1-\ell_i(\theta))} + \frac{\kappa_a}{\chi_a}\right)\tilde{p}_a^*, \end{aligned} \tag{89}$$

$$\tilde{\ell}_i(\theta) = \frac{1}{\sigma\vartheta+\phi\ell_i(\theta)/(1-\ell_i(\theta))}\tilde{w}_i^* - \frac{\zeta(p_a^*)}{\sigma\vartheta+\phi\ell_i(\theta)/(1-\ell_i(\theta))}\tilde{p}_a^*, \tag{90}$$

$$\tilde{C}(c_{a,i}(\theta), c_{n,i}(\theta)) = \frac{\vartheta}{\sigma\vartheta+\phi\ell_i(\theta)/(1-\ell_i(\theta))}\tilde{w}_i^* - \frac{\vartheta\zeta(p_a^*)}{\sigma\vartheta+\phi\ell_i(\theta)/(1-\ell_i(\theta))}\tilde{p}_a^*. \tag{91}$$

Here, the terms that multiply $\tilde{w}_i^*$ and $\tilde{p}_a^*$ are the compensated elasticities with respect to w_i^* and p_a^*, respectively.

To proceed, consider again equation (64). Use equation (68) to substitute out for $p_a c_{a,i}(\theta)$. This gives

$$\begin{aligned} & -\frac{\partial p_a}{\partial\tau_a}\times\mathcal{W}_{p_a}^* - \sum_i\frac{\partial w_i}{\partial\tau_a}\times\mathcal{W}_{w_i}^* = \sum_i\mu_i\int_{\Theta_i}(1-g_i^*(\theta))\left[\frac{\eta(p_a^*)}{1+\tau_a} + \frac{\zeta(p_a^*)}{1+\tau_a}M_i(\theta)\right]dK_i(\theta) \\ & + \tau_a\sum_i\mu_i p_a\int_{\Theta_i}\frac{\partial c_{a,i}^c(\theta)}{\partial p_a^*}p_a dK_i(\theta) + \tau_y\sum_i\mu_i w_i\int_{\Theta_i}\theta\frac{\partial\ell_i^c(\theta)}{\partial p_a^*}p_a dK_i(\theta). \end{aligned} \tag{92}$$

Substituting out for $M_i(\theta)$ using equation (71) and collecting terms gives

$$
\begin{aligned}
&-\frac{\partial p_a}{\partial \tau_a}\times \mathcal{W}^*_{p_a} - \sum_i \frac{\partial w_i}{\partial \tau_a}\times \mathcal{W}^*_{w_i} = \left(\frac{\eta(p_a^*)}{1+\tau_a} + \frac{\zeta(p_a^*)T}{1+\tau_a}\right)\sum_i \mu_i \int_{\Theta_i}(1-g_i^*(\theta))dK_i(\theta) \\
&+\sum_i \mu_i \int_{\Theta_i}(1-g_i^*(\theta))\frac{\zeta(p_a^*)w_i(1-\tau_y)}{1+\tau_a}\theta\ell_i(\theta)dK_i(\theta) \\
&+\tau_a \sum_i \mu_i p_a \int_{\Theta_i}\frac{\partial c^c_{a,i}(\theta)}{\partial p_a^*}p_a dK_i(\theta) + \tau_y \sum_i \mu_i w_i \int_{\Theta_i}\theta \frac{\partial \ell_i^c(\theta)}{\partial p_a^*}p_a dK_i(\theta). \qquad (93)
\end{aligned}
$$

Use the optimal tax formula (67) to substitute out for the first term on the right-hand side:

$$
\begin{aligned}
&-\frac{\partial p_a}{\partial \tau_a}\times \mathcal{W}^*_{p_a} - \sum_i \frac{\partial w_i}{\partial \tau_a}\times \mathcal{W}^*_{w_i} - \left(\frac{\eta(p_a^*)}{1+\tau_a} + \frac{\zeta(p_a^*)T}{1+\tau_a}\right)\left[\frac{\partial p_a}{\partial T}\times \mathcal{W}^*_{p_a} + \sum_i \frac{\partial w_i}{\partial T}\times \mathcal{W}^*_{w_i}\right] \\
=&\sum_i \mu_i \int_{\Theta_i}(1-g_i^*(\theta))\frac{\zeta(p_a^*)w_i(1-\tau_y)}{1+\tau_a}\theta\ell_i(\theta)dK_i(\theta) \\
&+\tau_a \sum_i \mu_i p_a \int_{\Theta_i}\frac{\partial c^c_{a,i}(\theta)}{\partial p_a^*}p_a dK_i(\theta) + \tau_y \sum_i \mu_i w_i \int_{\Theta_i}\theta \frac{\partial \ell_i^c(\theta)}{\partial p_a^*}p_a dK_i(\theta). \qquad (94)
\end{aligned}
$$

Next, multiply and divide the right-hand side by $(1-\tau_y)\zeta(p_a^*)/(1+\tau_a)$ and collect terms to get

$$
\begin{aligned}
&-\frac{\partial p_a}{\partial \tau_a}\mathcal{W}^*_{p_a} - \sum_i \frac{\partial w_i}{\partial \tau_a}\mathcal{W}^*_{w_i} - \left(\frac{\eta(p_a^*)}{1+\tau_a} + \frac{\zeta(p_a^*)T}{1+\tau_a}\right)\left[\frac{\partial p_a}{\partial T}\mathcal{W}^*_{p_a} + \sum_i \frac{\partial w_i}{\partial T}\mathcal{W}^*_{w_i}\right] = \qquad (95) \\
&\frac{(1-\tau_y)\zeta(p_a^*)}{1+\tau_a}\sum_i \mu_i w_i \int_{\Theta_i}\left[1-g_i^*(\theta) + \frac{\tau_a}{1+\tau_a}\frac{\varepsilon^c_{ap}(\theta)}{\zeta(p_a^*)}\beta_{a,i}(\theta) + \frac{\tau_y}{1-\tau_y}\frac{\varepsilon^c_{\ell p}(\theta)}{\zeta(p_a^*)}\right]\theta\ell_i(\theta)dK_i(\theta),
\end{aligned}
$$

where $\varepsilon^c_{ap}(\theta)$ and $\varepsilon^c_{\ell p}(\theta)$ are compensated elasticities of agricultural consumption and labor supply with respect to p_a^* (see equations (88) and (90)):

$$
\varepsilon^c_{ap}(\theta) = \frac{\partial c^c_{a,i}(\theta)}{\partial p_a^*}\frac{p_a^*}{c_{a,i}(\theta)}, \quad \varepsilon^c_{\ell p}(\theta) = \frac{\partial \ell^c_i(\theta)}{\partial p_a^*}\frac{p_a^*}{\ell_i(\theta)} \qquad (96)
$$

and $\beta_{a,i}(\theta)$ captures the spending on agricultural goods as a fraction of after-tax labor income (excluding the lump-sum transfer):

$$
\beta_{a,i}(\theta) = \frac{p_a(1+\tau_a)c_{a,i}(\theta)}{w_i(1-\tau_y)\theta\ell_i(\theta)}. \qquad (97)
$$

To proceed, write the optimal tax formula (66) for τ_y as

$$
\begin{aligned}
&-\frac{\partial p_a}{\partial \tau_y}\mathcal{W}^*_{p_a} - \sum_i \frac{\partial w_i}{\partial \tau_y}\mathcal{W}^*_{w_i} = \qquad (98) \\
&\sum_i \mu_i w_i \int_{\Theta_i}\left[1-g_i^*(\theta) - \frac{\tau_a}{1+\tau_a}\varepsilon^c_{aw}(\theta)\beta_{a,i}(\theta) - \frac{\tau_y}{1-\tau_y}\varepsilon^c_{\ell w}(\theta)\right]\theta\ell_i(\theta)dK_i(\theta),
\end{aligned}
$$

where $\varepsilon^c_{aw}(\theta)$ and $\varepsilon^c_{\ell w}(\theta)$ are compensated elasticities of agricultural consumption and labor supply with respect to w_i^* (see equations (88) and (90)):

$$\varepsilon^c_{aw}(\theta) = \frac{\partial c^c_{a,i}(\theta)}{\partial w_i^*}\frac{w_i^*}{c_{a,i}(\theta)}, \quad \varepsilon^c_{\ell w}(\theta) = \frac{\partial \ell^c_i(\theta)}{\partial w_i^*}\frac{w_i^*}{\ell_i(\theta)}. \tag{99}$$

Combine equation (95) and (98) to obtain

$$\begin{aligned}
&\mathcal{W}^*_{p_a}\left[\frac{\partial p_a}{\partial \tau_y} - \frac{1+\tau_a}{(1-\tau_y)\zeta(p_a^*)}\frac{\partial p_a}{\partial \tau_a} - \frac{1+\tau_a}{(1-\tau_y)\zeta(p_a^*)}\left(\frac{\eta(p_a^*)}{1+\tau_a} + \frac{\zeta(p_a^*)T}{1+\tau_a}\right)\frac{\partial p_a}{\partial T}\right] \\
&+ \sum_i \mathcal{W}^*_{w_i}\left[\frac{\partial w_i}{\partial \tau_y} - \frac{1+\tau_a}{(1-\tau_y)\zeta(p_a^*)}\frac{\partial w_i}{\partial \tau_a} - \frac{1+\tau_a}{(1-\tau_y)\zeta(p_a^*)}\left(\frac{\eta(p_a^*)}{1+\tau_a} + \frac{\zeta(p_a^*)T}{1+\tau_a}\right)\frac{\partial w_i}{\partial T}\right] = \\
&\sum_i \mu_i w_i \int_{\Theta_i}\left[\frac{\tau_a}{1+\tau_a}\left(\frac{\varepsilon^c_{ap}(\theta)}{\zeta(p_a^*)} + \varepsilon^c_{aw}(\theta)\right)\beta_{a,i}(\theta) + \frac{\tau_y}{1-\tau_y}\left(\frac{\varepsilon^c_{\ell p}(\theta)}{\zeta(p_a^*)} + \varepsilon^c_{\ell w}(\theta)\right)\right]\theta\ell_i(\theta)dK_i(\theta).
\end{aligned} \tag{100}$$

From equation (90), it follows that

$$\frac{\varepsilon^c_{\ell p}(\theta)}{\zeta(p_a^*)} + \varepsilon^c_{\ell w}(\theta) = 0. \tag{101}$$

Substituting this in equation (100):

$$\begin{aligned}
&\mathcal{W}^*_{p_a}\left[\frac{\partial p_a}{\partial \tau_y} - \frac{1+\tau_a}{(1-\tau_y)\zeta(p_a^*)}\frac{\partial p_a}{\partial \tau_a} - \frac{1+\tau_a}{(1-\tau_y)\zeta(p_a^*)}\left(\frac{\eta(p_a^*)}{1+\tau_a} + \frac{\zeta(p_a^*)T}{1+\tau_a}\right)\frac{\partial p_a}{\partial T}\right] \\
&+ \sum_i \mathcal{W}^*_{w_i}\left[\frac{\partial w_i}{\partial \tau_y} - \frac{1+\tau_a}{(1-\tau_y)\zeta(p_a^*)}\frac{\partial w_i}{\partial \tau_a} - \frac{1+\tau_a}{(1-\tau_y)\zeta(p_a^*)}\left(\frac{\eta(p_a^*)}{1+\tau_a} + \frac{\zeta(p_a^*)T}{1+\tau_a}\right)\frac{\partial w_i}{\partial T}\right] = \\
&\frac{\tau_a}{1+\tau_a}\sum_i \mu_i w_i \int_{\Theta_i}\left(\frac{\varepsilon^c_{ap}(\theta)}{\zeta(p_a^*)} + \varepsilon^c_{aw}(\theta)\right)\beta_{a,i}(\theta)\theta\ell_i(\theta)dK_i(\theta).
\end{aligned} \tag{102}$$

Next, multiply both sides by $-(1-\tau_y)\zeta(p_a^*)/(1+\tau_a)$ and use the equation (19). Rearranging gives

$$\begin{aligned}
&\frac{\tau_a}{1+\tau_a}\sum_i \mu_i w_i \int_{\Theta_i}(-\varepsilon^c_{ap}(\theta) - \zeta(p_a^*)\varepsilon^c_{aw}(\theta))\frac{\beta_{a,i}(\theta)(1-\tau_y)}{1+\tau_a}\theta\ell_i(\theta)dK_i(\theta) \\
&= \mathcal{W}^*_{p_a}\left[\frac{\partial p_a}{\partial \tau_a} + \frac{\partial p_a}{\partial \tau_y}\frac{\mathrm{d}\tau_y}{\mathrm{d}\tau_a} + \frac{\partial p_a}{\partial T}\frac{\mathrm{d}T}{\mathrm{d}\tau_a}\right] + \sum_i \mathcal{W}^*_{w_i}\left[\frac{\partial w_i}{\partial \tau_a} + \frac{\partial w_i}{\partial \tau_y}\frac{\mathrm{d}\tau_y}{\mathrm{d}\tau_a} + \frac{\partial w_i}{\partial T}\frac{\mathrm{d}T}{\mathrm{d}\tau_a}\right].
\end{aligned} \tag{103}$$

Using equation (97), we can simplify

$$\frac{\beta_{a,i}(\theta)(1-\tau_y)}{1+\tau_a} = \frac{p_a c_{a,i}(\theta)}{w_i\theta\ell_i(\theta)}. \tag{104}$$

Substituting this in equation (103) and rearranging gives

$$\begin{aligned}
&\frac{\tau_a}{1+\tau_a}\sum_i \mu_i \int_{\Theta_i} p_a c_{a,i}(\theta)(-\varepsilon^c_{ap}(\theta) - \zeta(p_a^*)\varepsilon^c_{aw}(\theta))dK_i(\theta) = \\
&\mathcal{W}^*_{p_a}\left[\frac{\partial p_a}{\partial \tau_a} + \frac{\partial p_a}{\partial \tau_y}\frac{\mathrm{d}\tau_y}{\mathrm{d}\tau_a} + \frac{\partial p_a}{\partial T}\frac{\mathrm{d}T}{\mathrm{d}\tau_a}\right] + \sum_i \mathcal{W}^*_{w_i}\left[\frac{\partial w_i}{\partial \tau_a} + \frac{\partial w_i}{\partial \tau_y}\frac{\mathrm{d}\tau_y}{\mathrm{d}\tau_a} + \frac{\partial w_i}{\partial T}\frac{\mathrm{d}T}{\mathrm{d}\tau_a}\right],
\end{aligned} \tag{105}$$

which coincides with equation (23) from Proposition 1. By combining the first-order conditions for τ_a, τ_y and T in this specific way, the final equation gives the optimality condition obtained from equating to zero the welfare impact of the tax reform R from Definition 2.

D Derivation of equation (28)

Because firms in the agricultural and non-agricultural sector make zero profits, the following conditions must hold:

$$p_a F_a(L_a, H_a) = w_L L_a + w_H H_a, \tag{106}$$

$$F_n(L_n, H_n) = w_L L_n + w_H H_n, \tag{107}$$

where we used the normalization $p_n = 1$. Taking changes on both sides of equation (107):

$$F_{L,n}(L_n, H_n)\mathrm{d}L_n + F_{H,n}(L_n, H_n)\mathrm{d}H_n = w_L \mathrm{d}L_n + w_H \mathrm{d}H_n + L_n \mathrm{d}w_L + H_n \mathrm{d}w_H. \tag{108}$$

From equations (9)–(10), the terms on the left-hand side cancel against the first two terms on the right-hand side. Rearranging leads to the second result from equation (28):

$$\mathrm{d}w_H = -\left(\frac{L_n}{H_n}\right)\mathrm{d}w_L. \tag{109}$$

To derive the first result, take changes on both sides of equation (106):

$$\begin{aligned} &F_a(L_a, H_a)\mathrm{d}p_a + p_a F_{L,a}(L_a, H_a)\mathrm{d}L_a + p_a F_{H,a}(L_a, H_a)\mathrm{d}H_a \\ &= w_L \mathrm{d}L_a + w_H \mathrm{d}H_a + L_a \mathrm{d}w_L + H_a \mathrm{d}w_H. \end{aligned} \tag{110}$$

Equations (9)–(10) imply that the final two terms on the left-hand side cancel against the first two terms on the right-hand side. Using equation (109) to substitute out for $\mathrm{d}w_H$ and rearranging leads to

$$\mathrm{d}w_L \left(\frac{L_a}{H_a} - \frac{L_n}{H_n}\right) = \frac{F_a(L_a, H_a)}{H_a}\mathrm{d}p_a, \tag{111}$$

which coincides with the first result from equation (28), upon substituting $Y_a = F_a(L_a, H_a)$.